

DR. GABRIEL LINS DE HOLANDA COELHO (Orcid ID : 0000-0003-4744-3151)

DR. AUDREY DUNNGALVIN (Orcid ID : 0000-0002-1540-3959)

DR. MATTHEW GREENHAWT (Orcid ID : 0000-0002-2365-9372)

DR. GANG CHEN (Orcid ID : 0000-0002-8385-5965)

Article type : Original article

Psychometric Parameters of Food Allergy Quality of Life During an Allergen Immunotherapy Trial

Authors: Gabriel Lins de Holanda Coelho¹, Audrey DunnGalvin^{1,2}, Matthew Greenhawt³, Jonathan O'B Hourihane⁴, David M. Fleischer³, Gang Chen⁵, Marcus Shaker^{6,7}, Dianne E. Campbell^{8,9}, Todd D. Green^{8,10}, Philippe Bégin¹¹

¹School of Applied Psychology, University College Cork, Cork, Ireland.

²Royal College of Surgeons in Ireland, Children's Health Ireland, Temple St Hospital, Dublin, Ireland.

³Children's Hospital Colorado, University of Colorado School of Medicine, Aurora, CO, USA.

⁴Department of Paediatrics, Royal College of Surgeons in Ireland and Children's Health Ireland, Temple St Hospital, Dublin, Ireland.

⁵Centre for Health Economics, Monash Business School, Monash University, Caulfield East, VIC, Australia.

⁶Section of Allergy and Immunology, Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA.

⁷Dartmouth Geisel School of Medicine, Hanover, NH, USA.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ALL.15323](https://doi.org/10.1111/ALL.15323)

This article is protected by copyright. All rights reserved

⁸DBV Technologies SA, Montrouge, France.

⁹The Children's Hospital at Westmead, Sydney, NSW, Australia.

¹⁰UPMC Children's Hospital of Pittsburgh and University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.

¹¹CHU Sainte-Justine, Montreal, QC, Canada.

Corresponding author: linshc@gmail.com or glhcoelho@ucc.ie (Dr. Gabriel Coelho)

Acknowledgments: This study was, in part, funded by DBV Technologies.

Conflicts of Interest: **Dianne E. Campbell** and **Todd D. Green** are employees of DBV Technologies. **Audrey DunnGalvin** has received research grants from Aimmune Therapeutics, National Children's Research Centre, DBV Technologies, and Food Allergy Research and Resource Program, as well as other research support from Safefood Ireland, and has served as a consultant and/or advisory board member for Aimmune Therapeutics, Atlantia Clinical Trials in Food Ireland, and Anaphylaxis Ireland. **Matthew Greenhawt** is a consultant for Aquestive; is a member of physician/medical advisory boards for DBV Technologies, Sanofi/Regeneron, Nutricia, Novartis, Aquestive, Allergy Therapeutics, AstraZeneca, ALK-Abello, and Protalix, with all activity unrelated to vaccines/vaccine development or COVID-19 treatment; is an unpaid member of the scientific advisory council for the National Peanut Board and medical advisory board of the International Food Protein Induced Enterocolitis Syndrome Association; is a member of the Brighton Collaboration Criteria Vaccine Anaphylaxis 2.0 working group; is the senior associate editor for the Annals of Allergy, Asthma, and Immunology, and is member of the Joint Taskforce on Allergy Practice Parameters. He has received honorarium for lectures from ImSci, MedLearningGroup, RMEI Medical Education and multiple state/local allergy societies. He received past research support ending in 2020 from the Agency for Healthcare Quality and Research (K08-HS024599). **Jonathan O'B Hourihane** reports advisory board fees, Aimmune Therapeutics; speakers bureau, Aimmune Therapeutics, Nutricia, Mead Johnson; grants to

institution/research funding, and clinical trials within past 2 years, Aimmune Therapeutics, Johnson& Johnson, National Childrens' Research Centre Ireland, Dublin Skin and Cancer Hospital Charity, Temple St Foundation and Clemens von Pirquet Foundation. **David M. Fleischer** has received research support to his institution from DBV Technologies and Aimmune Therapeutics; reports clinical medical advisory board membership with DBV Technologies; has served as a consultant for DBV Technologies, Allergenics, Aquestive Therapeutics, Aravax, Genentech, Nasus Pharma, and Intromune Therapeutics; and has received honorarium for lectures from DBV Technologies and royalties from UpToDate. **Marcus Shaker** has participated in research funded by DBV Technologies, is a member of the Joint Task Force on Practice Parameters, is an associate editor for *Annals of Allergy, Asthma, and Immunology*, and serves on the Editorial Boards for the *Journal of Allergy and Clinical Immunology In Practice* and the *Journal of Food Allergy*. **Philippe Bégin** has received grants through his institution from DBV Technologies during the conduct of the study; personal fees from ALK, Novartis, Sanofi, Bausch Health, Aralez, and Pfizer; and grants from Regeneron, Sanofi, Novartis and the Canadian Allergy, Asthma, and Immunology Foundation outside the submitted work. The following authors have no conflicts of interest to disclose: **Gang Chen** and **Gabriel Lins de Holanda Coelho**.

Author contribution: Gabriel Coelho and Audrey DunnGalvin contributed to study design, analysis, and manuscript preparation. Todd Green and Dianne Campbell contributed to study design and manuscript preparation. All the other authors contributed to manuscript preparation and review.

Abstract

Background: The Food Allergy Quality of Life Questionnaire–Parent Form (FAQLQ-PF) is a commonly used patient-reported outcome measure in food allergy (FA) research. It was developed before FA treatment clinical trials were commonplace and is used as a secondary outcome measure in pivotal FA treatment trials. We examined the psychometric properties of the FAQLQ-PF and its relevance to children with peanut allergy engaged in an epicutaneous immunotherapy (EPIT) clinical trial.

Methods: Analysis was performed on 26 universally answered items of the FAQLQ-PF, from assessments undertaken during the phase 3 PEPITES study (baseline, Month 12), which examined the safety and efficacy of EPIT for children with peanut allergy aged 4–11 years. Item response theory (IRT) was used to assess psychometric parameters of the FAQLQ-PF (ie, discrimination, difficulty, information). Confirmatory factor analysis was also employed; reliability was assessed using McDonald's omega (ω) and Cronbach's alpha (α).

Results: 23 of 26 items presented very high discrimination levels (> 1.7), and all 26 fell within the recommended difficulty threshold (between -1.5 and 1.5). The items contributed a reasonable information level for their respective factors/subdomains. The measure also presented a marginally acceptable model fit for the 3-factor structure (eg, comparative fit index = 0.88, Tucker-Lewis index = 0.87) and good reliability levels across timepoints (ω and $\alpha > 0.90$).

Conclusion: Herein, we present a novel re-analysis of the FAQLQ-PF items using IRT. The longitudinal performance of individual items and subscales was corroborated, and items with the highest discrimination were identified, showing that the tool is suitable for longitudinal measurements in FA treatment trials.

