SportRxiv

Part of the Society for Transparency, Openness and Replication in Kinesiology (STORK)

Preprint
not peer reviewed

Optimizing Research Methodology for the Detection of Individual Response Variation in Resistance Training

For correspondence:
zrobinson2019@fau.edu

Zac P. Robinson[1], Eric R. Helms[1,2], Eric T. Trexler[1,3], James Steele[1,4], Michael E. Hall[1], Chun-Jung Huang[1], Michael C. Zourdos[1]

[1] Department of Exercise Science and Health Promotion, Muscle Physiology Laboratory, Florida Atlantic University

[2] Sports Performance Research Institute New Zealand (SPRINZ), Auckland University of Technology, Auckland, New Zealand

[3] Department of Evolutionary Anthropology, Duke University, Durham, NC, USA

[4] Faculty of Sport, Health, and Social Sciences, Solent University, Southampton, UK

Please cite as: Robinson ZP, Helms ER, Trexler ET, Steele J, Hall ME, Huang CJ, and Zourdos MC, (2023). Optimizing Research Methodology for the Detection of Individual Response Variation in Resistance Training

All authors have read and approved this version of the manuscript. This article was last modified on Month October, 2023.

ABSTRACT

Most resistance training research focuses on group-level outcomes (i.e., group A versus group B). However, many practitioners are more interested in training responses on the individual level (i.e., intervention A versus intervention B for individual X). In order to properly examine individual response variation, multiple confounding sources of variation (e.g., random sampling variation, measurement error, biological variation) must be addressed. Novel study designs where participants complete both interventions and at least one intervention twice can be leveraged to account for these sources of variation. Specifically, the appropriate statistical methods can separate variability into the signal (i.e., participant-by-training interaction) versus the noise (i.e., within-participant variance). This distinction can allow researchers to detect evidence of individual response variation. If evidence of individual response variation exists, researchers can explore potential predictors of the more favorable intervention, thereby improving exercise prescription. This review outlines the methodology necessary to explore individual response variation to resistance training, predict favorable interventions, and the limitations thereof.

INTRODUCTION

It is well established that, on the group-level, resistance training (RT) leads to robust increases in muscle size and strength [1]. These neuromuscular adaptations are thought to be influenced, in part, by training variables such as volume (number of sets per week for a muscle group) and load (percentage of one-repetition maximum; %1RM). Indeed, meta-analyses establish general training recommendations for volume [2,3] and load [4,5] based on group-level data. However, some argue that uniform training recommendations derived from such data may be inappropriate for some individuals due to the variability observed in RT outcomes1

The variability in training outcomes following a standardized RT intervention is well illustrated in widely cited studies from Hubal et al. [6] and Erskine et al. [7]. Specifically, Hubal et al. reported large mean increases in elbow flexor muscle cross-sectional area (CSA) (3.2 ± 0.1 cm^2; +18.9%), maximal elbow flexion isometric force (7.5 ± 0.3 kg; +19.5%) and 1RM preacher curl (+3.9 ± 0.1 kg; +54.1%) among 585 untrained men and women following a 12-week training program. However, the individual observations for cross-sectional area, isometric force, and 1RM strength ranged from -2.5% to 59.3%, -31.5% to +148.5%, and 0% to +250%, respectively. Similarly, in untrained individuals, Erskine et al. [7] observed mean changes of +33.9 kg and +14 cm^2 for knee extension 1RM and quadriceps CSA accompanied by a wide range of individual outcomes 18% to 113% (1RM) and +0% to +16% (CSA). Further, this outcome heterogeneity seems to extend to trained participants. To illustrate, Figure 1 displays the individual observations of four similarly designed longitudinal studies from our lab [8–11] (Figure 1).

1 There is a critical distinction between individual-level variability in RT *outcomes* and *responses*. The former purely describes the result of the training intervention but fails to contextualize the data relative to crucial confounding variables (e.g., random sampling variation, measurement error, biological variation, etc). Individual *"responses"*, and variation thereof, can only be determined in the latter case.

