Screening Avoidant/Restrictive Food Intake Disorder (ARFID) in Children: Outcomes from utilitarian versus specialist psychometrics.

Terence M. Dovey[1]*., Victoria K. Aldridge[2]., Clarissa I. Martin[3]., Markus Wilken[4]., & Caroline Meyer[5,6].

[1]Institute of Environment, Health & Societies, Department of Life Sciences, Marie Jahoda Building, Brunel University London, Uxbridge, UB8 3PH, United Kingdom.

[2]UCL Institute of Child Health, 30 Guilford Street, London, WC1N 1EH, United Kingdom.

[3]Midlands Psychology, Midland House, 19 Newport Road, Stafford, Staffordshire, ST16 1BA. United Kingdom.

[4]Institute for Pediatric Feeding Tube Management and Weaning, Siegburg, Germany.

[5] WMG, University of Warwick, Coventry, Warwickshire, CV4 7AL, United Kingdom.

[6] Coventry and Warwickshire Partnership NHS Trust, Coventry, UK.

*Author for correspondence email terence.dovey@brunel.ac.uk.

1

ABSTRACT

This study assessed the specificity and sensitivity of two commonly used psychometric methods to assess ARFID in children. To achieve this, a sample of 329 mothers and one father completed the Behavioral Pediatrics Feeding Assessment Scale (BPFAS) and the Child Food Neophobia Scale (CFNS). A Receiver Operating Characteristic (ROC) analysis indicated that both measures were able to successfully differentiate a known clinical sample from those of typically developing population. Although the BPFAS was more accurate at differentiating ARFID from the general population, the CFNS was acceptable and on some metrics better than its longer counterpart. The ability of a food neophobia scale to differentiate clinical and population samples, and detect gradation of food avoidance within the population sample, suggests that the multitude of psychometric measures available may be measuring similar constructs. Therefore, confidence can be expected in cross-site comparisons despite each using different psychometric measures of food avoidance in children.

KEY WORDS: Childhood; Clinical; Avoidant/Restrictive Food Intake Disorder

WORDS: 4636

INTRODUCTION

Specialist paediatric feeding clinics in different geographical locales use different psychometrics to screen for Avoidant/Restrictive Food Intake Disorder (ARFID - previously termed *feeding disorders*). Of the various psychometric scales that have been used to assess ARFID, only the Behavioral Pediatrics Feeding Assessment Scale (BPFAS; Crist et al, 1994) has shown consistent reliability and validity metrics, the ability to differentiating samples of clinical relevance from the general population (Crist & Napier-Philips., 2001; Dovey et al., 2013) and sensitivity to changes following intervention (Dovey & Martin, 2012a; Dovey & Martin, 2012b). Despite its favourable characteristics, the BPFAS is currently the longest, in terms of items to analyse, psychometric questionnaire available to clinicians who wish to screen for ARFID. Furthermore, the questionnaire is required to be interpreted by the clinician, as multiple analytic processes have been offered (e.g. Crist & Naiper-Phillips, 2001). Recent investigations have found that only the core variables of the BPFAS show changes following successful intervention (Dovey & Martin, 2012a), and are subsequently able to discriminate between ARFID and typically developing children (Dovey et al., 2013), which has provided an evidence-base for a more simplified analytic process.

A common method for assessing feeding problems in community-based samples and services in the United Kingdom has been the Child Food Neophobia Scale (CFNS; e.g., Dovey et al., 2011). The CFNS (Pliner et al., 1994) was designed to measure levels of food neophobia in children. Typically, developmental food neophobia is a developmental stage that all children progress through at varying

rates, usually begins around eighteen months and progressively declines throughout childhood. Eventually, food neophobia settles to become a trait that represents an individual's underlying openness to novel foods (Aldridge et al., 2009). The presentation of developmentally appropriate food neophobia can vary. Some children refuse to eat both novel and previously accepted foods during this stage (e.g., Nicklaus, 2009), but in most cases the developmental stage is characterised by a reluctance to try new foods. Typically, this reluctance results in a lower dietary variety in children during this phase. In most cases, children in the food neophobic phase go on to accept the foods that they become familiar with through repeat exposure (Wardle et al., 2003; Williams et al., 2008). In contrast, food avoidance in children with ARFID is generally stable and more consistently includes refusal of both novel and familiar foods (e.g., Dahl & Sundelin, 1992; Schmid et al., 2010; Wolke et al., 2009). ARFID results in a diet that is low in calories and/or does not contain sufficient micronutrients to maintain growth or health. The problem any clinician has in identifying children with ARFID is that they are screening for the eating disorder in a population who are currently transitioning through a developmental stage characterised by food neophobia.