Keywords: Food Allergy Quality of Life; FAQLQ; item response theory; IRT; psychometrics

Introduction

Peanut allergy is associated with reduced health-related quality of life,¹⁻⁴ which can be assessed with disease-specific instruments such as the well-validated Food Allergy Quality of Life Questionnaire (FAQLQ) instruments.⁵⁻⁷ These instruments were designed and validated for the cross-sectional assessment of food allergy quality of life (FAQL). Now, with rapidly expanding treatment options being investigated for individuals with food allergies, these instruments are increasingly being used to examine potential changes in FAQL over time during clinical trials of food allergy immunotherapy.^{8,9} However, these tools were not explicitly designed or validated to be used in this setting, given their development predated all but one food therapy trial (TNX-901).

Because a scale's measurement properties can influence the interpretation of clinical trials and research results and determine its appropriateness for use in such contexts, it is essential to understand the psychometric properties of the FAQLQ more precisely. Item response theory (IRT) is a method for nuanced item analysis that analyzes the performance of individual items within patient-reported outcomes (PROs) assessments (such as those measured by the FAQLQ). Using IRT compared with traditional classical test theory (eg, exploratory and confirmatory factor analysis [CFA]) provides more detailed and more accurate descriptions of the item and scale-level performance. This is because classical test theory focuses on the structure of the measure and its external/internal validity, whereas an IRT approach offers more specific information about the items (eg, discrimination, difficulty, information). This information is obtained through mathematical formulas, expressing the relationship between observed and hypothetical variables, called latent traits. IRT assumes that the latent construct (in this case, FAQL) and items of a measure (eg, the FAQLQ-PF [FAQLQ–Parent Form]) are organized along a continuum.¹⁰ One of the technique's primary purposes is to establish the individual's position on that continuum. The idea of a continuum is helpful in the context of PROs for 2 main reasons: (1) the intervals between scores on a measure's response scale cannot necessarily be assumed to be equivalent (eg, 1–2 vs 4–5 may not represent equal degrees of change); (2) the weight, importance, or impact of each question in the questionnaire may differ depending on the particular respondent.¹¹ It also demonstrates equivalence across different subgroups, which is valuable for measures that are used in different settings.

In this post hoc analysis, we assessed the discrimination, difficulty, and information levels of each item of the FAQLQ-PF using IRT to understand its relevance to children with peanut

allergy involved in immunotherapy trials.¹² The longer-term objective is to provide data to improve and streamline the assessment of health-related quality of life during food allergy treatment. Data sets were derived from PEPITES, a global phase 3, pivotal, double-blind, placebo-controlled trial of epicutaneous immunotherapy for peanut allergy in children aged 4–11 years that evaluated the safety and efficacy of Viaskin™ Peanut (DBV712) 250 µg for peanut allergy.¹³

Method

Details of study designs, subject characteristics, and primary outcomes of the PEPITES study have been published previously.¹³ Briefly, PEPITES was a phase 3, multicenter, randomized, double-blind, placebo-controlled clinical trial of Viaskin Peanut, conducted between January 2016 and August 2017 in 4- to 11-year-old children with peanut allergy. Three hundred fifty-six participants (238 Viaskin Peanut 250 µg, 118 placebo) were enrolled from 31 sites in Australia, Canada, Germany, Ireland, and the United States. Overall, 79.5% of subjects were from North America, 13.8% from Europe (Ireland and Germany), and 6.7% from Australia. A total of 81.5% of subjects identified as Caucasian, 7.6% as Asian, and 0.8% as African American. Among other inclusion criteria, participants were required to react to an eliciting dose of ≤ 300 mg peanut protein on baseline double-blind, placebo-controlled food challenge (DBPCFC) and meet dual minimum peanut skin-prick test and serum peanut immunoglobulin E criteria.

Participant FAQL was assessed using the FAQLQ-PF for all subjects and the FAQLQ–Child Form (FAQLQ-CF) for subjects 8 years and older at entry and following 12 months of treatment as a prespecified endpoint. The FAQLQ-CF is only validated in persons 8 years of age and older. The FAQLQ assessment at Month 12 was performed prior to the DBPCFC that was used to assess the desensitization effect of Viaskin Peanut treatment, meaning that all assessments were obtained both prior to unblinding of study arm and challenge outcome.

We chose to assess 26 of the 30 possible items of the FAQLQ-PF because items 1–26 are completed by the parent (from the child’s perspective) in children aged 4–12 and applied to all participants in the sample. The remaining items, 27–30, are completed only on behalf of children aged 7–12 years and were only completed by the older age group in the sample. The FAQLQ-PF has previously been shown to have 3 domains (Food-Related Anxiety [FRA], Social and Dietary Limitations [SDL], Emotional Impact [EI]) derived through exploratory factor analysis during the initial validation of the index.¹⁴

In this study, we assess item quality within each of these domains, considering the following parameters: (1) **Discrimination** parameter (α) is an index of how well an item in the

FAQLQ can differentiate participants with varying levels of FAQL; the higher the value, the higher the item is discriminative between those with low and high FAQL. For discrimination, Baker's classification was adopted: very high (> 1.70), high (1.35–1.69), moderate (0.65–1.34), low (0.35–0.64), and very low (0.01–0.34) discrimination.¹² (2) **Difficulty** parameter (b) indicates the FAQL level that the individual must endorse to select the next higher response option category on the response scale. More difficult items tend to be endorsed only by those individuals who report higher scores of FAQL. In contrast, easier items tend to be endorsed by a broader range of individuals. Items that are neither too easy nor too difficult are recommended (ie, means across b1–b6 between -1.5 and 1.5).^{15,16} (3) **Item/test information curves** evaluate how much information an item shares with the total information of the measure. Items with a higher level of information are more informative, indicating higher measurement precision, less measurement error, and higher scale reliability.¹⁷

Data were analyzed using R software. All the analyses were performed twice—at baseline and Month 12. IRT analysis (ie, discrimination, difficulty, and informative curves) was performed using the multidimensional IRT package.¹⁸ The graded response model was considered in this package, as the FAQLQ uses a scale with more than 2 response categories (eg, 0–6 scale). Additionally, we performed CFA with the weighted least square mean and variance adjusted (WLSMV) estimator, considering the following model fit indices^{19,20}: the comparative fit index (CFI) and Tucker-Lewis index (TLI), which are both recommended to be higher than 0.90, and the root mean square error approximation (RMSEA), which must be lower than 0.08. Also, we assessed the reliability levels of the measure using McDonald's omega (ω) and Cronbach's alpha (α). Both ω and α are considered adequate when above 0.70.²¹ Missing data were not imputed as in IRT; a missing response does not contribute to the item's measurement.

Ethics approval for the post hoc analysis of prespecified data from the PEPITES study was not required.

Results

Subjects

Data from fully completed FAQLQ-PF assessments were available for 326 (106 placebo and 220 Viaskin Peanut 250 μ g subjects) of 356 randomized subjects' parents/caregivers at baseline and for 306 subjects (97 placebo and 209 Viaskin Peanut 250 μ g subjects), which were completed in full at both time points (baseline and Month 12). The median (interquartile range)

age of the 326 subjects included in this analysis was 7 years (6–9), and 62% were male. The analysis population consisted of 91% and 86% of the randomized study population.