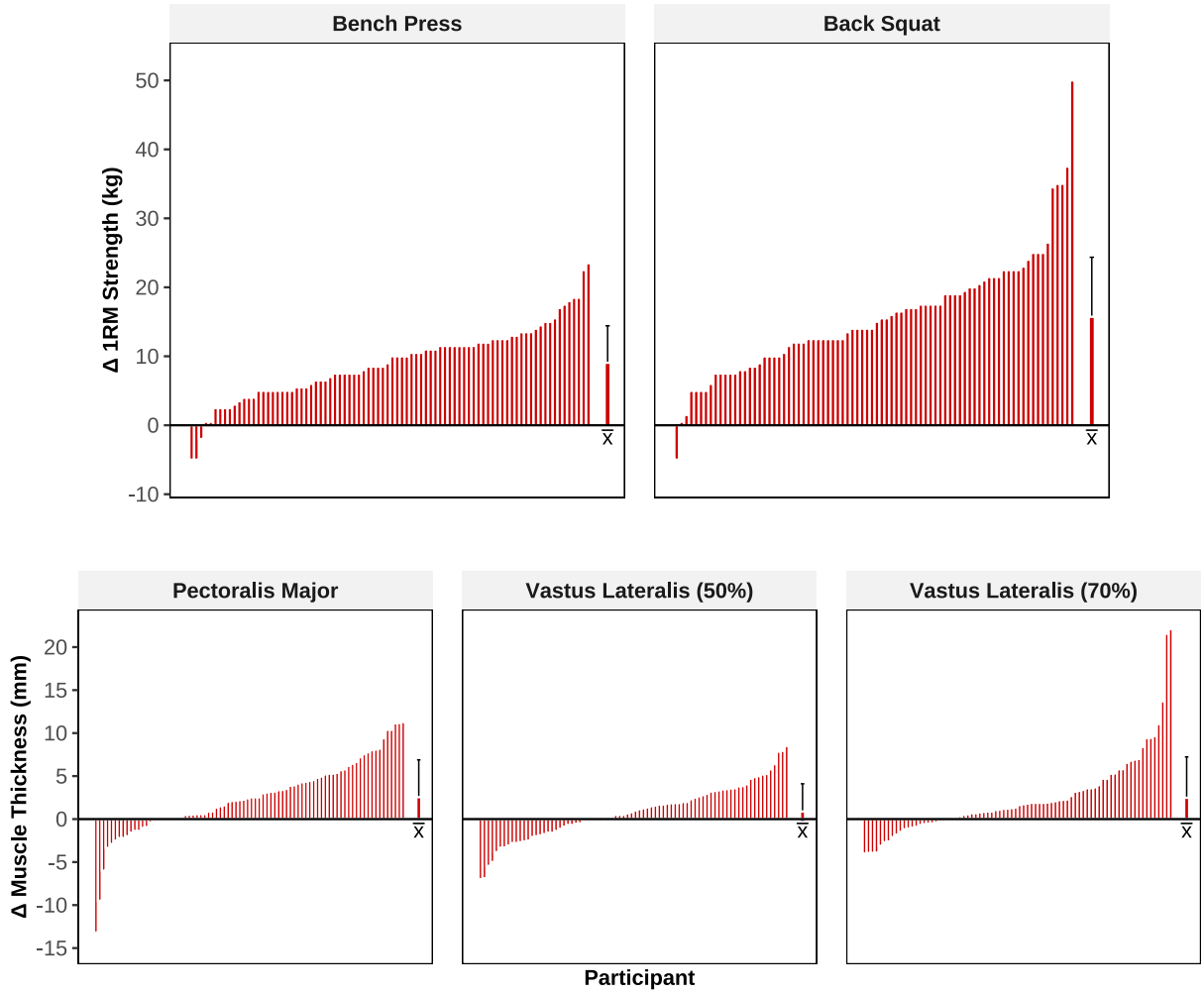

Figure 1: *Visual summary of response heterogeneity from four similarly designed longitudinal training studies [8–11]. Data are absolute change in outcome with each bar representing a single participant. $\bar{x}$= the mean ± the standard deviation of the outcome. A) n = 84; $\bar{x}$ = +9.06 kg ± 5.36; range = -5.0 to +23.5 kg B) n = 82; $\bar{x}$ = +15.73 kg ± 8.63; range = -5.0 to 50.0 kg C) n = 80; $\bar{x}$ = 2.56 mm ± 4.33; range = -13.2 to 11.3 mm D) n = 78; $\bar{x}$ = 0.88 mm ± 3.22; range = -7.0 to 8.5 mm E) n = 78; $\bar{x}$ = 2.48 mm ± 4.75; range = -4.0 to 22.1 mm*

Despite this observed outcome heterogeneity, targeted methods are necessary to confirm the causal link between a training intervention and the observed variation in outcomes, rather than other potential sources of variation (e.g., random sampling variation, measurement error, biologiclar variation, etc.). Indeed, Hecksteden et al. [12] described that various factors that can interfere with the investigation of individual response heterogeneity in typical study designs and outlined potential modifications for researchers to appropriately explore this phenomena.

Moreover, the authors discuss methods to predict individual response variation (if there is sufficient evidence to suggest it may exist) and the limitations thereof. In a novel approach that integrates some of these modifications within the field of exercise science, Steele et al. found conflicting evidence of individual response heterogeneity upon comparing the observed variances in groups completing a RT intervention versus untrained controls [13]; however, the analysis is incapable of delineating within-participant variance from the participant-by-training interaction. Moreover, whether these findings apply to the comparisons between different interventions (e.g., high versus low volumes), unexplored intervention duration (e.g., years rather than months or weeks), and in trained individuals remains unclear.

RT research that utilizes the appropriate methodology is needed to investigate individual response heterogeneity. With sufficient data, recommendations based on individual-level outcomes could be created. Therefore, the purpose of this review is to describe the limitations of typical RT study designs for examining individual response variation and suggest potential solutions. Additionally, we will explore methods to identify any predictors of individual response variation. To facilitate the application of these concepts, we will use examples that investigate individual responses in strength and hypertrophy to different RT volumes; although the concepts may apply to all training variables (e.g., frequency, load, proximity to failure, etc.) and outcomes (e.g., power, speed, etc.). Finally, although the existence of meaningful individual response variation remains unclear, this review will focus on strategies of examining it if sufficient evidence is observed.

EPISTEMOLOGY OF TRAINING RESPONSE: ASKING THE RIGHT QUESTION

To effectively investigate individual response variation, the research question of interest must be clearly defined due to implications for study design. Most RT studies are designed to answer questions related to group-level outcomes (i.e, average treatment effect). For example, a typical parallel-group study may randomize participants into one of two groups: low volume (e.g., 8 sets/wk) or high volume (e.g., 16 sets/wk). Muscle size and strength will be measured before and after the 12-week training intervention and the mean change scores of each condition will be compared. This study is designed to answer:

What training volume will produce the best average outcome when performed by many individuals?

That question may be appropriate when uniform program design is logistically necessary, such as strength and conditioning coaches managing many athletes simultaneously. However, the potentially more useful question for many RT practitioners would be:

What training volume will produce the best outcome when performed by a specific individual?

