**SUPPLEMENTARY MATERIAL: REPORTING OF METHODS**

This supplementary file details the methodology and analyses used in our multilevel study, aligned with the LEVEL (Logical Explanations & Visualizations of Estimates in Linear mixed models) reporting guidelines (Monsalves et al., 2020). The LEVEL checklist contains 12 items, each consisting of several subitems. This document covers items 4–12 relevant to methods and results reporting.

**STRUCTURE OF DATA AND MODELING (Items 4a, 4b, 6a, 6b)**

Table 1 within this Supplementary File, describes the multilevel data structure, including hierarchical notations, model diagrams, and fixed/random effect equations. We also specify the origin of predictor variables across levels and justify the analytical levels selected.

**SAMPLE AND STUDY DESIGN (Item 5)**

*5a. Participant Characteristics and Design.* Fourteen elite male team-sport athletes (mean ± SD: age = 22.6 ± 2.8 yrs; height = 180.3 ± 5.9 cm; mass = 72.3 ± 3.8 kg) participated. All had 10–12 years of training experience, no neuromuscular disorders or recent injuries, and refrained from vigorous activity, alcohol, and caffeine 48 hours before testing. A randomized two-condition crossover design was used, with each participant completing both protocols one week apart. Each protocol involved 4 sets × 6 maximal voluntary isometric contractions (MVICs) and differed only in MVIC duration (3 seconds vs. 6 seconds).

*5b. Repeated Measures and Time Spacing.* Twitch torque was assessed 2 seconds after each MVIC via tibial nerve stimulation. MVIC Repetitions occurred every 15 seconds within sets, sets were spaced by 2 minutes of rest. Sessions were separated by a 1-week washout.

*5c. Missingness and Imbalances.* No missing data or imbalances across levels were observed due to the tightly controlled laboratory settings.

**STUDY SIZE (Item 7)**

*7a. Sample Size and VPC/ICC.* No a priori sample size calculation was performed for this secondary analysis. The original study included a power analysis relevant to its primary aims.

*7b. Justification for ICCs.* No prior studies on Post-activation Potentiation (PAP) have investigated individual variability using multi-level models and this study was a secondary analysis of already available data, therefore no ICCs from previous literature or pilot data were available for use.

**STATISTICAL METHODS (Item 8)**

*8a. Statistical Approach and Handling of Correlation.* We used linear mixed models with a nested random-effects structure to account for within-subject dependence. Intervention predictors (sets, repetitions) were mean-centered to reduce multicollinearity. Model design details are presented in Table 1 within this Supplementary File. We computed model parameter estimate p values using their t-values and Satterthwaite’s method of degrees of freedom through the “lmerTest” R package (Kuznetsova et al., 2017). We conducted case bootstraps and estimated 95% confidence intervals of model parameter estimates using the “lmeresampler” R package (Loy & Korobova, 2021). The case bootstrapping procedure allows re-sampling only from the highest-level cluster, or in our case the participant’s level, and therefore preserves the within-participant correlation structure across repeated observations. This type of bootstrap is appropriate for hierarchical models with repeated measurements and provides consistent results under minor model assumption violations (e.g., heteroskedasticity), which are common in mixed effects models (Leeden et al., 2008). We performed model variance decomposition by computing the variance contribution of each random effect to the total variance and reported the Variance Partition Coefficient (VPC) for each random effect (Monsalves et al., 2020).

*8b. Estimation Procedure.* Initial models were estimated using Maximum Likelihood (ML) allowing for fixed-effect model comparisons and final models used Restricted ML (REML) for unbiased random effect estimations (Bates et al., 2015; Bolker et al., 2009; Harrison et al., 2018).

*8c. Variance Components and VPCs.* Variance Partition Coefficients (VPCs) for null and final models are presented in the results section of our study and Supplementary Files 3 and 4.