Item Parameters: Discrimination, Difficulty, and Information

Twenty-three FAQLQ-PF items assessed at baseline showed very high levels of discrimination ($\alpha > 1.7$) (**Table 1**) both at baseline and Month 12. The other 3 items had either high (1.35–1.69) or moderate (0.65–1.34) discrimination. At baseline, Item 20 (FRA) and Item 18 (SDL) were the most discriminative, whereas Item 26 (EI) and Item 08 (SDL) were the least discriminative. At Month 12, Items 20 and 21 (FRA) were the most discriminative, whereas, once again, Items 26 and 08 were the least discriminative. Only Item 26 showed a difficulty level below the recommended threshold at baseline and Month 12. The full discrimination and difficulty values tables are available as **Supplemental Material** (See **Supplementary Tables 1 and 2**).

Next, we assessed each item's information level, as shown in the **Supplemental Material** (See **Supplementary Figures 1 to 6**). The information curves with a higher $I(\theta)$ (ie, higher curve) are more informative. With a few exceptions (Items 08 and 26, at both baseline and Month 12), results showed that the items shared reasonable information with their respective domains. Also, the sum of the contribution of all items that compose the FAQLQ factors generates the total information curve (**Figures 1a, 1b, and 1c**). The EI and FRA domains were the most informative, with a reasonable spread of information across their respective factors. All the domains were associated with reasonable (and very similar) information levels across the 2 time points, which show them to be reliable factors.

[FIGURES 1a, 1b and 1c ABOUT HERE]

FIGURE 1a. Test Information Curves (Food-Related Anxiety).

Unbroken line=baseline; dashed line=Month 12.

FIGURE 1b. Test Information Curves (Social and Dietary Limitations).

Unbroken line=baseline; dashed line=Month 12.

FIGURE 1c. Test Information Curves (Emotional Impact).

Unbroken line=baseline; dashed line=Month 12.

Confirmatory Factor Analysis

Using the WLSMV estimator, baseline FAQLQ-PF responses had the following model fit: CFI = 0.88, TLI = 0.87, and RMSEA = 0.074, indicating a marginally acceptable fit for the 3-factor structure. The results were similar for Month 12 responses: CFI = 0.88, TLI = 0.87, and

RMSEA = 0.077. All factorial weights (lambdas) were statistically different from zero ($\lambda \neq 0$; $z > 1.96$, $P < .01$).

Reliability

We used McDonald's omega (ω) and Cronbach's alpha (α) to assess the reliability levels. Results indicated good internal consistency for all domains across both stages²¹: FRA (Stages 1 and 2, ω , and $\alpha = 0.93$), SDL (Stage 1, ω , and $\alpha = 0.90$; Stage 2, ω , and $\alpha = 0.91$), and EI (Stage 1 and 2, ω , and $\alpha = 0.92$). The full measure also was associated with high reliability levels (Stage 1, ω , and $\alpha = 0.96$; Stage 2, ω , and $\alpha = 0.97$).

Discussion

IRT, used extensively in educational testing applications, has gained popularity and acceptance in PRO-related research, which is vital to assessing treatment outcomes.²² Although US regulatory authorities will not accept PRO data as primary outcome measures for product approval, they are increasingly recognized by regulators, clinicians, and patients as valuable tools to collect patient-centered data.^{23,24} Indeed, the US Food and Drug Administration has encouraged their incorporation in clinical trials and has recently issued guidance for the industry on the use of core PRO measures in clinical cancer trials in 2020 and 2021.²⁵

IRT analyzes the individual parameters of items to ensure measurement quality,¹⁰ providing a deeper and more accurate description of item-level PRO questionnaire data.²⁶ Specifically, the technique creates a link between the items (or questions) on a measure (ie, the FAQLQ), the individuals responding to these items, and the underlying construct being measured (ie, FAQL). As a result of this synergy, in any particular context (eg, food immunotherapy), items with the best parameters will provide more robust results. The technique has been applied successfully in many settings, including in health outcomes research in oncology,²⁶ multiple sclerosis,^{27,28} and diabetes.²⁹ This success led the Patient-Reported Outcomes Measurement Information System initiative to recommend the use of IRT as a practical approach to examining the construct validity of PROs.³⁰

We considered 3 crucial individual parameters for our analyses: Discrimination, Difficulty, and Information. More specifically, for an item to be adequate to the composition of the measure, it must (1) be able to discriminate participants between different levels on the latent trait or construct (ie, FAQL)¹²; (2) be neither too easy nor too difficult (suitable for patients with different levels of FAQL, from poor to excellent)^{15,16}; and (3) provide sufficient information to the total score of the measure (ie, the FAQLQ).¹⁷ Unlike classical test theory (which is typically used to

develop or refine health questionnaires), IRT considers the number of questions answered correctly *and* the difficulty of the questions answered. For example, 2 respondents (A and B) received a total score of 3.2 but had higher scores for some of the questions and lower scores for others. Although they had the same score, their food allergy burden may not be the same if the questions that make up the total score are “easier” for A and more “difficult” for B.

The results were consistent across both baseline and Month 12 data sets. Most items had difficulty within the recommended level and contributed a reasonable amount of information to their respective domains. Also, 23 of the 26 items in the FAQLQ-PF had very high discrimination levels. Item 08 had moderate discrimination levels at baseline and Month 12, and Item 26 had difficulty levels slightly beyond the recommended threshold. It is essential to highlight that these results are expected due to the general content of these items. For instance, Item 26 states that “... *wishes his/her food allergy would go away,*” which is something with which all those affected by the disease would likely agree, logically. Therefore, this item would be too “easy.” Despite the results in this item’s parameters, it is unlikely that its inclusion to measure FAQL will influence the research outcome when all findings are taken together.

Furthermore, CFA showed a marginally acceptable model fit for the 3-factor structure. Such findings confirm that the 3 domains of the FAQLQ (FRA, SDL, and EI) provide a solid theoretical framework, making the FAQLQ-PF a robust tool for assessing the impact of immunotherapy on the concept of FAQL. The reliability levels provide further evidence of the quality of the measure across time points for both the factors in isolation and for the full measure itself.²¹ Limitations of the analysis include that not all parents completed in full the FAQLQ at either baseline or follow-up at Month 12. However, as our sample is originated in one of the most extensive food allergy immunotherapy trials conducted to date, the results are likely generalizable across the whole study population. Given that this population sample was obtained in the context of a phase 3 clinical treatment trial, offered only at particular institutions and geographic regions, which may cluster to more severe or difficult-to-manage cases, there may be some selection bias. Furthermore, the majority of the subjects were from North America, and the relatively smaller contribution of subjects from Europe and Australia did not allow for analysis of variability of findings across regions. Additional selection bias may result from a clinical trial population that may differ considerably from the general population with food allergy. IRT is a theory of content analysis and not an exclusive method of doing so, and, thus, other analytic techniques may not

lead to the same conclusions reached herein. This IRT was performed in the context of peanut allergy for this index, and, in theory, this may differ when used in the context of other allergens.