In a parallel-group design comparing low volumes (e.g., 8 sets per week) to high volumes (e.g., 16 sets per week), between group analyses may suggest that one condition leads to superior outcomes, on average. However, that finding would not imply that the same condition would optimize outcomes for every individual. In order to answer the question related to individual-level outcomes, two additional sub-components must be evaluated; i) *Is there sufficient evidence of individual response variation? (i.e., does the variance associated with the training intervention exceed that of within-participant variance?)* ii) *Can individual response variation be predicted? (i.e., if there is sufficient evidence of response variation, can practical characteristics identify when a given individual will likely respond better to one intervention versus another)* To effectively evaluate these sub-components, appropriate study design is necessary.

STUDY DESIGNS TO EXAMINE INDIVIDUAL RESPONSE VARIATION

When researchers intend to investigate a specific question, they should choose the study design best fit to answer that question. As described earlier, traditional parallel-group designs, which randomize or counterbalance participants to only one intervention group (e.g., high or low volumes), are limited in investigating individual response variation. However, modifications can be made to a parallel-group design to improve its ability to investigate individual response variation, and other study designs can be utilized to examine this topic (Table 1).

Study Designs for Investigating Individual-Level Response to Resistance Training

Type	*Separates Training Associated Variance*	*Separates Participant by Training Interaction*	*Supplies Individual Difference for Prediction*	*Limitations*	*Example References*
Parallel Groups	Only with a non-training control or repeated testing	Only with repeated testing	No	-Need to recruit more total participants -Lower control of external variables	Aube et al. (2022)
Crossover	Only with a non-training control or repeated testing	Only with repeated testing	Yes	-Lower control of external variables -Second-order effects -Time intensive for researchers and participants	Fortes et al. (2018)
Within-participant Unilateral	Only with a non-training control or repeated testing	Only with repeated testing	Yes	-Cannot experimentally manipulate bilateral exercises -Possibility of the cross-education effect	Scarpelli et al. (2022)
Crossover with Replication	Yes	Yes	Yes	-The most time intensive design -Second-order effects	Simoneau et al. (1987)
Within-participant Unilateral with Replication	Yes	Yes	Yes	-Cannot experimentally manipulate bilateral exercises -Possibility of the cross-education effect	N/A

Table 1: *Features of potential study designs to investigate the individual response to resistance training. [14–17]*

Parallel-Group Design with Non-Training Control

Although parallel-group designs typically offer extremely limited information pertaining to individual response variation, modifications can be made to give researchers increased access to individual-level information. In most RT studies, researchers do not include a non-training control group. However, inclusion of a non-training control would allow researchers to, in part, evaluate the first sub-component (i.e., is there sufficient evidence of individual response variation?) Specifically, by comparing the variance of the change scores between the RT group and non-training control, it can be determined if there is evidence for individual response variation. Most sources of variability (e.g., measurement error, random sampling variation, biological variation, etc.) would be shared by both groups, with additional variability attributed to the RT intervention if individual response variation does exist. This approach was demonstrated meta-analytically in the previously mentioned paper by Steele et al. [13]. Despite a non-training control improving a parallel-group design, it still cannot fully address the first sub-component due to the inability of this study design to delineate between different sources of training associated variance (i.e., within-participant variance versus the participant-by-training interaction). It merely demonstrates whether or not the introduction of the

intervention has introduced additional variance over and above the test-retest variation over the period of observation.

Further, the inclusion of a control group also comes with the logistical limitation of trained participants consenting to not exercise for the duration of the study (which in this population also represents an 'intervention' albeit one of de-training). Moreover, a parallel-group design with a non-training control cannot evaluate the response variation between different RT interventions; thus, it is incapable of addressing the second sub component to improve individual-level RT prescription (i.e., can we predict individual response variation?) To rectify these limitations, study designs in which participants complete multiple training interventions must be leveraged.

Crossover Designs

A longitudinal crossover design offers multiple advantages over a parallel-group design when investigating the individual response to training. In a crossover design, participants complete one training condition (i.e., period), enter a washout period intended to remove any residual effects from the first condition, and then complete the other condition. Further, the order in which conditions are completed is randomized to control for the order effect, which stipulates that an individual's response throughout the second condition may be impacted by effects from the first condition (i.e., potentiation or inhibition). In a crossover design, participants complete both training conditions and serve as their own control, which allows researchers to compare individual outcomes between conditions. For example, consider a study in which participants are initially randomized to complete either a low (i.e., 8 sets per week) or high volume (i.e., 16 sets per week) 8-week training program. Following the initial 8 weeks, each participant would enter a washout period to revert the adaptations back to baseline, then complete the opposite condition.

Despite the benefits of crossover designs, their limitations also prevent addressing the first sub-component (i.e., is there sufficient evidence of individual response variation?). Specifically, following a washout period, it is ideal if increases in muscle size and strength return to baseline. However, gains in muscle size and strength are resilient following ~30 weeks of training cessation [18], and also exhibit variability in outcomes [19]. Thus, the return to baseline within a washout period may be unrealistic in the context of RT due to the lack of predictable trends in these outcomes. Further, trained individuals may not participate in

studies requiring such washout periods. Additionally, while the aforementioned order effect can be controlled via randomization on the group-level, the response of a given individual to the second training period could be impacted by second-order effects (i.e., potentiation or inhibition) of the first training period. For example, adaptations (or lack thereof) from the first training period (e.g., hypertrophy) could result in more (or less) robust responses in the second training period (e.g., strength gain) that could confound comparisons between conditions.

