A familiarization protocol: Repetitions in reserve

Elvert A. Wise
University of Northern Iowa

A Familiarization Protocol: Repetitions in Reserve

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of Master of Arts

Elvert A. Wise

University of Northern Iowa

December 2022

Abstract

Research suggests RIR is a valid and reliable method to prescribe and manipulate RT intensities. However, the literature reviewed revealed that a protocol had not been developed or introduced that allows practitioners to educate athletes or recreational populations to implement the RIR scale. **Purpose:** Thus, this study aimed to develop a familiarization protocol through synthesized research-study findings and evaluate its efficacy in resistance-trained populations.

Methods: Participants were males (n = 9) with > 3 years of RT experience. The study consisted of 5 sessions. Participants performed the bench press and v-squat for four sets each at varying intensities. The PG followed the familiarization protocol, and the NPG was only provided the prescribed intensities and assistance to ensure safety. eeRIR, over and underestimations, and the predicted probability of being accurate were compared across groups to evaluate protocol efficacy. **Results:** Data analyses revealed the FP was effective in increasing RIR estimation accuracy for the PG. eeRIR was significantly less for the PG (BP = 0.144 repetitions, VS = 0.112 repetitions) when compared to the NPG (BP = 0.953 repetitions, VS = 1.472 repetitions). Additionally, the predicted probability of accurately estimating RIR was significantly higher for the PG (BP = 85.8%, VS = 89.7%) than the NPG (BP = 28.2%, VS = 12.3%) – suggesting the familiarization protocol was effective in increasing RIR estimation accuracy.

This Study by: Elvert A. Wise

Entitled: A Familiarization Protocol: Repetitions in Reserve

has been approved as meeting the thesis requirements for the

Degree of Master of Arts

Date	Dr. Jacob Reed, Chair, Thesis Committee
Date	Dr. Fabio Fontana, Thesis Committee Member
Date	Dr. Mick Mack, Thesis Committee Member
Date	Dr. Gabriela Olivares, Interim Dean, Graduate College

Dedication

I want to dedicate this thesis to my parents. To my dad – although you passed one month before I started graduate school – I knew you were proud of me and believed in me more than anybody. Losing you made graduate school challenging, and even though you weren't here in person, I knew you were with me the entire time. To my mom, thank you so much for being the best mom ever! Thanks for supporting me through everything and never giving up on me! You've always encouraged me to chase my dreams, even when you didn't understand them. Most importantly, I want to thank both of you for loving me through it all.

Acknowledgements

My research project would not have been possible without the help of my committee chair, Dr. Jacob Reed. Thank you for believing in this research project from day one. From my undergraduate career to my thesis, you've encouraged me from the beginning. You've been a great help throughout this project and my academic career – and I am genuinely grateful for that. Your "but why?" mentality has inspired me to continue a career in academia. Also, I owe a huge thank you to Dr. Fabio Fontana (committee member and professor). You have expected nothing but the best, which has motivated me to be the best version of myself. I appreciate your guidance, support, and advice throughout my time at UNI.

Dr. Mick Mack (committee member and professor), you never hesitated to help whenever I asked, and I cannot thank you enough. Your attention to detail has taught me that the small details are just as important as the big picture. Also, I want to thank Dr. Justine Radunzel for lending her invaluable expertise in analyzing the data. Lastly, a special thank you to Dr. Wilfred "Mickye" Johnson, you've pushed me to be the best version of myself since high school, and I honestly wouldn't be where I am today without your mentorship. Your actions taught me never to give up, regardless of the circumstance.

Table of Contents

List of Tables **ix**

List of Figures **x**

List of Charts **xi**

Chapter 1: Introduction **1**

Purpose Statement 4

Research Hypothesis 4

Significance of the Study 4

Assumptions 7

Limitations 7

Delimitations 7

Operational Definitions 8

Chapter 2: Literature Review **11**

Resistance Training 11

Load 12

Load Prescription 12

Velocity: Objective Measures 14

Subjective Measures: RPE, RIR-based RPE, and RIR 16

RPE 16

RPE (CR10) & Velocity 17

RPE (CR10) Scoring 18

RPE (CR10) & Resistance Training Experience 18

RIR-Based RPE 20

RIR 21

aRIR and RIR Estimation Correlations 22

RIR Estimation Accuracy: Upper body versus Lower body 23

RIR Estimation Accuracy and RT Experience 24

RIR Estimation Accuracy: Across-Sets 25

RIR Estimation Accuracy and Proximity to MF 26

RIR Estimation Accuracy: Across-Sessions 26

Protocol Development 28

Category A: Participant History 28

Category B: Protocol Development 30

Chapter 3: Methods **34**

Participants 34

Instrumentation 34

Procedures 34

Session 1: The Initial Assessment 36

Session 2-4: Non-Protocol Group 36

Session 2-4: Protocol Group 37

Session 5: The Final Assessment 40

Data Analysis 40

Statistical Results 42

First set of analyses – Group comparisons, including Session 1 and Session 5 data only 42

Second set of analyses – Session comparisons for the Protocol Group 46

Third set of analyses – Set comparisons within session for the Protocol Group 49

Discussion 50

PG vs. NPG (absolute mean eeRIR) 50

PG vs. NPG (Probability of being accurate) 51

PG vs. NPG (Over and Underestimations) 52

PG Absolute mean eeRIR (pRIR – aRIR) 52

PG: Across set (eeRIR) 54

Limitations: 56

Future Research 58

Practical Applications: 58

References 60

Appendix A: Vocabulary Handout 69

Appendix B: The RIR scale, Handout 70

List of Tables

Table 1. Protocol Group (PG): Session Training Intensities 71

Table 2. Non-Protocol Group (NPG): Session Training Intensities 72

Table 3. Pairwise Session Comparisons of Absolute Difference of pRIR - aRIR for the Protocol Group 73

Table 4. Pairwise Session Comparisons of Probability of Estimating RIR for the Protocol Group 75

Table 5: Model Estimates for Absolute Difference of pRIR - aRIR 77

Table 6: Model Estimates for Probability of Being Accurate in Estimating RIR (i.e., pRIR = aRIR) 78

Table 7. Model Estimates for Absolute Difference of pRIR – aRIR by Session for Protocol Group 79

Table 8. Model Estimates for Probability of Being Accurate (pRIR = aRIR) by Session for Protocol Group 80

List of Figures

Figure 1. Mean Absolute Difference of pRIR – aRIR by Group, Session, and Movement 81

Figure 2. Mean Difference of pRIR – aRIR by Group, Session, and Movement 82

Figure 3. Probability of Being Accurate (pRIR = aRIR) by Group, Session, and Movement 83

Figure 4. Mean Absolute Difference of pRIR – aRIR by Session and Movement for the Protocol Group 84

Figure 5. Probability of Being Accurate (pRIR = aRIR) by Session and Movement for the Protocol Group 85

Figure 6. Mean Absolute Difference of pRIR – aRIR by Session and Set for the Protocol Group (Bench press) 86

Figure 7. Mean Absolute Difference of pRIR – aRIR by Session and Set for the Protocol Group (V-Squat) 87

Figure 8. Mean Difference of pRIR – aRIR by Session and Set for the Protocol Group (Bench press) 88

Figure 9. Mean Difference of pRIR – aRIR by Session and Set for the Protocol Group (V-Squat) 89

List of Charts

Chart 1: Mean Absolute Difference of pRIR - aRIR by Movement and Session for Protocol Group 90

Chart 2: Probability of Being Accurate (i.e., pRIR = aRIR) by Movement and Session for Protocol Group 91

Chapter 1: Introduction

Traditionally, resistance training (RT) has been implemented to improve one or more of the following: muscular strength, power, local muscular endurance, and/or muscular hypertrophy (American College of Sports Medicine [ACSM], 2009). The desired outcome of an RT program should dictate program design. An effective RT program is dependent on the proper periodization of acute RT variables (Deschenes & Kraemer, 2002). Acute RT variables are exercise intensity (absolute and relative load), repetitions, repetition velocity with submaximal loads, rest periods, training volume (repetitions x sets), and frequency (ACSM, 2009). The systematic manipulation or periodization of acute RT variables should be based on the following RT principles: progressive overload, specificity, variation, and individualization (ACSM, 2002; 2009).

Prescribing intensity is a crucial variable when designing an RT program. Percent-based training (PBT), repetition maximum (RM), and maximum repetition zones (RM-zones) are frequently used when prescribing RT intensities. However, the methods mentioned above have their fair share of limitations. Thus, the utilization of subjective methods, such as ratings of perceived exertion (RPE), repetitions in reserve-based rating of perceived exertion (RIR-based RPE), and repetitions in reserve (RIR), have increased in popularity. Previous research has compared the subjective scores of RPE, RIR-based RPE, and RIR to the objective measures of the mean concentric velocities (MCV) of various lifts in an effort to validate their use as reliable methods to prescribe and manipulate RT intensities (Balsalobre-Fernandez et al., 2018; Helms, Storey, et al., 2017; Ormsbee et al., 2017).

RPE, RIR-based RPE, and RIR were found to have a strong inverse relationship with MCV (Balsalobre-Fernandez et al., 2018; Helms, Storey et al., 2017; Ormsbee et al., 2017; Zourdos et al., 2015). Findings suggest that RIR-based RPE and RIR are valid and reliable methods to prescribe and manipulate RT intensities. However, due to the subjective nature of the RIR scale, the effectiveness and benefit of utilizing RIR largely depends on the user's ability to accurately assess training intensities (Helms et al., 2016).

Although the RIR scale seems to be a reliable method to regulate acute RT variables, to my knowledge – there does not seem to be a protocol in existence that educates one to implement the RIR scale into a RT program. Thus, a review of the current literature was conducted to develop a familiarization protocol (FP) that allows practitioners to educate athletes to utilize the RIR scale accurately. The review of study findings that evaluated participant accuracy in identifying RIR or estimating repetitions to failure (muscular, momentary, and technical), allowed research findings and the variables that influenced RIR estimation accuracy to be identified and synthesized to construct this familiarization protocol.

It was found that the load lifted (Cavarretta et al., 2019), proximity to muscle failure (Hackett et al., 2016; 2019; Mansfield et al., 2020; Zourdos et al., 2019), training experience (Helms, Storey et al., 2017; Ormsbee et al., 2017; Steele et al., 2017; Zourdos et al., 2015), sex (Hackett et al., 2016; Hackett et al., 2018), prior experience training to failure or with heavy loads (Helms, Storey et al., 2017), muscle group(s) used (Cavarretta et al., 2019), the ambiguity of associated vocabulary (Halperin & Emanuel, 2019), and the level of familiarization with the RIR scale all affected the accuracy of RIR estimations.

Furthermore, RIR estimation accuracy was found to improve intrasession, across sets – in that RIR estimation accuracy increased as participants progressed through sets within a session (Hackett et al., 2012; 2018; Mansfield et al., 2020). Previous studies also found that accuracy estimating RIR improved across sessions, suggesting that as one becomes more familiar with utilizing the RIR scale, their ability to accurately estimate RIR increases (Hackett et al., 2018; Helms, Brown, et al., 2017; Mansfield et al., 2020).

Considering the purpose of this study is to develop a protocol that teaches one to implement RIR – heightened attention was given to the method sections of the literature reviewed. Familiarization protocols, or the lack thereof, helped direct the development of this familiarization protocol. Consistent procedures across studies were noted. Additionally, variations in study procedures were evaluated to determine if the differences in procedures may have affected the accuracy of RIR estimations. A collective consideration of the variables mentioned above ultimately directed the design of this familiarization protocol.

Changes in accuracy observed across sets influenced the number of sets prescribed per session, inherently allowing the session's length to be partially defined. Whether accuracy improved across sessions helped determine the number of sessions, therefore, the duration of the study protocol. Some of the findings mentioned above may be manipulated to create an optimal setting that lends itself to increasing the accuracy of RIR estimations. For instance, prescribed intensities (pRIR), muscle group(s) used, and participant selection.

Purpose Statement

1. This study aims to develop a protocol that educates one to effectively implement the Repetitions in Reserve scale (RIR).
2. An additional purpose of this study is to evaluate the efficacy of the developed familiarization protocol in improving one's accuracy in estimating RIR.
3. Thirdly, this study aims to identify the number of sessions and sets per session necessary to elicit a significant increase in the ability to identify the number of repetitions in reserve accurately.

Research Hypothesis

1. Utilizing a familiarization protocol will increase one's accuracy in estimating RIR.
2. Memory anchoring – the ability to rely on acute memory anchors, thus, recalling past training experiences, will improve RIR estimation accuracy.
3. Increased exposure to the RIR scale – will increase accuracy in estimating RIR.

Significance of the Study

Research findings support the use of the RIR scale as a valid and reliable method to prescribe, assess, and manipulate RT intensities (Hackett et al., 2012; Lovegrove et al., 2021; Mangine et al., 2022). The RIR scale may be implemented into RT programs to maximize muscular strength and hypertrophy – with relatively heavy to heavy loads (Helms et al., 2016). Furthermore, the subjective nature coupled with the immediate feedback provided by the RIR scale may allow coaches and athletes to autoregulate (AR) acute RT variables (Hackett et al., 2012; Mann et al., 2010; Mansfield et al., 2020). Precisely, autoregulatory progressive resistance exercise (APRE) allows one to manipulate RT variables based on current performance and readiness to train (Mann et al., 2010), which increases in importance as one accumulates greater training experience (ACSM, 2009; Greig et al., 2020; Helms et al., 2020). Autoregulation also allows one to account for factors such as sleep, nutrition, and life stressors – which inherently affects one's readiness to train (Helms et al., 2016). Considering the capacity of the RIR scale to autoregulate RT programs – the RIR scale may provide the end user with the benefits of AR.

Being able to identify a decrease or increase in performance as it occurs allows the coach or athlete to adjust acute training variables with greater accuracy. For instance, observing a reduction in the number of repetitions completed to reach the desired number of prescribed RIR (pRIR) while training at similar intensities as previous training sessions may indicate a need to adjust acute training variables. Intrasession adjustments may include decreasing intensity (load) or increasing the duration of rest intervals to maintain the desired number of repetitions and pRIR. Additionally, the RIR scale allows one to adjust acute training variables intrasession due to performance increases (i.e., strength gains). In that, a coach or athlete may observe an increase in performance through an increase of repetitions necessary to achieve the desired number of pRIR, even though training intensities may be similar to previous training sessions (Hackett et al., 2012; Mansfield et al., 2020). This observation may indicate a need to increase load or decrease rest interval(s) duration to maintain the desired training intensity or pRIR. The increased capacity to accurately adjust acute training variables will enable one to maximize training efforts, thus, neuromuscular and physiological adaptions.

The RIR scale was also suggested to account for inter-individual differences when prescribing RT intensities in team settings (Mansfield et al., 2020). Without the ability to account for inter-individual differences between athletes in team sports – some athletes may train at too high of an intensity, increasing the risks of injury and unnecessary fatigue. Contrary, training intensities may be too low for some, thus, failing to elicit the desired improvements in performance. However, utilizing the RIR scale to autoregulate RT variables may enable coaches and athletes to better manage the stimulus-to-fatigue ratio and account for improvements or decrements in performance more frequently (Hackett et al., 2012; Helms et al., 2016; Mangine et al., 2022) – permitting intensities to be prescribed with greater precision across athletes. In that, manipulating acute training variables through AR should allow athletes to train closer to their threshold more frequently (Graham & Cleather, 2019).

Furthermore, the autoregulation of acute training variables while using the RIR scale produced greater strength increases compared to traditional methods (i.e., block periodization) (Graham & Cleather, 2019). The findings of Graham and Cleather (2019) seem to be consistent with Mann et al. (2010) and Zhang et al. (2021) – in that the AR of RT variables produced superior strength gains. AR was also found to yield greater increases in muscle mass versus fixed exercise selections (Rauch et al., 2017).

Although research supports the use of the RIR scale to prescribe, assess, and manipulate RT intensities, to my knowledge, there has not been a readily available protocol designed to familiarize or teach one to implement and utilize the RIR scale. Hence, this research aims to develop and evaluate the efficacy of this developed familiarization protocol.

Assumptions

1. Increased resistance training experience enhances one's ability to accurately identify training intensities with subjective scales, such as the RIR scale.
2. Sex influences the ability to accurately determine or estimate RIR, in that males were found to estimate RIR more accurately than females.
3. The ability to accurately estimate RIR is affected by the muscle group(s) used during the lift (upper body versus lower body).
4. The proximity to muscle or technical failure affects RIR estimation accuracy; the closer one is to failure, the greater the accuracy of estimating RIR.

Limitations

1. No objective measures will be utilized to evaluate the accuracy of RIR estimations (i.e., the velocity of lifts; linear position transducer).
2. Subjective nature of repetitions in reserve – solely relying on the participant's description of intensities.
3. Technical failure vs. absolute failure – considering technical failure will be used, identifying the number of RIR may not truly reflect the actual repetitions in reserve (aRIR).

Delimitations

1. All participants will have a high level of resistance training experience. Particularly increased familiarity training with heavy loads and near muscle failure ($\geq$ three years of RT experience).

2. Exercise selection(s): v-squat and barbell bench press. The v-squat and barbell bench press will be used in an effort to account for potential differences between upper and lower body exercises and accuracy estimating RIR.
3. Considering RIR was reported with greater accuracy when training within 0-3 RIR, exercise intensities will be prescribed between 0 to 3 RIR (pRIR).
4. All participants will be males. An all-male pool will eliminate potential differences due to sex.

Operational Definitions

1. Technical failure: observed when the performance of the lift can no longer be completed with proper technique; another repetition would result in incorrect form or technique.
2. (Absolute) muscle failure: complete muscle failure during a lift, in which the muscle(s) is fatigued beyond the point of recovery. Another repetition cannot be completed.
3. Repetitions in reserve (RIR): the ratio of repetitions completed in relation to the maximum number of repetitions that "can" be completed. For example, selecting a load that is consistent with a 10RM, but the set is stopped short at the eighth repetition, the participants may have completed two more repetitions at this particular load. Hence, in this scenario, two repetitions were left in reserve.
4. Prescribed Repetitions in Reserve (pRIR): the prescribed number of RIR. I.e., 10 repetitions, with 2 RIR – the pRIR in this scenario is two.
5. Rating of perceived exertion scale (RPE): designed to assess aerobic exercise due to the linear relationship between exercise intensity and aerobic energy demands

(i.e., increasing heart rate) (Borg, 1998; Halperin & Emanuel, 2019). RPE scoring (6 - 20) should closely reflect an increasing heart 60 - 200 BPM with increasing exercise intensity. In that, an RPE score of 6 should indicate a heart rate of 60 bpm, and an RPE score of 20 should be consistent with a heart rate of 200 bpm.

   a. A subjective evaluation – "subjective effort, strain, discomfort and fatigue during exercise" (ACSM, 2009).