Research on food neophobia in children has indicated that there are three known underlying components (Dovey et al., 2008). These are: a lack of dietary variety (Cooke et al., 2006; Falciglia et al., 2000); sensory sensitivity or defensiveness specific to food (Coulthard & Blissett, 2009; Farrow & Coulthard, 2012; Nederkoorn et al., 2015; Smith et al., 2005); and problematic behaviour during mealtimes (e.g. spitting out food, hand batting food away, packing) (Carruth & Skinner, 2000; Lewinshon et al., 2005). A lack of dietary variety (Bryant-Waugh & Piepenstock,

4

2008), sensory defensiveness (Dovey & Martin., 2012b) and food avoidance behaviour (Piazza et al., 2003) have also been attributed to ARFID in children (Bryant-Waugh et al., 2010; Dovey et al., 2009; Field et al., 2003; Kreipe & Palomaki, 2012) and appear as important characteristics under the development and course subsection of the Diagnostic and Statistical Manual version 5 (DSM-V) for ARFID. These similarities suggest that a measure for either ARFID or food neophobia may be able to adequately screen for clinically-relevant food avoidance in the general population.

Recent methodological investigations into the psychometric properties of children's eating questionnaires have highlighted several key criticisms that would be important for paediatric clinics to consider. Particularly relevant criticisms that have been levied include different item number, slightly different purposes to what is being measured (i.e., proneness to obesity, expected dietary variety, or food avoidance), and the inability of current measures to distinguish between food neophobia and developmentally atypical food avoidance (de Lauson-Guillain et al., 2012). Effectively, the available measures are potentially anchored to similar observable behaviours that result in refusal to eat. Despite this insight, few studies have attempted to systematically and statistically explore the similarities and differences between different measures purportedly used for similar purposes within children's eating.

Based on the available data, two possible assertions concerning screening for ARFID could be held. The first would be that different measures offer different findings. This first assertion would lead to a conclusion that a more comprehensive

assessment would be necessary and that the number of items a psychometric contains is of paramount importance. Alternatively, different measures may have similar behavioural anchors and thus offer similar findings irrespective of the number of questions asked. This would place the emphasis on the quality of the items. The aim of the current study was to replicate the ability of the comprehensive BPFAS measure to differentiate a cohort of children with a diagnosis of ARFID from a large community sample. The secondary aim was to extend knowledge by comparing the ability of the brief CFNS measure of food neophobia (de Lauson-Guillain et al., 2012) to also screen for ARFID. The objective of the current study was to assess the relative sensitivity and specificity of different psychometrics measures of food avoidance in children to identify those with ARFID. The measures for food neophobia and food avoidance have varying item numbers from a minimum of 6 for the CFNS to a maximum of 70 for the BPFAS (including responses to both frequency and problem subscales). With these two questionnaires sitting on opposite ends of the item number continuum, comparisons between these two measures will provide novel insight into the need for comprehensive psychometric testing for food avoidance in children. Specifically, it will provide insight about the utility of item number in a psychometric and its relative properties concerning screening for ARFID. Moreover, as ARFID in children is primarily assessed in a target population transitioning through developmentally appropriate food neophobia, clinicians would have increased confidence from the screening tools if they are able to discriminate between high levels of food neophobia and ARFID. A discrimination analysis between the BPFAS and the CFNS would provide further insights concerning the validity of an ARFID screening tool. Due to limited data on the comparisons between measures on sensitivity and specificity in children, no directional hypotheses could

be held. It was predicted that there would be a difference between the longest and shortest psychometric measure on its specificity and sensitivity to differentiate a known clinical sample of children with ARFID from a sample of typically developing children.

METHODS

Participants

Three hundred and one mothers of children (141 males; 154 females; 6 not reported) were recruited face-to-face from a variety of locations including schools, nurseries and play centres across the UK, and were directed to complete the study questionnaires online to create a normative sample of scores. The ages of the children ranged from 24 to 84 months (mean = 42.4±15.8 months). The inclusion criteria for the lower age range was determined by the validity/reliability limitations of the questionnaires, which are only suitable for parents of children aged two years and above (Crist & Napier Phillips, 2001). A further sample of 28 parents (27 mothers and 1 father) of children (21 male; 7 female) with avoidant and restrictive food intake disorder (Mean 55.6±23.4 months; Range 25-96 months) were recruited from a peripatetic paediatric ARFID psychology service in the UK and embedded into the large normative sample. The inclusion of the clinical group allowed certainty in the analytic process. This group allowed statistical comparisons to identify a cut-off score for the psychometric measures. This was achieved by making a known clinical sample 'pop-out' of an opportunistic sample recruited from the general population. To overcome the limitations of people completing the measures in different formats

(Dovey et al., 2013), participants in both the population sample and the clinical sample completed the forms using an online programme in their own home. The diagnostic process used in the service has been explained in detail elsewhere (see Dovey et al., 2013) and in the current study the DSM-V was used to diagnose ARFID. All of the children in the clinical sample met the criteria for ARFID through dependence on oral liquid nutritional supplements to achieve sufficient caloric intakes. In all cases, children were referred to the national specialist feeding service, through their local clinical commissioning groups, following unsuccessful attempts to improve their food intake within local hospital services.