*8d. Model Selection and Justification.* Null models included fixed effects of intervention parameters (sets and repetitions) and their interactions (plus polynomial trends), as justified by PAP’s simultaneous theoretical dependency on both fatigue and potentiation (Xenofondos et al., 2018), as well as model comparisons. The final null model was determined as the final model when the further addition of any other potentially plausible interactions or polynomial trends between the intervention parameters did not further improve model fit based on predefined model fit comparison criteria. To then specify baseline models, we first examined potentially plausible interactions and polynomial trends between baseline torque variables and intervention parameters (sets, repetitions), using visualizations (Supplementary File 2). Baseline models were then specified by including baseline torque variables with interactions and polynomial trends that were further supported by visual evidence (Fife, 2020). A separate baseline model was built for each baseline torque variable (baseline twitch torque, MVIC torque, and TT/MVIC ratio), considering the small N at the participant level (n = 14). The improvement in fit of baseline models compared to the null models was evaluated with model comparisons. Model comparisons relied on Likelihood Ratio Tests (LRTs; p < .05) (Bates et al., 2015), change in Akaike Information Criterion corrected for small sample sizes (ΔAICc; < -2) and Marginal R2 computed through the Multi-Model Inference R package “MuMIn” (Barton et al., 2021). Model assumption checks (e.g., residuals, VIF, homoscedasticity) were assessed using the “performance” R package (Figures 1, 2 and 3 within this Supplementary File) (Lüdecke et al., 2021).

*8e. Random Effects Specification.* The objective of the study was to assess individual variability in response to exercise and the role that baseline torque variables have in explaining this variability. Specifically, we aimed to determine whether baseline variables could account for inter-individual differences in responses. These were modeled using random intercepts (capturing variability in PAP responses between participants), random slopes for sets (capturing variability in PAP responses that results from variability in inter-individual linear effects of sets), and random slopes for repetitions (capturing variability in PAP responses that results from variability in inter-individual linear effects of repetitions). A secondary goal was to assess whether baseline variables could also explain intra-individual variability in responses. This was modeled using random intercepts for each set within each participant (capturing variability in PAP response that results from intra-individual variability of PAP responses across sets), random slopes for repetitions within each set (capturing variability in PAP response that results from intra-individual variability in linear effects of repetitions across sets), and random slopes for quadratic trends in repetitions within each set (capturing variability in PAP response that results from intra-individual variability in quadratic effects of repetitions across sets). Following recommendations for specifying maximal random effects structures (Barr et al., 2013), we initially aimed to include all plausible random effects and their correlations, accounting for the nested and correlated data structure. However, during the initial case bootstrap estimations, we encountered substantial convergence and singularity issues (>20%). To address this, we simplified the random effects by specifying them as uncorrelated (Harrison et al., 2018). This reduced convergence and singularity issues to <13%, suggesting that the small sample size at the participant level (n = 14) and the nature of case resampling (which resamples only at the participant level), was problematic when computing random effect correlations. Table 2 within this Supplementary File, provides information on the final encountered bootstrap model estimation errors, after simplifying random effects.

**RESULTS REPORTING (Items 10, 11, 12)**

We present the model equation in Table 1 within this Supplementary File, while model parameter estimates are presented within the Supplementary Files 3 and 4. In the study we report VPCs for random effects of models, model comparison goodness of fit statistics, fixed-effect adjusted variance components and all other results conducted from multi-level analyses.