6. Category ratio scale (CR-10): an alternative to the original RPE scale (6 - 20). More frequently used for anaerobic exercise or testing, including resistance training. The CR-10 scale allows one to quantify their perceived level of effort or exertion to identify training intensity, scoring intensity between 1 - 10 (Borg, 1998).
7. RIR-based RPE: a method of determining the number of repetitions in reserve through a corresponding RPE score (CR10). For instance, an RPE score of 10 indicates 0 RIR, while an RPE score of 7 is equivalent to 3 RIR.
8. Familiarization protocol: developed through synthesized research-study findings – a tool designed to educate practitioners and competitive and recreational athletes to effectively implement the RIR scale into RT programs while increasing RIR estimation accuracy.
9. Training intensity: may reference training load, or proximity to failure (pRIR).
10. Training volume: repetitions x sets x load
11. Perception of discomfort: "physiological and unpleasant sensations associated with exercise." (Steele et al., 2016).

12. Perception of fatigue: "a feeling of diminishing capacity to cope with physical or mental stressors, either imagined or real." (Micklewright et al., 2017)

Chapter 2: Literature Review

Resistance Training

RT has been found to improve muscular strength, power, local muscular endurance, and muscular hypertrophy (ACSM, 2009). Additionally, RT has been found to offer health benefits such as decreased blood pressure, "improved blood lipid profiles, glucose tolerance, and insulin sensitivity" (Kraemer et al., 2002). The desired outcome of an RT program should dictate program design. An optimal RT program must consider the fundamental RT principles – progressive overload, specificity, variation, and individualization (ACSM, 2002; 2009). Manipulating acute training variables in accordance with the basic RT principles will promote continued progression. Acute training variables include exercise intensity (absolute and relative load), repetitions, repetition velocity (with submaximal loads), rest periods, training volume (repetitions x sets), and training frequency (ACSM, 2009).

Achieving overload requires the stress or stimulation [intensity] placed on the musculoskeletal system to be greater than it is accustomed to (Ormsbee et al., 2017; Zourdos et al., 2015). Furthermore, overload should be progressive to limit the risk of overtraining and reduce the likelihood of non-functional overreaching and injury (Cavarretta et al., 2019; Hackett et al., 2016; 2018). Manipulating acute training variables in accordance with specificity and individualization principles should ensure training is consistent with the individual's program goals. Furthermore, strategic variations of acute training variables are essential in promoting continued adaptations and avoiding potential plateaus.

Load

The ACSM (2009) identified load as the weight lifted or the resistance selected. Load is a primary driver of adaptation in RT. Effective load prescription and manipulation should consider the training status and goals of the individual. The relationship between load and exercise order, volume, frequency, muscle action, repetition velocity, and rest intervals illustrates the importance of the prescribed load in RT programs (Kraemer & Ratamess, 2004).

Inherently, load selection affects training volume (Gonzalez-Badillo et al., 2011; Kraemer & Ratamess, 2004). For instance, continued increases in load typically require a decrease in repetitions, consequently decreasing (set x repetition) volume. Additionally, multiplying the load lifted, repetitions performed, and the number of sets completed identifies the load volume. Thus, if the goal is to maintain training volume while increasing the load (i.e., a hypertrophy mesocycle), the number of sets must increase to maintain training volume. Thus, load selection must acknowledge the resulting changes in repetitions and load volume as a byproduct of load alterations.

Load Prescription

Load has traditionally been prescribed using one of the following: PBT, RM, or training within an RM-zone (i.e., 10-12 RM-zone) (ACSM, 2009). Conventionally, maximal strength is assessed at the start of an RT program or mesocycle (Seo et al., 2012). The one-repetition maximum assessment (1RM) is one of the most frequently used methods to determine an individual's maximal strength across varying lifts. Typically, 1RM testing consists of incremental load increases until the lift attempt is failed or when an increase in load will result in a failed lift attempt. The 1RM is then used to define RT intensities by prescribing load(s) based on a percentage of the 1RM (%1RM), also known as percent-based training (PBT) (Bird et al., 2005).

Although PBT is frequently used when prescribing and manipulating RT intensities, PBT has limitations. For instance, defining training intensities based on %1RM requires training intensities to be dependent on past performances, thus, failing to acknowledge the daily fluctuations of strength due to fatigue, readiness to train, and changes in strength. Considering PBT does not account for acute strength improvements, current levels of fatigue, or one's readiness to train – the manipulation of acute training variables may not be optimal. Consequently, reducing the effectiveness of the training program (Helms et al., 2020). Increasing the accuracy of load selection with PBT requires 1RM to be regularly assessed, which may cause an accumulation of unnecessary fatigue. Considering the shortcomings of PBT coupled with its limited feedback, PBT may not be the ideal method to prescribe RT intensities (Gonzalez-Badillo et al., 2011).

Hass et al. (2001) suggested that RM protocols may be a better approach to prescribing and manipulating RT intensities when compared to PBT. The utilization of RM protocols entails performing the maximum number of repetitions relative to the load selected (i.e., 10RM). Similarly, training within an RM zone involves selecting a load that limits repetitions to a target RM zone (i.e., 5-8RM). However, one must consider the fatiguing nature of the trial-and-error strategy associated with the procedures utilized to identify RM loads that correspond with the assigned number of repetitions (intensities). Consequently, RM protocols may require some to train to failure frequently, resulting in an accumulation of fatigue at a heightened rate – increasing the risk of non-functional overreaching and the potential of overtraining and injury.

Velocity: Objective Measures

In an attempt to overcome the limitations of PBT, RM, and RM-zones, the relationship between the velocity of the lift (movement) completed and the load lifted was evaluated as a potential approach to quantify, thus, prescribe and manipulate RT intensities objectively. The use of mean concentric velocity (MCV) to prescribe and manipulate RT intensity is often referred to as velocity-based training (VBT) (Gonzalez-Badillo & Sanchez-Medina, 2010; Moore & Dorrell, 2020). The MCV of lifts were found to have a very strong inverse relationship with the load lifted (Carroll et al., 2012; Dorrell et al., 2019; Gonzalez-Badillo & Sanchez-Medina, 2010). In that, as the load increased, MCV decreased. The inverse relationship between MCV and the load lifted is recognized in literature as the load-velocity relationship. The load-velocity relationship allows one to create a load velocity profile (LVP).

The load-velocity relationship allows performance to be assessed as the lift occurs, giving the athlete or practitioner the capacity to adjust training intensity intrasession based on this acute feedback (Shattock & Tee, 2020). The ability to manipulate training intensity (i.e., load lifted) based on the athlete's current performance allows one to optimize training (Dorrell et al., 2019). VBT enables the athlete to train closer to their actual threshold while avoiding non-functional overreaching and decreasing the risk of overtraining and injury. The findings of Gonzalez-Badillo and Sanchez-Medina (2010) and Sanchez-Medina and Gonzalez-Badillo (2011) suggest that changes in velocity during a lift indicate the level of fatigue and effort. Furthermore, Dorrell et al. (2019) noted similar findings – decreasing MCV was indicative of fatigue accumulating within a set. Additional observations revealed that MCV decreased as sets progressed intrasession when using submaximal loads (Carroll et al., 2012). Thus, VBT should allow the stimulus-to-fatigue ratio to be manipulated, allowing stimulus levels to be maintained while limiting unnecessary fatigue.

VBT protocols may be implemented through one or more of the following approaches. The MCV of the first repetition of a set can be used to prescribe RT intensity (load), considering MCV was found to be indicative of training intensity (i.e., %RM) (Balsalobre-Fernandez et al., 2018; Gonzalez-Badillo et al., 2011). In addition to prescribing load, VBT allows the number of repetitions, thus, training volume to be altered by utilizing maximum percent velocity loss protocols (Banyard et al., 2019; Gonzalez-Badillo et al., 2011; 2017). The variable sets-velocity loss threshold approach uses a pre-determined load and number of repetitions. The set is stopped once the velocity loss threshold is reached – thus, the number of sets may vary. The variable sets-velocity loss threshold approach allows the number of repetitions completed to determine the number of sets, ensuring optimal training stimulus (volume), pending the athlete's current performance and fatigue level. The fixed sets-velocity loss threshold approach requires the load and number of sets to be known, while the number of repetitions per set is dictated by performance. Although sets are fixed, repetitions are completed if they remain above the pre-determined velocity-loss threshold (Banyard et al., 2019). The implementation of VBT allows the practitioner to assign load and repetitions to manipulate volume through objective measures.

The literature reviewed revealed that VBT protocols seem to produce greater strength gains when compared to PBT. This trend was observed during the completion of varying exercises, including the strict overhead press, deadlift, and countermovement jump (Dorrell et al., 2019), bench press (Gonzalez-Badillo & Sanchez-Medina, 2010; Gonzalez-Badillo et al., 2017), and back squat (Banyard et al., 2019; Carroll et al., 2012). Furthermore, a loss in velocity intraset was found to have a strong correlation with the number of repetitions completed and the maximum number of repetitions that could be completed at that load (Gonzalez-Badillo et al., 2017). Essentially, the velocity loss across repetitions in a set indicates the proximity to muscle failure – suggesting that velocity loss may be reliable for determining RIR (Gonzalez-Badillo et al., 2017).

Subjective Measures: RPE, RIR-based RPE, and RIR

The objective nature of VBT seems more reliable than traditional methods (i.e., PBT and RM methods) and their ability to accurately identify optimal training intensities. However, one must acknowledge that VBT requires the use of additional tools (i.e., linear position transducers) and the knowledge of load velocity profiles. The required use of tools and expertise needed to implement VBT protocols suggests VBT may not be practical for all populations. Although VBT may not be a practical approach for all populations, research has identified a strong inverse relationship between the objective measures of the load-velocity relationship and subjective scales (Helms, Storey et al., 2017; Rodriguez-Rosell et al., 2020). Specifically, subjective scales that were found to have a strong inverse relationship with the LVP were the RPE scale (CR10) and the RIR scale (Helms, Storey, et al., 2017).

RPE

RPE was initially introduced as a method to measure the intensity of aerobic exercise (Borg, 1998; Ormsbee et al., 2017; Zourdos et al., 2015). The BORG 15-point scale was constructed with the idea that exercising heart rate (HR) is indicative of exercise intensity (Borg, 1998). In that, RPE scoring would correspond to one's HR. For example, an RPE score of 6 should correspond to an HR of 60 bpm, and an RPE score of 20 should approximate an HR of 200 bpm. However, the Borg 15-point scale didn't seem to correlate well with RT intensities. In an attempt to better estimate the intensity of RT, a modified version of the Borg 15-point scale was developed, the RPE (CR10) scale. The CR10 scale may be one of the most frequently used scales to assess RT intensity. An RPE score assigns the level of intensity to a number on the scale (1-10). An RPE score of 10 indicates maximal effort, while a score of one indicates minimal effort exerted (Borg, 1998).

RPE (CR10) & Velocity

Helms, Storey, et al. (2017) observed a strong inverse correlation between RPE scores and average concentric velocity (ACV) when completing the bench press, squat, and deadlift amongst powerlifters. In that, as RPE scores increased, the ACV of lifts decreased. Essentially, velocity decreased with increasing effort, exertion, or load. The correlation between RPE and velocity suggests RPE may be a reliable and valid tool to prescribe and manipulate RT intensity. The findings of Helms, Storey, et al. (2017) seem to be consistent with Balsalobre-Fernandez et al. (2018), which reported objective measures (velocity) improved in accuracy when combined with subjective measures (RIR or RPE) when gauging RT intensity. The observations of Balsalobre-Fernandez et al. (2018) and Helms, Storey et al. (2017) support the use of subjective measures to prescribe and manipulate RT intensity.

RPE (CR10) Scoring

Although the RPE scale (CR10) seems reliable to gauge RT intensity, one must consider its subjective nature. Consequently, RPE scores can be expected to vary across participants due to their subjective nature. Study findings suggest training experience (Helms, Brown, et al., 2017; Ormsbee et al., 2017; Steele et al., 2017; Zourdos et al., 2015), muscle groups used (movement) (Cavarretta et al., 2019; Hackett et al., 2016; 2018), sex (Hackett et al., 2016; 2018), familiarity with MF (Helms, Storey, et al., 2017), proximity to momentary failure (Hackett et al., 2016; 2018; 2019; Mansfield et al., 2020; Zourdos et al., 2019), and load lifted (Hackett et al., 2012; Helms, Storey, et al., 2017) may affect the accuracy of utilizing RPE to identify RT intensities.

RPE (CR10) & Resistance Training Experience

Notably, training experience has been identified as a variable that consistently impacts the accuracy of RPE scores (Ormsbee et al., 2017; Zourdos et al., 2015). Study findings suggest that populations with lower levels of RT experience have decreased levels of accuracy in identifying RPE scores (Ormsbee et al., 2017; Steele et al., 2017; Zourdos et al., 2015). Hackett et al. (2012) found that novice lifters often reported submaximal RPE scores at maximal intensities. Training experience trends observed by Ormsbee et al. (2017) were consistent with previously mentioned studies (Steele et al., 2017; Zourdos et al., 2015; 2019). In that, 71.43% of experienced bench-pressers (n = 12) reported an RPE score of 10 when performing a 1RM; in comparison, only 23.08% of novice bench-pressers (n = 11) reported an RPE score of 10 when completing a 1RM (Ormsbee et al., 2017). Furthermore, 100% of experienced bench-pressers reported an RPE score greater than 9.5, while 30.77% of novice bench-pressers reported RPE scores less than 9 when completing a 1RM. Ormsbee et al. (2017) classified participants as experienced when they possessed at least two years of RT experience and performed the bench press at least once per week. Participants were considered novices when training experience was greater than or equal to three months but less than two years while bench-pressing once every two weeks.

Zourdos et al. (2015) noted similar findings between experienced (*n* = 15) and novice participants (*n* = 14) when completing a back squat. Experienced participants had two years or more of RT experience and performed the back squat at least once per/week, while novice participants had less than one year of RT experience and performed the back squat at least once every two weeks. Zourdos et al. (2015) found that experienced participants, on average, reported significantly higher RPE scores (9.80 $\pm$ 0.18) when compared to novice participants (8.96 $\pm$0.43). In that, 93.34% of experienced participants reported an RPE value greater than 9.5 at 1RM. Yet, only 57.14% of novice participants reported RPE scores equal to or less than 9 at loads equivalent to a 1RM when completing the back squat. The findings of Zourdos et al. (2015) suggest that accuracy in determining RPE scores increase as RT experience increases.

Contrary to similar RPE studies in which participants generally reported submaximal scores such as Ormsbee et al. (2017) and Zourdos et al. (2015) when training with maximal loads (i.e., 1RM), Helms, Storey, et al. (2017) found RPE scores to be accurate and consistent across participants, with no significant differences between lifts. Additionally, near-maximal RPE scores were observed during 1RM lifts (9.6 – 9.7). However, one must consider the participant selection (powerlifters) and the relationship between RT experience and accuracy estimating RPE. Experienced powerlifters typically possess higher levels of RT experience, coupled with an increased familiarity training with heavy loads. Thus, the increased accuracy and consistency of RPE scoring across participants may result from the increased training experience of all participants – further supporting the notion that the level of RT experience may affect RPE scoring accuracy. Although RPE scores seem to vary with varying RT experience, strong relationships were observed between RPE and the actual percentage of 1RM for the squat, deadlift, and bench press (Helms, Storey, et al., 2017).

RIR-Based RPE

Taking into account the limitations of RPE, further investigation into alternate subjective measures was warranted. As a result, in addition to solely using RPE to gauge RT intensity, RPE is frequently used in tandem with RIR – commonly referred to as RIR-based RPE. RIR-based RPE is used to predict RIR based on given RPE scores. For example, an RPE score of 10 corresponds to 0 RIR, and an RPE score of 8 corresponds to 2 RIR.

However, one must consider the externality of linking the two scales, CR10 and the RIR scale. Although both scales were found to be valid and reliable in gauging RT intensity – combining the two may be futile. One should consider the potential risk of reduced accuracy due to the limitations of combining both scales into one. Additionally, utilizing one scale to predict the rating of another "secondary" scale leaves room for the "original" subjective rating to become distorted. Hence, the increased risk of misinterpretation may not be worth combining the two scales. In addition, Hackett et al. (2016) suggested that the suspected correlations between CR10 and RIR may not perfectly represent each other. For instance, an RPE score of 8 may not indicate precisely

2 RIR. Hence, using RIR or RPE alone may be more accurate when determining and/or assigning RT intensity versus using the RIR-based RPE scale.

RIR was suggested to be more practical in determining RT intensity (Hackett et al., 2012). Furthermore, Hackett et al. (2012) found that participants reported submaximal RPE scores even though MF was achieved. More specifically, RPE scores for the bench press and squat were 8.9 $\pm$ 0.8 and 9.0 $\pm$ 0.7, respectively. However, RIR was reported at 0 for the same set (set five). These findings may suggest that RIR is more accurate in identifying RT intensities, particularly during maximal loads.

RIR

The findings of Balsalobre-Fernandez et al. (2018) revealed an improvement in the accuracy of load velocity profiles when combined with the subjective measures, RPE and RIR. Yet, variables considered independently were found to predict relative intensity with an acceptable level of accuracy, with slight variation between the three (RPE, RIR, and velocity). Balsalobre-Fernandez et al. (2018) also noted the inverse relationship between MCV and RIR followed similar trends as RPE and velocity. Furthermore, Rodriguez-Rosell et al. (2020) found that velocity loss was indicative of the completed repetitions and the number of repetitions left in reserve. In light of the relationship between RIR, MCV, and load velocity profiles – RIR seems to be a reliable method to prescribe and manipulate RT intensities.

Additionally, Hackett et al. (2012) found a negative correlation between the estimation of repetitions to failure (ERF) and RPE; and actual repetitions to failure (ARF) and RPE. However, the correlations between RPE, ERF, and ARF were not as strong as those between ERF and ARF alone. The similarities between RIR and RPE suggest they are effective in identifying RT intensities; however, the "stronger" correlations noted between ERF and ARF alone seem to suggest RIR may be a better fit to prescribe and manipulate RT intensities.