The parents of children with ARFID were slightly older (35±5.4 years old), on average, than the normative sample (32.5±5.2 years old). The study contained several exclusion criteria for each of the samples. Parents who did not speak English as their first language, and parents suffering from any known current mental health problem were excluded from the current samples. Parents of in the community sample who had ever sought professional help for their child's behaviour were also excluded (n=9). All of the participants freely volunteered to take part in the research and did not receive any form of remuneration for their participation. Ethical clearance was given from two separate sources depending on the sample obtained. Clearance for the community sample was obtained from an independent university ethics committee and the clinical sample was cleared by an independent NHS ethics committee.

Measures

Behavioral Pediatric Feeding Assessment Scale

The BPFAS (Crist et al., 1994) is a 35 item standardised and validated psychometric self-report questionnaire designed to assess mealtime behaviour associated with poor nutritional intake in children (9 to 84 months old). Respondents, usually the primary caregiver, are required to provide two answers for each question. The first answer concerns the frequency that the problematic feeding behaviour occurs. These answers are recorded on a 5 point Likert scale anchored with 1 never to 5 always. For the second answer, parents must indicate, on a binary yes/no scale, whether or not the issue is a problem for them. The 35 items are separated into two sections. The first 25 questions refer explicitly to the child's behaviour (e.g., takes longer than 20 min to finish a meal; enjoys eating; has problems chewing foods). This subscale contains a total of 50 answers - 25 child problematic feeding behaviour frequency answers (often referred to as child frequency score) and 25 answers concerning if the parent finds the behaviour problematic (often referred to as child problem score). The last 10 items concern how the parent feels about their child's behaviour or how they respond to the child's behaviour during mealtimes (e.g., I get frustrated and/or anxious when feeding my child; I feel confident my child gets enough to eat). These items provide a total of 20 answers for analysis. Ten items relate to how parents are likely to respond (often referred to as parent frequency score) and another 10 related to whether the parents perceive their responses as problematic (often referred to as parent problem score). The questionnaire has four stable domain scores referred to as child frequency score, parent frequency score (each derived from the sum of Likert responses), child problem score, and parent problem score (each derived from the sum of yes/no responses). It has previously been used in both clinical and non-clinical samples with

the primary aim to differentiate those children and families that have clinically significant feeding problems from those that do not (Crist & Napier-Philips, 2001). Previous research has suggested that scores on the BPFAS above 81 for the total frequency score (61 for the child and 20 for the parent subscales) and 8 for the total problem score (6 for the child and 2 for the parent subscales) would differentiate a clinical sample from a community sample (Dovey et al., 2013).

Child Food Neophobia Scale

The CFNS (Pliner et al., 1994) is a six item questionnaire used to assess children's responses to novel foods (e.g. My child doesn't trust new foods). Parents completed the questionnaire by answering a seven point Likert scale anchored with 1 strongly disagree to 7 strongly agree. A total score was calculated after two of the reversed items had been corrected. The reliability and validity of the CFNS is good in normative samples (Cooke et al., 2004; Drewenowski, 1997; Koivisto & Sjoden, 1996; Russell & Worsley, 2008) and it has been used extensively within the child feeding literature. The term and the measure have been listed as a characteristic of ARFID under the sensory aspects of food avoidance (Bryant-Waugh et al., 2010; Chatoor & Ganiban, 2003). To date, no clinical cut off scores have been offered for this psychometric measure and the current paper offered an analysis of the measures ability to assess clinical samples using this measure.

Procedure

Potential participants were approached in a variety of locations in order enquire if they would take part in the research. The locations included, soft play centres, cafes, schools, nurseries, and a variety of workplaces. The questionnaire and intention of

the study was explained to potential participants and a link was provided for them to complete the questionnaires online. Participants were asked that only one parent complete the questionnaires and that they not spend too long considering each answer. It was also requested that the most suitable parent to complete the questionnaires would be the primary care-giver who feeds their child the majority of their meals. In the current sample collected, this recruitment strategy resulted in all but one participant being the mother of the child.

Statistical Analysis.

To assess the sensitivity and specificity of the CFNS and BPFAS at discriminating children with ARFID from the general population, a Receiver Operating Characteristic (ROC) analysis was employed (Metz, 1978). The purpose of this analysis was two-fold. Firstly, it provided a discriminative score allowing the observer to differentiate between individuals with clinically relevant symptoms from those that do not. Secondly, to aid the clinician in estimating the likelihood of a correct assignment of a child with a diagnosis of ARFID, the ROC analysis provides scores for sensitivity (the amount of the known clinical sample that was above a specific cut off score on a questionnaire) and specificity (the number of the community sample that were below the cut off criteria). The critical score for this analysis has been generally accepted to be around 80% sensitivity and specificity. However, any score above 70% has been deemed acceptable (Mond et al., 2008). The ability to discriminate between clinical and typically developing scores, within the context of current analysis, is referred to as the accuracy of the model.

