

Background

Problem:

- ~20–25% of adults report fatigue in population surveys.
- PROMIS Fatigue short forms are widely used, but misinterpretation of items across health literacy levels can introduce item bias (Differential Item Functioning, DIF).
- Differential item functioning (DIF) may bias group comparisons by reflecting measurement artifacts rather than true differences in fatigue.

Our Aim: To evaluate whether PROMIS Fatigue items function equivalently across health literacy groups, ensuring fair and accurate measurement.

Methods

- Participants:** Adults (≥18 years) were recruited from Boston a Chicago. Surveys were completed in English or Spanish, either by paper or touchscreen.
- Health Literacy** was measured using the Health Literacy Assessment Using Talking Touchscreen Technology (Health LiTT), with scores above 55 indicating adequate health literacy.

Analysis:

- Unidimensionality** was assessed with Confirmatory Factor Analysis (CFA). Parallel analysis and exploratory bi-factor analysis were used to confirm factor structure.
- DIF detection:** Conducted using the *lordif* R package, using ordinal logistic regression and graded response models. DIF was flagged if Δ McFadden pseudo- R^2 exceeded item-specific thresholds from 5,000 Monte Carlo simulations.
- Assessment:** Item characteristic curves (ICCs) identified uniform vs. non-uniform DIF; test characteristic curves (TCCs) examined score-level impact.

Results

- Sample size:** 706 participants
- Health literacy:** 412 (58.4%) low, 294 (41.6%) adequate
- PROMIS Fatigue T-score:** Mean = 49.4 in both groups (no difference at the scale level)
- Unidimensionality:** The bifactor analysis confirms a dominant general factor: Omega hierarchical (ω) = 0.86 and Explained common variance (ECV) = 0.84.
- Flagged items for DIF:** 5 out of 8.