Additionally, crossover designs cannot fully account for the influence of external variables (e.g., sleep, nutrition, stress, etc.). For example, a participant who is a University student may have a more difficult semester during training period B than during period A; thus, may have higher anxiety and less sleep during period B which could impair performance [20]. Finally, while crossover designs may require fewer total participants than parallel-group designs, a longitudinal crossover design is more time-intensive for both the researchers and participants. Therefore, longitudinal crossover designs may not always be feasible.

Within-Participant Unilateral Designs

A within-participant unilateral design offers some advantages over both parallel-group and crossover designs for investigating the individual response to RT. First, a within-participant unilateral design, unlike a parallel-groups design, randomizes an individual's limbs (often while accounting for limb dominance) to different training conditions; thus, each participant performs both training protocols. Since, similar to a crossover design, each participant performs both protocols serving as their own control, researchers can compare individual outcomes from both training conditions (e.g., high versus low volumes). Second, performing both protocols at the same time rectifies the influence of external factors (sleep, stress, nutrition, etc.) that limit a crossover design. Third, a within-participants unilateral design likely increases the ability to detect a given effect size of interest (i.e, statistical power) compared to both a parallel-group and crossover designs, assuming an equivalent correlation of the outcome. Specifically, because each participant completes each protocol simultaneously, each person is effectively serving as two (i.e., one in each protocol); thereby, increasing recruitment efficiency.

Nonetheless, within-participant unilateral designs have limitations. Due to the nature of training each limb independently, these designs inherently exclude bilateral exercises (e.g., barbell bench press) from being experimentally manipulated, limiting the ecological validity of

the RT intervention. Another limitation is the cross-education effect, where strength increases in the limb opposite to the one being trained; however, this phenomenon does not seem to extend to muscle hypertrophy [21]. Due to the cross education effect, it would be difficult to determine the exact magnitude of difference in strength increases between protocols on the individual-level using a within-participants unilateral design. That said, if the limb with the superior training effect also results in a larger cross-education effect, any differences between limbs could be even larger than the observations indicate.

Introducing Replication: Establishing the Participant-By-Training Interaction

While crossover and within-participant unilateral designs improve upon typical parallel-group designs, they are still not fully capable of addressing both sub-components of interest with respect to individual response variation. Because both conditions completed by participants in crossover and within-participant unilateral designs feature a training intervention, any observed variance in the outcomes cannot be causally linked to a training intervention rather than other sources of variance (e.g., random sampling variation, measurement error, etc.). However, the introduction of replication into crossover and within-participant unilateral designs can overcome this limitation and effectively investigate individual response variation.

To conceptually explain why replication is necessary to evaluate individual-level response variation to RT, let's revisit an example of a crossover design in which a participant is initially randomized to a low volume training program, followed by a washout period, and then a high volume training program. Upon completing the crossover design, each participant will have change scores for outcome measures in both study phases. While an individual's change scores can be directly compared to evaluate which training period leads to more favorable outcomes, the *cause* of these outcomes cannot be fully justified. For example, if a participant increases their bench press 1RM by 10kg during a low volume training program and by 5kg during the high volume training program, it would be inappropriate to assign causality to the change in training intervention.

However, if the same participant repeated the same low-volume training intervention multiple times, the expected change score variance - completely unrelated to the training intervention (i.e., within-participant variance) - could be quantified. If the difference between change scores of the two conditions (i.e., high and low volumes) does not exceed that which is expected via within-participant variance, it is unjustified to claim that true individual response variation is

present. This concept, when comparing the variance of groups performing a training intervention versus a non-training control, has been coined the participant-by-training interaction [12] (Figure 2).