In addition to the ROC analysis, an area under the curve (AUC) and positive predictive value (PPV) analysis were also included. This method of analysis is considered the most appropriate method for assessing the ability for a psychometric measure to discriminate between clinical and non-clinical scores (Metz, 1978) and has been used in both eating disorder (Cotton, Ball, & Robinson, 2003; Mond et al., 2008; Parker, Lyons, & Bonner, 2005) and ARFID research (Dovey et al., 2013) in the past. To assess cut-off points, a simple cross-tabulation between sensitivity and specificity using the 80% criterion was used. Differences between the two groups were assessed using non-parametric t-tests due to the large differences in sample sizes between the population and the clinical sample. To assess the relationship between the two measures, correlations were reported. All analytics were subject to controls for multiple comparisons. All analyses were undertaken in IBM SPSS version 20 for Windows.

RESULTS

Characteristics of the sample

As expected, the descriptive statistics revealed large significant differences between the clinical and non-clinical groups across all subscales of the BPFAS (all $p<0.001$) and the CFNS ($U_{(329)}=4.62$; $p<0.001$). Table 1 offers a breakdown of the descriptive statistics.

Associations between CFNS and BPFAS

Additional analysis of the whole sample indicated that the relationship between the CFNS and the BPFAS was strongest for the child frequency score of the BPFAS ($r_{(329)}=0.62$; $p<0.001$). Other subscales of the BPFAS demonstrated weak to

moderate, yet significant, correlations with the CFNS (Parent frequency: $r_{(329)}=0.50$; $p<0.001$. Child problem: $\rho_{(329)}=0.43$; $p<0.001$. Parent problem: $\rho_{(329)}=0.24$; $p<0.001$).

When the sample was separated by group and then reassessed, the relationships between the BPFAS and the CFNS altered. The child frequency scores maintained a significant moderate correlation with the CFNS in the community ($r_{(301)}=0.57$; $p<0.001$) and clinical ($r_{(28)}=0.58$; $p=0.001$) samples. Among the other subscales of the BPFAS, the parent frequency ($r_{(329)}=0.47$; $p<0.001$) and child problem ($\rho_{(329)}=0.38$; $p<0.001$) scores in the community sample both correlated with the CFNS, while the parent problem score did not. In contrast, no subscales aside from the child frequency score correlated with the CFNS in the clinical sample (all r and ρ were below 0.2).

Optimising the BPFAS and CFNS to predict clinical cut off scores.

Simple contingency tables were produced based on the optimum cut-off scores to calculate sensitivity, specificity and PPV. Table 2 offers a breakdown of potential diagnostic cut-off scores for the BPFAS subscales and the CFNS scale. The ROC analysis indicated that scores above 59 on the child frequency of the BPFAS would delineate between clinical and community groups (Se=0.64, Sp=0.89, PPV=0.32, AUC=0.85, $p<0.001$). For the parent frequency subscale, a score of 22 (Se=0.68, Sp=0.82, PPV=0.25, AUC=0.82, $p<0.001$) was found to be adequate for separating the two groups. The final two problem subscales suggested that cut of scores of 6 (Child: Se=0.96, Sp=0.91, PPV=0.60, AUC=0.97, $p<0.001$) and 3 (Parent: Se=0.39,

Sp=0.97, PPV=0.68, AUC=0.95, p<0.001) would provide the best scores for group differentiation.

The CFNS was also able to separate the two groups well. Scores above 25 removed over 80% of the community sample. However, 29 was found to be the optimal cut-off score for the CFNS, removing all but 7% of the community sample whilst retaining 68% of the clinical sample (Se=0.68, Sp=0.93, PPV=0.40, AUC=0.77, p<0.001).

Outcome of cut-off values and the impact on the population.

Analysis of the true and false positives within the assessment revealed variation in the applicability of the BPFAS and the CFNS. The best measure for discriminating groups was the child problem subscale of the BPFAS. This retained 96.4% of the clinical sample (27/28) and removed 94% of the community population (18/301 cases retained). The parent problem score was highly discriminative, removing all but 8 (2.7%) members of the community sample; however, it also removed most of the clinical sample leaving only 11 (39.3%).

With the BPFAS child frequency subscale, the score of 59 included 34 children from the population (11.3%) and retained 18 clinically relevant children (64.3%). The score of 21 on the parent frequency scale left 53 (17.6%) members of the general population and 19 (67.9%) of the clinical sample. This frequency measure returned the worst metrics for discriminating between the two groups overall. The CFNS set at a score of 29 removed all but 22 members of the population (7.3%) and retained 19 of the clinical sample (67.9%). This placed the metrics for the optimal cut-off score

on the CFNS between those of the child frequency score and the child problem score of the BPFAS.