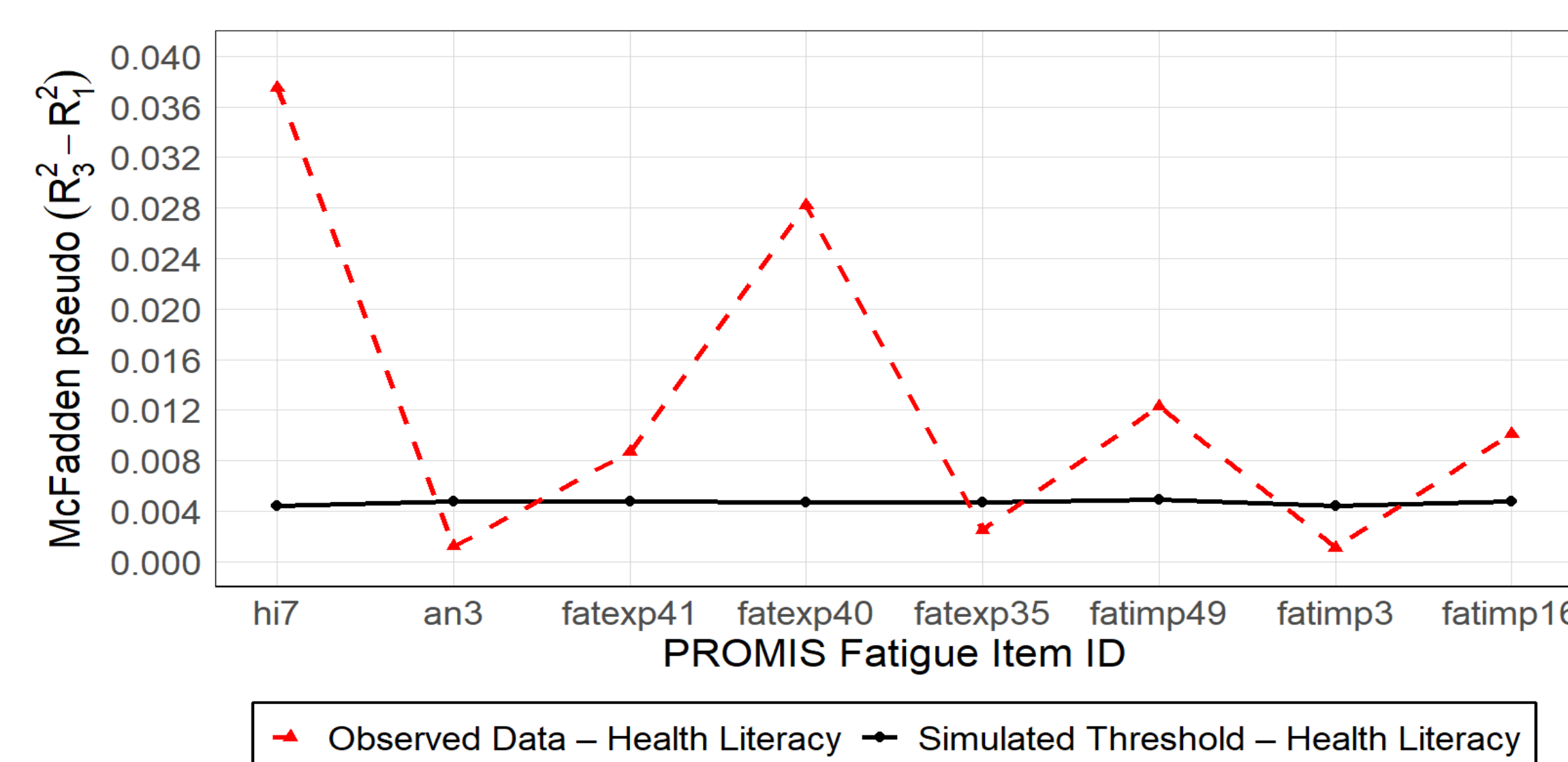


Figure 1. ΔR^2 values from observed data against 99% threshold values from simulated random data of the same size for health literacy (low vs. adequate).