Although meanings were similar, the verbiage used to identify the number of RIR varied across studies. RIR, estimation of repetitions to failure (ERF), and actual repetitions to failure (ARF) share similar meanings. RIR references the ratio of repetitions completed in relation to the maximum number of repetitions that can be completed within one set. Throughout this literature review, the estimation of repetitions to failure (ERF) will be referred to as the estimation of (RIR) to remain consistent. While ARF will be referred to as actual repetitions in reserve (aRIR). Additionally, the number of assigned RIR will be referred to as prescribed RIR (pRIR). For example, selecting a load that is consistent with an 8RM should only allow eight repetitions to be completed. Thus, if one completes 8 repetitions at this load, there are 0 aRIR. However, if the set is stopped at six repetitions, 2 repetitions are left in reserve (2 RIR). In essence, when someone estimates repetitions to failure, they identify how many repetitions are "left in reserve" until ARF or aRIR is achieved.

aRIR and RIR Estimation Correlations

Hackett et al. (2012) noted a positive correlation between aRIR and the estimation of RIR for all subjects (17 competitive male bodybuilders) and across lifts (squat and bench press). Additional observations revealed systematic increases in error estimating RIR (eeRIR) as aRIR increased across two sessions (Hackett et al., 2018). Such that the relationship between estimating RIR and aRIR was strong to very strong across two sessions. Furthermore, strong correlations were noted across all sets (three sets) and exercises (pin-loaded vertical chest press machine and pin-loaded horizontal leg press machine). Hackett et al. (2018) found that eeRIR and aRIR decreased as sets progressed. During session 1, eeRIR decreased from 2 repetitions at set 1 to 0.6 repetitions at set 3 for the chest press, while eeRIR decreased from 3.1 repetitions at set 1 to 1.6 at set 3 for the leg press. Although the leg press revealed no differences, the eeRIR for the chest press decreased from 2.4 repetitions at set one to 0.9 at set three for session two.

Furthermore, Hackett et al. (2019) reported a strong relationship between aRIR and eeRIR for the chest press and a very strong relationship for the leg press. In that, as aRIR increased, accuracy in identifying the number of RIR decreased. The findings of Hackett et al. (2016) are in agreeance with previous studies that found RIR estimation accuracy decreased as aRIR increased. Findings revealed that RIR estimations for the chest press were within 1 repetition when aRIR were within 0-5 repetitions. Yet, the leg press findings revealed accuracy estimating RIR was within 1 repetition when aRIR were between 0-3 repetitions (Hackett et al., 2016).

RIR Estimation Accuracy: Upper body versus Lower body

According to Hackett et al. (2016), upper-body exercises may allow one to estimate RIR with greater accuracy over a wider repetition range, suggesting RIR estimations for upper-body lifts may be more accurate when compared to lower-body lifts. The findings of Hackett et al. (2016) were further supported by the results of Hackett et al. (2018) – in that eeRIR was less during all sets and sessions for the chest press vs. the leg press during session 1. The differences observed in upper-body and lower-body RIR estimation accuracy may be associated with varying levels of complexity, the amount of musculature involved, the technique required, the effort exerted, or the fatigue accumulated during the lift.

RIR Estimation Accuracy and RT Experience

Similar to RPE study findings, RIR estimation accuracy improved as RT training experience increased. Such that greater RT experience seems to be indicative of increased accuracy estimating RIR (Helms et al., 2016). Steele et al. (2017) evaluated the relationship between RT experience and the accuracy of identifying RIR. The most experienced group had greater than 36 months of RT experience (experts, $n = 42$), and the participants with the least experience were identified as orientation ($n = 15$) and beginners ($n = 21$), who possessed less than 1.5 months to 5 months of RT experience. Experts were found to under-predict RIR by 1-2 repetitions. Contrary, participants with the least RT experience (orientation) under-predicted RIR by approximately 4-5 repetitions. Although the RIR estimations of novice lifters were consistent, they were not accurate. The standard error of measurement (SEM) for the combined samples were 2.64 to 3.38 repetitions, suggesting that participants were not highly accurate when estimating repetitions to failure for either group (Steele et al., 2017). However, one must acknowledge the limitations due to study procedures. In that, participants only performed one set for each exercise – considering findings from similar studies with multiple sets observed an increase in accuracy as sets progressed (Hackett et al., 2012; 2019; Mansfield et al., 2020); accuracy estimating RIR may have increased if multiple sets were prescribed.

RIR Estimation Accuracy: Across-Sets

Hackett et al. (2012) aimed to determine the validity of the RIR scale by comparing RIR estimations and aRIR across 5 sets of bench press and back squats at 70% 1RM. All participants were experienced bodybuilders, which minimized any potential differences in RIR estimation accuracy due to varying experience levels. The findings revealed participants underestimated RIR for both squat and bench press in earlier sets. Set 1 (small effect size) and set 2 (medium effect size) for the bench press and set 1 for the squat (medium effect size). Accuracy improved in later sets (3-5) for the bench press–analysis revealed no differences between sets 3-5. While sets 2-5 for the back squat followed the same trend. Additionally, the fifth set was removed from the data because all participants reported 0 RIR, further supporting the proposition that multiple sets may increase one's ability to estimate RIR accurately. The study findings suggest that accurately predicting RIR improves across sets. Furthermore, the findings of Hackett et al. (2012) seem to infer that four sets may be enough to familiarize one with the RIR scale, considering the fifth set was removed due to all participants reporting 0 RIR when aRIR were 0.

Hackett et al. (2018) found that the capacity to estimate RIR accurately followed similar trends as Hackett et al. (2012), with accuracy improving across sets during three sets of chest press at 70% 1RM and leg press at 80% 1RM. The eeRIR decreased as sets progressed from set 1 to set 3 for both the chest and leg press. For the chest press, the eeRIR decreased from 2.0 repetitions at set one to 0.6 repetitions at set 3, while the eeRIR at set 1 was 3.1 repetitions, decreasing to 1.6 repetitions at set 3 for the leg press. Similar trends were noted in both experimental sessions 1 and 2. Perceptual fatigue also increased as sets progressed when performing the chest press with a large effect size. The leg press followed a similar trend with a moderate to large effect size.

RIR Estimation Accuracy and Proximity to MF

Although not a prevalent finding, Mansfield et al. (2020) noted that their study was the first to observe that irrespective of load (%1RM), accuracy estimating RIR trended the same across exercises, particularly the bench press and prone row. The findings of Mansfield et al. (2020) may suggest that the proximity to MF has a greater impact on RIR estimation accuracy when compared to varying loads. Furthermore, RIR estimations may increase in accuracy if estimates are made closer to MF rather than at a pre-determined repetition landmark. The findings of Mansfield et al. (2020) are consistent with the assumptions that RIR estimations become more accurate as aRIR decreases – in other words, the closer one is to MF.

RIR Estimation Accuracy: Across-Sessions

A limited number of studies evaluated whether accuracy in estimating RIR improved across multiple sessions. However, in a study of powerlifters (n = 12), observations revealed that the accuracy of RIR-based RPE scoring of the back squat was the only lift that significantly improved across sessions (Helms, Brown, et al., 2017). From week one to week three, estimation accuracy increases approached statistical significance for squats completed during hypertrophy sessions (8 repetitions with an intensity of 8 RPE). Additionally, a significant improvement in estimation accuracy from week one to week three for squats during the power sessions (two repetitions with an intensity of 8 RPE) was noted. All other comparisons did not reach significance across sessions for the deadlift or bench press. Although the findings were not significant, as training progressed from week one to week three, RIR-based RPE estimation accuracy increased for the bench press and deadlift. Minimal increases in RIR accuracy may have been influenced by participant selection, considering all participants were experienced powerlifters and relatively accurate to begin with, which may have minimized the window of improvement. The observations of Helms, Brown, et al. (2017) suggest that three sessions may not be enough to improve accuracy in estimating RIR for experienced populations.

Hackett et al. (2018) assessed participant accuracy in estimating RIR over two sessions to determine whether eeRIR decreased. Across-session analysis revealed no differences in eeRIR for the chest press. Contrary, greater eeRIR was observed for the leg press in the first set of session two (3.1 repetitions) compared to eeRIR in session one, set 1 (1.9 repetitions). The findings mentioned above seem to imply that two sessions utilizing RIR protocols may not expose the athlete to RIR protocols long enough to elicit an increase in accuracy estimating RIR.

However, one must consider that the increased accuracy estimating RIR during session one may be a result of the study procedures (Hackett et al., 2018). In that, the 10RM assessment being conducted on the same day as the accuracy testing procedures may have affected RIR estimation accuracy. For instance, the fatigue accumulated during RM testing may be why fewer repetitions were completed during session one versus session two. In this scenario, the findings would be consistent with similar studies that noted RIR accuracy increased as aRIR decreased (Hackett et al., 2016; 2018; 2019).

Additionally, one could speculate that the fatigue induced during the 10RM assessment consequently caused participants to begin the following experimental sets closer to MF. If this is the case, it would be consistent with the findings of Mansfield et al. (2020) – in that RIR estimation accuracy improved as participants neared MF. Lastly, 10RM testing during the same session as the experimental procedure may have allowed the participants to create and utilize (acute) memory anchoring or "current exertional sensations" to identify RIR (Hackett et al., 2018).

Protocol Development

In an effort to develop a familiarization protocol to educate practitioners and/or athletes to implement the RIR scale, previous research findings will be synthesized and grouped into two separate categories, category A and B. Category A identifies research findings that were used to guide participant selection. Category B consists of research findings that will be used to develop the familiarization protocol (FP). Thus, category B will define the FP. Variables in category B will be strategically implemented to better familiarize and educate one to effectively implement the RIR scale into a RT training program. Furthermore, variables in category B will be manipulated across groups, the Protocol Group (PG) and the Non-protocol Group (NPG), to determine protocol efficacy.

Category A: Participant History

Findings suggest RT experience enhances one's ability to identify RT intensities through subjective scales; thus, RT experience will direct participant selection (Hackett et al., 2012; Helms, Brown et al., 2017; Mansfield et al., 2020; Ormsbee et al., 2017; Steele et al., 2017; Zourdos et al., 2015). Furthermore, observations revealed that competitive bodybuilders (Hackett et al., 2012) and powerlifters (Balsalobre-Fernandez et al., 2018) accurately estimated RPE, RIR-based RPE, and RIR. Likewise, Helms, Brown, et al. (2017) used powerlifters, and observations revealed that participants were accurate in self-selecting loads consistent with desired RPE scores; the average error was found to be 0.33 RPE. The increased accuracy noted in powerlifters and bodybuilders seems to support the relevance of RT experience when using subjective scales. Additionally, one must acknowledge the high-intensity training of bodybuilders and powerlifters – in that increased familiarity with training near or to MF and/or TF and training with heavy loads may assist participants in identifying RIR accurately.

Thus, it is hypothesized that as RT experience increases, so will one's ability to accurately identify RT intensities with subjective scales (i.e., the RIR scale). That being the case, relying on previous training experiences may allow one to associate past feelings of effort and perceived exertion to assist in identifying RIR (Hackett et al., 2018). Additionally, prior exposure to MF may enable one to determine proximity to MF with increased accuracy – as they know what to expect and what it should feel like as one nears failure. Thus, participants will be experienced ($\geq$ 3 years RT experience) as they should be familiar with training to failure and comfortable training at higher intensities.

Participant selection will also be influenced by sex – considering findings suggest males are more accurate when gauging RT intensity with subjective scales, such as RPE, RIR-based RPE RIR, and RIR (Hackett et al., 2016; 2018). Significant interactions were observed between sex, with males (n = 53) estimating RIR with greater accuracy when compared to females (n = 28) (Hackett et al., 2016). Furthermore, Hackett et al. (2018) observed a significant effect for the sex of participants; women participants reported RIR with greater error than men across all sets, sessions, and exercises. Thus, all participants will be males.

Category B: Protocol Development

Sessions. Studies with multiple sessions were reviewed to determine how many sessions would be necessary to improve accuracy in estimating RIR. Cross-session findings revealed two (Hackett et al., 2018) to three (Helms, Brown, et al., 2017) sessions may not be enough to improve one's accuracy when estimating RIR, simply due to minimal improvements noted across sessions. As previously mentioned, Hackett et al. (2018) found that participants did not improve their accuracy in estimating RIR for the chest press across two sessions. Additionally, participant eeRIR increased across sessions, from 1.9 (session one) to 3.1 repetitions (session two) for the first set of the leg press.

However, when considering the findings of Hackett et al. (2012), Helms, Brown, et al. (2017), and Hackett et al. (2018), one must acknowledge the experience of participants used and the study procedures. Hackett et al. (2012) used experienced bodybuilders, Helms, Brown, et al. (2017) used experienced powerlifters, and 42 of 48 participants had greater than one-year resistance training experience in a study by Hackett et al. (2018). Hence, participants may have already been relatively accurate in estimating RIR and may not have had much room for improvement. One may also speculate that the lack of improvement noted in experienced lifters may result from limited exposure to the RIR scale. Therefore, the familiarization protocol will consist of five sessions, with two sessions serving as assessments (initial and final, sessions 1 and 5). Three sessions (2-4) will follow the familiarization protocol.

Sets. The accuracy of estimating RIR improved as sets progressed (Hackett et al., 2012; 2016; 2019). Increased accuracy across sets may be a result of the fatigue accumulated during the completion of each set, causing the following set to begin closer to MF due to a reduction of aRIR. Considering RIR estimations improved as aRIR decreased and aRIR tended to decrease across sets, utilizing multiple sets may provide a better chance to increase accuracy in estimating RIR. It was also noted that perceptual fatigue had moderate correlations with aRIR (Hackett et al., 2019). In that, as aRIR decreased, perceptual fatigue increased. Additionally, Hackett et al. (2019) suggested that familiarity with using the RIR scale may improve one's ability to rate exertional sensations accurately.

Based on research findings, multiple sessions and sets will be prescribed – increasing participant exposure to the RIR scale. Increased exposure will allow for increased familiarity, which should hypothetically increase one's ability to estimate RIR accurately. Hackett et al. (2012) used five sets; however, at the fifth set, all participants identified RIR as 0. That being the case, four sets will be used during the familiarization protocol, considering no further improvements in accuracy were noted beyond the fourth set.

Intensity. Training intensity (load) varied across studies. Some studies prescribed load using PBT methods. For example, Hackett et al. (2012) assigned loads at 60% 1RM with eight repetitions and 80% 1RM with three repetitions. In comparison, Mansfield et al. (2020) prescribed loads at 70% 1RM with 10 repetitions. Other studies, such as Cavarretta et al. (2019), specified the training load through RM protocols, for instance, using a 10RM. Furthermore, Helms, Brown, et al. (2017) study design required participants to select loads based on RPE (8 repetitions at 8 RPE or 2 RIR, two repetitions at 8 RPE or 2 RIR, and three repetitions at 9 RPE or 1 RIR). However, the literature reviewed didn't report any changes in accuracy estimating RIR due to varying loads.

Yet, one may consider that increased accuracy was observed when RIR were between (0-3 RIR) (Hackett et al., 2016). Suggesting that the load specifically may not directly affect RIR estimation accuracy, but the proximity to MF seems to affect RIR estimation accuracy (Hackett et al., 2012; 2016; 2019; Mansfield et al., 2020). Nonetheless, intensities selected will vary across sessions: session one: 8 repetitions with 1 pRIR, session two - 10 repetitions with 2 pRIR, session three - 12 repetitions with 2 pRIR, session four - 6 repetitions with 1 pRIR, and session five – 8 repetitions with 2 pRIR.

Exercise (movement-lift). Cavarretta et al. (2019) found that lower-body RIR-based RPE scores were higher than upper-body scores. However, no differences were noted between free-weight and machine exercises, suggesting that the musculature used and not the exercise modality affects the accuracy of estimating RIR. Furthermore, Hackett et al. (2018) observed RPE and RIR mean scores to be significantly higher for lower-body exercises (7.9 $\pm$ 2.0) when compared to upper-body exercises (7.3 $\pm$ 2.2). Helms, Brown, et al. (2017) observed greater accuracy when scoring RPE for the bench press vs. the squat during hypertrophy training (8 repetitions with 2 RIR). Thus, the lifts chosen will target both the upper (bench press) and lower body (v-squat) to account for potential differences. The use of lower and upper-body exercises may identify a need to vary RIR familiarization protocols for lower-body vs. upper-body lifts.

Memory Anchors. Memory anchors are frequently used to familiarize participants with the ratings or estimations of subjective scales, such as RPE, RIR-based

RPE, and RIR (Hackett et al., 2012; 2016; 2018; 2019; Helms, Brown et al., 2017). Memory anchors aim to connect past RT session intensities with current RT session intensities. The effectiveness of memory anchors relies on the ability to recall training at minimal (1 RPE or 10 RIR) and maximal intensities (10 RPE or 0 RIR). Acute memory anchors seem to be more effective when compared to memories anchored in (older) training sessions (Hackett et al., 2018).

Hackett et al. (2018) found that RIR estimations were more precise during session one when participants assessed their 10RM, in addition to the experimental protocol. Completing the 10RM assessment may have allowed participants to experience and create acute memory anchors to reference. The findings of Hackett et al. (2018) seem to highlight the effectiveness of acute memory anchors. In light of the findings in this study, the FP will test RM at varying intensities (12 RM, 15 RM, 6 RM). Thus, creating an opportunity for participants to develop and implement acute memory anchors to be referenced. Consequently, the overall session volume will vary across groups due to the differences in procedures between the PG and NPG. (Table 1; Table 2)

Chapter 3: Methods

Participants

Participants were healthy males familiar with resistance training (*n* = 9). Participants possessed high levels of resistance training experience ($\geq$ 3 years). Participants were also familiar with and comfortable training with moderately heavy to heavy loads at near-maximal intensities. Participants were screened for risk factors and required to be free from musculoskeletal injuries. All participants were familiar with the lifts chosen: barbell bench press (BP) and variations of the v-squat (VS). Additionally, the ability to perform the exercises without pain or limitations was required. Following the initial session, participants were randomly assigned to their respective groups, protocol (PG) (*n* = 5) vs. non-protocol (NPG) (*n* = 4).