DISCUSSION

The ROC analysis indicated a favourable outcome for the BPFAS. High AUC scores were reported for three of the subscales with high scores for sensitivity and specificity, suggesting that the psychometric was accurate. The outcomes obtained in the current analyses mirror those of the original study on the discriminative properties of the BPFAS within a different sample (Dovey et al., 2013) with some minor variation. For the child frequency subscale, the optimal cut-off indentified in the current study was two points below the expected cut-off of 61. The other three subscales all indicated that a one to two point increase was required to reach optimal separation. These results suggest that the discriminative properties of the BPFAS were relatively stable in a UK population for three of the subscales. It was found that the sensitivity scores for the parent problem subscale of the BPFAS were too low to provide confidence in its discriminatory capabilities and could potentially be dropped from future assessments if consistent corroborating data is found.

The CFNS was also able to differentiate the clinical group from the community sample. Although it was not as good as the BPFAS child problem subscale, the CFNS did achieve acceptable results and was indeed better than the BPFAS on some metrics. In particular, the CFNS achieved a similar sensitivity and specificity result to the child frequency scores of the BPFAS, and like the BPFAS, was more specific than it was sensitive at the optimal cut off point. The only observable difference between the BPFAS and the CFNS was the area under the curve

analysis, which suggested that the CFNS was not as accurate as the BPFAS on measuring food avoidance. This data would therefore suggest that the number of items is of less importance compared to item quality.

The differential correlations between the clinical and community samples in this data offer an interesting conundrum for clinicians screening for ARFID. The ultimate arbitrator in psychometric screening process, based on this data, was the BPFAS child problem subscale. However, simply asking questions around problematic feeding has, historically, been problematic itself, as many parents report their child's eating is a problem (Aldridge et al., 2010). It is for this reason that caution should be applied when using the child problem score in isolation. The child frequency and neophobia scales did differentiate clinical and community samples too. The correlations suggest that the frequency of food avoidance, as measured by the BPFAS and neophobia scores, are related. Moreover, the AUC differences favouring the BPFAS is likely due to its inclusion of the child problem scale; although it cannot be discounted that specific items within the BPFAS child frequency scores do not have subtle differences that are responsible for the findings.

The first potential explanation based on these correlations would be asking an equivalent question of "is feeding your child problematic" is not an accurate discriminator. Rather, it could be the subtle differences obtained through recording multiple questions concerning problematic feeding that is used in the BPFAS. An alternatively explanation would be, the analytic procedure used within the current psychometric doctrine of creating scales and subscales may artificially separate the frequency of the problematic behaviour from the perception of it as a problem.

Equally, this process of creating scales does not consider the relative importance of items in its prediction of behaviour. The repercussions of the second explanation would suggest that the traditional analytic procedures only reveal a component of the food avoidant behaviour and are unable to assess item quality. With development of new analytic processes, especially those on big data analytics, that are not reliant on combining items into factors/scales, may be a suitable future direction to measure item quality.

The similarities in the analysis between the BPFAS and the CFNS in differentiating typically developing and clinical samples was likely due to one of two potential explanations. The first potential interpretation was that the frequency of the refusing food exists to the same level in the community, but some parents report that it is not a problem. The reason they do not report it as a problem could be because they share similar traits and dietary variety as their child (Galloway et al., 2003), have a higher self-efficacy in their parenting skills (Campbell et al., 2010), or perceive the frequency of the problematic mealtime behaviour as transient (Dahl & Sundelin, 1992; Dahl et al., 1994) or some other unidentified factor relating to parental engagement or responsibility. Further research into this perception is merited to form definitive conclusions concerning predictors of the child's behaviour and/or parental perceptions of food avoidance as a problem. The second interpretation could be that the measured frequency of problematic behaviour within the BPFAS does not capture the magnitude of the observable behaviours. Perhaps the difference between groups is that the high scorers in the clinical sample resist mealtimes with more veracity than high scorers in the non-clinical sample (Sanders et al., 1993).

More behavioural observation studies would be necessary to corroborate this interpretation.

The current psychometrics reported for the two measures have, at least on a functional level, homogeneity. Homogeneity of items measured in children's eating/feeding questionnaires has been previously reported (de Lauzon-Guillain et al., 2012; Dovey et al., 2011), but little data has been offered concerning comparative homology on their function and accuracy. This study was one of the first to do this. Data in the current study suggests that there are some shared characteristics between the measures of typically developing food neophobia and ARFID in children. The interpretation was that children with both normative levels of food avoidance and with ARFID can be successfully assessed and differentiated with a brief and specific measure of child food neophobia. This may suggest that specific questions, rather than whole questionnaires, better screen for ARFID. A form of item analysis would be necessary to uncover which particular questions are important to screen for ARFID. Irrespective of the outcome of this future analysis, it is clear that only a few questions are needed to adequately screen for ARFID in children.