Final Considerations

The FAQLQ was developed before most current research into treatments was performed. The index themes may be more reflective of the world before recent labelling laws, stock epinephrine legislation, and the advent of more treatment trials and clinical practice oral immunotherapy. Thus, the clinical impact of certain items in the final rating of the initial sample at the time may differ from how a current sample may rate the item's clinical impact, which would affect item inclusion/exclusion. The use of IRT may be a strategy to ensure the items included in a PRO measure have suitable parameters that can be reliably used to assess the FAQL construct, ensuring robust results. Items with unacceptable parameters may bias the analyses, leading to questionable results. IRT can also help find causative agents underlying the score differences and match them to the subgroup/group characteristics.

We plan to use the current analysis to validate further a short form of the FAQLQ that may be more easily administered across treatment settings than the current full-length tool. Our findings provide a basis for selecting the most sensitive items to streamline the FAQL assessment process and create a more sensitive and robust tool for evaluating the population with peanut allergy in future clinical trials and clinical practice of food allergen immunotherapy. Moreover, our IRT findings might support the development of computerized adaptive testing (CAT) applications using the FAQLQ.²² Whereas questionnaires use the same questions for all respondents, CAT uses an algorithm to tailor the items presented based on respondents' level of FAQL. That is, they will answer a specific initial question and, based on their level of agreement with it, will be presented with different (and reduced) sets of items. The use of IRT to support CAT has been increasing over the years, especially for PRO domains.

References

1. Walkner M, Warren C, Gupta RS. Quality of life in food allergy patients and their families. *Pediatr Clin North Am*. 2015;62(6):1453-1461.
2. Roy KM, Roberts MC. Peanut allergy in children: relationships to health-related quality of life, anxiety, and parental stress. *Clin Pediatr (Phila)*. 2011;50(11):1045-1051.
3. Shaker MS, Schwartz J, Ferguson M. An update on the impact of food allergy on anxiety and quality of life. *Curr Opin Pediatr*. 2017;29(4):497-502.

4. Stensgaard A, Bindslev-Jensen C, Nielsen D, Munch M, DunnGalvin A. Quality of life in childhood, adolescence and adult food allergy: patient and parent perspectives. *Clin Exp Allergy*. 2017;47(4):530-539.
5. Muraro A, Dubois AE, DunnGalvin A, et al. EAACI food allergy and anaphylaxis guidelines. Food allergy health-related quality of life measures. *Allergy*. 2014;69(7):845-853.
6. DunnGalvin A, Flokstra-de Blok BM, Burks AW, Dubois AE, Hourihane JO. Food allergy QoL questionnaire for children aged 0-12 years: content, construct, and cross-cultural validity. *Clin Exp Allergy*. 2008;38(6):977-986.
7. Flokstra-de Blok BM, DunnGalvin A, Vlieg-Boerstra BJ, et al. Development and validation of a self-administered Food Allergy Quality of Life Questionnaire for children. *Clin Exp Allergy*. 2009;39(1):127-137.
8. DunnGalvin A, Fleischer DM, Campbell DE, et al. Improvements in quality of life in children following epicutaneous immunotherapy (EPIT) for peanut allergy in the PEPITES and PEOPLE studies. *J Allergy Clin Immunol Pract*. 2021;9(1):216-224.e1.
9. Fernandez-Rivas M, Vereda A, Vickery BP, et al. Open-label follow-on study evaluating the efficacy, safety, and quality of life with extended daily oral immunotherapy in children with peanut allergy. *Allergy*. Published online July 28, 2021. doi:10.1111/all.15027.
10. Pasquali L, Primi R. Fundamentos da teoria da resposta ao item: TRI. *Aval Psicológica*. 2003;2(2):99-110.
11. Liddell TM, Kruschke JK. Analyzing ordinal data with metric models: What could possibly go wrong? *J Exp Soc Psychol*. 2018;79:328-348.
12. Baker FB. *The Basics of Item Response Theory*. 2nd ed. ERIC Clearinghouse on Assessment and Evaluation; 2001.
13. Fleischer DM, Greenhawt M, Sussman G, et al. Effect of epicutaneous immunotherapy vs placebo on reaction to peanut protein ingestion among children with peanut allergy: the PEPITES randomized clinical trial. *JAMA*. 2019;321(10):946-955.
14. DunnGalvin A, Cullinane C, Daly DA, Flokstra-de Blok BM, Dubois AE, Hourihane JO. Longitudinal validity and responsiveness of the Food Allergy Quality of Life Questionnaire - Parent Form in children 0-12 years following positive and negative food challenges. *Clin Exp Allergy*. 2010;40(3):476-485.

- Accepted Article
15. Lins de Holanda Coelho G, H P Hanel P, J Wolf L. The very efficient assessment of need for cognition: developing a six-item version. *Assessment*. 2020;27(8):1870-1885.
 16. Rauthmann JF. Investigating the MACH-IV with item response theory and proposing the trimmed MACH*. *J Pers Assess*. 2013;95(4):388-397.
 17. Cappelleri JC, Jason Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin Ther*. 2014;36(5):648-662.
 18. Chalmers RP. mirt: A multidimensional item response theory package for the R environment. *J Stat Soft*. 2012;48:1-29.
 19. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 6th ed. Pearson; 2013.
 20. Hair JF Jr, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. 7th ed. Prentice Hall; 2009.
 21. Kline P. *Handbook of Psychological Testing*. 2nd ed. Routledge; 2013.
 22. Nguyen TH, Han HR, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. *Patient*. 2014;7(1):23-35.
 23. Arciero V, Delos Santos S, Koshy L, et al. Assessment of Food and Drug Administration– and European Medicines Agency–approved systemic oncology therapies and clinically meaningful improvements in quality of life: a systematic review. *JAMA Network Open*. 2021;4(2):e2033004. doi:10.1001/jamanetworkopen.2020.33004.
 24. Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas*. 2018;9:353-367.
 25. Food and Drug Administration. Core patient-reported outcomes in cancer clinical trials. Guidance for industry. Accessed March 3, 2022. <https://www.fda.gov/media/149994/download>.
 26. Schindler E, Friberg LE, Lum BL, et al. A pharmacometric analysis of patient-reported outcomes in breast cancer patients through item response theory. *Pharm Res*. 2018;35(6):122.
 27. Novakovic AM, Krekels EH, Munafo A, Ueckert S, Karlsson MO. Application of item response theory to modeling of expanded disability status scale in multiple sclerosis. *AAPS J*. 2017;19(1):172-179.