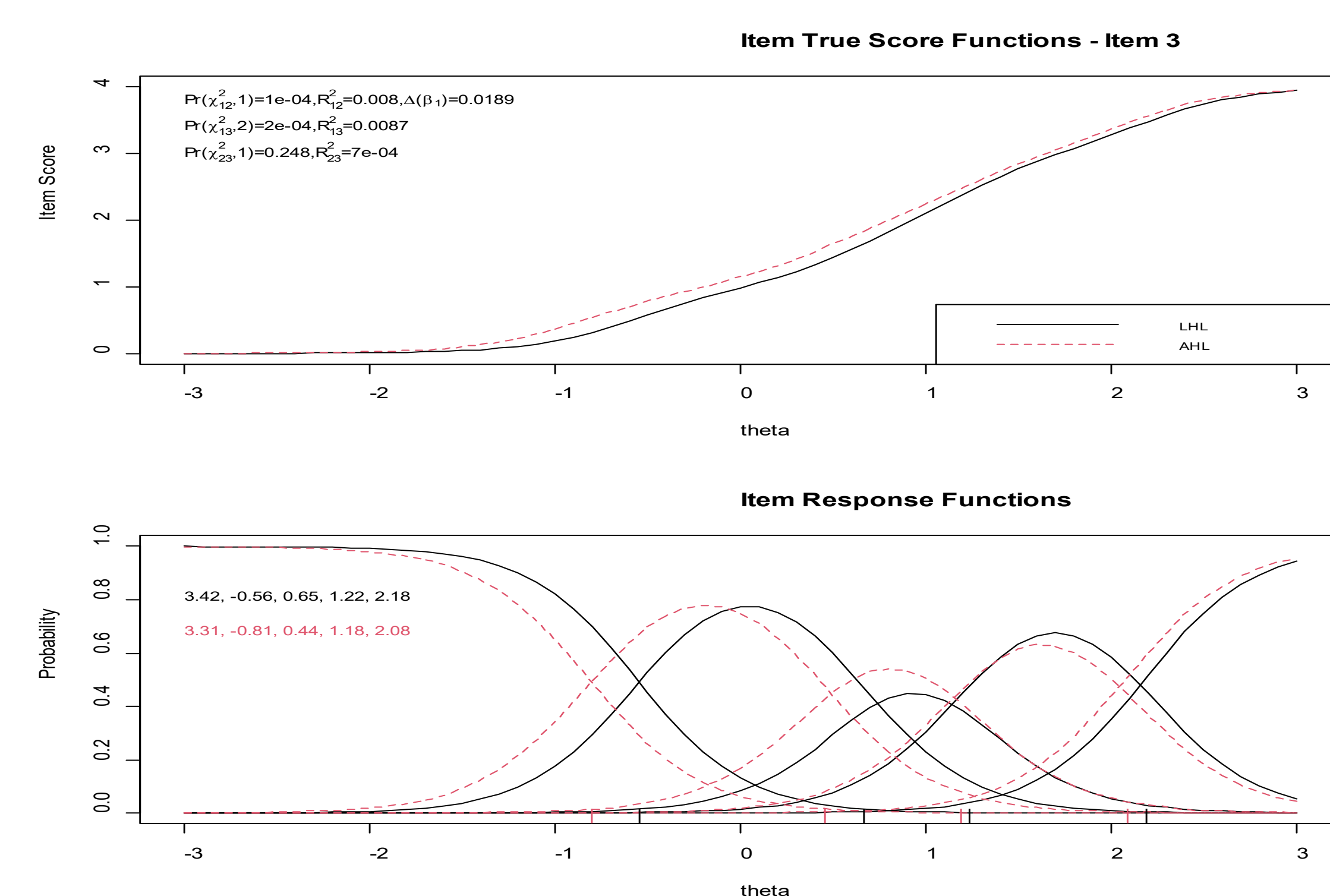


Figure 3. ICC of FATEXP41 *In the past 7 days: How run-down did you feel on average?*