It is important to acknowledge some limitations of the current design. This study contained a large sample of children in order to provide suitable metrics and means for the ROC analysis; however, it must be acknowledged that these children all came from one country. Although the scores on the psychometric measures are similar to those reported in other countries (e.g., Crist & Napier-Phillips, 2001), cross-cultural and country data collection is still needed. Moreover, the children with ARFID included here were those that were resistant to generic interventions and were in a

specialist service. There is likely to be a group of children that are responsive to more generic interventions and support within the community. These children would be seen by allied health professionals in the community and would be responsive to their interventions. This distinct group were missing from the current study and would be a viable future research direction in this field. All but one of the respondents in this study were the mothers of the child. Although the respondents reflect the overwhelming majority of familial division of labour, it does not reflect all children. Exploring the outcomes and metrics of alternative division of labour within the family situation, such as a grandparent or father as the primary-caregiver may be of interest. There is currently no data to suggest there would be any differences between who is the primary caregiver and the child's propensity to engage in food avoidance.

To conclude, the current study was the first to demonstrate similarity between the 6-item CFNS questionnaire for assessing child food neophobia and the longer BPFAS questionnaire in their ability to differentiate a known clinical sample with ARFID from a normative population. This has two clear conclusions. First, the similarities between the two questionnaires allowed both to differentiate a clinical sample, and secondly, that a the specific screening tool was very accurate, specifically the child-related problem score, and could differentiate ARFID cases from extreme CFNS scorers. Differences between the most and least comprehensive measures for food avoidance appear to be relatively small. Therefore, confidence can be expected in cross site comparisons despite each using different psychometric measures of food avoidance in children, as well as in the BPFAS as a screening tool.

REFERENCES:

Aldridge, V., Dovey, T. M. & Halford, J. C. G. (2009). The role of familiarity in dietary development. *Developmental Reviews*, *29*, 32-44.

Aldridge, V., Dovey, T. M., Martin, C. I., & Meyer, C. (2010). Identifying clinically relevant feeding problems and disorders. *Journal of Child Health Care, 14,* 261-270.

Bryant-Waugh, R. J., & Piepenstock, E. H. C. (2008). Childhood disorders: Feeding and related disorders of infancy or early childhood. In: Tasman A, Kay J, Lieberman JA, First MB, Maj M, editors. *Psychiatry*. 3rd ed. New York: Wiley. p. 830-46.

Bryant-Waugh, R., Markham, L., Kreipe, R. E., & Walsh, T. (2010). Feeding and eating disorders in childhood. *International Journal of Eating Disorders, 43*, 98-111.

Campbell, K., Hesketh, K., Silverii, A., & Abbott, G. (2010). Maternal self-efficacy regarding children's eating and sedentary behaviours in the early years: Associations with children's food intake and sedentary behaviours. *International Journal of Pediatric Obesity, 5*, 501-508.

Carruth, B. R., & Skinner, J. D. (2000). Revisiting the Picky Eater Phenomenon: Neophobic Behaviors of Young Children. *Journal of the American College of Nutrition, 19,* 771-780.

Chatoor, L., Ganiban, J., Harrison, J., & Hirsch, R. (2001). Observation of feeding in the diagnosis of posttraumatic feeding disorder of infancy. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 595-602.

Cooke, L., Carnell, S., & Wardle, J. (2006). Food neophobia and mealtime food consumption in 4-5 year old children. *International Journal of Behavioral Nutrition and Physical Activity, 3*, 14-20.

Cooke, L. J., Wardle, J., Gibson, E. L., Sapochnik, M., Sheiham, A., & Lawson, M. (2004). Demographic, familial and trait predictors of fruit and vegetable consumption by preschool children. *Public Health Nutrition, 7,* 295–302

Cotton, M. A., Ball, C., & Robinson, P. (2003). Four simple questions can help screen for eating disorders. *Journal of General and Internal Medicine, 18*, 53–56.

Coulthard, H., & Blissett, J. (2009). Fruit and vegetable consumption in children and their mothers. Moderating effects of child sensory sensitivity. *Appetite, 52*, 410-415.

Crist W, McDonnell P, Beck M, Gillespie C, Barrett P, & Mathews J. (1994). Behaviour at mealtimes and the young child with cystic fibrosis. *Journal of Developmental and Behavioral Pediatrics, 15*, 157-161.

Crist, W., & Napier-Philips, A. (2001). Mealtime behaviours of young children. A comparison of normative and clinical data. *Journal of Developmental and Behavioral Pediatrics, 22*, 279–286.