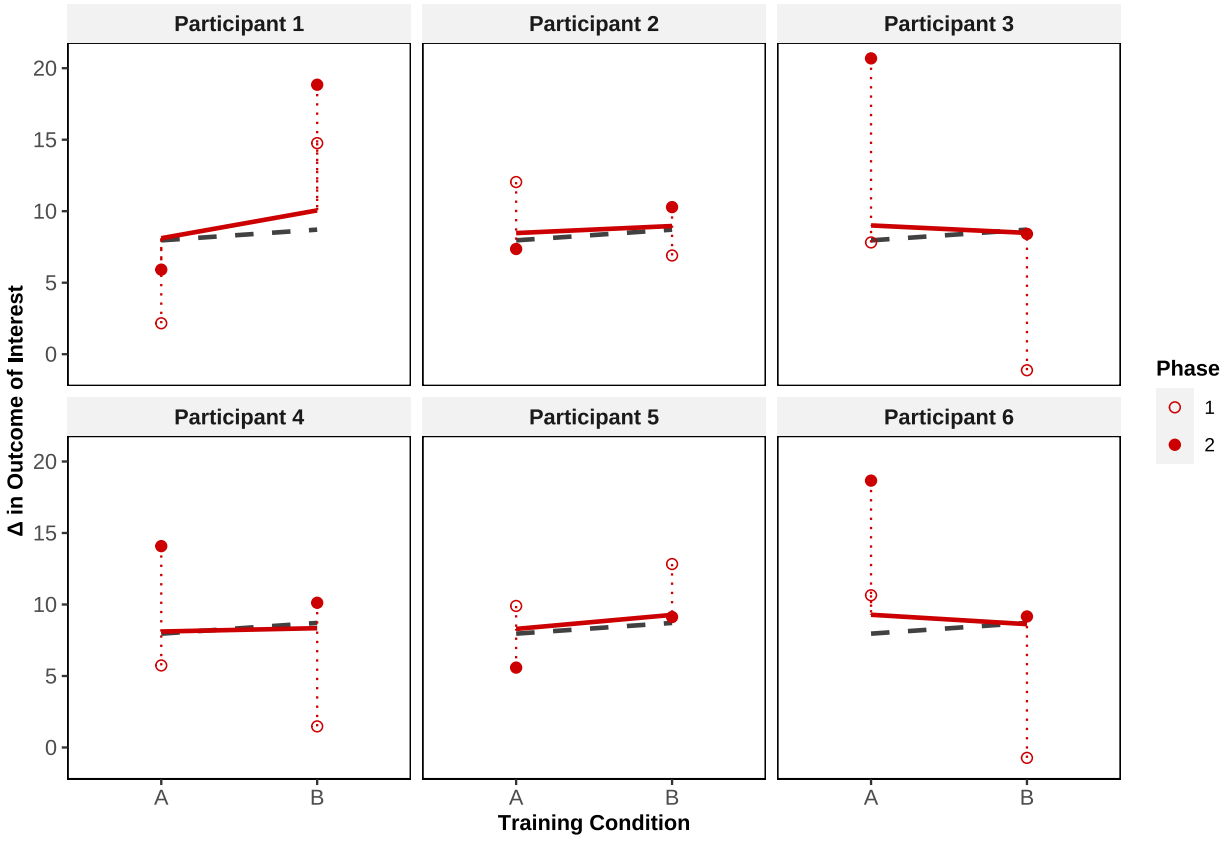

Figure 2: *Visualization of variance components derived from a linear mixed effects model with random intercepts introduced for each participant and random slopes for the participant-by-training interaction. This analysis is for a theoretical within-participant unilateral design with replication and the data of only 6 participants is shown for simplicity. The dashed black line represents the mean slope of condition in each panel. The dark red line represents the slope of condition for each participant. The dotted red lines going from each dot to the dark red line represents the remaining within-participant (i.e., residual variance) after accounting for other sources of variance. The vertical distance of each participant's intercept from the mean response represents the between-participant variance. Finally, the variation between participants in the slopes of condition represents the participant-by-training interaction.*

Thus, in order to establish the within-participant variance of a sample necessary to compare against the participant-by-training interaction, a training intervention must be completed at least twice (i.e., replication). Indeed, replication is advocated for in the biomedical literature to

investigate individual response variation. For example, Senn et al. [22] discussed a replicated crossover design, in which each participant completes one of the conditions twice. One obvious limitation of study designs that include replication is that it can substantially increase study duration, making it more challenging to obtain a desired sample size. However, compared to biomedical interventions that often elicit systemic effects, one advantage of RT is that the outcomes of interest are largely localized to the limb trained. Thus, by utilizing a within-participants unilateral design where each limb completes a training condition simultaneously, study duration can be decreased in comparison to a crossover design where conditions must be completed independently.

If it is not feasible to replicate the training intervention as a part of study design, more frequent measurements of the outcome variable (e.g., muscle size) can be taken to approximate response heterogeneity. For example, if a participant completes a 12-week training intervention and their muscle size is assessed every 2 weeks (i.e., 6 occasions), each 2-week interval could be considered an "intervention" that is replicated; thereby, allowing for the quantification of the participant-by-training interaction. However, this approach is not without limitations as there are multiple potential confounders introduced due to the accumulation of testing in a relatively short period of time [12].

STATISTICAL METHODS TO INVESTIGATE INDIVIDUAL RESPONSE VARIATION

In tandem with using an appropriate experimental design, it is crucial to consider the appropriate statistical analyses to examine individual response variation and its potential predictors. While a comprehensive discussion of the necessary statistical methods to investigate individual response variation in RT is outside the scope of this review, the following sections conceptually describe the appropriate methods and highlight useful resources.

Quantifying Evidence of Individual Response Variation

As previously mentioned, evidence for individual response variation can be quantified by the participant-by-training interaction and its comparison to within-participant variance. While there are likely multiple approaches that give researchers access to the estimands of interest, Heckstenden et al. [12] and Senn et al. [22] suggest that linear mixed effect models fit with random intercepts per participant and either random intercepts or slopes for the participant-

by-training interaction are the best fit approach when replication is utilized. Conceptually, this modeling strategy is comparable to ordinary least squares regression; however, variance of the estimates obtained for the independent variables included in the model (e.g., training condition) are partitioned across multiple sources (i.e., residual, participant, and participant-by-training interaction). Upon fitting these models, variance estimates can be obtained and consequently compared in order to quantify evidence of individual response variation. If the variance associated with the participant by training interaction exceeds that of within-participant variance, there is evidence of individual response variation, on average.

Taken a step further, the likelihood of individual response variation for a given participant could then be evaluated. Specifically, researchers could calculate a "sufficient variance threshold" by taking into account the smallest effect size of interest (SEOSI). While the SESOI is ultimately subjective, a reasonable suggestion may be to divide the standard deviation of the within participant variance by $\sqrt{2}$, similar to a standard error of measurement [23,24]. Then, by comparing each the point estimate and uncertainty interval, or the posterior distribution in a Bayesian framework, of each participants' marginal effect to this threshold [24,25], one could determine if a given participant exhibits meaningful response variation, probabilistically.