**TABLES AND FIGURES**

Table . Multi-level model diagrams, model notations and justifications

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1. **Initial analyses considering each intervention as independent** | | | | | | |
| *Justification for the level of the analysis* | | | | | | |
| We initially analysed each intervention independently because our earlier investigation (Xenofondos et al., 2018) showed that exercise parameters such as sets and repetitions may have distinct effects on twitch torque responses. Thus, to ensure that the built null models accurately captured the influence of these intervention parameters, we chose to analyse each intervention separately. This approach allowed us to build precise null model specifications for each intervention, which was essential for testing the relationships between baseline torque variables and twitch responses within each intervention.  Therefore, an initial multi-level analysis was conducted to investigate if the participant baseline torque variables can explain inter-/intra- individual variability in Post-Activation Potentiation responses as a function of exercise protocols including multiple sets and repetitions (multi-level repeated measurements). The repeated measurements at the 1st level (repetitions) were thought to be correlated within the 2nd level (sets) and within the 3rd level (participants). Thus, modelling and accounting for this hierarchical repeated measurement structure was essential to investigate the value that baseline torque variables have in explaining variability at each of these hierarchically dependent levels. | | | | | | |
| *Levels and Observations* | | | | | | |
| *Level* | *Subindex* | *Number of Units in each level* | *Total Observations per Unit within each level* | *Total Observations* | *Predictor Variables taken from each level* |
| *Participant (Part\_id)* | i = 1, …,14 | 14 | 24 | 336 | -Baseline torque variables (between participant, constant within participants) | |
| *Set (Set\_id)* | J = 1, …, 4 | 56 (4 sets per participant) | 6 | 336 | -Sets (1-4; varies within participant) | |
| *Repetitions (Rep\_id)* | t = 1, …, 6 | 336 (6 reps per set per participant) | - | 336 | -Reps (1-6; varies within set within participant) | |
| *Multi-level diagram and model notation* | | | | | | |
| Participant (Part\_id) i = 1, …,14  Set (Set\_id) j = 1, …,4  Set\_id = 4  Set\_id = 1  Rep\_id = 1  Repetition (Rep\_id) t= 1, …,6  Rep\_id = 6  Repetition (Rep\_id) t = 1,…,6 …,6  Rep\_id = 6  Rep\_id = 1 | | | | | | |
| *Levels* | *Nature* | *Model Notation:* | | | | |
| *3 – Level* | *-Clustered*  *-Repeated measurements* | Yijt =  *Fixed effects*  β0​ (intercept)  + β1 \* ​Repijt ​  + β2 \* ​Setij  + β3 \* (Repijt)2  + β4 \* (Repijt)3  + β5 \* ​Repijt ​\* Setij ​  + β6 \* ​(Repijt)2 ​\* Setij ​  + β7 \* ​(Repijt)3 ​\* Setij ​  + β8 \* ​ Baseline TT/MVICi ​  + β9 \* ​ Baseline TT/MVICi \* Repijt  + β10 \* ​ Baseline TT/MVICi \* Setij  + β11 \* Baseline TT/MVICi \* (Repijt)2  + β12 \* ​ Baseline TT/MVICi \* Repijt ​\* Setij  + β13 \* ​ Baseline TT/MVICi \* (Repijt)2 ​\* Setij  *Random effects*  + u0i (Participant-level intercept)  + u1i \* Repijt ​ (Participant-level slope for repetitions)  + u2i \* ​Setij ​ (Participant-level slope for sets)  + v0ij​ (Set-within-Participant-level intercept)  + v1ij \* ​ Repijt (Set-within-Participant-level slope for repetition)  ​+ v2ij \* ​(Repijt)2​ (Set-within-Participant-level slope for repetition2)  + ϵijt​​ (Residual error)  *Where…*  *i=1,…,14 (participants)*  *j=1,…,4 (sets)*  *t=1,…,6 (repetitions)*  *And assumptions…*   * u0i *,* u1i*,* u2i *~ N(0, ​Σu) - Normally distributed participant-level random effects* * v0ij, v1ij, v2ij ~ *N(0, ​Σu) - Normally distributed set-within-participant-level random effects* * ϵijt ~ *N(0, ​Σu) - Normally distributed residual error* | | | | |