- Accepted Article
28. Jin X, Liu GG, Gerstein HC, et al. Item reduction and validation of the Chinese version of diabetes quality-of-life measure (DQOL). *Health Qual Life Outcomes*. 2018;16(1):78.
 29. Borg S, Eeg-Olofsson K, Palaszewski B, Svedbo Engström M, Gerdtham UG, Gudbjörnsdottir S. Patient-reported outcome and experience measures for diabetes: development of scale models, differences between patient groups and relationships with cardiovascular and diabetes complication risk factors, in a combined registry and survey study in Sweden. *BMJ Open*. 2019;9(1):e025033. doi:10.1136/bmjopen-2018-025033.
 30. PROMIS. Instrument Development and Validation Scientific Standards. Published 2013. Accessed March 3, 2022. https://www.mcgill.ca/can-pro-network/files/can-pro-network/promisstandards_vers2.0_final.pdf.

Table 1. Item parameters of the FAQLQ – PF at baseline and Month 12.

Item	Discrimination— <i>a</i>		Difficulty— <i>b(m)</i>	
	<i>TM=0</i>	<i>TM=12</i>	<i>TM=0</i>	<i>TM=12</i>
Because of food allergy, my child...				
Food-Related Anxiety				
1. ... feels anxious about food	2.769	2.405	-0.08	0.14
4. ... feels afraid to try unfamiliar foods	2.105	2.253	-0.26	-0.12
5. ... feels concerned that I am worried that he/she will have a reaction to food	1.839	1.955	0.08	0.39
16. ... feels anxious when going to new places	3.334	3.189	0.50	0.68
17. ... feels concerned that he/she must always be cautious about food	2.912	3.06	-0.33	-0.15
20. ... feels anxious about accidentally eating an ingredient to which he/she is allergic	3.516	3.751	-0.21	-0.03
21. ... feels anxious when eating with unfamiliar adults/children	2.966	3.443	0.10	0.15
Social and Dietary Limitations				
3. ... feels frustrated by dietary restrictions	2.025	1.998	-0.02	0.07
8. ... has a lack of variety in his/her diet	1.095	1.117	0.86	1.07
12. ... 's social environment is restricted because of limitations on restaurants we can safely go to as a family	2.509	2.217	-0.60	-0.42
13. ... 's social environment is restricted because of limitations on holiday destinations we can safely go to as a family	1.891	2.169	-0.24	-0.10
14. ... 's ability to take part has been limited in social activities in other people's houses (sleepovers, parties, playtime)	2.067	2.121	-0.43	-0.07
15. ... 's ability to take part has been limited in preschool/school events involving food (class parties/treats/lunchtime)	2.18	2.254	-0.39	-0.24
18. ... feels "left out" in activities involving food	3.454	3.162	-0.22	-0.10
22. ... feels frustrated by social restrictions	2.781	3.369	0.14	0.23
Emotional Impact				
2. ... feels different from other children	1.974	1.906	0.02	0.24
6. ... experiences physical distress	1.9	1.946	1.25	1.45
7. ... experiences emotional distress	3.128	3.025	0.41	0.54
9. ... has been negatively affected by receiving more attention than other children of his/her age	2.151	2.045	0.70	0.91

10. ... has been negatively affected by having to grow up more quickly than other children of his/her age	2.055	2.195	0.61	0.71
11. ... has been negatively affected by his/her environment being more restricted than other children of his/her age	2.24	2.188	-0.02	0.22
19. ... feels upset that family social outings (eating out, celebrations, days out) have been limited by food allergy	2.024	2.113	0.31	0.48
23. ... is more anxious in general than other children of his/her age	2.232	2.713	0.47	0.61
24. ... is more cautious in general than other children of his/her age	1.681	1.661	-0.33	-0.08
25. ... is not as confident as other children of his/her age in social situations	1.795	2.188	1.04	1.12
26. ... wishes his/her food allergy would go away	1.418	1.22	-1.63	-1.77

Note: a =discrimination levels; $b(m)$ =means between $b1 - b6$.

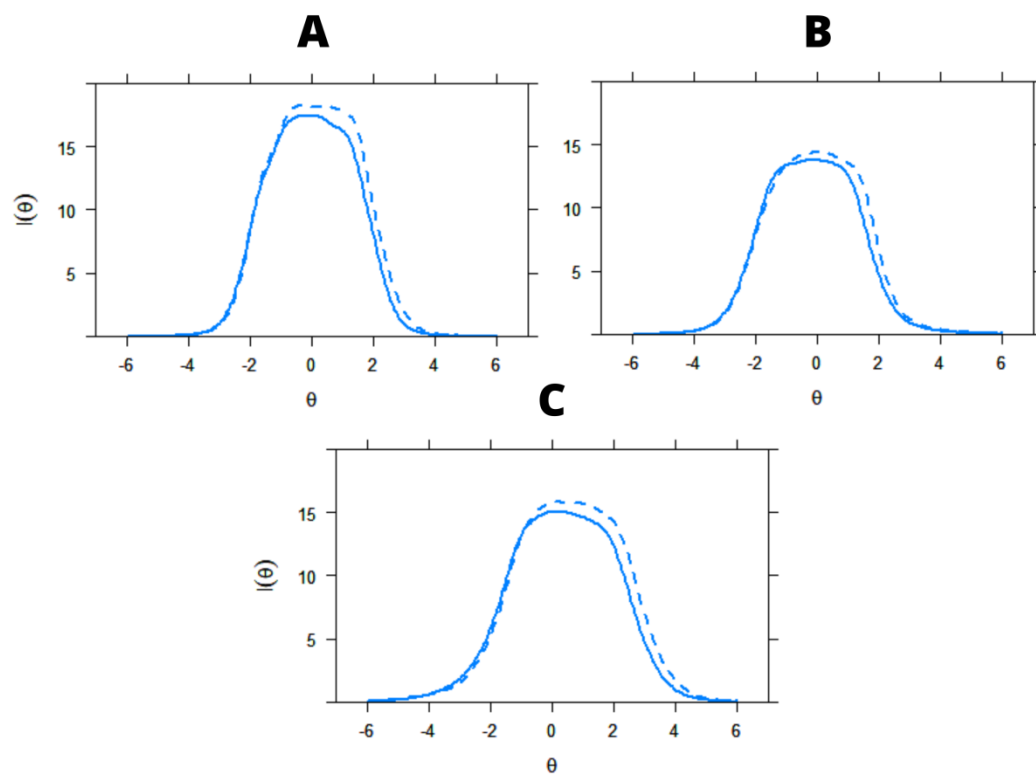


Figure 1 (A, B, C) - Coelho et al.