Instrumentation

The initial session occurred in the University of Northern Iowa Exercise Physiology Laboratory (WRC 126), where the study purpose and procedures were explained. Furthermore, participants signed the informed consent form and completed a PAR-Q and the necessary paperwork. Experimental sessions occurred in the University of Northern Iowa, Wellness Recreation Center Free Weight Room, WRC 153. Experimental sessions required the utilization of a standard weight bench, a standard 45-pound barbell, Olympic free weight plates, and a plate-loaded v-squat machine. RIR scores were documented in Excel.

Procedures

The study purpose, procedures, and time requirements were discussed in detail during session one. Participant requirements, including training intensities, RT experience, and potential risks, were reiterated, and consent forms were signed. Participants were required to avoid any exercise that could impact the performance of the BP or VS 72 hours before the experimental sessions. Sessions were at least 72 hours apart (3 days) to account for the accumulated fatigue of the PG. Participants were encouraged to complete the lift through the full range of motion with proper technique. The BP followed USPA protocol; the VS was completed per machine instructions (limited to machine ROM).

Participants were required to verbally report the pRIR at the top of the concentric portion of the lift – elbows extended on the BP, knees extended on the VS. RIR was reported as participants neared TF instead of a pre-determined landmark (repetition). TF was identified by the participant's inability to maintain proper technique. The study consists of five total sessions. Intensities vary across sessions but are consistent across groups (PG vs. NPG). Session one - 8 repetitions with 1 pRIR, session two - 10 repetitions with 2 pRIR, session three - 12 repetitions with 2 pRIR, session four - 6 repetitions with 1 pRIR, and session five - 8 repetitions with 2 pRIR.

The PG completed RM assessments during familiarization sessions (sessions 2-4) in addition to the four sets of BP and VS. Consequently, increasing overall session intensity and volume for the PG. RM testing for sessions 2-4 are as follows: session two (12RM), session three (15RM), and session four (6RM).

Warm-ups were consistent across groups (PG and NPG). All sessions (1-5) begin with a general warm-up, 5 minutes on a stationary bike, followed by lift-specific warm-ups. The lift-specific warm-up: required participants to complete the exercises at lighter intensities. Lift-specific warm-up intensities (estimated %1RM) for the BP and VS are as follows: 30% 1RM at 8 repetitions, 40 % 1RM at 8 repetitions, and 50% 1RM at 6 repetitions. Following the warm-up, procedures vary across groups (PG vs. NPG) for sessions 2-4.

Session 1: The Initial Assessment

Session one was designed as an initial assessment to determine the participant's current ability to estimate RIR accurately. The data obtained during session one was used as a baseline to assess improvements made throughout the study resulting from the use or non-use of the familiarization protocol. All participants followed the warm-up protocol mentioned above. Following the lift-specific warm-up, participants completed four sets of each exercise (BP and VS) at the prescribed intensity. Based on the participants' prior lifting experience – participants self-selected a load that allowed them to complete eight repetitions with 1 RIR. Per Haff and Triplett (2016), the load of an 8 RM should be relatively consistent with 80% 1RM. As participants neared TF, they verbally identified when they had 1 RIR. Upon identification of the pRIR, the participants continued to TF. The difference in prescribed RIR (pRIR) and actual RIR (aRIR) was used to identify error estimating RIR (eeRIR).

During session one, all participants received assistance (spotters) to ensure proper lifting mechanics were used and that the participants remained safe. Furthermore, all participants received verbal encouragement during session one to ensure participants achieved TF.

Session 2-4: Non-Protocol Group

Following the general and lift-specific warm-ups, procedures varied across groups (PG vs. NPG). Four sets were completed for each exercise, BP and VS. Based on the participant's prior lifting experience – participants self-selected a load consistent with the current session's intensity (Table 2). Participants of the NPG received assistance (spotters) to ensure proper lifting mechanics and safety. However, the NPG did not follow the FP. Thus, the NPG did not receive verbal encouragement nor assistance in identifying loads during sessions 2-4.

Session 2-4: Protocol Group

Differing from the NPG, the PG followed the familiarization protocol. Including the creation of acute memory anchors and the review and application of relevant vocabulary. Additionally, participants were instructed to consciously identify the interactions between the perceptions of discomfort, fatigue, and effort and their relationship when determining RIR. The PG also completed RM testing following the lift-specific warm-up. The RM assessment served as a tool to create memory anchors and an opportunity for participants to apply the given vocabulary and instructions.

Memory anchor application. According to Hackett et al. (2017; 2018), when using subjective scales (i.e., RIR), it is essential that one anchor the upper limits of the scale to previous performances. Additionally, research seems to insinuate acute memory anchors may be more effective (Hackett et al., 2018). Thus, instead of relying on verbal explanations and the anchoring (linking) of past performances, RM assessments were used as a tool to create "acute" memory anchors. Such that as participants progressed through their set, they received verbal feedback/cues in which to associate the perception of the current repetition with relevant vocabulary and ratings on the RIR scale. For example, identifying one's 8 RM also identifies 0 RIR at that particular load. Per the FP, participants were instructed to recall their perceptions during the 7^{th} repetition (1 RIR), 6th repetition (2 RIR), and 5th repetition (3 RIR) – ultimately creating acute memory anchors for 0-3 pRIR.

Vocabulary Application. Coupled with memory anchoring procedures, participants were instructed to identify the differences in the perception of discomfort, fatigue, and effort and their impact on one's ability to complete repetitions within a set. Steele et al. (2016) defined discomfort as the "physiological and unpleasant sensations associated with exercise." Related findings revealed participants who completed a higher volume (more repetitions with more time under tension) versus a lower volume (fewer repetitions with less time under tension) were able to identify increased levels of discomfort independent of effort (Fisher et al., 2018; Steele et al., 2016; Stuart et al., 2018). Thus, it seems one can differentiate between the perceptions of effort and discomfort if one receives proper instruction. Considering discomfort increased with increasing repetitions, participants completed a 15RM assessment intending to elicit "discomfort."

Additionally, fatigue and effort were suggested to have a strong positive correlation as they both seem to increase together. According to Halperin and Emanuel (2019), as muscular fatigue increases, the ability to continually produce similar levels of force becomes progressively more difficult – thus, one must put forth more effort. Considering the above, it is clear the reasons one may fail to distinguish the differences between the two. Therefore, participants were provided with the following definition put forth by Micklewright et al. (2017), based on the rating of fatigue (perceived fatigue), "a feeling of diminishing capacity to cope with physical or mental stressors, either imagined or real." Considering the above, for application purposes – participants were given cues and instructed to pay particular attention to the amount of effort required to complete repetitions within a set as the session progressed. For example, the amount of effort needed to complete set one versus set four to overcome increasing and/or accumulating fatigue. The goal is for participants to differentiate between the perception of effort and fatigue when identifying the number of RIR.

RM assessment procedure. RM assessments were consistent with the current session's prescribed intensities. Session 2 identified participants 12RM, session 2 identified participants 15RM, and session 4 identified participants 6RM. See Table 1 for detail. During RM assessments, participants received assistance determining the corresponding load. RM assessment procedures Haff and Triplett (2016) – if the lift attempt was successful, the load was increased by 2.5 - 5% for the BP or 5 - 10% for the VS. If the lift attempt was unsuccessful, in that the participant could not complete the assigned number of repetitions, the load was decreased by 2.5 - 5% for the BP or 5 - 10% for the VS.

Following RM assessments, participants completed four sets of bench press and v-squats at the current session's prescribed intensities. See Table 1 for detail. Participants received assistance in using acute memory anchors from RM assessments and previously completed sets to help identify the pRIR. For example, recalling levels of discomfort, fatigue, and effort experienced during RM assessments. Furthermore, all participants within the PG received verbal encouragement to ensure TF was achieved and assistance (spotters) to ensure proper lifting mechanics and safety.

Session 5: The Final Assessment

The fifth session was designed as a final assessment. Session five consists of the general warm-up, lift-specific warm-up, and four working sets of each exercise. All participants (PG and NPG) followed the same procedures during session five. In that, the PG no longer received assistance with RIR estimations, selecting corresponding load(s), or feedback on performance. However, both groups had a spotter and received verbal encouragement to ensure safety, proper lifting mechanics, and the attainment of TF.

Following the general and lift-specific warm-up, participants began the assessment session. Based on the participants' prior lifting experience – participants self-selected a load that allowed them to complete 8 repetitions with 2 pRIR. Participants identified the pRIR as they approached 2 RIR. After the participant verbalized the RIR, the participant continued to TF. Essentially, session five assessed how well the PG learned to implement RIR without the aid of the researcher.

Data Analysis

The error estimating RIR (eeRIR) was identified by the absolute difference between prescribed RIR (pRIR) and actual RIR (aRIR) (calculated as the absolute value of pRIR - aRIR). eeRIR indicates the proximity of RIR estimations to the pRIR. In that, decreasing absolute eeRIR is indicative of increasing accuracy estimating RIR. Additionally, whether or not RIR was estimated accurately (i.e., pRIR = aRIR; 0 for not equal, 1 for equal) was used as the outcomes in the analyses. The generalized estimating equations using a normal distribution with the identity link function and autoregressive 1 (AR(1)) correlation structure with robust standard errors were used to analyze the absolute difference between pRIR and aRIR (eeRIR) to test to see how accurately participants estimated RIR between groups and across sessions (Liang & Zeger, 1986). The generalized estimating equations using a binomial distribution with the logit link function and AR(1) correlation structure were used to analyze the dichotomous outcome of whether or not participants accurately estimated RIR and to estimate the predicted probabilities of doing so. The AR(1) correlation structure for the repeated measures of sessions across time and sets within sessions was used because this correlation structure fit the data better than the alternative correlation structures available that included exchangeable (or compound symmetry), independence, or unstructured. Separate models were estimated for each movement.

To evaluate the sensitivity of the results associated with using the normal distribution for the absolute difference between pRIR and aRIR (eeRIR): generalized estimating equations using a Poisson distribution for count data with the log link function were also examined (results not shown). Results from these supplemental analyses were consistent with those based on the normal distribution. Given that the count data from this study does not fit with the typical definition of a Poisson random variable, namely the number of events occurring in a fixed interval of time or space, results for the normal distribution are presented. The analyses based on the binomial distribution estimating the probability of accurately estimating RIR provides an alternative view of the data and serve as another set of sensitivity analyses.

Pairwise comparisons within the generalized estimating equation framework were based on Wald Chi-squared tests with Bonferroni-corrected p values (*the Bonferroni adjustment controls the type I error rate by correcting for the number of pairwise comparisons being made in a set of analyses*). There were convergence issues with trying to estimate generalized estimating equation models for both outcomes when comparing differences across sets within sessions. Instead, for these analyses, a one-way repeated measures ANOVA with Greenhouse Geiser correction was used to examine the mean absolute differences of pRIR – aRIR across sets within a session. To help describe whether there was a tendency for over- or under-estimation of RIR, the means and standard errors for the differences between pRIR – aRIR were computed by session and group, as well as by session and set for the PG. Analyses were conducted using SPSS 28.0. A significance level of .05 was used.

Statistical Results

First set of analyses – Group comparisons, including Session 1 and Session 5 data only

Absolute difference of pRIR – aRIR (eeRIR) Model-Based Results. Figure 1 shows the absolute mean eeRIR by group (PG vs. NPG), session (1 vs. 5), and movement (BP and VS), and corresponding standard error bars estimated from the generalized estimating equation models shown in Table 5. For both movements, there was a significant interaction term between group and session (Wald Chi-squared test statistic = 20.317, df = 1, $p < .001$ for bench press and Wald Chi-squared test statistic = 7.715, df = 1, $p = .005$ for v-squats). The significant interaction indicates that the difference in absolute mean eeRIR between groups depends on the session, and the difference between sessions depends on the group.

More specifically, for the BP, there was not a significant difference in absolute mean eeRIR between the two groups (PG vs. NPG) at session 1 (baseline; mean difference of -0.347 for the PG compared to the NPG, SE = 0.145, Bonferroni-corrected $p = .068$), though the difference was nearly significantly different from 0. On the other hand, there was a significant difference in the absolute mean eeRIR between the two groups at session 5 (mean difference of -0.808, SE = 0.106, Bonferroni-corrected $p < .0001$), with the PG tending to have a lower absolute mean eeRIR on average than the NPG by .81 repetitions (Figure 1). Additionally, for each group, there was a significant difference in the absolute mean eeRIR between session 1 and session 5. For the PG, the absolute mean eeRIR was significantly lower at session 5 compared to session 1 (mean difference of -0.321, SE = 0.094, Bonferroni-corrected $p = .004$), suggesting an improvement in accuracy (Figure 1). For the NPG, the absolute mean eeRIR was significantly higher at session 5 compared to session 1 (mean difference of 0.140, SE = 0.041, Bonferroni-corrected $p = .004$), suggesting there was not an improvement in accuracy for this group; instead, they did slightly worse at estimating RIR at session 5 than at session 1 (Figure 1).

For v-squats, there was no significant difference in absolute mean eeRIR between the two groups (PG vs. NPG) at session 1 (baseline; mean difference of -0.536 for the PG compared to the NPG, SE = 0.434, Bonferroni-corrected $p = .868$). On the other hand, there was a significant difference in absolute mean eeRIR between the two groups at session 5 (mean difference of -1.361, SE = 0.295, Bonferroni-corrected $p < .0001$), with the PG tending to have a lower absolute mean eeRIR on average than the NPG by 1.36 repetitions (Figure 1). In terms of session comparisons within group, findings suggest that there was a significant difference in the absolute mean error eeRIR between sessions 1 and 5 for the PG, where the absolute mean eeRIR was significantly lower at session 5 compared to session 1 (mean difference of -0.848, SE = 0.283, Bonferroni-corrected $p = .012$), suggesting an improvement in accuracy (Figure 1). On the other hand, for the

NPG, there was no significant difference in the absolute mean eeRIR between session 5 and session 1 (mean difference of -0.023, SE = .092, Bonferroni-corrected $p > .999$).

Difference of pRIR – aRIR (eeRIR) Descriptive Results. Figure 2 shows the mean difference of pRIR – aRIR (eeRIR) by group (PG vs. NPG), session (1 vs. 5), and movement (BP and VS) to help describe the over- or under-estimation of RIR. The smaller symbols are the means across sets for a participant, and the larger symbols are the group means. As illustrated in Figure 2, participants in the PG tended to overestimate RIR on average, while the NPG tended to underestimate RIR on average for the BP at baseline (session 1). In comparison for the BP, some individuals from each group tended to overestimate RIR during session 5, and others tended to underestimate RIR so that, on average, the group means balanced out near zero during session 5. This result highlights the appropriateness of using the absolute difference of pRIR – aRIR to evaluate RIR estimation accuracy for the outcome of this study.

For v-squats, most participants in both groups tended to underestimate RIR on average during session 1 (Figure 2). Again, during session 5 for VS, some individuals from each group tended to overestimate RIR during session 5, while others tended to underestimate RIR; thus, on average, the group means balanced out near zero during session 5.

Probability of Being Accurate in Estimating RIR (that is, pRIR = aRIR) Model-Based Results. Figure 3 shows the estimated predicted probability of being accurate (POBA) (i.e., pRIR = aRIR) by group (PG vs. NPG), session (1 vs. 5), and movement (BP and VS), and corresponding standard error bars estimated from the generalized estimating equation models shown in Table 6. For both movements, there was a significant interaction term between group and session (Wald Chi-squared test statistic = 9.684, df = 1, $p = .002$ for bench press and Wald Chi-squared test statistic = 12.654, df = 1, $p = .0004$ for v-squats). The significant interaction indicates that the difference in the predicted probabilities of estimating RIR accurately between the two groups depends on the session, and the difference between sessions depends on the group.

More specifically, for the BP, there was not a significant difference in the predicted POBA in estimating RIR between the two groups (PG vs. NPG) at session 1 (baseline; difference in probabilities of 0.220 for the PG compared to the NPG, SE = 0.091, Bonferroni-corrected $p = .068$), though the difference was nearly significantly different from 0. On the other hand, there was a significant difference in the predicted probabilities of being accurate between the two groups at session 5 (difference in probabilities of 0.58, SE = 0.053, Bonferroni-corrected $p < .0001$), with the PG tending to estimate RIR more accurately on average than the NPG by 58 percent (Figure 3).

Furthermore, for the PG, there was a significant difference in the predicted probabilities of accurately estimating RIR between sessions 1 and 5 for BP. For the PG, the predicted probability of accurately estimating RIR was significantly higher at session 5 compared to session 1 (difference in probabilities of 0.33, SE = 0.092, Bonferroni-corrected $p = .001$), suggesting a 33 percent improvement in accuracy estimating RIR for BP (Figure 3). For the NPG, there was no significant difference in the predicted probabilities between session 5 and session 1 (difference in probabilities of -0.030, SE = 0.052, Bonferroni-corrected $p > .999$; Figure 3).

For VS, there was not a significant difference in the predicted probabilities of estimating RIR accurately between the two groups (PG vs. NPG) at session 1 (baseline; difference in probabilities of 0.04 for protocol compared to non-protocol, SE = 0.175, Bonferroni-corrected $p > .999$). On the other hand, there was a significant difference in the predicted probabilities of accurately estimating RIR between the two groups at session 5 (difference in probabilities of 0.77, SE = 0.125, Bonferroni-corrected $p < .0001$), with the PG tending to have a greater predicted probability to accurately estimate RIR (Figure 3).

In terms of session comparisons within group for v-squats, findings suggest that there was a significant difference in the predicted probabilities of estimating RIR accurately between sessions 1 and 5 for both groups. For the PG, the predicted probability of estimating RIR accurately was significantly greater at session 5 than at session 1 (difference of probabilities of 0.54, SE = 0.152, Bonferroni-corrected $p = .001$), suggesting an improvement in accuracy (Figure 3). For the NPG, there was also a significant difference in the predicted probability of estimating RIR accurately between session 5 and session 1 (mean difference of -0.19, SE = .055, Bonferroni-corrected $p = .002$); the predicted probability was significantly lower at session 5 compared to session 1, suggesting there was not an improvement in accuracy for this group instead participants were less accurate at session 5 (Figure 3).

Second set of analyses – Session comparisons for the Protocol Group

Absolute difference of pRIR – aRIR (eeRIR) Model-Based Session Results for the Protocol Group. Figure 4 shows the absolute mean eeRIR by session (1 through 5) and movement (BP and VS) for the PG; and corresponding standard error bars estimated from the generalized estimating equation models shown in Table 8. For both movements, there was a significant session effect on absolute mean eeRIR (Wald Chi-squared test statistic = 9482, df = 1, $p < .001$ for bench press and Wald Chi-squared test statistic = 104.6, df = 1, $p < .001$ for v-squats).