| **Note:** Interaction terms (e.g., Baseline × Reps, Baseline × Sets, Reps × Sets, Baseline × Reps × Sets) are cross-level interactions that combine variables measured across Levels 1, 2, and 3. Although their components are drawn from multiple levels, these interactions are modeled at the lowest level where they explain variability typically the repetition level (Level 1 – observation level). For example, Baseline × Set interactions combine Level 3 and Level 2 predictors but are modeled within the Level 1 outcome structure. | | | | | | |
| 1. **Subsequent analyses considering each intervention as nested within participants** | | | | | | |
| *Justification for the level of the analysis* | | | | | | |
| After first analyzing each intervention and estimating the explanatory value of baseline torque variables within each intervention separately, we then implemented a combined multi-level analysis model that accounted for the fact that the same participants underwent both interventions. In this combined analysis, observations from each intervention were modeled as part of a crossover design, recognizing that intervention observations were potentially correlated within participants and thus not independent. Accounting for this more complex hierarchical dependency was critical to formally test whether the explanatory contribution of baseline torque variables differed between interventions. This analysis specifically aimed to test whether baseline variables held differential explanatory value in accounting for individual variability across the two intervention conditions. | | | | | | |
| *Levels and Observations* | | | | | | |
| *Level* | *Subindex* | *Number of Units in each level* | *Total Observations per Unit within each level* | *Total Observations* | *Predictor Variables taken from each level* | |
| Participant (Part\_id) | i = 1, …,14 | 14 | 48 | 672 | - | |
| Intervention (Intervention\_id) | l = 1, 2 | 28 (2 interventions per participant) | 24 | 672 | -Interventions (1-2; varies within participant)  -Baseline torque variables (measured between participants, but vary between interventions and constant within interventions) | |
| Set (Set\_id) | j = 1, …,4 | 112 (4 sets per intervention per participant) | 6 | 672 | -Sets (1-4; varies within intervention within participant) | |
| Repetitions (Rep\_id) | t = 1, …,6 | 672 (6 repetitions per set per intervention per participant) | - | 672 | -Reps (1-6; varies within set within intervention within participant) | |
| *Multi-level diagram and model notation* | | | | | | |
| Participant (Part\_id) i = 1, …,14  Intervention 2  Intervention 1  A diagram of a set  AI-generated content may be incorrect.A diagram of a set  AI-generated content may be incorrect. | | | | | | |
| Levels | Nature | Model Notation | | | | |
