Towards identifying a method of screening for autism amongst women with restrictive eating disorders

James Adamson1 | Janina Brede1 | Charli Babb2 | Lucy Serpell1,3 | Catherine R. G. Jones2 | John Fox2 | Will Mandy1

Abstract

Objective: Up to 37% of patients with anorexia nervosa score above cut-off on autism screening measures. These individuals typically have poorer outcomes from standard eating disorder interventions and could therefore benefit from adaptations. Accurately identifying these individuals is important for improving autism referral processes and clinical pathway decisions. This study's aim was to identify subscales of questionnaires measuring constructs associated with either autism or eating disorders that, when combined with traditional autism screening measures, would improve the ability to identify women with restrictive eating disorders who might benefit from a full autism assessment.

Method: One hundred and sixty women with restrictive eating disorders, with (n = 42) or without (n = 118) an autism diagnosis completed a battery of questionnaires. Using conditional stepwise binary logistic regression, we attempted to improve the autism spectrum quotient 10 item's (AQ-10) ability to discriminate between autistic and non-autistic women in a restrictive eating disorder sample.

Results: In a binary logistic regression model, the AQ-10 reliably discriminated between autistic and non-autistic women with an accuracy rate of 85% but had relatively low (69%) sensitivity, reflecting a high rate of false negatives. Adding three subscales to the model (Glasgow Sensory Questionnaire Auditory, Camouflaging Autistic Traits Questionnaire Compensation and Toronto Alexithymia Scale Externally Orientated Thinking) significantly improved its differentiating ability (accuracy = 88%, sensitivity = 76%, specificity = 92%).

Conclusions: We have identified three subscales that, when used in combination with the AQ-10, may help clinicians understand the pattern of autistic traits in their patients with a restrictive eating disorder. This can inform treatment planning and referral decisions.

Abbreviations: AAN, Atypical Anorexia Nervosa; AN, Anorexia Nervosa; ARFID, Avoidant Restrictive Food Intake Disorder; AUC, Area Under the Curve; ED, Eating Disorder; NICE, The National Institute for Health and Care Excellence; RED, Restrictive Eating Disorder; ROC, Receiver Operating Characteristic Curve; VIF, Variation Inflation Factor.
clinical decisions about whether to refer for a full autism assessment and whether to adapt standard eating disorder treatments to accommodate autistic traits. Future studies are needed to test the model in samples where participants have undergone a full autism assessment.

**KEYWORDS**
anorexia nervosa, autism, autism spectrum disorder, eating disorders, screening

**Highlights**
- In a restrictive eating disorder sample, the AQ-10 accurately identified 85% of autistic women, but had a sensitivity of only 69%, indicating that it leads to many false negatives.
- Adding questions about auditory sensitivity, social compensation and externally orientated thinking, in combination with the AQ-10, led to an improved autism screening model (sensitivity = 76%, specificity = 92%).
- The model indicates additional autistic characteristics that when supplemented with the AQ-10 could improve autism screening tools for a restrictive eating disorder population.

# INTRODUCTION

Up to 37% of patients with Anorexia Nervosa (AN) score above cut-off on autism screening measures (Boltri & Sapuppo, 2021; Huke et al., 2013). ‘Autism spectrum disorder’ (hereafter ‘autism’) is a neurodevelopmental condition that is associated with differences in social communication, and the presence of restrictive and repetitive patterns of behaviour (American Psychiatric Association, 2013). In this paper, in line with recommendations from the autism community, we will use the terms ‘autism’ and ‘autistic person’ (Bury et al., 2020; Kenny et al., 2016). AN is an eating disorder (ED) characterised by low body weight, an intense fear of gaining weight and extreme weight and shape concerns (American Psychiatric Association, 2013). According to epidemiological research, AN largely affects women, with estimates up to a 18:1 female to male ratio, whereas autism is a condition that is more common in boys and men with a 3:1 male to female ratio (Loomes et al., 2017; Nicholls et al., 2011; Raevuori et al., 2014). Autistic women are likely to be identified and diagnosed later in life than men, potentially because of a different clinical presentation that is missed by standard assessment tools (Gould & Ashton-Smith, 2011; Sedgewick et al., 2019).

There is an increased recognition that autism research is skewed towards more male-typical presentations, partially due to the diagnostic bias against girls and women. Therefore, more research is needed to improve the recognition of autism in women (Milner et al., 2019).