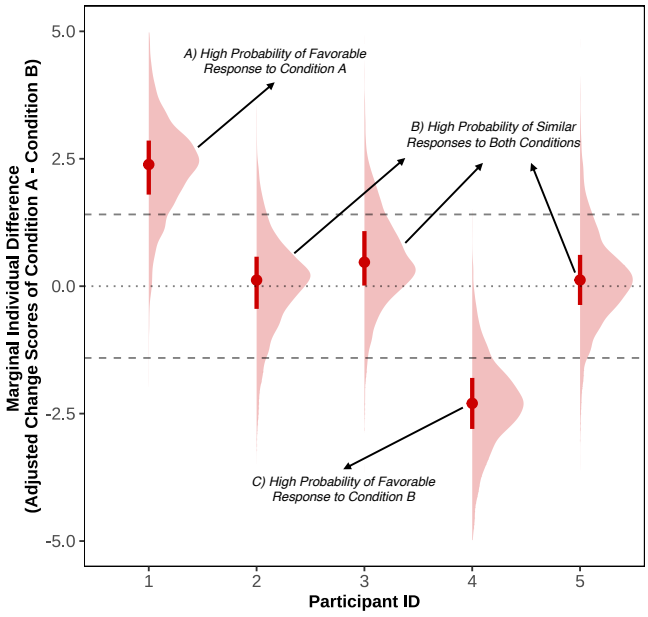

Figure 3: *Visualization of participant-level marginal effects derived from a linear mixed effects model with random intercepts introduced for each participant and random slopes the participant-by-training interaction. This analysis is for a theoretical within-participant unilateral design with replication and the data of only 5 participants is shown for simplicity. The red dots and intervals show the mean and 50% quantile interval for each participant. The lighter shade density plots visualize the draws from the posterior distribution. The dashed thresholds represent the sufficient variance threshold that can be calculated by diving the standard deviation of the within-participant variance by the square root of 2. AC) Examples where a participant exhibits a high probability of a favorable response to one of the two conditions B) Examples where the participants exhibit a low probability of a favorable response to one of the two conditions.*

Predicting Individual Response Variation

Evaluating the presence of individual response variation is only the first step to improve individual-level RT prescription and should be followed by attempting to identify predictors of individual-level outcomes. For example, Hammarstrom et al. [26] observed that individuals who benefited from higher training volumes demonstrated higher total RNA levels at week 2 of the training intervention, indicating greater rates of ribosomal biogenesis. However, since this information is not available at baseline and requires invasive techniques (i.e., muscle biopsies and biochemical assays), it is not available to most practitioners. Ideally, RT practitioners would be able to use widely available assessments (Table 2) to individualize training prescriptions. Specifically, characteristics that could be obtained apriori could help identify participants that would respond favorably to a given training intervention. For example, if the number of repetitions performed at a given percentage of 1RM is predictive of the volume dose-response relationship, then volume could be prescribed depending on how many repetitions someone performed at 80% of 1RM in a set to momentary failure.

Potential Predictive Assessments for Resistance Training Program Design

Anthropometric	*Performance*	*Physiological*	*Qualitative*
-Limb Lengths	-Repetitions to Failure	-Cardiovascular Fitness	-Training History
-Body Composition	-Velocity Loss	-Lactate	-Psychology Questionnaires
-Body Mass	-Vertical Jump	-Salivary Testosterone	-Perception of Training Quality

Table 2: *Categorized list of potential predictors of the individual response to resistance training.*

To examine predictors in each of the aforementioned study designs, researchers could collect a variety of potential predictive tools at baseline and, if there is sufficient evidence of individual response variation, their predictive capacity could be explored further. In the pursuit of potential tools that may aid individual-level training prescription, the distinction between goals

of *prediction* versus *explanation* [27] become important. While potentially counterintuitive, a more explanatory model (i.e., a model that accurately reflects causality observed in nature) may actually result in less accurate predictions and vice versa. Practitioners are most likely willing to favor making accurate predictions even if it means slight error in the explanation of the phenomenon; thus, this review will focus on prediction, rather than explanation.

The first step in exploring potential predictors of individual response variation is defining the dependent variable of interest. To quantify the difference in longitudinal outcomes between conditions on the individual-level, the "*individual difference*" can described as:

$$\text{individual difference} = \left(post^b - pre^b\right) - \left(post^a - pre^a\right)$$

pre^a = pre-test value from the training condition A
$post^a$ = post-test value from the training condition A
pre^b = pre-test value from the training condition B
$post^b$ = post-test value from the training condition B

Note: Model-adjusted estimates (e.g., participant-level marginal effects in Figure 3) could be used rather than calculating the individual differnce from raw data.

In the event that an individual difference exceeds typical measurement error and/or within-participant variance (as in Figure 3AC), we can cautiously conclude the participant experienced a more favorable response from one of the training conditions. If sufficient evidence is present to suggest there is meaningful individual response variation on average or in a subset of participants, prediction can be further explored.

With the combination of the individual difference and baseline participant characteristics from a variety of potentially predictive tools, the predictive capacity can be statistically assessed. Ultimately, while there are many ways one could go about analyzing these data, fitting general linear models with the individual difference as the outcome variable and data from the predictive tools of interest as fixed effects (e.g., number of repetitions performed in a set to momentary failure with 80% of 1RM) may be the simplest approach:

$$Y_{individual\ difference} = \beta_0 + \beta_{predictor_1} X_{predictor_1} + \ldots + \beta_{predictor_n} X_{predictor_n} + \varepsilon$$

This method of analysis could help identify promising candidate predictor variables for follow-up study, potentially confirming the predictive nature of these associations. A mock example of

a visualization of the relationship between leg press repetitions performed to failure at 80% of 1RM (predictor) with the individual difference (dependent variable) can be seen in Figure 3.