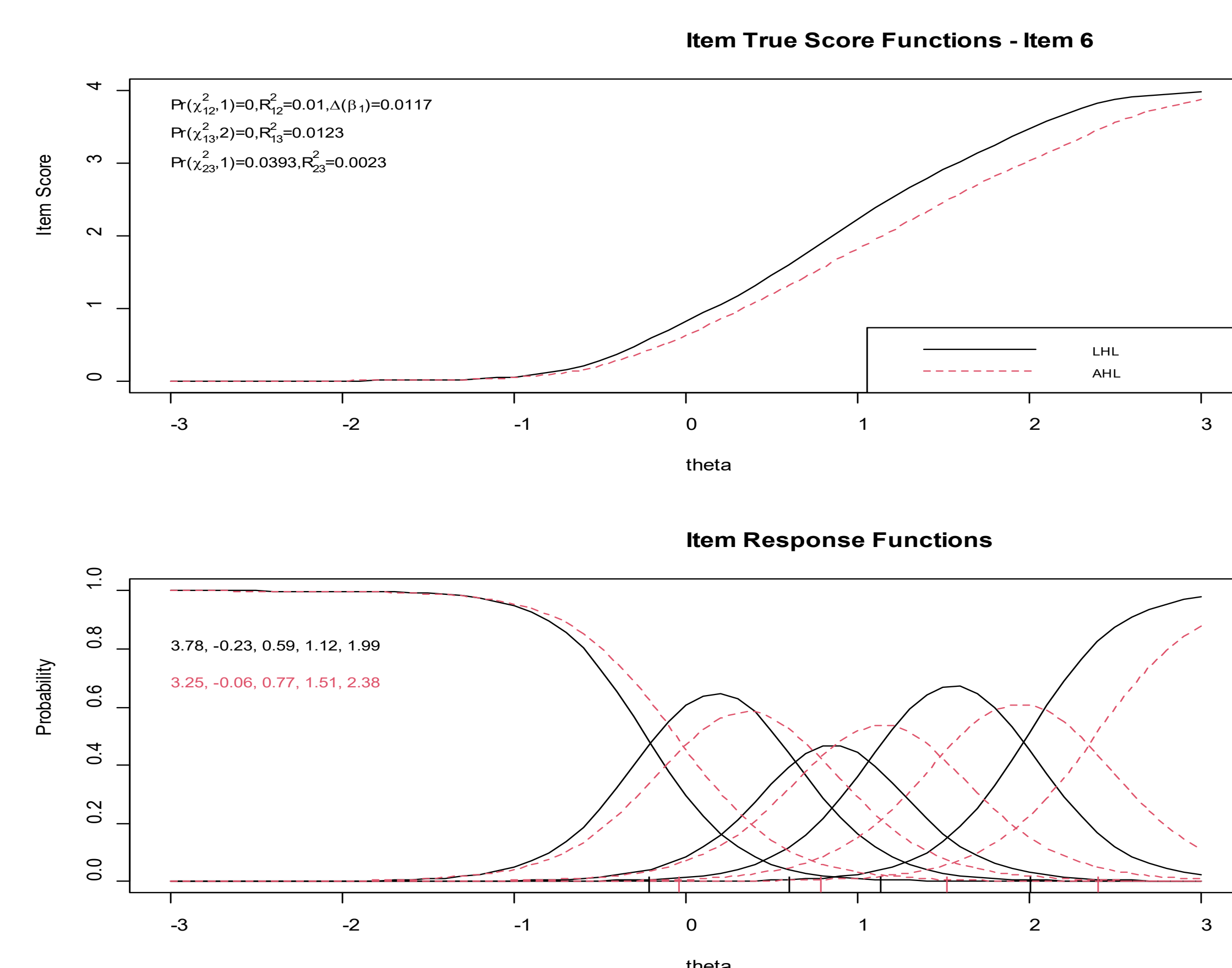


Figure 5. ICC of FATIMP49 *To what degree did your fatigue interfere with your physical functioning?*