Carruth, B. R., & Skinner, J. D. (2000). Revisiting the 'picky/fussy' eater phenomenon: Neophobic behaviours of young children. *Journal of the American College of Nutrition, 19*, 771–780.

Dahl, M., & Sundelin, C. (1992). Feeding problems in an affluent society. Follow-up at four years of age in children with early refusal to eat. *Acta Paediatrica*, *81*, 575-579.

Dahl, M., Rydell, A. M., & Sundelin, C. (1994). Children with early refusal to eat: follow-up during primary school. *Acta Paediatrica, 83,* 54-58

de Lauzon-Guillain, B., Oliveira, A., Charles, M. A., Evangelia, G., Jones, L., Rigal., N. et al., (2012). A review of methods to assess parental feeding practices and preschool children's eating behaviour: the need for further development of tools. *Journal of the Academy of Nutrition and Dietetics, 10,* 1578-1602.

Dovey, T. M., & Martin, C. I. (2012). A quantitative psychometric evaluation of an intervention for poor dietary variety in children with a feeding problem of clinical significance. *Infant Mental Health Journal, 33(2),* 148-162.

Dovey, T. M., Jordan, C., Aldridge, V., & Martin, C. (2013). Screening for feeding disorders: Creating critical values using the behavioral paediatrics feeding assessment scale. *Appetite, 69,* 108-113.

Dovey, T. M., Martin, C. I., Aldridge, V. K., Haycraft, E. & Meyer, C. (2011). Measures, measures everywhere, but which ones should I use? *The Feeding News,6(1),* 1-13.

Dovey, T. M. & Martin, C. I. (2012). A parent-led contingent reward desensitisation invention for children with a feeding problem resulting from sensory defensiveness. *Infant, Child & Adolescent Nutrition, 4(6),* 384-393.

Dovey, T. M., Aldridge, V. K., Dignan, W., Staples, P. A., Gibson, E. L. & Halford, J. C. G. (2012). Developmental differences in sensory decision-making involved in deciding to try a novel fruit. *British Journal of Health Psychology*, *17*, 258-272.

Drewenowski, A. (1997). Taste preferences and food intake. *Annual Reviews in Nutrition, 17*, 237-253.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A 'transdiagnostic' theory and treatment. *Behaviour Research and Therapy, 41,* 509–528.

Farrow, C. V., & Blissett, J. (2006). Maternal cognitions, psychopathologic symptoms, and infant temperament as predictors of early infant feeding problems: A longitudinal study. *International Journal of Eating Disorders, 39,* 128-134.

Farrow, C. V., & Coulthard, H. (2012). Relationships between sensory sensitivity, anxiety and selective eating in children. *Appetite, 58,* 842-846.

Field, D., Garland, M., & Williams, K. (2003). Correlates of specific childhood feeding problems. *Journal of Paediatric and Child Health, 39,* 299-304.

Galloway, A. T., Lee, Y., & Birch, L. L. (2003). Predictors and consequences of food neophobia and pickiness in children. *Journal of the American Dietetic Association, 103*, 692–698.

Koivisto, U-K., Sjödén, P-O. (1996). Food and general neophobia in Swedish families: parent-child comparisons and relationships with serving specific foods *Appetite, 26*, 107–118.

Lewinsohn, P. M., Holm-Denoma, J. M., Joiner, T. E., Striegel-Moore, R., Bear, P., & Lamoureux, B. (2005). Problematic eating and feeding behaviours of 36-month-old children. *International Journal of Eating Disorders, 38,* 208-219.

Loewen, R., & Pliner, P. (1999). Effects of prior exposure to palatable and unpalatable novel foods on children's willingness to taste other novel foods. *Appetite, 32,* 351–366.

Martin, C. I., Dovey, T. M., Coulthard, H., & Southall, A. (2013). Maternal stress and problem-solving skills in a sample of children with non-organic feeding disorders. *Infant Mental Health Journal, 34,* 202-210

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, 8*, 283–297.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Temporal stability of the eating disorder examination questionnaire. *The International Journal of Eating Disorders, 36*, 195–203.

Mond, J. M., Myers, T. C., Crosby, R. D., Hay, P. J., Rodgers, B., Morgan, J. F., Lacey, H., & Mitchell, J. E. (2008). Screening for eating disorders in primary care. EDE-Q versus SCOFF. *Behaviour Research and Therapy, 46*, 612–622.

Nederkoorn, C., Jansen, A., & Havermans, R. C. (2015). Feel your food. The influence of tactile sensitivity on picky eating in children. *Appetite, 84*, 7-10.

Nicklaus, S. (2009). Development of food variety in children. *Appetite, 52*, 253-255.

Piazza, C. C., Wayne Fisher, W. W., Brown, K. A., Shore, B. A., Patel, M. R., Katz, R. M., Sevin, B. M., Gulotta, C. S., & Blakely-Smith, A. (2003). Functional analysis of inappropriate mealtime behaviors. *Journal of Applied Behavior Analysis, 36*, 187-204.