| 4 – Level | -Clustered  -Repeated measurements | Yiljt =  *Fixed effects*  β0​ (Intercept of reference intervention)  + β1 \* ​ Baseline TT/MVICil  + β2 \* ​ Interventionil  + β3 \* ​ Setilj  + β4 \* ​ Repiljt ​  + β5 \* ​(Setilj)2  + β6​ \* (Repiljt)2  + β7​ \* (Repiljt)3  + β8 \* Repiljt ​\* Setilj  + β9​ \* (Repiljt)2 ​\* Setilj  + β10​ \* (Repiljt)3 ​ \* Setilj  + β11 \* ​ Interventionil \* ​ Baseline TT/MVICil  + β12 \* ​ Setilj \* ​ Baseline TT/MVICil  + β13 \* ​ Repiljt ​\* ​ Baseline TT/MVICil  + β14 \* ​(Setilj)2 \* ​ Baseline TT/MVICil  + β15 \* (Repiljt)2 \* ​ Baseline TT/MVICil  + β16​ \* (Repiljt)3 \* ​ Baseline TT/MVICil  + β17 \* Repiljt ​\* Setilj \* ​ Baseline TT/MVICil  + β18 \* (Repiljt)2 ​\* Setilj \* ​ Baseline TT/MVICil  + β19​ \* (Repiljt)3 ​ \* Setilj \* ​ Baseline TT/MVICil  + β20 \* ​ Setilj \* ​ Interventionil  + β21 \* ​ Repiljt ​\* ​ Interventionil  + β22 \* ​(Setilj)2 \* ​ Interventionil  + β23 \* (Repiljt)2 \* ​ Interventionil  + β24 \* (Repiljt)3 \* ​ Interventionil  + β25 \* Repiljt ​\* Setilj \* ​ Interventionil  + β26​ \* (Repiljt)2 ​\* Setilj \* ​ Interventionil  + β27​ \* (Repiljt)3 ​ \* Setilj \* ​ Interventionil  + β28 \* ​ Setilj \* ​ Baseline TT/MVICil \* ​ Interventionil  + β29 \* ​ Repiljt ​\* ​ Baseline TT/MVICil \* ​ Interventionil  + β30 \* ​(Setilj)2 \* ​ Baseline TT/MVICil \* ​ Interventionil  + β31 \* (Repiljt)2 \* ​ Baseline TT/MVICil \* ​ Interventionil  + β32​ \* (Repiljt)3 \* ​ Baseline TT/MVICil \* ​ Interventionil  + β33 \* Repiljt ​\* Setilj \* ​ Baseline TT/MVICil \* ​ Interventionil  + β34 \* (Repiljt)2 ​\* Setilj \* ​ Baseline TT/MVICil \* ​ Interventionil  + β35​ \* (Repiljt)3 ​ \* Setilj \* ​ Baseline TT/MVICil \* ​ Interventionil  *Random effects*  + u0i (Participant-level intercept)  + u1i \* Repiljt ​ (Participant-level slope for repetitions)  + u2i \* ​Setilj ​ (Participant-level slope for sets)  + w0il (Intervention within Part. -level intercept)  + w1il \* Repiljt ​ (Intervention within Part. -level slope for repetitions)  + w2il \* ​Setilj ​ (Intervention within Part. -level slope for sets)  + v0ilj​ (Set-within intervention-Part. -level intercept)  + v1ilj \* ​ Repiljt (Set-within intervention-Part. -level slope for repetition)  + v2ilj \* ​(Repiljt)2​ (Set-within intervention-Part. -level slope for repetition2)  + ϵiljt​​ (Residual error)  *Where…*  *i=1,…,14 (participants)*  *l=1, 2 (interventions)*  *j=1,…,4 (sets)*  *t=1,…,6 (repetitions)*  *And assumptions…*   * u0i *,* u1i*,* u2i *~ N(0, ​Σu) - Normally distributed participant-level random effects* * *w0il, w1il, w2il ~ N(0, Σw) – Normally distributed intervention within participant level random effects* * v0ij, v1ij, v2ij ~ *N(0, ​Σv) - Normally distributed set-within-participant-level random effects* * ϵijt ~ *N(0, ​σ2) - Normally distributed residual error* | | | | |
| **Note:** Interaction terms (e.g., *Baseline × Repetitions × Sets × Intervention*) represent cross-level interactions that combine within-participant variables measured at Level 3 (intervention) with within-intervention predictors measured at Levels 2 (sets) and 1 (repetitions). For example, the *Baseline × Repetitions × Sets × Intervention* interaction combines Level 3 within-participant variables (baseline measurements and intervention condition) with Level 2 and Level 1 predictors, but the interaction is modeled at the Level 1 outcome level (repetition-level observations). Although baseline torque variables are conceptually between-participant (Level 4) variables, taking them separately within each intervention (Level 3) and modelling them within the same model, effectively treats these baseline variables as within-participant, between-intervention variables. Modeling these interactions in such a way (e.g., *Intervention × Baseline*) enables us to investigate whether the explanatory value of the baseline variable differs across interventions, while simultaneously accounting for within-participant between-session variability and potential measurement error in baseline measurements across sessions. | | | | | | |