In patients with AN, higher levels of autistic traits as measured by screening tools are associated with poorer treatment outcomes, more severe presentations and longer stays in in-patient settings (Nielsen et al., 2015; Tchanturia et al., 2019). Furthermore, individuals with high autistic traits respond to standard ED treatment differently than those with low autistic traits (Li et al., 2021; Tchanturia et al., 2019; Westwood & Tchanturia, 2017). For example, patients with AN and high autistic traits show little clinical change after group psychotherapy interventions (Adamson et al., 2018; Baron-Cohen et al., 2013) but show significant improvements if the same intervention is delivered individually (Adamson et al., 2018; Dandil et al., 2020). It is not surprising that co-occurrence affects treatment outcomes, given that interventions are often designed and validated on non-autistic people. Reflecting evidence that autistic individuals may respond to elements of treatment differently, they could potentially benefit from adaptations to standard AN treatment (Babb et al., 2021; Tchanturia, 2021). Considering the different treatment trajectories for autistic individuals with AN, it is clinically important to be able to recognise autistic traits in an accurate and timely manner. Consequently, clinical services could adapt existing treatment to provide tailored interventions that take account of autistic traits, leading to potential improvements in outcomes (Adamson et al., 2020; Babb et al., 2021).

Most autistic women with an ED do not have an autism diagnosis when they first present to ED services (Brede et al., 2020; Kinnaird et al., 2019; Solmi et al., 2021). Currently, ED services struggle to identify those with undiagnosed autism (Babb et al., 2021). Crane and Hill (2016) surveyed 1000 parents in the UK whose children had gone through autism assessments and found the average wait
time was three and a half years. Wait times are similar in adult settings, partly reflecting the resource-intensive nature of autism assessment, which often takes two trained clinicians and a full day (Crane et al., 2016). Considering typical ED treatment lengths range from a few months to more than a year (Tchanturia et al., 2019), it is often not feasible for patients to receive a full autism diagnostic assessment within the timeframes of their ED treatment. Furthermore, clinicians in child and adolescent ED services report low confidence in identifying and referring on children for an autism assessment (Kinnaird et al., 2017). It is also the case that there will be people with an ED and with sub-clinical levels of autistic traits who, although not autistic, would still benefit from their distinct profile of strengths and difficulties being recognised and considered in treatment programmes (Sauré et al., 2021). Clinical services need an efficient and reliable way to identify individuals with high levels of autistic traits at the start of their treatment to make timely clinical decisions around treatment and referral. However, there are currently no established screening methods that have been validated for use in ED services and instead services typically use the autism quotient screening questionnaires such as the AQ-10 (Allison et al., 2012) to indicate autism traits.

A study of 476 adults seen at a national adult autism diagnostic service found that, in this setting, the AQ-10 was a poor predictor of clinical diagnosis, with scores not significantly predicting diagnosis (Ashwood et al., 2016). In this clinical sample the AQ-10, whilst having acceptable sensitivity (77%), demonstrated poor specificity (28%), reflecting high rates of false positives. There is also a concern about the use of the AQ-10 within ED populations due to the gender differences in the characteristics of autism, with the AQ-10 being originally validated in a mainly male sample (Allison et al., 2012). Clinicians commonly use the AQ-10 in ED settings because there are currently no alternatives of proven accuracy for use with ED patients; and because it is the only measure currently recommended by NICE for screening adults for autism (NICE, 2021).

With knock-on effects for the appropriateness of treatment and referrals (Westwood et al., 2017). Other self-report autism screening tools have been developed, for example, the Ritvo Autism Asperger Diagnostic Scale (RAADS) (Ritvo et al., 2011) and used as a screening tool within ED services (Vagni et al., 2016). However, short-ened screening tools such as the RAADS-14 (Eriksson et al., 2013) are developed to have high sensitivity to reduce the chance of false negatives, but this is often at the cost of low specificity in clinical samples, leading to high rates of false positives, that is, scoring above threshold but not actually being autistic. Developing a screening method that has both high sensitivity and specificity for possible autism in ED populations will ensure that appropriate diagnostic referrals can be made and that ED clinicians have reliable and timely insight into the profile of autistic traits in their patients. Ultimately, this can lead to better treatment and outcomes for these patients.

One major issue that screening measures face when used with people with ED is that autism and AN have many overlapping features including cognitive (i.e., rigidity and attention to detail), social and behavioural difficulties, and atypical eating behaviours (Kinnaird et al., 2019; Tchanturia et al., 2013). These shared features make it difficult for standard autism screening measures and diagnostic tools to distinguish between characteristics of autism versus those reflecting AN. Low specificity within some screening measures combined with an overlap between conditions mean that our current estimates of prevalence may be overinflated. The more severe the illness state of AN, such as requiring inpatient treatment, the higher the incidence of autistic traits, as assessed using screening measures (Westwood & Tchanturia, 2017). Due to the limited research using gold-standard diagnostic tools in these settings, it is difficult to ascertain whether these individuals have undiagnosed autism or whether the illness state of AN combined with the limitations of current self-report screening measures, is leading to some false positives.