Table 3 shows the Bonferroni-corrected pairwise session comparisons for each movement. Mean differences that are statistically significantly different at the .05 significance level are bolded. Seven of the ten unique pairwise comparisons for BP are statistically significant at the .05 significance level. Specifically, for BP, the absolute mean eeRIR for session 5 is significantly lower than the means for sessions 1, 2, and 3 but not lower than session 4 (Table 3). Additionally, the absolute mean eeRIR for session 4 is significantly lower than the means for sessions 2 and 3, but it is not lower than that for session 1. The absolute mean eeRIR for session 3 is significantly different from session 2 but not from session 1. While the absolute mean eeRIR for session 2 is significantly higher than for session 1.

Four of the ten unique pairwise comparisons for VS are statistically significant at the .05 significance level. Specifically, for VS, the absolute mean eeRIR for session 5 is significantly lower than the means for each session (Table 3). None of the other pairwise comparisons are statistically significant at the .05 level. As shown in Figure 4, the absolute mean eeRIR decreases across the sessions for VS, suggesting that participants in the PG tended to get more accurate on average with each new session and seemed to benefit from the 5 sessions.

Probability of Being Accurate in Estimating RIR (that is, pRIR = aRIR) Model-Based Session Results for the Protocol Group. Figure 5 shows the predicted probabilities of estimating RIR accurately by session (1 through 5) and movement (BP and VS) for the PG and corresponding standard error bars estimated from the generalized estimating equation models shown in Table 8. For both movements, there was a significant session effect on the predicted probabilities of estimating RIR accurately (Wald Chi-squared test statistic = 359.85, df = 1, $p < .0001$ for bench press and Wald Chi-squared test statistic = 22.5, df = 1, $p = .0002$ for v-squats).

Table 4 shows the Bonferroni-corrected pairwise session comparisons for each movement. Differences in probabilities that are statistically significant at the .05 significance level are bolded. Five of the ten unique pairwise comparisons for the BP are statistically significant at the .05 significance level. More specifically for the BP, the predicted probability of estimating RIR accurately at session 5 is significantly greater than the corresponding probabilities at sessions 1, 2, and 3, but it is not significantly greater than that for session 4 (Table 4). Additionally, the predicted probability of being accurate for session 4 is significantly greater than those for sessions 2 and 3, but it is not significantly greater than that for session 1.

Four of the ten unique pairwise comparisons for VS are statistically significant at the .05 significance level. More specifically, for VS, the predicted probability for estimating RIR accurately at session 5 is significantly greater than those for each of the other sessions (Table 4). None of the other pairwise comparisons are statistically significant at the .05 level. As shown in Figure 5, the predicted probabilities of estimating RIR accurately increases across sessions for VS, suggesting that participants in the PG tended to become more accurate on average with each new session and seemed to benefit from the 3 sessions.

Third set of analyses – Set comparisons within session for the Protocol Group

Absolute difference of pRIR – aRIR (eeRIR) Descriptive Results. Figure 6 shows the absolute mean eeRIR by set (1-4) within each session BP among the PG and corresponding standard error bars. Figure 7 shows the absolute mean eeRIR by set (1-4) within each session for VS among the PG and corresponding standard error bars. The statistics shown in Figure 6 and Figure 7 are based on descriptive statistics. Based on one-way repeated measures ANOVA for each session, there was not a statistically significant difference among sets within each session for bench press: *session 1*, $F(2.3,9.3) = .595$, $p = .595$; *session 2*, $F(2.8,11) = 1.522$, $p = .263$; *session 3*, $F(2.08,8.36) = 2.667$, $p = .126$; *session 4*, $F(3,12) = 0.000$, $p > .999$; *session 5*, $F(1,4) = 6.000$, $p = .070$. or v-squat: *session 1*, $F(1.56,6.24) = 1.540$, $p = .278$; *session 2*, $F(1.63,6.51) = 2.429$, $p = .165$; *session 3*, $F(2,8) = .524$, $p = .611$; *session 4*, $F(1.40,5.60) = 0.828$, $p = .442$; *session 5*, $F(1,4) = 2.667$, $p = .178$.

Difference of pRIR – aRIR (eeRIR) Descriptive Results. Figure 8 shows the absolute mean eeRIR for the BP by session (1-5) and set (1-4) for the PG. The statistics reported are based on descriptive statistics to help describe over- or under-estimation across the sessions and sets. No statistical tests were conducted on these analyses. Note that no bar appears in the figure for sets 2, 3, and 4 within session 5 due to all participants estimating the RIR accurately so that the mean difference of pRIR – aRIR is 0 with a standard error of 0 (no error bars). As illustrated in Figure 8, participants in the PG tended to overestimate RIR for each set within sessions 1 and 4 (BP). However, the participants (PG) tended to underestimate RIR – on average for each set within session 3. On the other hand, participants in the PG tended to both underestimate and overestimate on average across sets within session 2. For session 5, participants in the PG underestimated RIR on average for the first set but then accurately estimated RIR for sets 2-4.

Figure 9 shows the absolute mean eeRIR by session (1-5) and set (1-4) for the PG (VS). The statistics reported are based on descriptive statistics to help describe over- or under-estimation across sessions and sets. No statistical tests were conducted on these analyses. Note that no bars appear in the figure for sets 2, 3, and 4 within session 5 due to all participants estimating the RIR accurately so that the mean difference of pRIR – aRIR is 0 with a standard error of 0 (no error bars). As illustrated in Figure 9, participants in the PG tended to underestimate RIR on average for each set within session 1 and most of the sets in session 2. For sessions 3 and 4, the PG tended to underestimate and overestimate RIR (on average) across sets. For session 5, participants in the PG tended to estimate RIR across sets accurately.

Discussion

PG vs. NPG (absolute mean eeRIR)

Consistent with the proposed research hypotheses (1-3), the PG estimated RIR with greater precision than the NPG at session 5 for both movements (BP and VS). Specifically, a significant difference was noted between groups as the absolute mean eeRIR was 0.144 repetitions (PG), compared to 0.953 repetitions (NPG) for the BP at session 5. Similar trends were noted for the VS, with a significant difference between groups at session 5, as the absolute mean eeRIR was 0.112 repetitions (PG) vs. 1.496 repetitions (NPG). Considering RIR estimation accuracy was consistent between groups at session 1, with no significant differences in absolute mean eeRIR or the predicted

POBA, the above observations suggest that the FP effectively increases one's capacity to estimate RIR accurately.

Furthermore, an increase in absolute mean eeRIR was observed for the NPG from session 1 (0.813 repetitions) to session 5 (0.953 repetitions). This finding may be associated with the increase in pRIR. For instance, (Hackett et al., 2016; 2018; Mansfield et al., 2020) found that estimation accuracy decreased as aRIR increased, which would be consistent with the above observations. On the other hand, participants in the PG still had an increased capacity to estimate RIR accurately despite the increased pRIR (1 to 2 RIR). This finding further supports the effectiveness of the FP. Additionally, the consistency in absolute mean eeRIR between movements, 0.144 repetitions (BP), and 0.112 repetitions (VS), suggest that the FP effectively increases RIR estimation accuracy for both upper- and lower-body lifts. The above observations support the hypothesis that utilizing a familiarization protocol will increase one's accuracy in estimating RIR.

PG vs. NPG (Probability of being accurate)

To further evaluate the differences between groups – the predicted probability of accurately estimating RIR was calculated and compared across groups. The PG was 58% more likely to estimate RIR accurately for the BP and 77% more likely to estimate RIR accurately for the VS. Specifically, at the fifth session (BP), the PG had an 85.8% chance of accurately identifying RIR vs. the NPG, with a 28.2% probability. For the VS, the PG had an 89.7% predicted probability of accurately identifying RIR vs. the NPG with 12.3%, further supporting the idea that the differences noted in session 5 may result from increased exposure to the developed FP.

PG vs. NPG (Over and Underestimations)

Although only descriptive, Figure 2 revealed identifiable trends – such that RIR estimations tightened around 0 at the fifth session, producing similar average means across groups. However, the magnitude of outliers varied across groups. More specifically, for the fifth session of v-squats, the outliers include an overestimation of 0.75 repetitions and an underestimation of 0.50 repetitions (NPG), compared to an overestimation of 0.25 repetitions and an underestimation of 0.25 repetitions (PG). Furthermore, for the BP at session 5, the identified outliers included an overestimation of 2.0 repetitions and an underestimation of 0.75 repetitions (NPG) vs. an overestimation of 0.25 repetitions and an underestimation of 0.25 repetitions (PG). Although the group means of over- and underestimations were similar across groups, the magnitude of "individual participant means" were greater for the NPG. Based on these findings, the FP prepared participants to estimate RIR with less variation in over and under-estimations.

PG Absolute mean eeRIR (pRIR – aRIR)

Across-session data analyses revealed that the absolute mean eeRIR at session 5 was significantly lower than all previous sessions except session 4 (BP). However, the absolute mean eeRIR for session 4 was significantly lower than the absolute mean eeRIR for sessions 2 and 3, but not session 1. First, addressing the variation from sessions 1 to 2, observations revealed a significant decrease in RIR estimation accuracy through an increase in absolute mean eeRIR. In light of past research – this finding may be associated with the increased number of pRIR and aRIR at session 2 (10 repetitions with 2 RIR) compared to session 1 (8 repetitions with 1 RIR). Such that as aRIR increased, RIR estimation accuracy was found to decrease (Hackett et al., 2016; 2018; Mansfield et al., 2020; Zourdos et al., 2019).

Nonetheless, considering session 1 was designed to assess initial RIR estimation accuracy to identify baseline measurements – participants hadn't yet been exposed to the familiarization protocol. But, once participants were exposed to the familiarization protocol, RIR estimation accuracy improvement trends from sessions (2-4) were consistent in that accuracy estimating RIR increased through decreasing absolute means eeRIR across sessions. The absolute mean eeRIR for sessions 2-4 are as follows: 0.949 repetitions, 0.501 repetitions, and 0.203 repetitions. Although no longer exposed to the FP at session 5, the absolute mean eeRIR continued to decrease (0.145 repetitions) (Chart 1) – consistent with the research hypothesis that increased exposure to the FP and the RIR scale – will increase accuracy estimating RIR.

Similar trends were observed in absolute mean eeRIR for BP and VS at sessions 3, 4, and 5. However, variations in absolute mean eeRIR across sessions 1 and 2 varied across movements (Figure 4). One may speculate that the novelty of the VS was responsible for the increased absolute mean eeRIR at session 1 compared to the BP. In spite of the previous research findings of Hackett et al. (2016; 2018), which noted decreasing RIR estimation accuracy with increasing aRIR, VS RIR estimation accuracy increased from session 1 (1 RIR) to session 2 (2 RIR), varying from this trend. This observation suggests that despite the v-squats novelty, the vocabulary review and application, coupled with the creation of acute memory anchors, from the FP still allowed participants to increase accuracy in estimating RIR.

The novelty of the VS may also be responsible for the lack of significant interactions observed between sessions 1-4 for absolute mean eeRIR. As participants followed the specifics of the FP, they also had to become familiar with the mechanics and perceptions of the VS. Thus, one must acknowledge that the novelty of the VS may have distorted the effectiveness of the FP, influencing the absolute mean eeRIR. Furthermore, the lack of familiarity with the VS required more sets to identify load selection (intended training intensity) during RM assessments compared to the BP. Thus, the increased sets performed during RM assessments may have allowed participants to continue to improve accuracy in estimating RIR regardless of the novelty of the VS. Another possibility – is that the participants may have become familiar with the movement (VS) and the FP by session 5 to illustrate the significant decrease in absolute mean eeRIR. Further supporting the FP's effectiveness in that despite the novelty of the movement being assessed, RIR estimation accuracy still increased (VS).

PG: Across set (eeRIR)

When analyzing RIR estimation accuracy, no significant interactions were observed across sets (1-4) for either movement (BP or VS) during all sessions (1-5). However, when analyzing Figure 6 and Figure 7, there were identifiable trends. For instance, the average of absolute means eeRIR for sets (1-4) decreased across sessions for both movements, suggesting FP effectiveness (Chart 1). Additionally, participants began each session (sessions 2-4) more accurately than the previous – in that the mean absolute eeRIR of the first set decreased as sessions progressed – 1.40 repetitions, 0.8 repetitions, 0.2 repetitions (BP); 0.8 repetitions, 0.2 repetitions, 0.2 repetitions (VS).

Furthermore, as sessions progressed for the BP, eeRIR decreased, with an increasing number of sets becoming more accurate intrasession. More specifically, at session 2, the absolute mean eeRIR was within 0.4 repetitions of the pRIR for one set, while at session 3, the absolute mean eeRIR was within 0.2 repetitions of the pRIR for 2 of 4 sets. By session 4, the absolute mean eeRIR was within 0.2 repetitions of the pRIR for all 4 sets. Suggesting as exposure to the FP increased, accuracy estimating RIR increased through a decrease in absolute mean eeRIR. Additionally, even though participants of the PG were no longer receiving the assistance of the researcher or the FP, eeRIR continued to decrease at session 5, set 1 (0.6 repetitions eeRIR), set 2-4 (pRIR = aRIR), which provides further support for the practical applicability of the FP.

Additionally, at sessions 3 and 4 of the VS**,** the absolute mean eeRIR was within 0.6 repetitions of the pRIR for all sets. In comparison, at session 2, sets 1 and 2 were within 0.8 repetitions of the pRIR, set 3 was within 0.2 repetitions of the pRIR, while set 4 was within 1.2 repetitions of the pRIR, which suggests an improvement in accuracy across sets, as sessions progressed – supporting the familiarization protocols efficacy. Additional observations revealed one set was within 0.2 repetitions of the pRIR at sessions 2 and 3, yet 3 out of 4 sets were within 0.2 repetitions of pRIR for session 4 – suggesting a continued increase in RIR estimation accuracy with the ongoing application of the FP. The final assessment (session 5) – RIR estimation accuracy continued to increase, with the first set being within 0.4 repetitions of pRIR, while the remaining sets (2-4) revealed no error in identifying RIR (pRIR = aRIR). Even though participants were no longer receiving assistance from the researcher or the FP, accuracy continued to improve during session 5. The increases in estimation accuracy observed for the VS at the fifth session suggests that participants were able to apply knowledge gained through the FP. This supports the hypothesis that the ability to rely on acute memory anchors will improve RIR estimation accuracy.

Although participants began the study relatively accurate in identifying RIR, findings suggest the FP effectively increased one's accuracy in estimating RIR. However, all findings combined – an increase in the number of sessions may increase the effectiveness of the FP. For instance, the predicted probability of accurately estimating RIR increased across sessions of the FP (2-4). For the VS, 40%, 55%, and 70% (sessions 2-4, respectively); for the BP, 29.5%, 50.3%, and 79.4% (sessions 2-4, respectively). When considering the pattern of the predicted probabilities of being accurate for the VS, 1-2 additional sessions may allow participants to reach a probability nearing 100%. Considering the predicted POBA during session 5 was 90%, 1-2 additional sessions seem sufficient. The BP trend was similar to VS in that 1-2 additional sessions may allow participants to achieve a probability nearing 100%. This is also supported by the fifth session having an 85.6% predicted probability of accurately estimating RIR.

Limitations:

Going into the study, some limitations were known. Such as the small sample size (n = 9). Five participants were in the PG, and four were in the NPG. Additional limitations were observed during study procedures, such as the novelty of the v-squat. In that none of the participants in the study were familiar with the particular VS machine used. The lack of familiarity with the v-squat may have extended the familiarization curve.

Additional limitations with the study procedures were known during the study design, including the use of TF to identify aRIR instead of absolute MF. Thus, aRIR may not reflect the "true" number of RIR. Additionally, the lack of objective measures (i.e., linear position transducers) required the identification of aRIR to rely solely on the participant and researcher. Also, it may have been beneficial to document the number of sets completed during RM assessments. Considering the accumulation of fatigue from RM assessments may have caused subsequent sets to begin closer to TF (less aRIR) – identifying the number of sets completed for RM assessments may provide more insight into RIR estimation accuracy trends. As previous findings suggest, the closer one is to TF; the more accurate RIR estimations are (Mansfield et al., 2020).

Furthermore, although RM assessments were used to create acute memory anchors relative to the current session's training intensities, the prescribed RM assessments weren't precisely equivalent to the prescribed training intensities – which may have impacted RIR estimation accuracy. For example, the prescribed training intensities for session 3 was 12 repetitions (pRIR = 2); however, the RM assessment was 15, which is a 1 repetition difference from the assigned training intensity. Additionally, the randomization or order of movements (BP and VS) may have affected RIR estimation accuracy. Such that the fatigue accumulated during the VS may have impacted BP performance, thus, RIR estimation accuracy. For instance, completing the VS first may have decreased the lower body (leg drive) when completing the BP, potentially leading to an overestimation of RIR. In order to maximize observations while still allowing movement randomization, one should document the order of movement (exercise) completion.

Future Research

Considering findings suggest the FP is effective, future research should consider broadening the participant population. For instance, future studies should include female participants to determine if the FP maintains its effectiveness. Furthermore, one should consider novice and intermediate athletes – to determine if the protocol elicits similar increases in RIR estimation accuracy. Future research should consider additional sessions when assessing the FP effectiveness with less experienced populations, as the learning curve may differ from advanced trained populations. Future research should also evaluate whether the FP is effective in increasing RIR estimation accuracy for other movements (i.e., barbell row, shoulder press, barbell back squat). Additionally, future research should account for the order of exercise to account for its potential impact on estimation accuracy.

Practical Applications:

The findings of this study suggest the FP effectively increases RIR estimation accuracy among resistance-trained populations (experience, $\geq$ 3 years). Although participants began the study relatively accurate estimating RIR, the FP still generated a consistent improvement in RIR estimation accuracy. Thus, the FP may be implemented into the RT program of resistance-trained populations – allowing one to become familiar with and accurate using the RIR scale.

Furthermore, the prescription and manipulation of training intensities are crucial for RT programs to ensure the desired adaptations. Thus, the ability to identify training intensity is imperative, whether objectively or subjectively. Considering the FP promoted an increased predicted probability of being accurate, 85.8% (BP) and 89.7% (VS) –

participants are able to identify training intensities with an acceptable level of accuracy (BP and VS). In theory, participants are able to identify pRIR with a predicted accuracy nearing 90%, allowing one to maximize training efficiency.