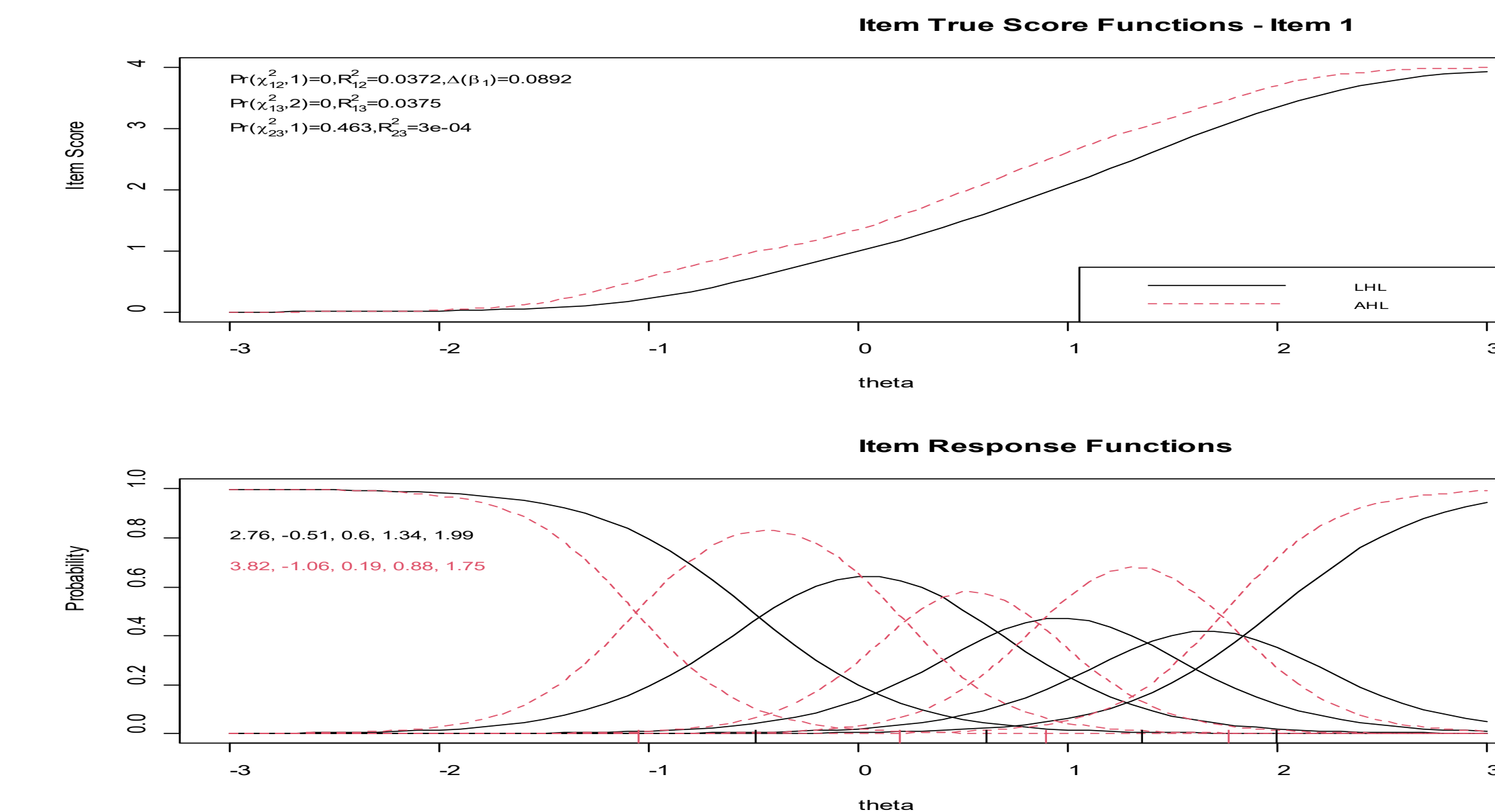


Figure 2. ICC of HI7 *During the past 7 days, I feel fatigued*