all_15323_f1.png

Table S1. *Item parameters of the FAQLQ - Parents (Stage 1).*

Item	<i>a</i>	<i>b1</i>	<i>b2</i>	<i>b3</i>	<i>b4</i>	<i>b5</i>	<i>b6</i>	<i>b(m)</i>
Food-related Anxiety								
<i>Item01</i>	2.769	-1.649	-0.989	-0.446	0.19	0.81	1.608	-.08
<i>Item04</i>	2.105	-1.636	-0.983	-0.527	-0.01	0.407	1.2	-.26
<i>Item05</i>	1.839	-1.45	-0.794	-0.25	0.333	0.991	1.677	.08
<i>Item16</i>	3.334	-0.84	-0.345	0.172	0.698	1.332	1.956	.50
<i>Item17</i>	2.912	-1.879	-1.049	-0.575	-0.036	0.489	1.089	-.33
<i>Item20</i>	3.516	-1.639	-0.833	-0.373	0.044	0.417	1.124	-.21
<i>Item21</i>	2.966	-1.245	-0.615	-0.136	0.339	0.818	1.426	.10
Social & Dietary Limitations								
<i>Item03</i>	2.025	-1.67	-0.813	-0.33	0.241	0.883	1.581	-.02
<i>Item08</i>	1.095	-1.029	-0.01	0.592	1.143	1.839	2.603	.86
<i>Item12</i>	2.509	-1.879	-1.451	-0.834	-0.359	0.131	0.806	-.60
<i>Item13</i>	1.891	-1.368	-0.905	-0.589	-0.132	0.39	1.154	-.24
<i>Item14</i>	2.067	-1.938	-1.199	-0.642	-0.173	0.346	1.011	-.43
<i>Item15</i>	2.18	-1.951	-1.219	-0.657	-0.108	0.491	1.116	-.39
<i>Item18</i>	3.454	-1.517	-1.013	-0.431	0.038	0.526	1.058	-.22
<i>Item22</i>	2.781	-1.169	-0.568	-0.113	0.361	0.864	1.492	.14
Emotional Impact								
<i>Item02</i>	1.974	-1.752	-0.994	-0.439	0.309	1.003	1.965	.02
<i>Item06</i>	1.9	-0.347	0.358	0.955	1.522	2.22	2.781	1.25
<i>Item07</i>	3.128	-0.989	-0.34	0.142	0.63	1.183	1.82	.41
<i>Item09</i>	2.151	-0.896	-0.19	0.34	0.965	1.679	2.297	.70
<i>Item10</i>	2.055	-0.907	-0.22	0.283	0.882	1.377	2.218	.61
<i>Item11</i>	2.24	-1.459	-0.862	-0.313	0.264	0.776	1.495	-.02
<i>Item19</i>	2.024	-1.27	-0.547	-0.087	0.494	1.147	2.107	.31
<i>Item23</i>	2.232	-0.973	-0.24	0.216	0.641	1.269	1.93	.47
<i>Item24</i>	1.681	-2.255	-1.272	-0.662	-0.07	0.716	1.577	-.33
<i>Item25</i>	1.795	-0.441	0.325	0.747	1.289	1.886	2.46	1.04
<i>Item26</i>	1.418	-3.108	-2.306	-1.935	-1.37	-0.83	-0.246	-1.63

Note. *a* = discrimination levels, very high (green), high (yellow), moderate (red); *b1* – *b4* = threshold; *b(m)* = means between *b1* – *b6*, acceptable levels in green.

Table S2. *Item parameters of the FAQLQ - Parents (Stage 2).*

Item	<i>a</i>	<i>b1</i>	<i>b2</i>	<i>b3</i>	<i>b4</i>	<i>b5</i>	<i>b6</i>	<i>b(m)</i>
Food-related Anxiety								
<i>Item01</i>	2.405	-1.677	-0.823	-0.267	0.466	1.166	1.954	.14
<i>Item04</i>	2.253	-1.55	-0.872	-0.336	0.09	0.606	1.369	-.12
<i>Item05</i>	1.955	-1.309	-0.606	-0.069	0.783	1.462	2.098	.39
<i>Item16</i>	3.189	-0.796	-0.144	0.337	0.933	1.508	2.22	.68
<i>Item17</i>	3.06	-1.8	-0.904	-0.449	0.117	0.735	1.404	-.15
<i>Item20</i>	3.751	-1.607	-0.749	-0.302	0.301	0.772	1.378	-.03
<i>Item21</i>	3.443	-1.224	-0.568	-0.106	0.393	0.884	1.519	.15
Social & Dietary Limitations								
<i>Item03</i>	1.998	-1.726	-0.9	-0.241	0.501	1.041	1.729	.07
<i>Item08</i>	1.117	-0.833	0.067	0.587	1.338	2.06	3.171	1.07
<i>Item12</i>	2.217	-1.904	-1.225	-0.616	-0.144	0.335	1.043	-.42
<i>Item13</i>	2.169	-1.363	-0.823	-0.388	0.092	0.571	1.315	-.10
<i>Item14</i>	2.121	-1.476	-0.798	-0.351	0.133	0.683	1.37	-.07
<i>Item15</i>	2.254	-1.988	-1.056	-0.531	0.152	0.7	1.305	-.24
<i>Item18</i>	3.162	-1.644	-0.806	-0.298	0.209	0.626	1.304	-.10
<i>Item22</i>	3.369	-1.067	-0.442	-0.049	0.395	0.972	1.585	.23
Emotional Impact								
<i>Item02</i>	1.906	-1.777	-0.85	-0.165	0.581	1.34	2.312	.24
<i>Item06</i>	1.946	-0.349	0.529	1.128	1.791	2.359	3.24	1.45
<i>Item07</i>	3.025	-1.005	-0.284	0.157	0.858	1.424	2.09	.54
<i>Item09</i>	2.045	-0.876	0.008	0.59	1.202	1.806	2.756	.91
<i>Item10</i>	2.195	-0.923	-0.114	0.365	0.845	1.508	2.559	.71
<i>Item11</i>	2.188	-1.55	-0.639	-0.036	0.523	1.131	1.863	.22
<i>Item19</i>	2.113	-1.064	-0.288	0.153	0.738	1.345	2.006	.48
<i>Item23</i>	2.713	-0.731	-0.135	0.306	0.81	1.291	2.128	.61
<i>Item24</i>	1.661	-1.975	-0.978	-0.45	0.156	0.906	1.866	-.08
<i>Item25</i>	2.188	-0.377	0.417	0.761	1.338	1.884	2.671	1.12
<i>Item26</i>	1.22	-3.759	-2.318	-1.863	-1.423	-1.032	-0.247	-1.77

Note. *a* = discrimination levels, very high (green), high (yellow), moderate (red); *b1* – *b4* = threshold; *b(m)* = means between *b1* – *b6*, acceptable levels in green.