References

American College of Sports Medicine. (2002). Position stand: Progression models in resistance training for healthy adults. *Medicine & Science in Sports & Exercise, 34*(2), 364-380.

American College of Sports Medicine. (2009). Position stand: Progression models in resistance training for healthy adults. *Medicine & Science in Sports & Exercise, 41*(3), 687-708. https://doi.org/10.1249/MSS.0b013e3181915670

Balsalobre-Fernandez, C., Munoz-Lopez, M., Marchante, D., & Garcia-Ramos, A. (2018). Repetitions in reserve and Rate of perceived exertion increase the prediction capabilities of the load-velocity relationship. *The Journal of Strength and Conditioning Research, 35*(3), 724 – 730. https://doi.org/10.1519/JSC.0000000000002818

Banyard, H. G., Tufano, J. J., Delgado, J., Thompson, S. W., & Nosaka, K. (2019). Comparison of velocity-based training methods and traditional 1RM-percent-based training prescription on acute kinetic and kinematic variables. *International Journal of Sports Physiology and Performance, 14*(2), 246-255. https://doi.org/10.1123/ijspp.2018-0147

Bird, S. P., Tarpenning, K. M., & Marino, F. E. (2005). Designing resistance training programs to enhance muscular fitness: A review of the acute program variables. *Sports Medicine, 35*(10), 841-851. https://doi.org/10.2165/00007256-200535100-00002

Borg, G. (1998). Borg's Perceived Exertion and Pain Scales. *The Borg RPE scale, The Borg CR10 scale* (pp. 29 – 43). Human Kinetics.

Carroll, K. M., Sato, K., Beckham, G. K., Triplett, N. T., Griggs, C. V., & Stone, M. H. (2012). Relationship between concentric velocities at varying intensity in the back squat using a wireless inertial sensor. *Journal of Trainology*, *6*(1), 9-12. https://doi.org/10.17338/trainology.6.1_9

Cavarretta, D. J., Hall, E. E., & Bixby, W. R. (2019). The effects of increasing training load on affect and perceived exertion. *Journal of Strength and Conditioning Research,* 1-6. Advance online publication. https://doi.org/10.1519/JSC.0000000000003393

Deschenes, M. R., & Kraemer, W. J. (2002). Performance and physiologic adaptions to resistance training. *American Journal of Physical Medicine and Rehabilitation, 81*(11), 3 - 16. https://doi.org/10.1097/00002060-200211001-00003

Dorrell, H. F., Smith, M. F., & Gee, T. I. (2019). Comparison of velocity-based and traditional percentage-based loading methods on maximal strength and power adaptions. *Journal of Strength and Conditioning Research, 34*(1), 46-53. https://doi.org/10.1519/JSC.0000000000003089

Fisher, J., Steele, J., Bruce-Low, S., & Smith, D. (2011). Evidence-based resistance training recommendations. *Medicina Sportiva*, *15*(3), 147-162. https://doi.org/10.2478/v10036-011-0025-x

Fisher, J., Stuart, C., Steele, J., Gentil, P., & Giessing, J. (2018). Heavier and lighter load isolated lumbar extension resistance training produce similar strength increases, but different perceptual responses in healthy males and females. *PeerJ*, *6*(4), E6001. https://doi.org/10.7717/peerj.6001

Fisher, J. P., & Steele, J. (2017). Heavier and lighter load resistance training to momentary failure produce similar increases in strength with differing degrees of discomfort. *Muscle Nerve, 56*(4), 797 – 803. https://doi.org/10.1002/mus.25537

Gonzalez-Badillo, J. J., Marques, M. C., & Sanchez-Medina, L. (2011). The importance of movement velocity as a measure to control resistance training intensity. *Journal of Human Kinetics, 29A*, 15-19. https://doi.org/10.2478/v10078-011-0053-6

Gonzalez-Badillo, J. J., & Sanchez-Medina, L. (2010). Movement velocity as a measure of loading intensity in resistance training. *International Journal of Sports Medicine, 31*(5), 347-352. https://doi.org/10.1055/s-0030-1248333

Gonzalez-Badillo, J. J., Yanez-Garcia, J. M., Mora-Custodio, R., & Rodriguez-Rosell, D. (2017). Velocity loss as a variable for monitoring resistance exercise. *International Journal of Sports Medicine, 38*(3), 217-225. https://doi.org/10.1055/s-0042-120324

Graham, T., & Cleather, D. J. (2019). Autoregulation by "repetitions in reserve" leads to greater improvements in strength over a 12-week training program than fixed loading. *Journal of Strength & Conditioning Research.* Advance online publication. https://doi.org/10.1519/JSC.0000000000003164

Greig, L., Stephens Hemmingway, B. H., Aspe, R. R., Cooper, K., Comfort, P., & Swinton, P. A. (2020). Autoregulation in resistance training: Addressing the inconsistencies. *Sports Medicine, 50*(11), 1873-1887. https://doi.org/10.1007/s40279-020-01330-8

Hackett, D. A., Cobley, S. P., Davies, T. B., Michael, S. W., & Halaki, M. (2016). Accuracy in estimating repetitions to failure during resistance exercise. *Journal of Strength and Conditioning, 31*(8), 2162-2168. https://doi.10.1519/JSC.0000000000001683

Hackett, D. A., Cobley, S. P., & Halaki, M. (2018). Estimation of repetitions to failure for monitoring resistance exercise intensity: Building a case for application. *Journal of Strength and Conditioning Research, 32*(5), 1352 – 1359. https://doi.org/10.1519/JSC.0000000000002419

Hackett, D. A., Johnson, N. A., Halaki, M., & Chow, C. (2012). A novel scale to assess resistance-exercise effort. *Journal of Sports Sciences, 30*(13), 1405 – 1413. https://doi.org/10.1080/02640414.2012.710757

Hackett, D. A., Selvanayagam, V. S., Halaki, M., & Cobley, S. P. (2019). Associations between perceptual fatigue and accuracy of estimated repetitions to failure during resistance exercises. *Journal of Functional Morphology and Kinesiology, 4*(56). https://doi.org/10.3390/jfmk4030056

Haff, G. G., & Triplett, N. T. (2016). Essentials of strength training and conditioning. In J.M. Sheppard & N.T. Triplett (Eds.), *Program design for resistance training* (4th ed., pp. 439 – 468). Human Kinetics. National Strength and Conditioning Association.

Halperin, I., & Emanuel, A. (2019). Rating of perceived effort: Methodological concerns and future directions. *Sports Medicine, 50*(4), 679-687. https://doi.org/10.1007/s40279-019-01229-z

Hass, C. J., Feigenbaum, M. S., & Franklin, B. A. (2001). Prescription of resistance training for healthy populations. *Sports Medicine*, *31*(14), 953-964. https://doi.org/10.2165/00007256-200131140-00001

Helms, E. R., Brown, S. R., Cross, M. R., Storey, A., Cronin, J., & Zourdos, M. C. (2017). Self-rated accuracy of rating of perceived exertion-based load prescription in powerlifters. *Journal of Strength and Conditioning Research*, *31*(10), 2938-2943. https://doi.org/10.1519/JSC.0000000000002097

Helms, E. R., Cronin, J., Storey, A., & Zourdos, M. C. (2016). Application of the Repetitions in Reserve–Based Rating of Perceived Exertion Scale for Resistance Training. *National Strength and Conditioning Association, 38*(4), 42 – 49. https://doi.org/10.1519/SSC.0000000000000218

Helms, E. R., Kwan, K., Sousa, C. A., Cronin, J. B., Storey, A. G., & Zourdos, M. C. (2020). Methods for regulating and monitoring resistance training. *Journal of Human Kinetics, 74,* 23-42. https://doi.org/10.1007/s40279-019-01229-z

Helms, E. R., Storey, A., Cross, M. R., Brown, S. R., Lenetsky, S., Ramsay, H., Dillen, C., & Zourdos, M. C. (2017). RPE and velocity relationships for the back squat, bench press, and deadlift in powerlifters. *Journal of Strength and Conditioning Research*, *31*(2), 292-297. https://doi.org/10.1519/JSC.0000000000001517

International Business Machines. (n.d.). IBM SPSS advanced statistics 28 [PDF]. https://www.ibm.com/docs/en/SSLVMB_28.0.0/pdf/IBM_SPSS_Advanced_Statistics.pdf

Kraemer, W. J., & Ratamess, N. A. (2004). Fundamentals of resistance training: Progression and exercise prescription. *Medicine and Science in Sports and Exercise, 36*(4), 674-688. https://doi.org/10.1249/01.MSS.0000121945.36635.61

Kraemer, W. J., Ratamess, N. A., & French, D. N. (2002). Resistance training for health and performance. *Current Sports Medicine Reports,* 165-171. https://doi.org/10.1007/s11932-002-0017-7

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13–22.

Lovegrove, S., Hughes, L. J., Mansfield, S. K., Read, P. J., Price, P., & Patterson, S. D. (2021). Repetitions in reserve is a reliable tool for prescribing resistance training load. *Journal of Strength and Conditioning*, *XX* (X): 000–000. https://doi.org/10.1519/JSC.0000000000004158

Mangine, G. T., Serafini, P. R., Stratton, M. T., Olmos, A. A., VanDusseldrop, T. A., & Feito, Y. (2022). Effect of the repetitions in reserve resistance training strategy on bench press performance, perceived effort, and recovery in trained me. *Journal of Strength and Conditioning*, *36*(1), 1 – 9. https://doi.org/10.1519/JSC.0000000000004158

Mann, J. B., Thyfault, J. P., Ivey, P. A., & Slayers, S. P. (2010). The effect of autoregulatory progressive resistance exercise vs. linear periodization on strength improvement in college athletes. *Journal of Strength and Conditioning Research*, *24*(7), 1718-1723. https://doi.org/10.1519/JSC.0b013e3181def4a6

Mansfield, S. K., Peiffer, J. J., Hughes, L. J., & Scott, B. R. (2020). Estimating repetitions in reserve for resistance exercise: An analysis of factors which impact on prediction accuracy. *Journal of Strength and Conditioning Research*, 1 – 7. Advance online publication. https://doi.org/10.1519/JSC.0000000000003779

Micklewright, D., St Clair Gibson, A., Gladwell, V., & Al Salman, A. (2017). Development and validity of the Rating of Fatigue Scale. *Sports Medicine, 47*(11), 2375 – 2393. https://doi.org/10.1007/s40279-017-0711-5

Moore, J. M., & Dorrell, H. F. (2020). Guidelines and resources for prescribing load using velocity-based training. *International Universities Strength and Conditioning Association Journal, 1*(1), 1-14. https://doi.org/10.47206/iuscaj.v1i1.4

Ormsbee, M. J., Carzoli, J. P., Klemp, A., Allman, B. R., Zourdos, M. C., Kim, J., & Panton, L. B. (2017). Efficacy of the repetitions in reserve-based rating of perceived exertion for the bench press in experienced and novice benchers. *The Journal of Strength and Conditioning Research, 33*(2), 337 – 345. https://doi.org/10.1519/JSC.0000000000001901

Rauch, J. T., Ugrinowitsch, C., Barakat, C. I., Alvarez, M. R., Brummert, D. L., Aube, D. W., Barshun, A. S., Hayes, D., Tricoli, V., & DeSouza, E. O. (2017). Auto-regulated exercise selection training regimen produces small increases in lean body mass and maximal strength adaptions in strength trained individuals. *Journal of Strength and Conditioning Research, 34(*4), 1133-1140. https://doi.org/10.1519/JSC.0000000000002272

Rodriguez-Rosell, D., Yanez-Garcia, J. M., Sanchez-Medina, L., Mora-Custodio, R., & Gonzalez-Badillo, J. J. (2020). Relationship between velocity loss and repetitions in reserve in the bench press and back squat exercises. *Journal of Strength and Conditioning Research, 34*(9), 2537-2547. https://doi.org/10.1519/JSC.0000000000002881

Sanchez-Medina, L., & Gonzalez-Badillo, J. J. (2011). Velocity loss as an indicator of neuromuscular fatigue during resistance training. *Medicine & Science in Sports & Exercise, 43*(9), 1725-1734. https://doi.org/10.1249/MSS.0b013e318213f880

Seo, D. I., Kim, E., Fahs, C. A., Rossow, L., Young, K., Ferguson, S. L., Thiebaud, R., Sherk, V. D., Loenneke, J. P., Kim, D., Lee, M. K., Choi, K. H., Bemben, D. A., Bemben, M. G., & So, W. Y. (2012). Reliability of the one-repetition maximum test based on muscle group and gender. *Journal of Sports Science & Medicine*, *11*(2), 221–225.

Shattock, K., & Tee, J. C. (2020). Autoregulation in resistance training: A comparison of subjective versus objective methods. *Journal of Strength and Conditioning Research*, Advance online publication. https://doi.org/doi:10.1519/JSC.0000000000003530

Steele, J., Endres, A., Fisher, J., Gentil, P., & Giessing, J. (2017). Ability to predict repetitions to momentary failure is not perfectly accurate, though improves with resistance training experience. *PeerJ, 5*, 1 – 14. https://doi.org/10.7717/peerj.4105

Steele, J., Fisher, J., McKinnon, S., & McKinnon, P. (2016). Differentiation between perceived effort and discomfort during resistance training in older adults: Reliability of trainee ratings of effort and discomfort, and reliability and validity of trainer ratings of trainee effort. *Journal of Trainology*, *6*(1), 1 – 8. https://doi.org/10.17338/trainology.6.1_1

Stuart, C., Steele, J., Gentil, P., Giessing, J., & Fisher, J. P. (2018). Fatigue and perceptual responses of heavier and lighter load isolated lumbar extension resistance exercise in males and females. *PeerJ*, *6*(8), e4523. https://doi.org/10.7717/peerj.4523

Zhang, X., Li, A., Bi, S., Luo, Y., Cao, Y., & Zhang, G. (2021). Auto-regulation method vs. fixed-loading method in maximum strength training athletes: A systematic review and meta-analysis. *Frontiers in Physiology, 12,* 651112. https://doi.org/10.3389/fphys.2021.651112

Zourdos, M. C., Goldsmith, J. A., Helms, E. R., Trepeck, C., Halle, J. L., Mendez, K. M., Cooke, D. M., Haischer, M. H., Sousa, C. A., Klemp, A., & Byrnes, R. K. (2019). Proximity to failure and total repetitions performed in a set influences accuracy of intraset repetitions in reserve-based rating of perceived exertion. *Journal of Strength and Conditioning Research, 30*(2), 158 – 165. https://doi:10.1519/JSC.0000000000002995

Zourdos, M. C., Klemp, A., Dolan, C., Quiles, J. M., Schau, E. J., Helms, E., Esgro, B., Duncan, S., Garcia Merino, S., & Blanco, R. (2015). Novel resistance training – specific rating of perceived exertion scale measuring repetitions in reserve. *Journal of Strength and Conditioning Research, 30*(1), 267 – 275. https://doi.org/10.1519/JSC.0000000000001049

Appendix A: Vocabulary Handout

Repetitions in reserve (RIR):

The ratio of repetitions completed in relation to the maximum number of repetitions that "can" be completed. For example, selecting a load that is consistent with a 10RM, but the set is stopped short at the eighth repetition, the participants may have completed two more repetitions at this particular load. Hence, in this scenario, two repetitions were left in reserve.

Technical Failure:

"Participant can no longer overcome the demands of the lift, regardless of effort (with proper technique). (Fisher & Steele 2017)

Discomfort:

"Physiological and unpleasant sensations associated with exercise." (Steele et al., 2016).

Fatigue:

"A feeling of diminishing capacity to cope with physical or mental stressors, either imagined or real." (Micklewright et al., 2017)

"Effort":

Fatigue and effort are suggested to have a strong correlation as they both seem to increase "together." As muscular fatigue increases, the ability to continually produce similar levels of force becomes progressively more difficult – thus, one has to put forth more effort. (Halperin and Emanuel, 2019)

Memory anchor:

The purpose of memory anchors are to connect past RT sessions with current RT sessions to assist one in identifying training intensities. Memory anchors rely on the ability to recall training at minimal (1 RPE or 10 RIR) and maximal intensities (10 RPE or 0 RIR) to help identify current RT intensities.

Appendix B: The RIR scale, Handout

Repetitions in Reserve (RIR) Scale	
0	**Max effort!** Could not have done any more reps.
1	Could have done **+1** more rep
2	Could have done **+2** more reps
3	Could have done **+3** more reps
4	Could have done **+4** more reps
5	Could have done **+5** more reps

Table 1. Protocol Group (PG): Session Training Intensities

Protocol Group – Training intensities					
	Session 1	**Session 2**	**Session 3**	**Session 4**	**Session 5**
RM	--	12	15	6	--
Repetitions	8 repetitions	10 repetitions	12 repetitions	6 repetitions	8 repetitions
pRIR	1	2	2	1	2

Table 2. Non-Protocol Group (NPG): Session Training Intensities

Non-Protocol Group – Training intensities					
	Session 1	**Session 2**	**Session 3**	**Session 4**	**Session 5**
Repetitions	8 repetitions	10 repetitions	12 repetitions	6 repetitions	8 repetitions
pRIR	1	2	2	1	2

Table 3. Pairwise Session Comparisons of Absolute Difference of pRIR - aRIR for the Protocol Group

Session I	Session J	Mean Difference (I-J)	Std. Error of Differences	Bonferroni-corrected p-value
Bench press				
1	2	**-0.493**	0.142	0.005
	3	-0.044	0.109	> .999
	4	0.254	0.125	0.429
	5	**0.312**	0.109	0.042
2	1	**0.493**	0.142	0.005
	3	**0.448**	0.101	< .0001
	4	**0.746**	0.117	< .0001
	5	**0.804**	0.080	< .0001
3	1	0.044	0.109	> .999
	2	**-0.448**	0.101	< .0001
	4	**0.298**	0.078	0.001
	5	**0.356**	0.050	< .0001
4	1	-0.254	0.125	0.429
	2	**-0.746**	0.117	< .0001
	3	**-0.298**	0.078	0.001
	5	0.058	0.042	> .999
5	1	**-0.312**	0.109	0.042
	2	**-0.804**	0.080	< .0001
	3	**-0.356**	0.050	< .0001
	4	-0.058	0.042	> .999
V-squat				
1	2	0.201	0.277	> .999
	3	0.494	0.221	0.256
	4	0.643	0.258	0.129
	5	**0.848**	0.268	0.016
2	1	-0.201	0.277	> .999
	3	0.292	0.167	0.792
	4	0.441	0.196	0.242
	5	**0.647**	0.196	0.010
3	1	-0.494	0.221	0.256
	2	-0.292	0.167	0.792
	4	0.149	0.056	0.081
	5	**0.354**	0.055	< .0001

(table continues)

Session I	Session J	Mean Difference (I-J)	Std. Error of Differences	Bonferroni-corrected p-value
4	1	-0.643	0.258	0.129
	2	-0.441	0.196	0.242
	3	-0.149	0.056	0.081
	5	**0.205**	0.046	< .0001
5	1	**-0.848**	0.268	0.016
	2	**-0.647**	0.196	0.010
	3	**-0.354**	0.055	< .0001
	4	**-0.205**	0.046	< .0001

Bolded mean differences are statistically significant at the .05 significance level after correcting for Bonferroni adjustment.
Pairwise tests based on Wald Chi-squared test statistics.