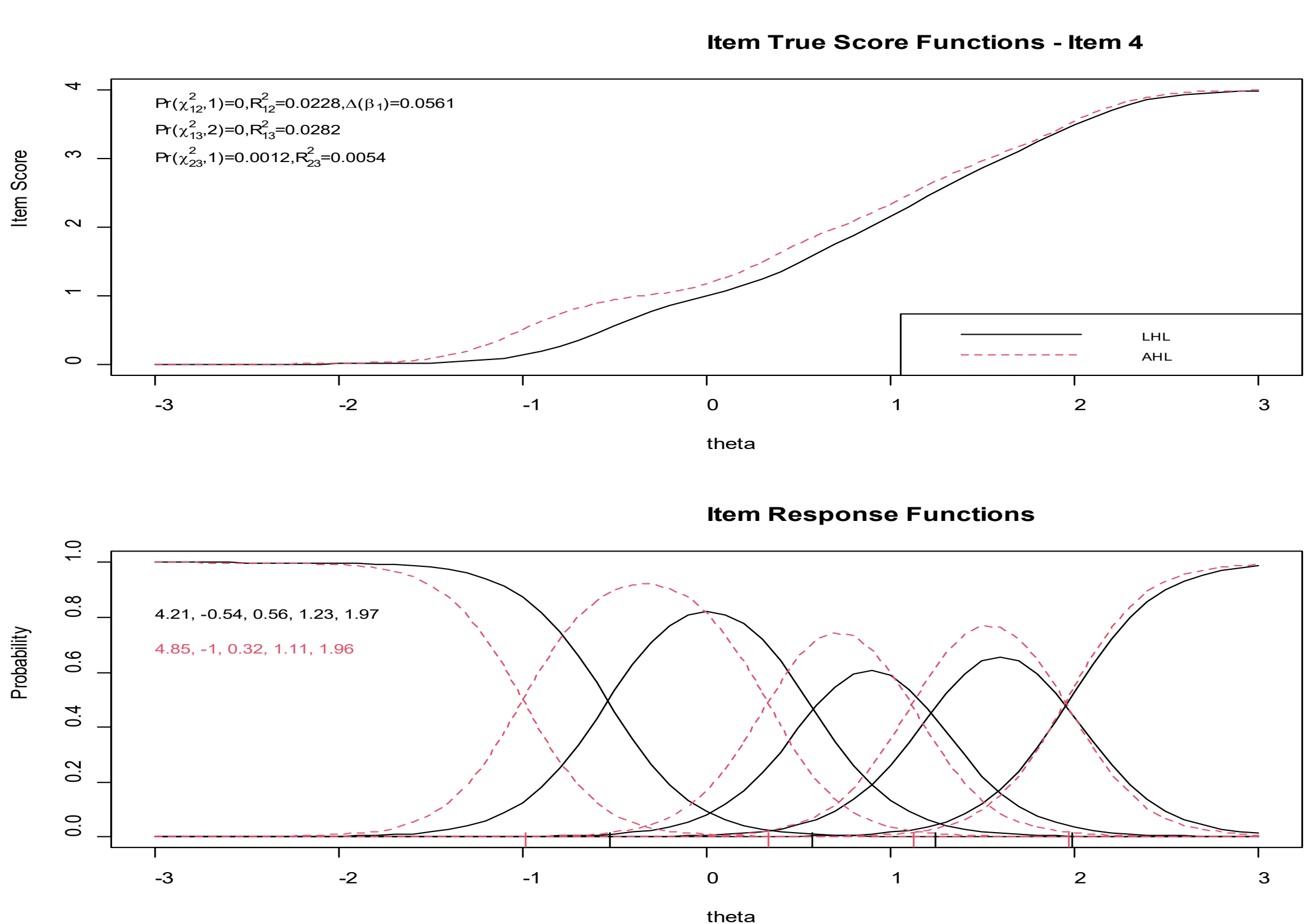


Figure 4. ICC of FATEXP40 *How fatigued were you on average?*

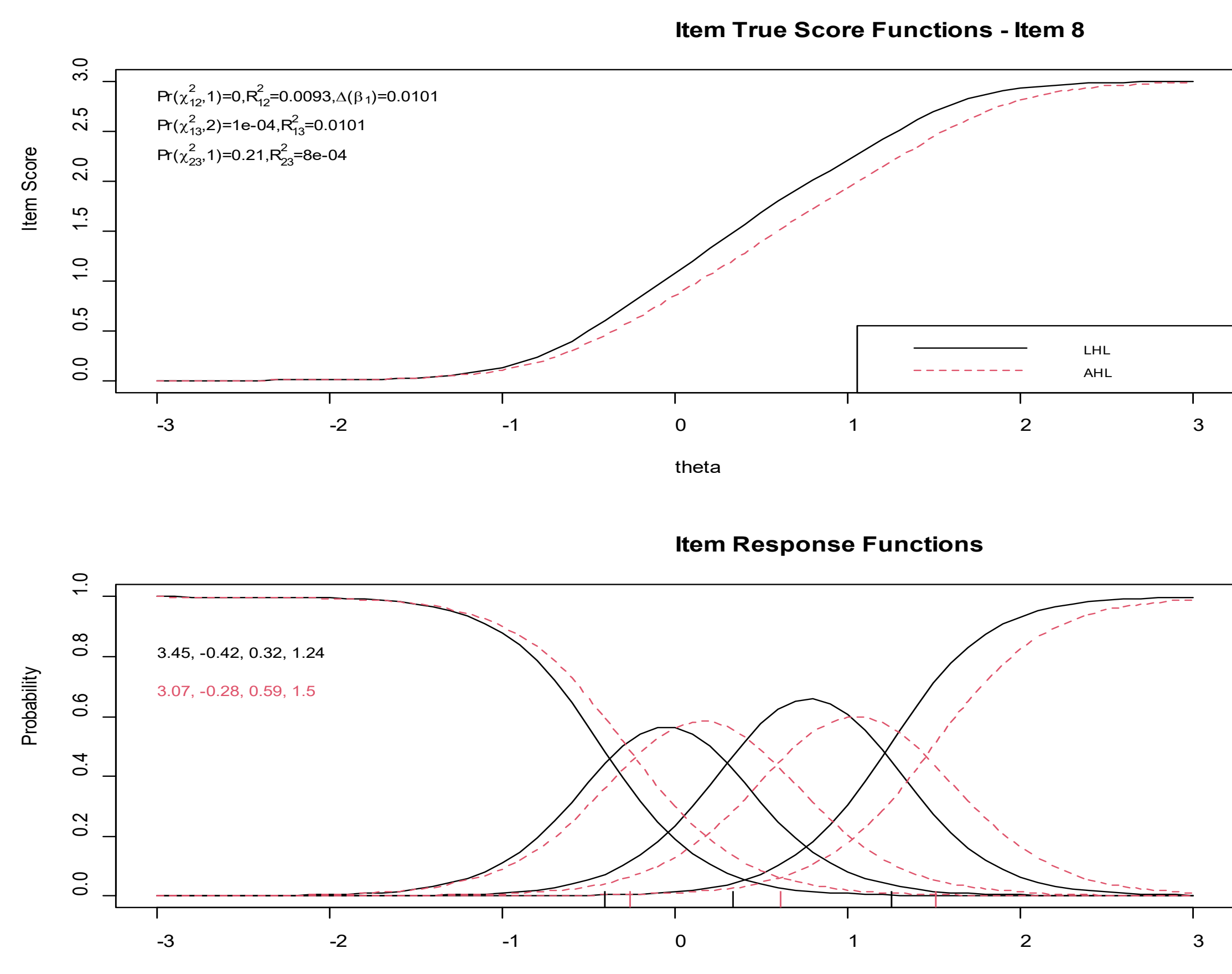


Figure 6. ICC of FATIMP16 *How often did you have trouble finishing things because of your fatigue?*