Researchers and clinicians in eating disorder services have called for a pragmatic autism screening tool that can be used with predominantly women with EDs to inform referrals and clinical decisions (Li et al., 2021; Westwood & Tchanturia, 2017). Such screening tool would need to be able to accurately discriminate between autism and EDs in mainly female populations without being too long or complex to be administered in routine clinical settings. The brevity of the AQ-10 gives it a narrow focus, especially in ED populations where clinical features can be confounding (Kinnaird et al., 2019; Tchanturia et al., 2013). The aim of the study was to generate a screening procedure that enhances the AQ-10’s ability to differentiate autistic and non-autistic individuals in an ED sample. We did this by building a statistical model that can, in addition to the AQ-10, draw on a range of additional self-report questionnaires that tap into diverse characteristics of both autism and EDs. The model can then help to inform future screening measure development by highlighting areas that are more likely to be specific to autism within an ED sample.

2 | METHOD

The following procedures set out in this cross-sectional study followed the recommendations from the
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (Vandenbroucke et al., 2007).

### 2.1 | Participants

Participants were recruited between October 2020 and April 2021 as part of a larger study (Babb et al., 2021; Brede et al., 2020) looking at eating difficulties in autistic women. Since data collection took place during the Covid-19 pandemic, recruitment was conducted predominantly online through existing clinical networks, autism and ED charities, and social media advertising. Participants were reimbursed £15 to complete a set of questionnaires taking approximately 1 h. Participants were initially recruited into three groups: (i) those with a current restrictive eating disorder (RED) diagnosis and an autism diagnosis and no ED diagnosis, (ii) those with autism diagnosis and no RED diagnosis, and (iii) those with both autism and a current RED diagnosis. For this study, to maximise generalisability of findings to ED services, we only included participants with a RED, thereby excluding from the analysis autistic women with no RED diagnosis. A RED was considered a diagnosis of either Anorexia Nervosa (AN), Atypical Anorexia Nervosa (AAN) or Avoidant Restrictive Food Intake Disorder (ARFID).

Individuals were screened for inclusion, given detailed information about the study and gave written informed consent before being asked to complete the battery of questionnaires. Inclusion criteria for the autistic group was a formal diagnosis of autism from a relevant clinician, as well as a current diagnosed RED. The non-autistic group consisted of individuals with a RED who self-disclosed they had never received a diagnosis of autism, were not currently referred for or currently undergoing an assessment for autism. Given the necessity to conduct remote data collection (due to the COVID-19 pandemic) we had to rely on the accuracy of participants self-reporting RED diagnosis, however, where clinical records were available and consent obtained, we cross-referenced with the current responsible clinician. Further inclusion criteria for both groups were the absence of an intellectual disability and being over the age of 18 at time of recruitment. Finally, considering the methodology used in this study, only participants who completed all questionnaires could be included.

### 2.2 | Measures and procedure

Most of the study was conducted online due to the COVID-19 pandemic with only 27% of the participants completing measures in person. Once participants were screened for eligibility and consent was obtained, they were sent a secure link to a questionnaire platform to first provide some demographic and clinical data followed by completing 18 self-report measures. The self-report AQ-50, the full version of the shortened AQ-10 was administered. The AQ-10 can be obtained by selecting only the questions included in the shortened version, as per Ashwood et al. (2016). The Ritvo Autism Asperger Diagnostic Scale-Revised Screen (RAADS-14 Screen, (Eriksson et al., 2013)) another autism screening tool was also included, as well as the Adult Repetitive Behaviours Questionnaire-2 (RBQ-2A, Barrett et al., 2018), which directly assesses restricted and repetitive behaviours. The Eating Disorder Examination Questionnaire (EDE-Q, Mond et al., 2004) and the Swedish Eating Assessment for Autism Spectrum Disorders (SWEAA, Sullivan & Karlsson, 1998) were used to assess type and severity of the RED. The Hospital Anxiety and Depression Scale (HADS, Zigmond & Snaith, 1983), was used to assess anxiety and depression levels. Further measures were used to tap into domains relevant to both autism and EDs including: the Camouflaging autistic Traits Questionnaire (CAT-Q, Hull et al., 2019), the Intolerance of Uncertainty Scale (IUS, Lauriola et al., 2018), the Interoception Sensory Questionnaire (ISQ, Fiene et al., 2018), the Glasgow Sensory Questionnaire (GSQ, Robertson & Simmons, 2013), the Toronto Alexithymia Scale (TAS, Bagby et al., 1994), the Brief Fear of Negative Evaluation (BFNE, Leary, 2016), the Social Phobia Inventory (SPIN, Connor et al., 2000), the Self-Compassion Scales (SCS, (Neff, 2003)), the Submissive Behaviour Scale (SBS, Allan & Gilbert, 1997), the Pride in Eating Pathology Scale (PEP-S, Faija et al., 2017), the Body Shape Questionnaire (BSQ, Arnow et al., 1995) and finally, the Sociocultural Attitudes Towards Appearance Questionnaire-3 (SATAQ, Thompson et al., 2004).