Table 4. Pairwise Session Comparisons of Probability of Estimating RIR for the Protocol Group

Session I	Session J	Difference in Probabilities (I-J)	Std. Error of Differences	Bonferroni-corrected p-value
Bench press				
1	2	0.250	0.119	0.367
	3	0.042	0.106	> .999
	4	-0.249	0.123	0.426
	5	**-0.310**	0.106	0.032
2	1	-0.250	0.119	0.367
	3	-0.207	0.085	0.147
	4	**-0.500**	0.072	< .0001
	5	**-0.560**	0.046	< .0001
3	1	-0.042	0.106	> .999
	2	0.207	0.085	0.147
	4	**-0.290**	0.077	0.002
	5	**-0.350**	0.052	< .0001
4	1	0.249	0.123	0.426
	2	**0.500**	0.072	< .0001
	3	**0.290**	0.077	0.002
	5	-0.063	0.041	> .999
5	1	**0.310**	0.106	0.032
	2	**0.560**	0.046	< .0001
	3	**0.350**	0.052	< .0001
	4	0.063	0.041	> .999
V-squat				
1	2	-0.050	0.148	> .999
	3	-0.200	0.109	0.672
	4	-0.351	0.152	0.207
	5	**-0.550**	0.148	0.002
2	1	0.050	0.148	> .999
	3	-0.150	0.114	> .999
	4	-0.301	0.130	0.206
	5	**-0.500**	0.122	0.0004
3	1	0.200	0.109	0.672
	2	0.150	0.114	> .999
	4	-0.150	0.055	0.059
	5	**-0.350**	0.055	< .0001

(table continues)

Session I	Session J	Difference in Probabilities (I-J)	Std. Error of Differences	Bonferroni-corrected p-value
4	1	0.351	0.152	0.207
	2	0.301	0.130	0.206
	3	0.150	0.055	0.059
	5	**-0.200**	0.045	< .0001
5	1	**0.550**	0.148	0.002
	2	**0.500**	0.122	0.0004
	3	**0.350**	0.055	< .0001
	4	**0.200**	0.045	< .0001

Bolded differences in probabilities are statistically significant at the .05 significance level after correcting for Bonferroni adjustment. Pairwise tests based on Wald Chi-squared test statistics.

Table 5: Model Estimates for Absolute Difference of pRIR - aRIR

Variable	Parameter estimate	Std. error	Wald Chi-Square	df	*p*-value
Model 1 for Bench press					
(Intercept)	0.953	0.0920	107.303	1	<.0001
[Session=1]	-0.140	0.0409	11.659	1	0.001
[Session=5]	0				
[Group=1]	-0.808	0.1062	57.879	1	<.0001
[Group=2]	0				
[Group=1] * [Session=1]	0.461	0.1022	20.317	1	<.0001
[Group=1] * [Session=5]	0				
[Group=2] * [Session=1]	0				
[Group=2] * [Session=5]	0				
Model 2 for V-squat					
(Intercept)	1.472	0.288	26.184	1	<.0001
[Session=1]	0.023	0.092	0.064	1	0.800
[Session=5]	0				
[Group=1]	-1.361	0.295	21.314	1	<.0001
[Group=2]	0				
[Group=1] * [Session=1]	0.825	0.297	7.715	1	0.005
[Group=1] * [Session=5]	0				
[Group=2] * [Session=1]	0				
[Group=2] * [Session=5]	0				

Based on generalized estimating equations using a normal distribution with identity link and AR(1) correlation structure.

Table 6: Model Estimates for Probability of Being Accurate in Estimating RIR (i.e., pRIR = aRIR)

Variable	Parameter estimate	Std. error	Wald Chi-Square	df	*p*-value
Model 1 for Bench press					
(Intercept)	-0.934	0.048	378.870	1	<.0001
[Session=1]	0.128	0.245	0.271	1	0.603
[Session=5]	0				
[Group=1]	2.729	0.430	40.260	1	<.0001
[Group=2]	0				
[Group=1] * [Session=1]	-1.820	0.585	9.684	1	0.002
[Group=1] * [Session=5]	0				
[Group=2] * [Session=1]	0				
[Group=2] * [Session=5]	0				
Model 2 for V-squat					
(Intercept)	-1.961	1.024	3.667	1	0.056
[Session=1]	1.169	0.523	5.004	1	0.025
[Session=5]	0				
[Group=1]	4.125	1.198	11.848	1	0.001
[Group=2]	0				
[Group=1] * [Session=1]	-3.942	1.108	12.654	1	0.0004
[Group=1] * [Session=5]	0				
[Group=2] * [Session=1]	0				
[Group=2] * [Session=5]	0				

Based on generalized estimating equations using a binomial distribution with logit link and AR(1) correlation structure. Logit is log(p/(1-p)) where p is the probability of being accurate (i.e., pRIR = aRIR).

Table 7. Model Estimates for Absolute Difference of pRIR – aRIR by Session for Protocol Group

Variable	Parameter estimate	Std. error	Wald Chi-Square	df	p-value
Model 1 for Bench press					
(Intercept)	0.145	0.054	7.268	1	0.007
[Session=1]	0.312	0.109	8.194	1	0.004
[Session=2]	0.804	0.080	99.872	1	<.0001
[Session=3]	0.356	0.050	49.828	1	<.0001
[Session=4]	0.058	0.042	1.920	1	0.166
[Session=5]	0				
Model 2 for V-squat					
(Intercept)	0.098	0.054	3.269	1	0.071
[Session=1]	0.848	0.268	9.982	1	0.002
[Session=2]	0.647	0.196	10.862	1	0.001
[Session=3]	0.354	0.055	41.029	1	<.0001
[Session=4]	0.205	0.046	19.653	1	<.0001
[Session=5]	0				

Based on generalized estimating equations using a normal distribution with identity link and AR(1) correlation structure.

Table 8. Model Estimates for Probability of Being Accurate (pRIR = aRIR) by Session for Protocol Group

Variable	Parameter estimate	Std. error	Wald Chi-Square	df	p-value
Model 1 for Bench press					
(Intercept)	1.786	0.434	16.959	1	<.0001
[Session=1]	-1.605	0.596	7.243	1	0.007
[Session=2]	-2.655	0.217	149.858	1	<.0001
[Session=3]	-1.775	0.324	30.091	1	<.0001
[Session=4]	-0.438	0.316	1.917	1	0.166
[Session=5]	0				
Model 2 for V-squat					
(Intercept)	2.195	0.608	13.026	1	0.0003
[Session=1]	-2.815	0.956	8.672	1	0.003
[Session=2]	-2.602	0.741	12.346	1	0.0004
[Session=3]	-1.994	0.559	12.704	1	0.0004
[Session=4]	-1.345	0.325	17.108	1	<.0001
[Session=5]	0				

Based on generalized estimating equations using a binomial distribution with logit link and AR(1) correlation structure. Logit is log(p/(1-p)) where p is the probability of being accurate (i.e., pRIR = aRIR).

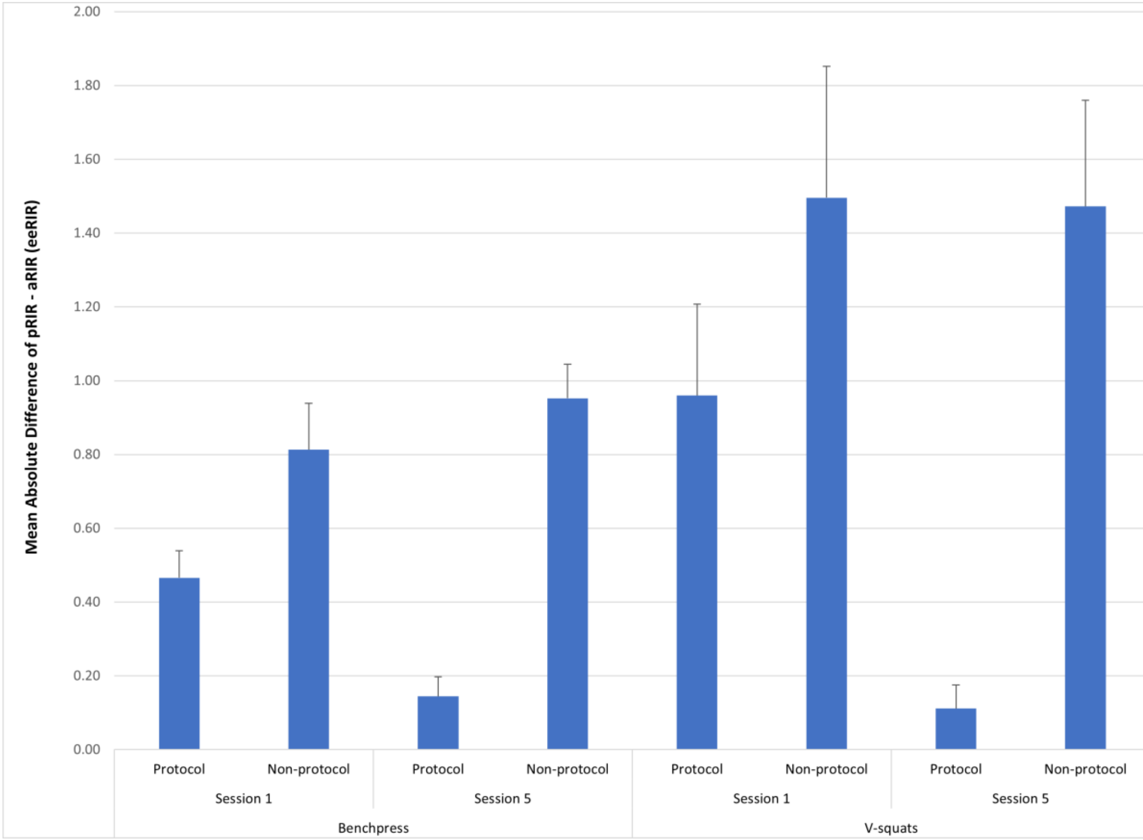

Figure 1. Mean Absolute Difference of pRIR – aRIR by Group, Session, and Movement

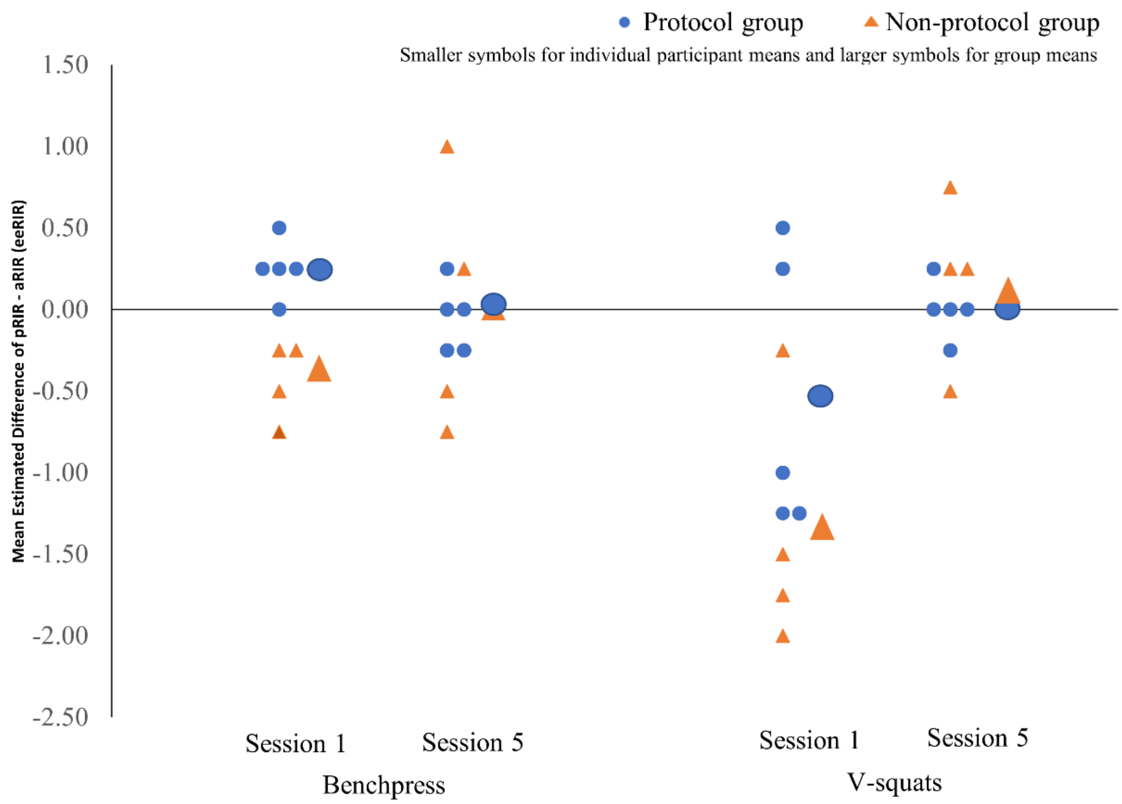

Figure 2. Mean Difference of pRIR – aRIR by Group, Session, and Movement

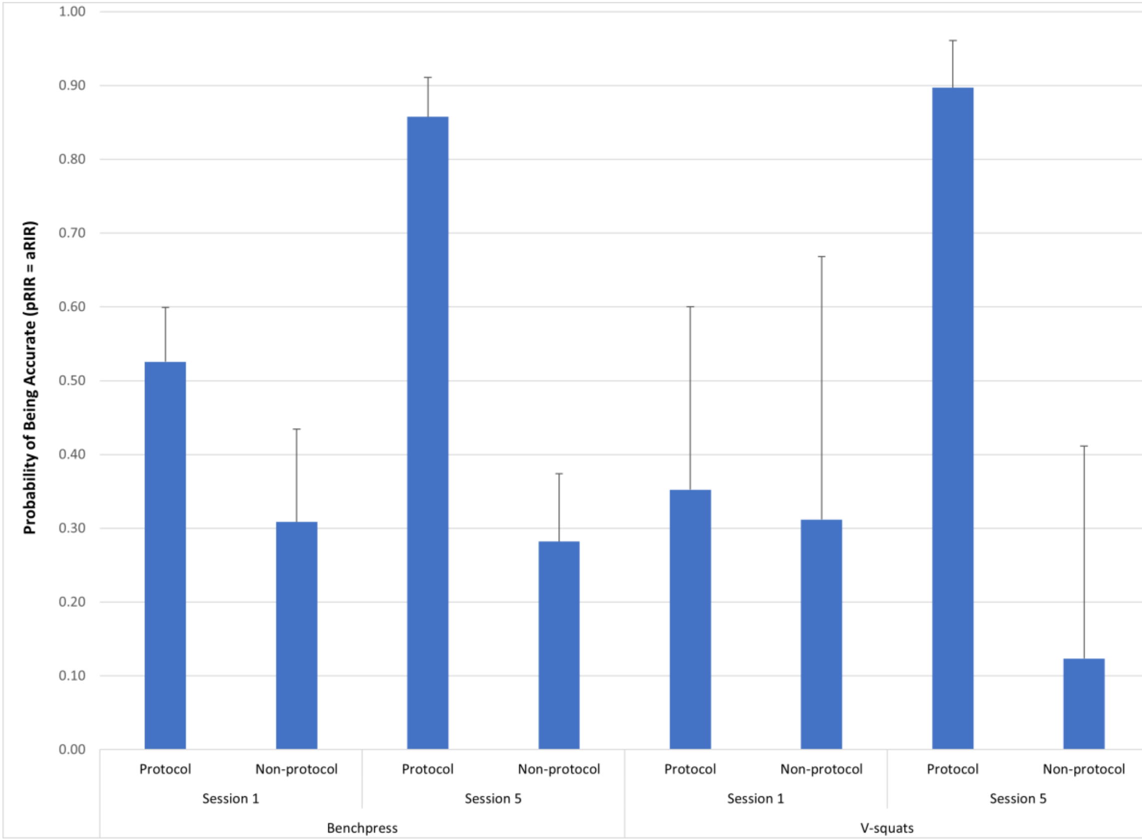

Figure 3. Probability of Being Accurate (pRIR = aRIR) by Group, Session, and Movement