Table 2. Convergence and singular fit warnings as percentages of bootstrapped models.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Singular fits  (%) | Convergence warnings (%) | (n) Models kept for CI computation |
| Intervention 3s | | | |
| Null model | 11.5 | 7.1 | 9262 |
| MVIC model | 11.5 | 7.3 | 9266 |
| TT model | 11.4 | 6.9 | 9312 |
| TT/MVIC model | 15.4 | 6.7 | 9326 |
| Intervention 6s | | | |
| Null model | > 0.01 | 11.7 | 8831 |
| MVIC model | 0.11 | 12.5 | 8755 |
| TT model | 0.4 | 12.4 | 8756 |
| TT/MVIC model | 0.4 | 11.4 | 8856 |
| Combined model | | | |
| Null model | > 0.01 | 3.1 | 9690 |
| MVIC model | 1.3 | 3.3 | 9675 |
| TT model | 0.5 | 3.3 | 9673 |
| TT/MVIC model | 0.4 | 3.4 | 9662 |

Percentages of models that produced singular fits and convergence issues from 10000 case bootstraps for each analysis. Only converged models were used to estimate bootstrap 95% confidence intervals. The count of these converged models remaining after excluding non-converged models is also reported.

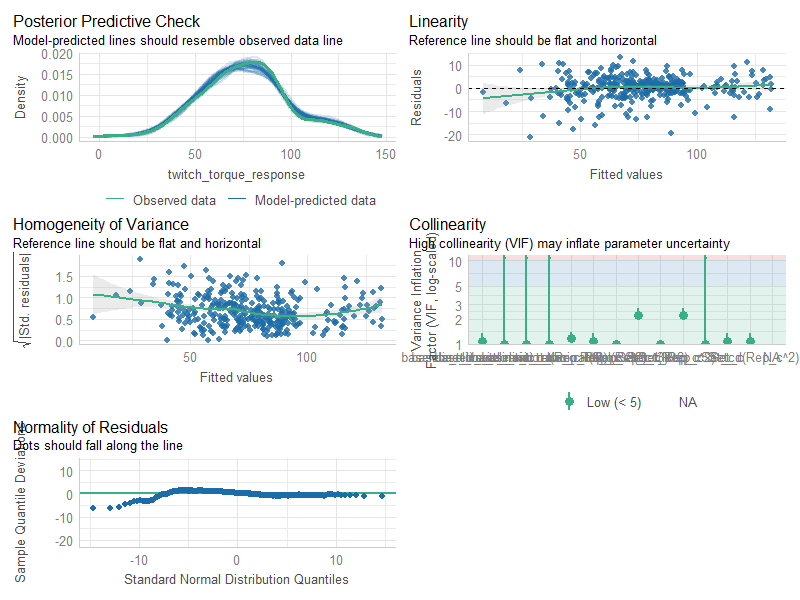
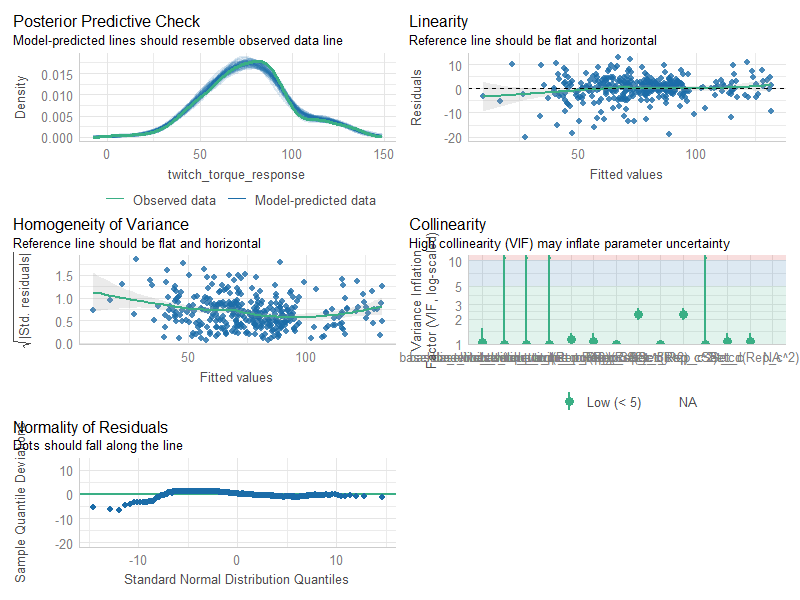


Figure . Model diagnostics of 3s intervention models. Diagnostic plots on the left refer to the baseline TT model and diagnostic plots on the right refer to the baseline TT/MVIC model. Diagnostic plots revealed the possibility of minor violations of homogeneity of variance and normality of residuals. The effects of these minor violations could be handled effectively by the implementation of case bootstraps confidence intervals which would correct any potential wrongful inferences based on minor miss-specification of error structures (Leeden et al., 2008).