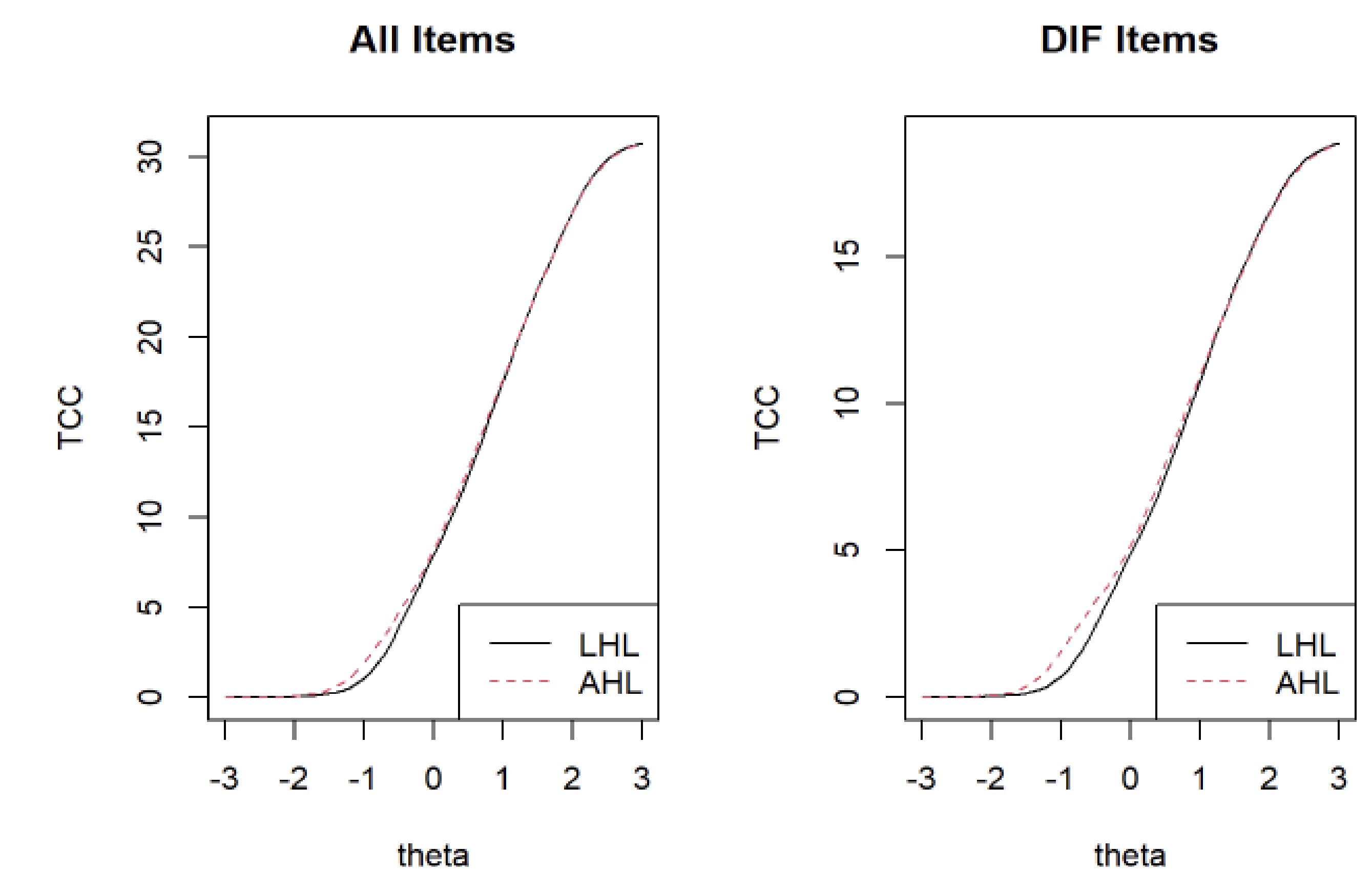


Figure 7. Test Characteristic Curves (TCCs)

Conclusion

- PROMIS Fatigue 8a showed sufficient unidimensionality.
- Observed DIF effects are not substantial enough to impact the overall scale score

Discussion

Strengths:

- Monte Carlo simulations used for item-specific DIF cutoffs
- Large sample size supporting robust analyses
- Diverse participants from two major metropolitan areas

Limitations:

- Non-clinical sample
- Boston and Chicago may not generalize to other places
- Reliance on Health LiTT only (no eHealth literacy or alternative measures)

References

• Buysse, D. J., Yu, L., Moul, D. E., Germain, A., Stover, A., Dodds, N. E., Johnston, K. L., Shablesky-Cade, M. A., & Pilkonis, P. A. (2010). Development and Validation of Patient-Reported Outcome Measures for Sleep Disturbance and Sleep-Related Impairments. *Sleep*, 33(6), 781–792. <https://doi.org/10.1093/sleep/33.6.781>

• Cella, D., Choi, S. W., Condon, D. M., Schalet, B., Hays, R. D., Rothrock, N. E., Yount, S., Cook, K. F., Gershon, R. C., Amtmann, D., DeWalt, D. A., Pilkonis, P. A., Stone, A. A., Weinfurt, K., & Reeve, B. B. (2019). PROMIS® Adult Health Profiles: Efficient Short-Form Measures of Seven Health Domains. *Value in Health*, 22(5), 537–544. <https://doi.org/10.1016/j.jval.2019.02.004>

• Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). lordif: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations. *Journal of Statistical Software*, 39(8). <https://doi.org/10.18637/jss.v039.i08>

• Hahn, E. A., Choi, S. W., Griffith, J. W., Yost, K. J., & Baker, D. W. (2011). Health Literacy Assessment Using Talking Touchscreen Technology (Health LiTT): A New Item Response Theory-Based Measure of Health Literacy. *Journal of Health Communication*, 16(sup3), 150–162. <https://doi.org/10.1080/10810730.2011.605434>

• Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., Johnston, K. L., & Pilkonis, P. A. (2012). Development of Short Forms From the PROMIS™ Sleep Disturbance and Sleep-Related Impairment Item Banks. *Behavioral Sleep Medicine*, 10(1), 6–24. <https://doi.org/10.1080/15402002.2012.636266>