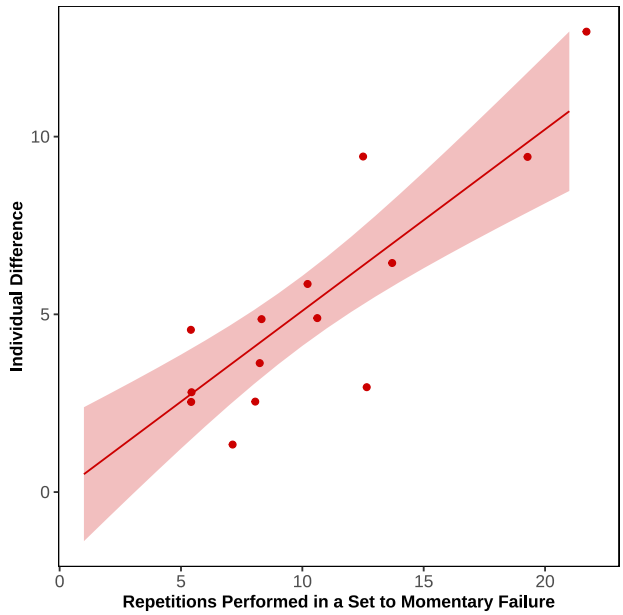

Figure 4: *Example analysis of the relationship between a predictor (i.e., repetitions performed in a set to momentary failure) and the individual difference.*

Countless potential predictors could be analyzed as the example in Figure 4. However, when examining numerous predictors in this fashion there is an increased risk of type-I error; thus, we urge authors to be transparent and explicitly delineate between exploratory versus confirmatory analyses [28].

A NOTE OF CAUTION

This papers aims to provide resistance training researchers with a simplified overview of the necessary methods to appropriately investigate individual response variation. However, it needs to be understood that these methods are not to be undertaken lightly. Exercise science is already riddled with under-powered/poor-precision, small sample research [29]. Compared to investigating group-level effects, studying individual response variation with a reasonable amount of statistical power/precision requires much larger sample sizes[13,30]. Thus, researchers interested in this topic might consider creative methods to improve the feasibility of larger sample sizes (e.g., multi-year data collection periods, multi-site data collection, less supervised training interventions, more frequent measurements, etc.). Otherwise, it may be the

best use of resources to better answer group-level research questions that will likely generalize to a majority of the target population anyway [31].

SUMMARY

In conclusion, modifications to common experimental designs may be needed to optimize individual-level RT recommendations. Presently, much of the current literature is limited in its ability to inform individual-level prescription, which is of paramount concern to practitioners. The proposed research designs are not without limitations; however, implementing replication alongside study designs in which participants complete both conditions (i.e., crossover and within-participant unilateral) will allow for a more complete investigation of individual response variation. Overall, it is essential for researchers to carefully consider the research question of interest, then design, carry out, and analyze their experiments accordingly.

REFERENCES

1. Currier BS, Mcleod JC, Banfield L, Beyene J, Welton NJ, D'Souza AC, et al. Resistance training prescription for muscle strength and hypertrophy in healthy adults: A systematic review and bayesian network meta-analysis. British Journal of Sports Medicine [Internet]. 2023;bjsports-2023-106807. Available from: https://doi.org/10.1136/bjsports-2023-106807
2. Schoenfeld BJ, Ogborn D, Krieger JW. Dose-response relationship between weekly resistance training volume and increases in muscle mass: A systematic review and meta-analysis. Journal of Sports Sciences [Internet]. 2016;35:1073–82. Available from: https://doi.org/10.1080/02640414.2016.1210197
3. Ralston GW, Kilgore L, Wyatt FB, Baker JS. The effect of weekly set volume on strength gain: A meta-analysis. Sports Medicine [Internet]. 2017;47:2585–601. Available from: https://doi.org/10.1007/s40279-017-0762-7
4. LOPEZ P, RADAELLI R, TAAFFE DR, NEWTON RU, GALVÃO DA, TRAJANO GS, et al. Resistance training load effects on muscle hypertrophy and strength gain: Systematic review and network meta-analysis. Medicine & Science in Sports & Exercise [Internet]. 2020;53:1206–16. Available from: https://doi.org/10.1249/mss.0000000000002585
5. Refalo MC, Hamilton DL, Paval DR, Gallagher IJ, Feros SA, Fyfe JJ. Influence of resistance training load on measures of skeletal muscle hypertrophy and improvements in maximal strength and neuromuscular task performance: A systematic review and meta-analysis. Journal

of Sports Sciences [Internet]. 2021;39:1723–45. Available from: https://doi.org/10.1080/02640414.2021.1898094
6. Hubal MJ, Gordish-Dressman H, Thompson PD, Price TB, Hoffman EP, Angelopoulos TJ, et al. Variability in muscle size and strength gain after unilateral resistance training. Med Sci Sports Exerc. 2005;37:964–72.
7. Erskine RM, Jones DA, Williams AG, Stewart CE, Degens H. Inter-individual variability in the adaptation of human muscle specific tension to progressive resistance training. European Journal of Applied Physiology [Internet]. 2010;110:1117–25. Available from: https://doi.org/10.1007/s00421-010-1601-9
8. Klemp A, Dolan C, Quiles JM, Blanco R, Zoeller RF, Graves BS, et al. Volume-equated high- and low-repetition daily undulating programming strategies produce similar hypertrophy and strength adaptations. Applied Physiology, Nutrition, and Metabolism [Internet]. 2016;41:699–705. Available from: https://doi.org/10.1139/apnm-2015-0707
9. Dolan C. THE INFLUENCE OF TIME-EQUATED TRAINING PROGRAMS ON MUSCLE HYPERTROPHY, STRENGTH, AND BODY COMPOSITION.
10. Helms ER, Byrnes RK, Cooke DM, Haischer MH, Carzoli JP, Johnson TK, et al. RPE vs. Percentage 1RM loading in periodized programs matched for sets and repetitions. Frontiers in Physiology [Internet]. 2018;9. Available from: https://doi.org/10.3389/fphys.2018.00247
11. Robinson ZP. The effect of resistance training proximity to failure on muscular adaptations and longitudinal fatigue in trained men. Master's Thesis. 2021;
12. Hecksteden A, Kraushaar J, Scharhag-Rosenberger F, Theisen D, Senn S, Meyer T. Individual response to exercise training - a statistical perspective. Journal of Applied Physiology [Internet]. 2015;118:1450–9. Available from: https://doi.org/10.1152/japplphysiol.00714.2014
13. Steele J, Fisher J, Smith D, Schoenfeld B, Yang Y, Nakagawa S. Meta-analysis of variation in sport and exercise science. 2022; Available from: https://doi.org/10.51224/srxiv.214
14. Aube D, Wadhi T, Rauch J, Anand A, Barakat C, Pearson J, et al. Progressive resistance training volume: Effects on muscle thickness, mass, and strength adaptations in resistance-trained individuals. Journal of Strength and Conditioning Research [Internet]. 2020;36:600–7. Available from: https://doi.org/10.1519/jsc.0000000000003524
15. Sousa Fortes L de, Cunha Costa M da, Perrier-Melo RJ, Brito-Gomes JL, Nascimento-Júnior JRA, Lima-Júnior DRAA de, et al. Effect of volume in resistance training on inhibitory control in young adults: A randomized and crossover investigation. Frontiers in Psychology [Internet]. 2018;9. Available from: https://doi.org/10.3389/fpsyg.2018.02028