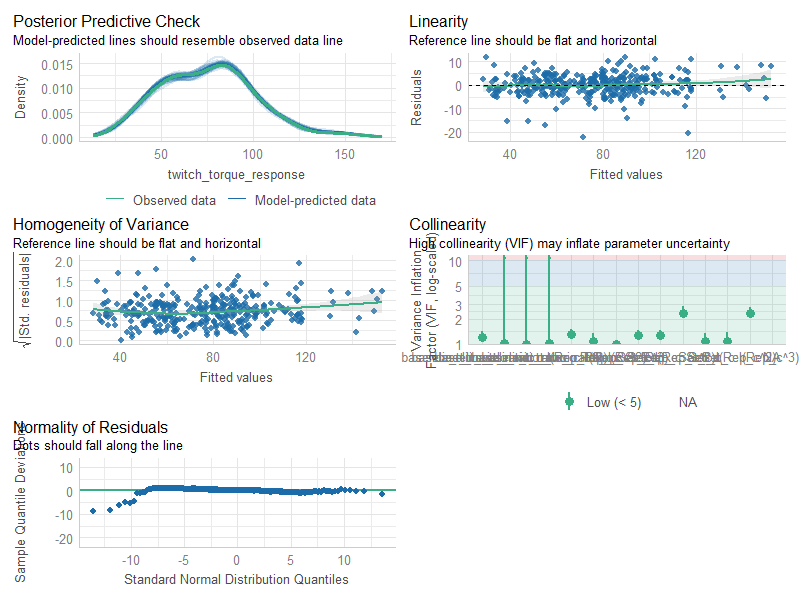
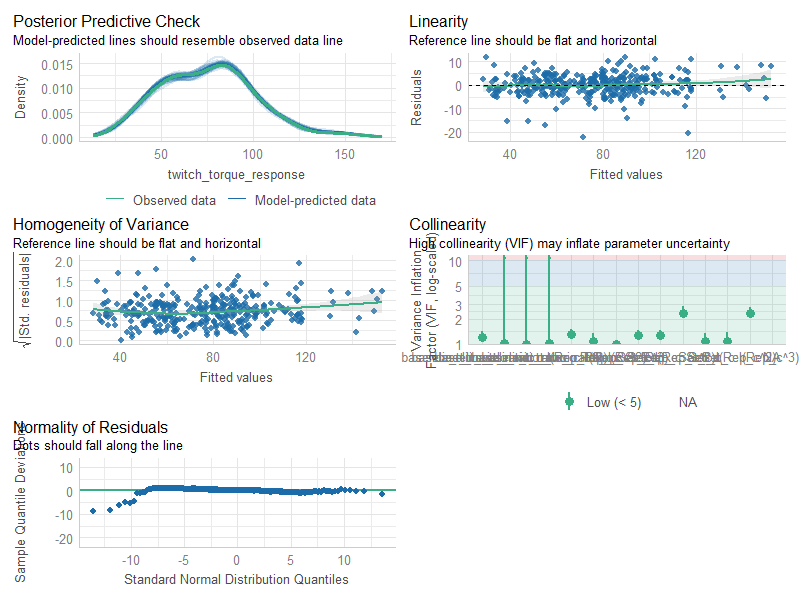


Figure . Model diagnostics of 6s intervention models. Diagnostic plots on the left refer to the baseline TT model and diagnostic plots on the right refer to the baseline TT/MVIC model. Diagnostic plots revealed the possibility of very minor violations of normality of residuals. These minor violations would have been handled effectively by our implementation case bootstrap confidence intervals which would correct any potential wrongful inferences based on minor miss-specification of error structures.