Item information trace lines

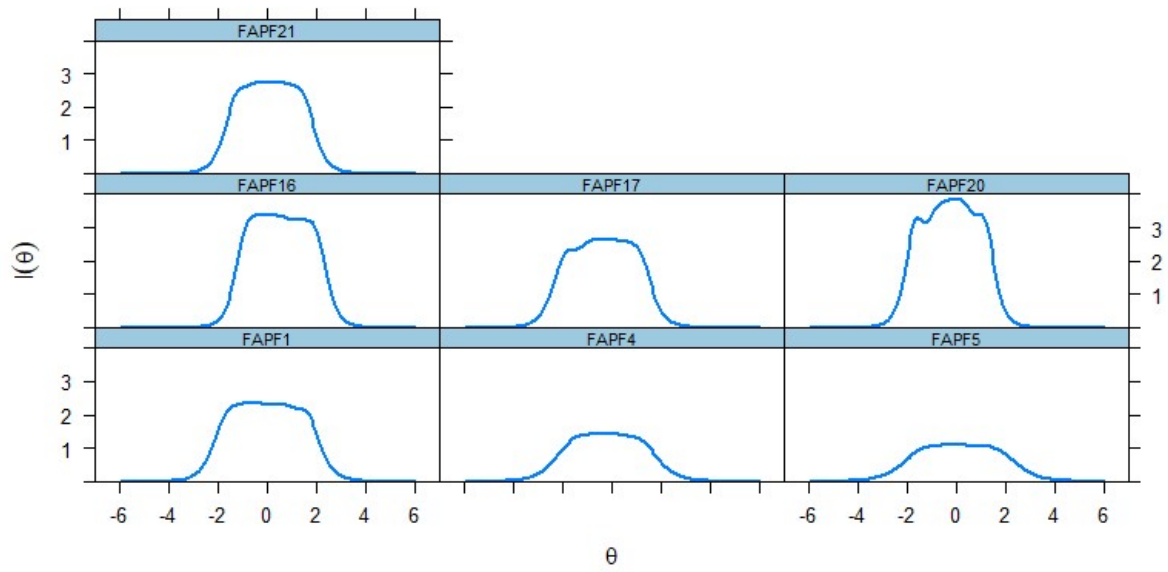


Figure S1. Item information curves (Food-related Anxiety; Stage 1)

Item information trace lines

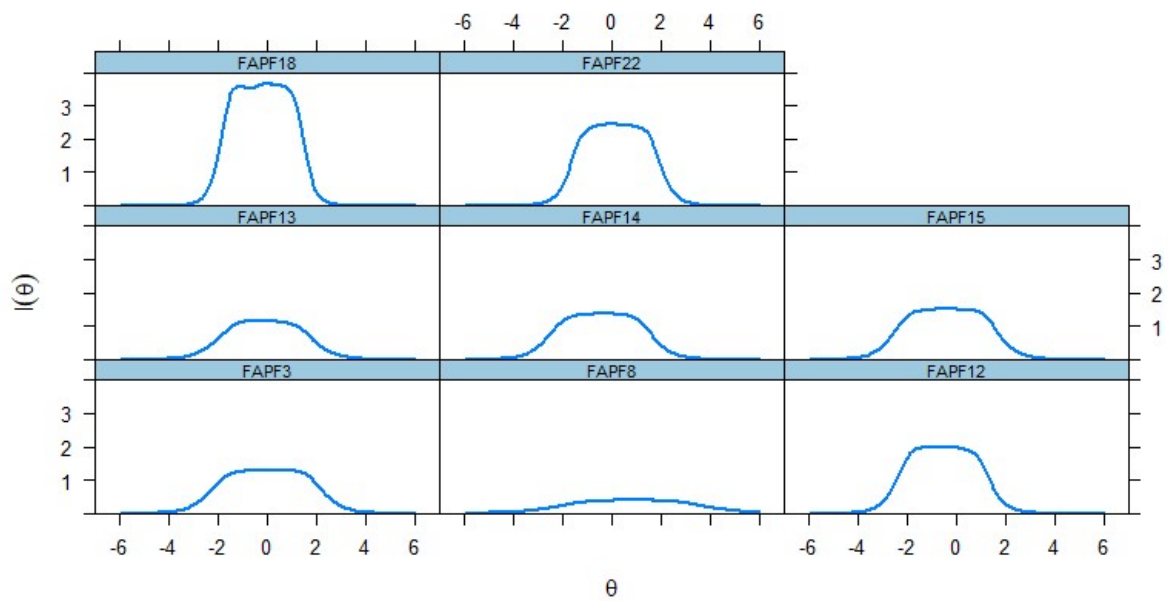


Figure S2. Item information curves (Social & Dietary Limitations; Stage 1)

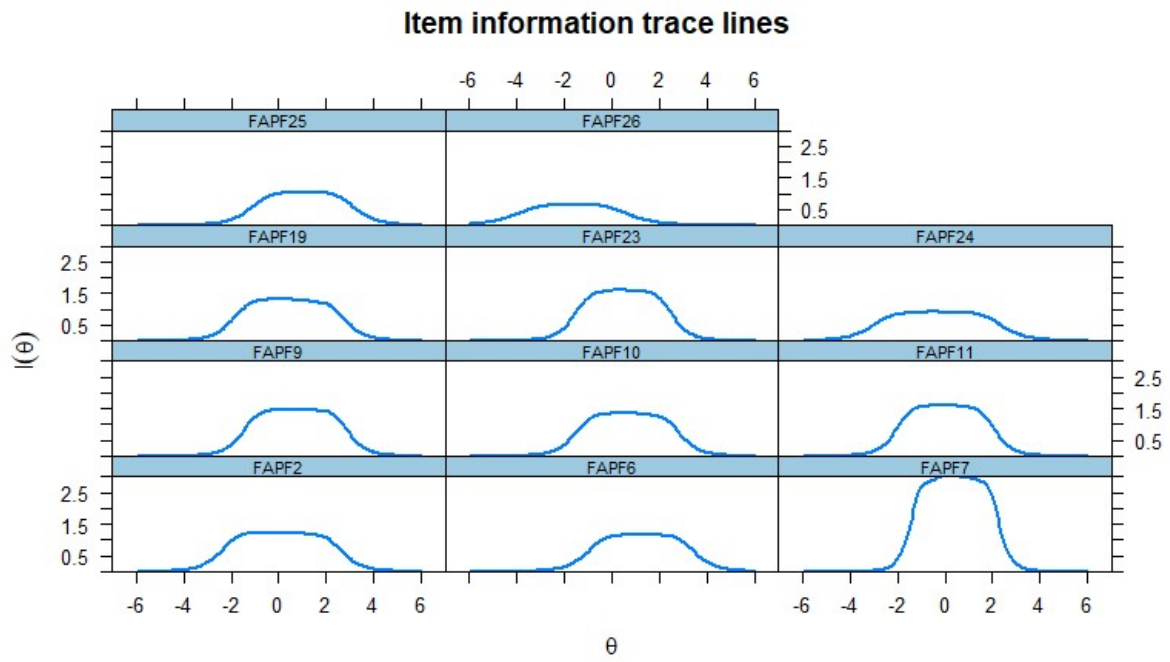


Figure S3. Item information curves (Emotional Impact; Stage 1)