16. Scarpelli MC, Nóbrega SR, Santanielo N, Alvarez IF, Otoboni GB, Ugrinowitsch C, et al. Muscle hypertrophy response is affected by previous resistance training volume in trained individuals. Journal of Strength and Conditioning Research [Internet]. 2020;36:1153–7. Available from: https://doi.org/10.1519/jsc.0000000000003558
17. Simoneau J-A, Lortie G, Boulay MR, Thibault M-C, Bouchard C. Repeatability of fibre type and enzyme activity measurements in human skeletal muscle. Clinical Physiology [Internet]. 1986;6:347–56. Available from: https://doi.org/10.1111/j.1475-097x.1986.tb00240.x
18. Staron RS, Leonardi MJ, Karapondo DL, Malicky ES, Falkel JE, Hagerman FC, et al. Strength and skeletal muscle adaptations in heavy-resistance-trained women after detraining and retraining. Journal of Applied Physiology [Internet]. 1991;70:631–40. Available from: https://doi.org/10.1152/jappl.1991.70.2.631
19. Räntilä A, Ahtiainen JP, Avela J, Restuccia J, Dawson K, Häkkinen K. High responders to hypertrophic strength training also tend to lose more muscle mass and strength during detraining than low responders. Journal of Strength and Conditioning Research [Internet]. 2021;Publish Ahead of Print. Available from: https://doi.org/10.1519/jsc.0000000000004044
20. Haischer MH, Cooke DM, Carzoli JP, Johnson TK, Shipherd AM, Zoeller RF, et al. Impact of cognitive measures and sleep on acute squat strength performance and perceptual responses among well-trained men and women. Journal of Strength and Conditioning Research [Internet]. 2021;35:S16–22. Available from: https://doi.org/10.1519/jsc.0000000000003202
21. Hendy AM, Lamon S. The cross-education phenomenon: Brain and beyond. Frontiers in Physiology [Internet]. 2017;8. Available from: https://doi.org/10.3389/fphys.2017.00297
22. Senn S, Rolfe K, Julious SA. Investigating variability in patient response to treatment a case study from a replicate cross-over study. Statistical Methods in Medical Research [Internet]. 2010;20:657–66. Available from: https://doi.org/10.1177/0962280210379174
23. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. The Journal of Strength and Conditioning Research [Internet]. 2005;19:231. Available from: https://doi.org/10.1519/15184.1
24. Swinton PA, Hemingway BS, Saunders B, Gualano B, Dolan E. A statistical framework to interpret individual response to intervention: Paving the way for personalized nutrition and exercise prescription. Frontiers in Nutrition [Internet]. 2018;5. Available from: https://doi.org/10.3389/fnut.2018.00041
25. Batterham AM. Is the intervention as good as (or not substantially worse than) a comparator? Experimental Physiology [Internet]. 2022;107:199–200. Available from: https://doi.org/10.1113/ep090321

26. Hammarström D, Øfsteng S, Koll L, Hanestadhaugen M, Hollan I, Apró W, et al. Benefits of higher resistance-training volume are related to ribosome biogenesis. The Journal of Physiology [Internet]. 2020;598:543–65. Available from: https://doi.org/10.1113/jp278455
27. Shmueli G. To explain or to predict? Statistical Science [Internet]. 2010;25. Available from: https://doi.org/10.1214/10-sts330
28. Caldwell AR, Andrew D. Vigotsky and, Tenan MS, Radel R, Mellor DT, Kreutzer A, et al. Moving sport and exercise science forward: A call for the adoption of more transparent research practices. Sports Medicine [Internet]. 2020;50:449–59. Available from: https://doi.org/10.1007/s40279-019-01227-1
29. Mesquida C, Murphy J, Lakens D, Warne J. Replication concerns in sports and exercise science: A narrative review of selected methodological issues in the field. Royal Society Open Science [Internet]. 2022;9. Available from: https://doi.org/10.1098/rsos.220946
30. Hopkins WG. Individual responses made easy. Journal of Applied Physiology [Internet]. 2015;118:1444–6. Available from: https://doi.org/10.1152/japplphysiol.00098.2015
31. Williamson PJ, Atkinson G, Batterham AM. Inter-individual responses of maximal oxygen uptake to exercise training: A critical review. Sports Medicine [Internet]. 2017;47:1501–13. Available from: https://doi.org/10.1007/s40279-017-0680-8