Parker, S. C., Lyons, J., & Bonner, J. (2005). Eating disorders in graduate students. Exploring the SCOFF questionnaire as a simple screening tool. *Journal of American College Health, 54,* 103–107.

Pliner, P. (1994). Development of measures of food neophobia in children. *Appetite, 23,* 147–163.

Robinson, J. R., Drotar, D., & Bounty, M. (2001). Problem-solving abilities among mothers of infants with failure to thrive. *Journal of Pediatric Psychology, 26*, 21-32.

Russell, C. G., & Worsley, A. (2008). A population-based study of preschoolers' food neophobia and its associations with food preferences. *Journal of Nutrition Education and Behavior, 40*, 11-19.

Sanders, M. R., Patel, R. L., le Grice, B., & Shepard, R. W. (1993). Children with persistent feeding difficulties: an observational analysis of the feeding interactions of problem and non-problem eaters. *Health Psychology, 12*, 64-73.

Schmid, G., Schreier, A., Meyer, R., & Wolke, D. (2010). A prosepctive study on the persistence of infant crying, eing and feeding problems and preschool behaviour. *Acta Paediatrica, 99*, 286-290.

Shisslak., C. M., Crago, M., & Estes, L. A. (1995). The spectrum of eating disturbances. *International Journal of Eating Disorders, 18*, 209-219.

Smith, A. M., Roux, S., Naidoo, N. T. R., & Venter, D. J. L. (2005). Food choices of tactile defensive children. *Nutrition, 21*, 14–19.

Waller, G. (2008). A 'trans-transdiagnostic' model of the eating disorders: A new way to open an egg. *European Eating Disorder Review, 16*, 165-172.

Wardle, J., Herrera, M. L., Cooke, L. (2003). Modifying children's food preferences: the effect of exposure and reward on accpetance of an unfamiliar vegetable. *European Journal of Clinical Nutrition, 57*, 341-348.

Werle, M. A., & Murphy, T. B. (1998). Broadening the parameters of investigation in treating young children chronic food refusal. *Behavior Therapy, 29,* 87-104.

Wilken, M., Cremer, V., Berry, J., & Partmann, P. Rapid home-based weaning of small children with feeding tube dependency: positive effects on feeding behaviour without deceleration of growth. *Archives of Disease in Childhood, 98*, 856-861.

Williams, K. E., Paul, C., Pizzo, B., Riegel, K. (2008). Practice does make perfect. A longitudinal look at repeat taste exposure. *Appetite, 51*, 739-742.

Wolke, D., Schmid, G., Schreier, A., & Meyer, R. (2009). Crying and feeding problems in infancy and cognitive outcome in preschool children born at risk: a prospective population study. *Journal of Developmental & Behavioral Pediatrics, 30*, 226-238.

Table 1 – Descriptive statistics for the clinical and population groups included in the ROC analysis.

Scale	Population (n=301) Means±Standard Deviation	Clinical (n=28) Means±Standard Deviation
BPFAS Total Score	62.19±16.34	91.18±21.66***
BPFAS Child Frequency	45.26±12.10	66.39±15.97***
BPFAS Parent Frequency	16.92±5.12	24.79±7.19***
BPFAS Total Problem Score	1.90±3.37	13.43±3.82***
BPFAS Child Problem Score	1.45±2.55	10.11±3.26***
BPFAS Parent Problem Score	0.45±1.09	3.32±1.52***
Child Food Neophobia Score	19.08±7.40	27.68±9.41***

*** = $p < 0.001$ on Mann-Whitney U-tests of Difference

Table - 2 Scores on the Behavioral Pediatrics Feeding Assessment Scale and the Child Food Neophobia Scale with their respective sensitivity and specificity ratings.

Scale	Questionnaire Score	Sensitivity	True +ve Clinical detected	95% CI	Specificity	95% CI	False +ve Control Detected
BPFAS Total	81	.750	75.0%	(56.6, 87.3%)	.830	(79.1, 87.5%)	16.3%
BPFAS Child Frequency	59	.643	64.3%	(45.8, 79.3%)	.887	(84.6, 91.8%)	11.3%
BPFAS Parent Frequency	22	.679	67.9%	(49.3, 82.1%)	.824	(77.7, 86.3%)	17.6%
BPFAS Total Problem	9	.964	96.4%	(87.9, 100%)	.940	(81.3, 89.2%)	14.3%
BPFAS Child Problem	6	.964	96.4%	(82.3, 99.4%)	.917	(88.0, 94.3%)	8.3%
BPFAS Parent Problem	3	.393	39.3%	(23.6, 57.6%)	.973	(94.8, 98.7%)	2.7%
Child Food Neophobia	29	.679	67.9%	(49.3, 82.1%)	.927	(89.2, 95.1%)	7.3%