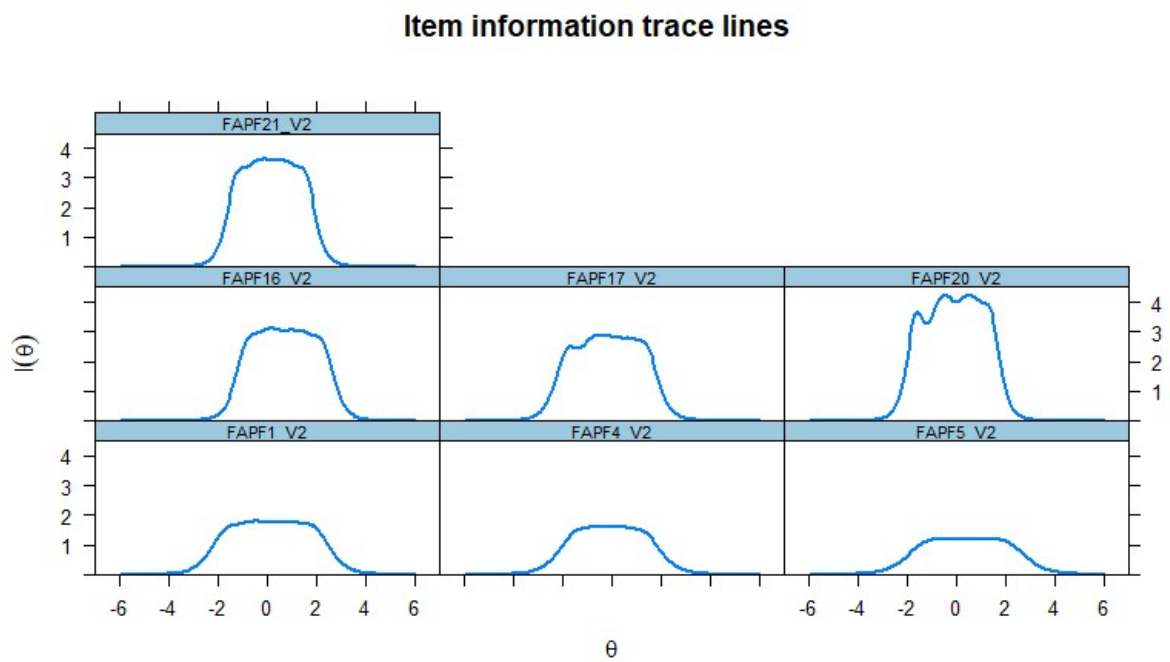


Figure S4. Item information curves (Food-related Anxiety; Stage 2)

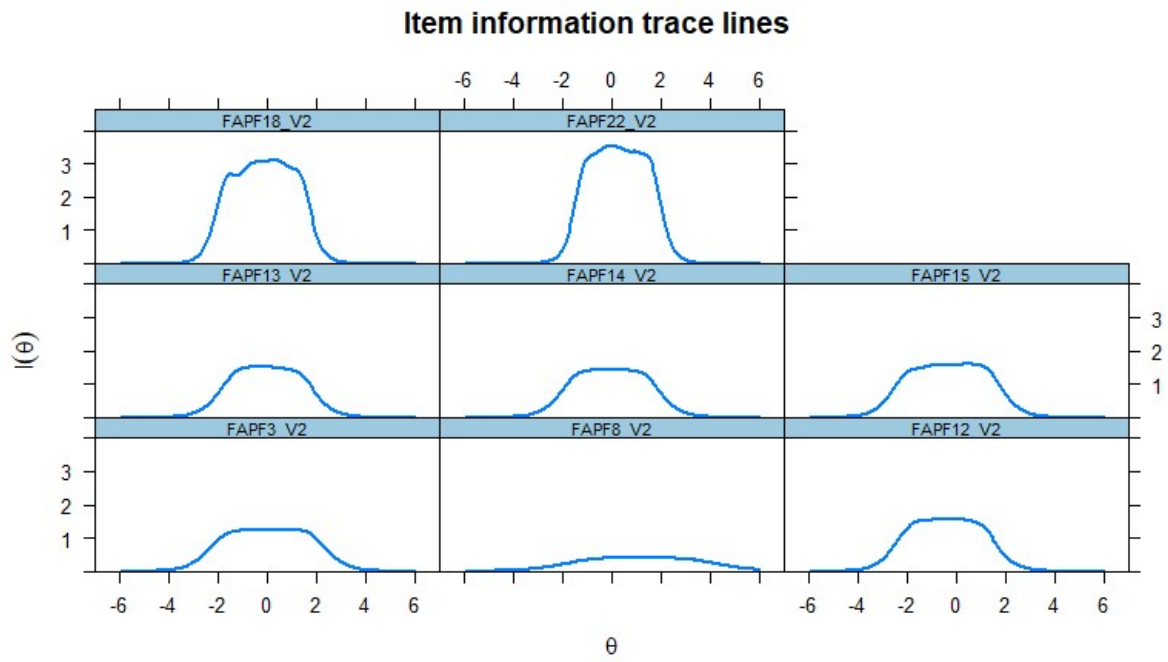


Figure S5. Item information curves (Social & Dietary Limitations; Stage 2)

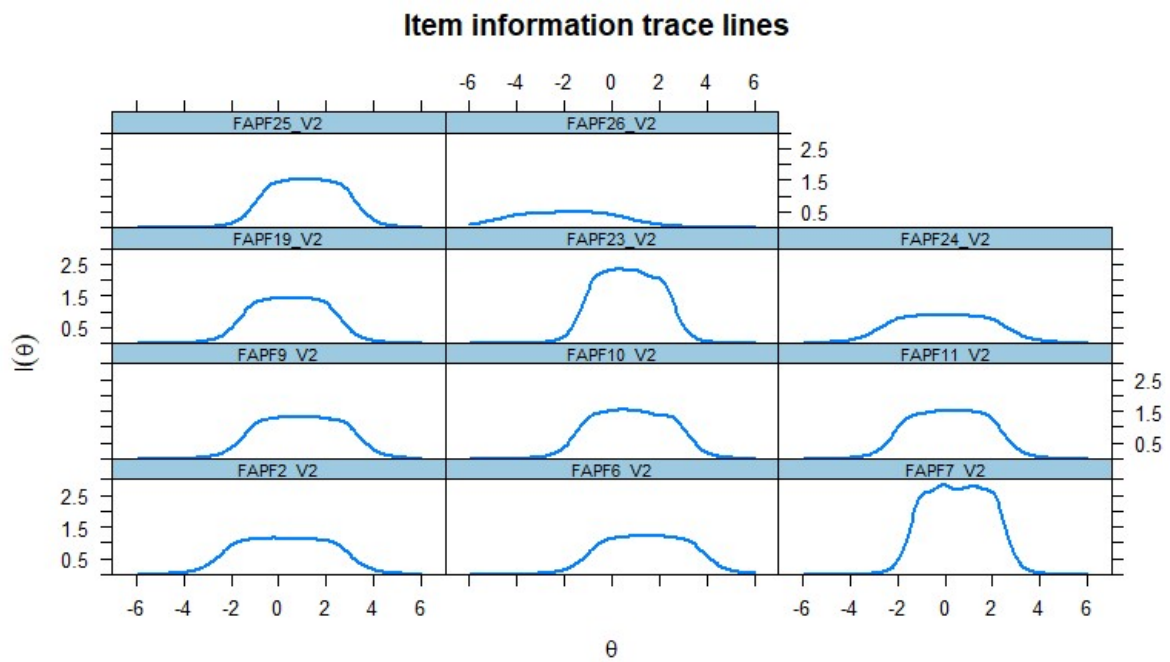


Figure S6. Item information curves (Emotional Impact; Stage 2)