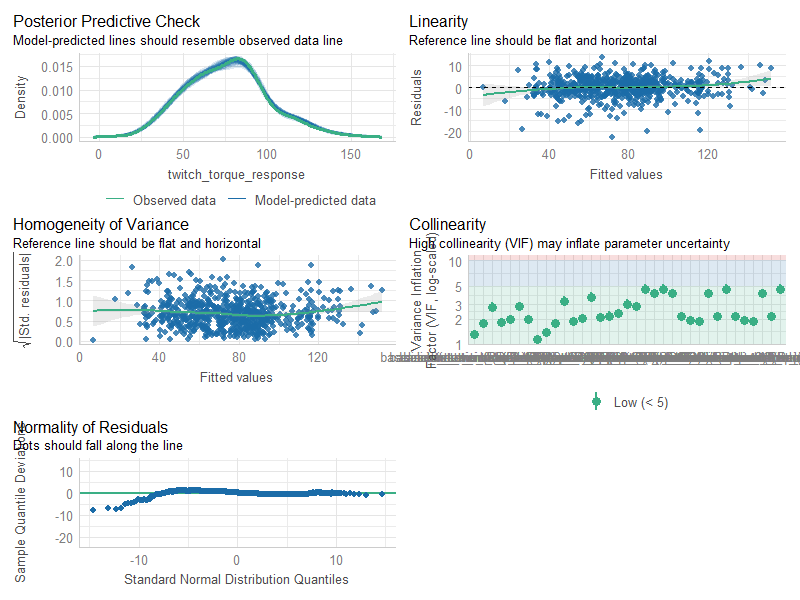
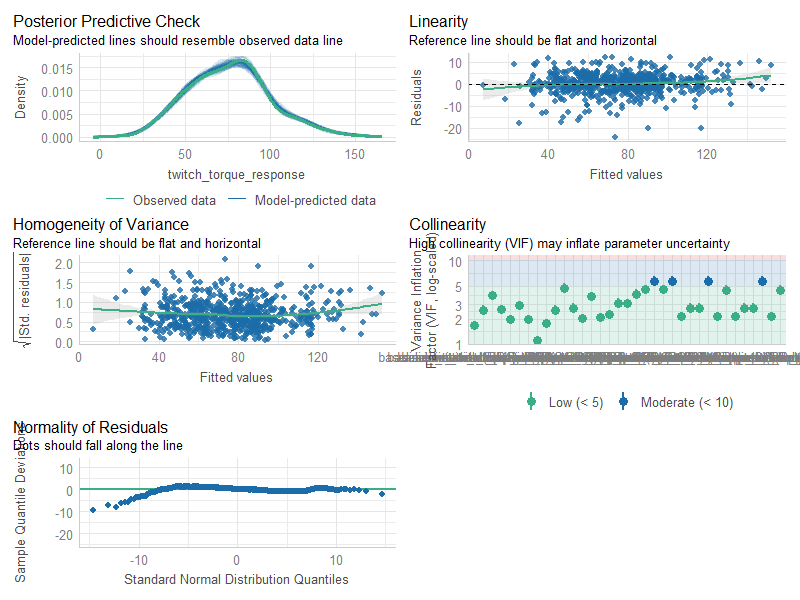


Figure . Model diagnostics of combined intervention models. Diagnostic plots on the left refer to the baseline TT model and diagnostic plots on the right refer to the baseline TT/MVIC model. Diagnostic plots revealed the possibility of minor violations of normality of residuals. These minor violations would have been handled effectively by our implementation case bootstrap confidence intervals which would correct any potential wrongful inferences based on minor miss-specification of error structures.

**Bibliography**

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). https://doi.org/10.1016/j.jml.2012.11.001

Barton, C. J., King, M. G., Dascombe, B., Taylor, N. F., de Oliveira Silva, D., Holden, S., Goff, A. J., Takarangi, K., & Shields, N. (2021). Many physiotherapists lack preparedness to prescribe physical activity and exercise to people with musculoskeletal pain: A multi-national survey. *Physical Therapy in Sport*, *49*, 98–105. https://doi.org/10.1016/j.ptsp.2021.02.002

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Fife, D. (2020). The Eight Steps of Data Analysis: A Graphical Framework to Promote Sound Statistical Analysis. *Perspectives on Psychological Science*, *15*(4), 1054–1075. https://doi.org/10.1177/1745691620917333

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Leeden, R. van der, Meijer, E., & Busing, F. M. T. A. (2008). Resampling Multilevel Models. In J. de Leeuw & E. Meijer (Eds.), *Handbook of Multilevel Analysis* (pp. 401–433). Springer. https://doi.org/10.1007/978-0-387-73186-5\_11

Loy, A., & Korobova, J. (2021). *Bootstrapping Clustered Data in R using lmeresampler* (arXiv:2106.06568). arXiv. https://doi.org/10.48550/arXiv.2106.06568

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Monsalves, M. J., Bangdiwala, A. S., Thabane, A., & Bangdiwala, S. I. (2020). LEVEL (Logical Explanations & Visualizations of Estimates in Linear mixed models): Recommendations for reporting multilevel data and analyses. *BMC Medical Research Methodology*, *20*(1), Article 1. https://doi.org/10.1186/s12874-019-0876-8

Xenofondos, A., Bassa, E., Vrabas, I. S., Kotzamanidis, C., & Patikas, D. A. (2018). Muscle Twitch Torque During Two Different in Volume Isometric Exercise Protocols: Fatigue Effects on Postactivation Potentiation. *The Journal of Strength & Conditioning Research*, *32*(2), 578. https://doi.org/10.1519/JSC.0000000000002311