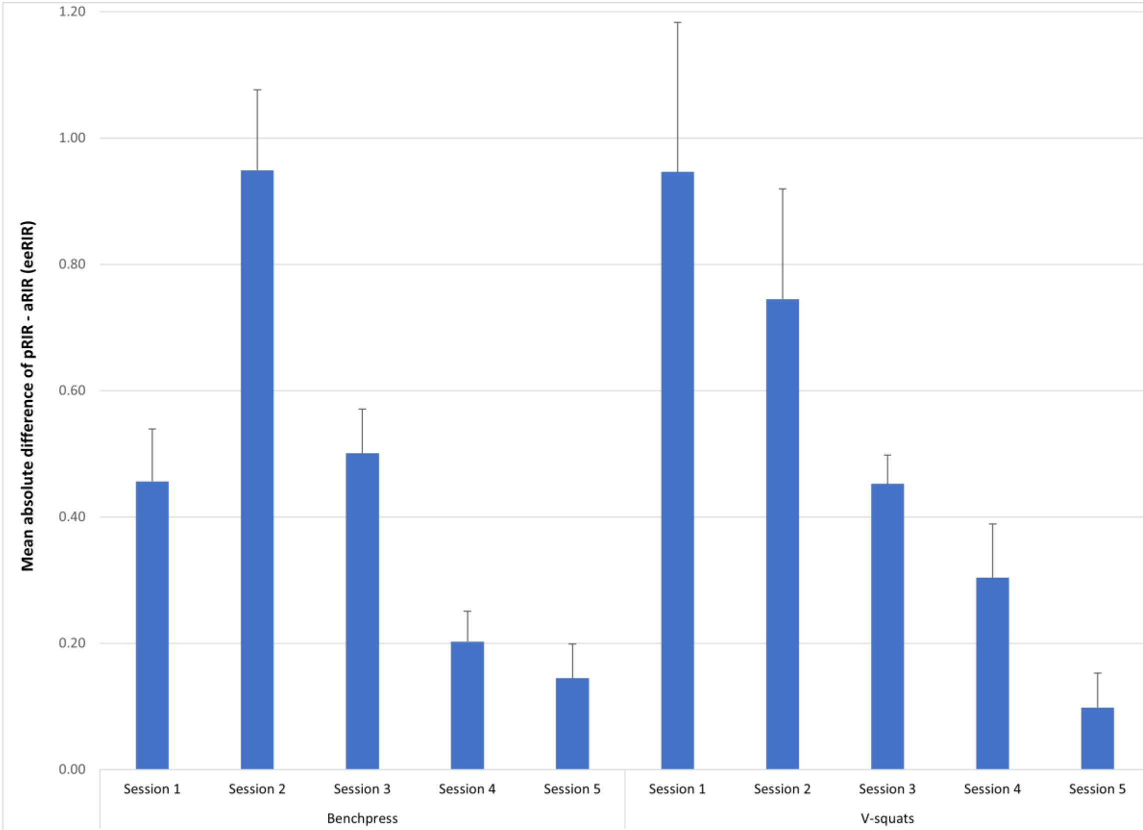

Figure 4. Mean Absolute Difference of pRIR – aRIR by Session and Movement for the Protocol Group

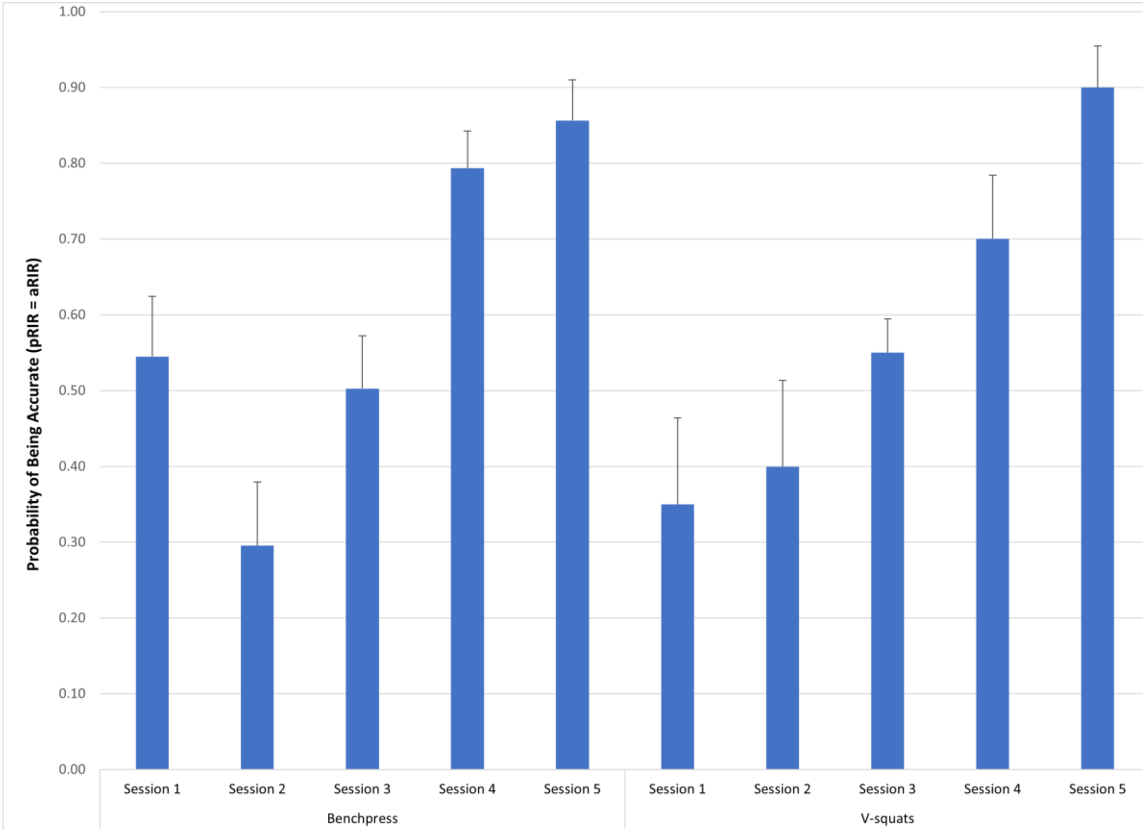

Figure 5. Probability of Being Accurate (pRIR = aRIR) by Session and Movement for the Protocol Group

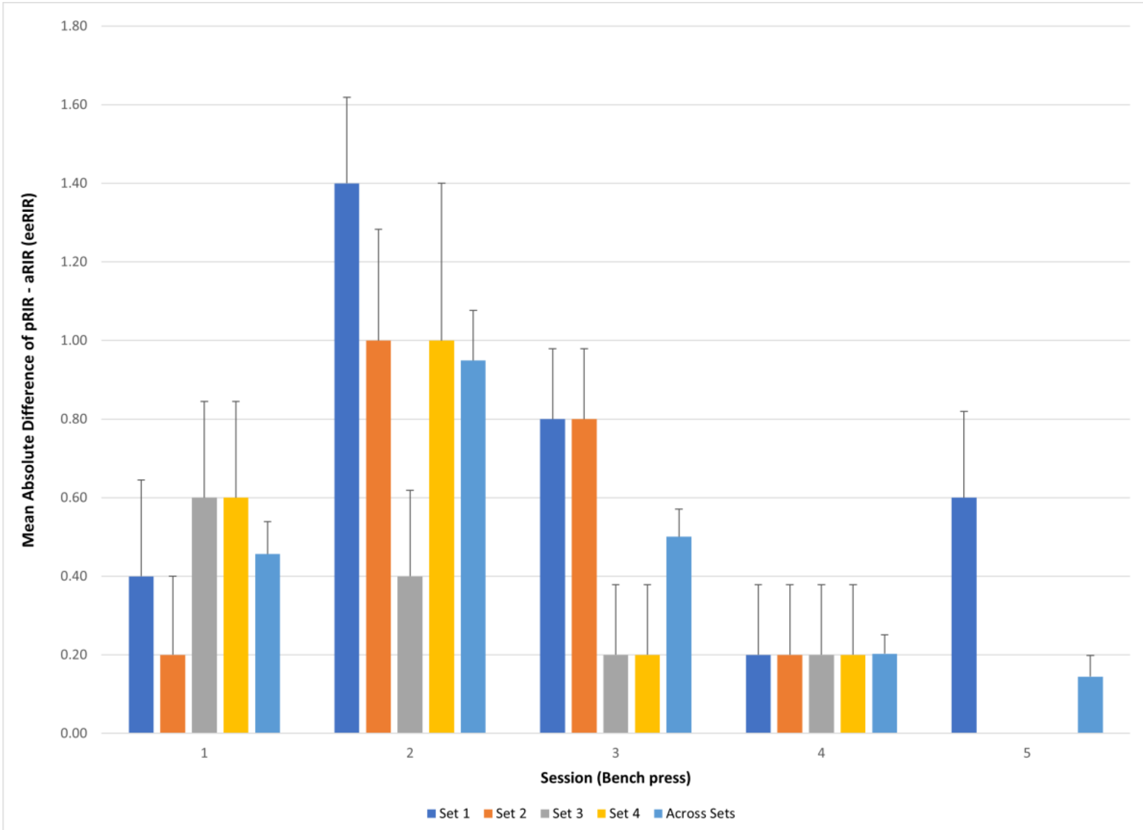

Figure 6. Mean Absolute Difference of pRIR – aRIR by Session and Set for the Protocol Group (Bench press)

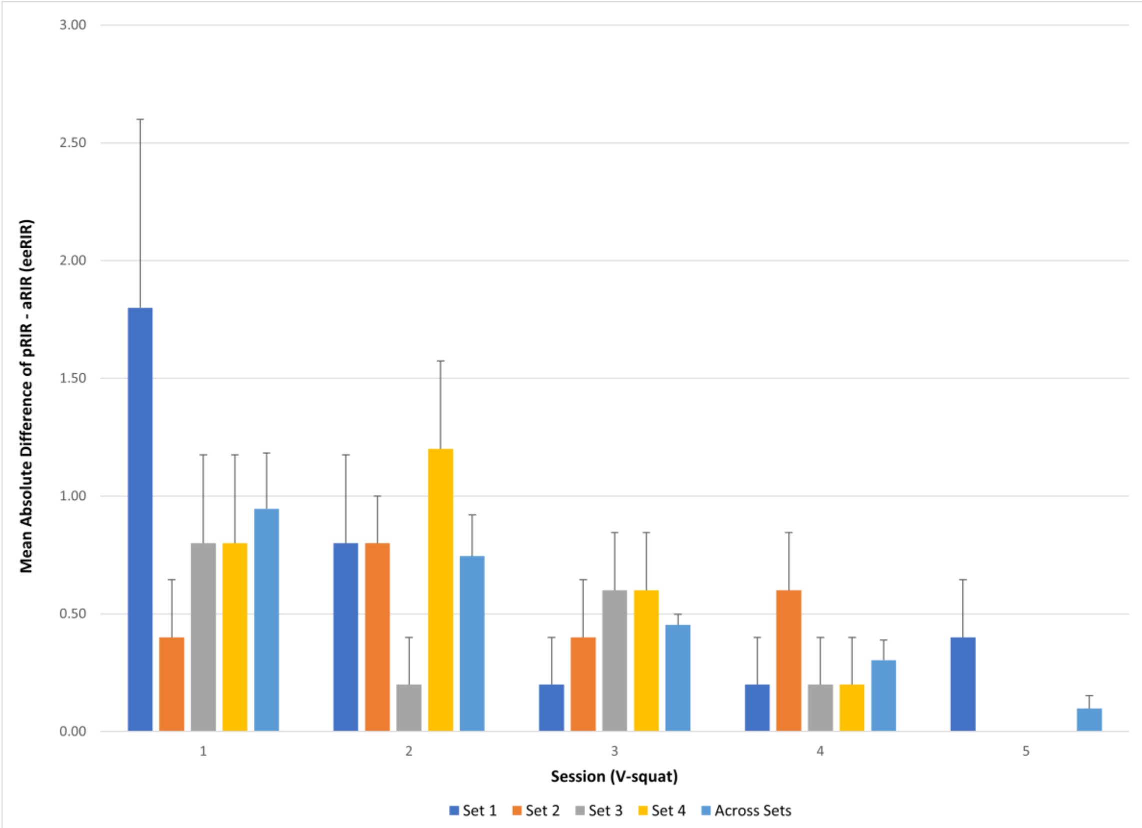

Figure 7. Mean Absolute Difference of pRIR – aRIR by Session and Set for the Protocol Group (V-Squat)

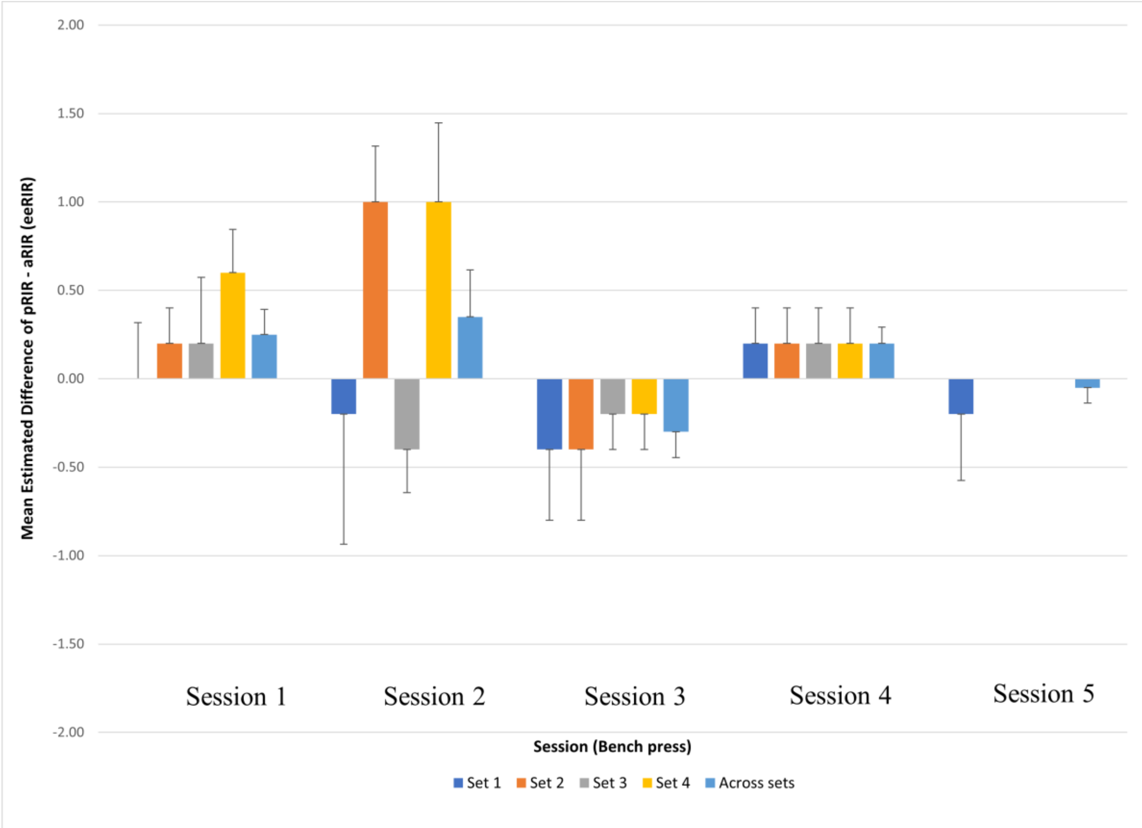

Figure 8. Mean Difference of pRIR – aRIR by Session and Set for the Protocol Group (Bench press)

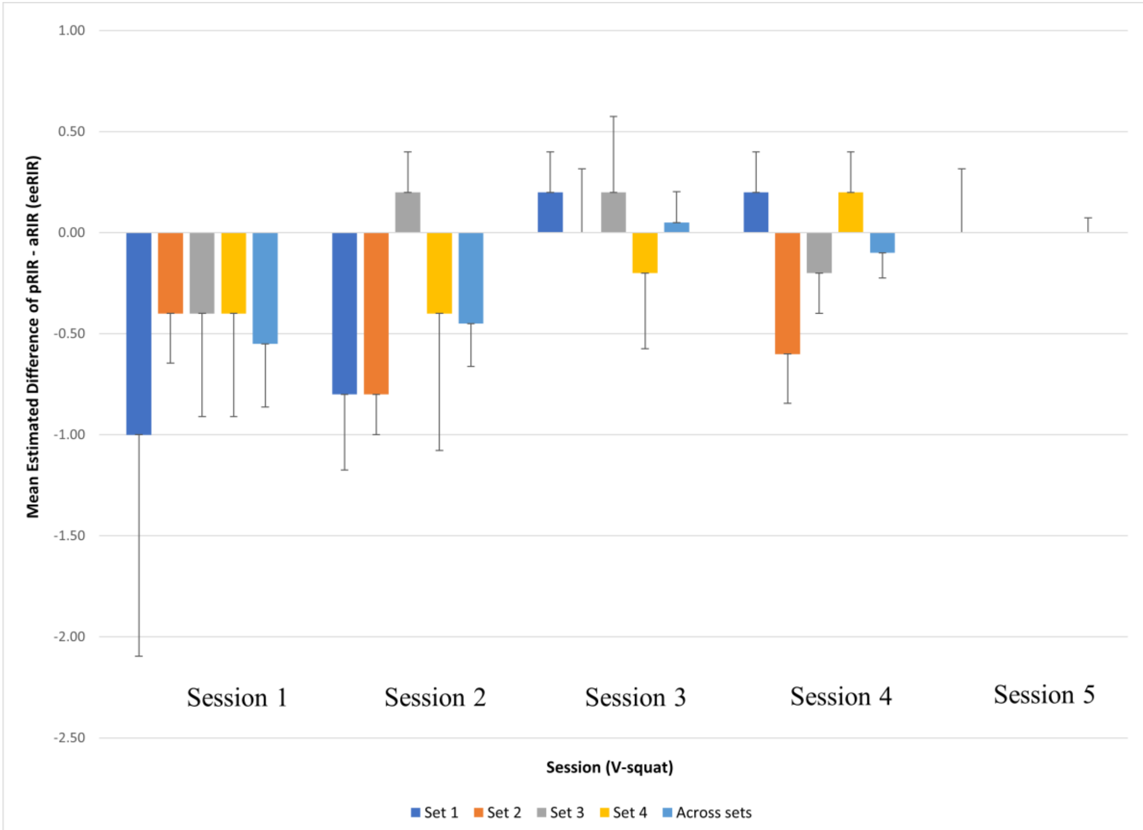

Figure 9. Mean Difference of pRIR – aRIR by Session and Set for the Protocol Group (V-Squat)

Chart 1: Mean Absolute Difference of pRIR - aRIR by Movement and Session for Protocol Group

Movement	Session	Mean	Std. Error
Bench press	Session 1	0.456	0.083
	Session 2	0.949	0.127
	Session 3	0.501	0.070
	Session 4	0.203	0.048
	Session 5	0.145	0.054
V-squats	Session 1	0.946	0.237
	Session 2	0.745	0.175
	Session 3	0.453	0.045
	Session 4	0.304	0.085
	Session 5	0.098	0.054

Chart 2: Probability of Being Accurate (i.e., pRIR = aRIR) by Movement and Session for Protocol Group

Movement	Session	Probability	Std. Error
Bench press	Session 1	0.545	0.079
	Session 2	0.295	0.084
	Session 3	0.503	0.070
	Session 4	0.794	0.049
	Session 5	0.856	0.053
V-squats	Session 1	0.350	0.114
	Session 2	0.400	0.114
	Session 3	0.550	0.045
	Session 4	0.701	0.083
	Session 5	0.900	0.055