### 2.3 | Statistical analysis

Analysis was performed using JASP (Version 0.16.1) for Macintosh (JASP Team, 2022), an open-source statistical software application. Demographic group differences were analysed using measures of central tendency and, where appropriate, student t-tests.

Normality of the distributions across all measures was assessed visually using histograms and homogeneity of variance was verified using the Levene test. Our method for generating a screening method followed two steps. First, in Step One, we identified measures that had the potential to contribute to the screening method. Second, in Step Two, we combined the identified measures from
Step One in a statistical model to develop an efficient way of combining scores to screen for autism.

For Step One, parametric comparative analysis was performed for all questionnaires' total scores, and subscale scores if available. Independent-samples student t-tests were conducted to compare scores for autistic and non-autistic participants with a RED. Only those measures that were significantly different between groups were taken forward to Step Two. Effects size calculations were conducted using Cohen’s d (Cohen, 1992) to inform variable selection.

For Step Two, conditional stepwise binary logistic regression (dependent variable = autistic (1)/non-autistic (0)) was conducted first with AQ-10 alone, to define the comparison model. Then each questionnaire identified in Step One was added to the model in order of effect size until a questionnaire was added that did not significantly contribute to the model. This questionnaire was then removed and the next one added, until all questionnaires had been tested. If the contribution of a questionnaire became non-significant due to the addition of a subsequent questionnaire, then it was removed from the model. Significance was assessed with an alpha error rate of 0.05 (two-tailed) and multicollinearity between the variables was assessed at each stage using the variation inflation factor (VIF). For comparability, a cut-off value of 0.5 was chosen for all models, including the initial comparative AQ-10 model which typically has a cut-off of 0.6 (Allison et al., 2012). Models were compared using both adjusted Nagelkerke’s $R^2$ and area under the curve (AUC) analysis derived from the models’ receiver operating characteristic curve (ROC).

### 2.4 Ethics

The study was approved by the University College London’s ethics committee (12973/002). Participants provided written informed consent after reading an information sheet, approved by the ethics committee. Participants were fully debriefed at the end of the study and offered the opportunity to get a summary of their results.

### 3 RESULTS

#### 3.1 Participants

All participants were women with a RED, with 42 in the autistic group with a mean age of 29.2 years ($SD = 9.4$), and 118 in the non-autistic group who had a mean age of 29.8 years ($SD = 9.1$). Thirty-six (85.7%) of the autistic group and 107 (90.7%) of the non-autistic group identified ethnically as White British. All participants were living in and registered to a General Practitioner (i.e., family doctor) in the UK. Thirty-one (75.6%) of the autistic group had a current diagnosis of AN, seven (17%) had AAN and three (7%) were diagnosed with ARFID. The non-autistic group was comprised of 100 (84.8%) of participants with a current AN diagnosis, 17 (14.4%) with AAN and one participant with ARFID. There was no significant difference between the groups as to what age they received their RED diagnosis, $t(156) = −2.1$, $p > 0.05$, with the autistic group receiving a diagnosis on average at the age of 18.8 years ($SD = 5.8$) and the non-autistic group receiving a diagnosis at the age of 21.9 ($SD = 8.1$). BMI data was available for 91% of participants and demonstrates that the autistic group had an average BMI of 18.3 ($SD = 3.2$) and the non-autistic group a BMI of 17.2 ($SD = 2.6$) but this was not a significant difference, $t(144) = 1.94$, $p > 0.05$. There were also no significant difference between the groups’ lowest ever weight, $t(148) = 0.81$, $p > 0.05$, suggesting comparable RED severity.

#### 3.2 Step one – Identifying measures of potential value for autism screening

Significant differences between the autistic and non-autistic groups with small to large effect sizes were found across 52 full and subscale scores, with the AQ-10 providing the largest difference. There were no significant group differences in EDEQ global scores or any of its subscales, suggesting that the groups were similar in ED symptom severity. Group comparison significance and effect sizes with 95% confidence intervals (CI) are displayed in Table 1 below.

#### 3.3 Step two – Identifying autism screening model

#### 3.3.1 Comparison model

Binary logistic regression was conducted to classify participants into either the autistic or non-autistic group using the AQ-10. The model was statistically significant ($\chi^2 (158) = 72.3 \ p < 0.01$), indicating that the AQ-10 significantly improves the model’s ability to discriminate between the two groups over random chance alone. The model explained 53.2% (Nagelkerke’s $R^2$) of the variance and correctly classified 85.0% of cases (Sensitivity 69.1% (95% CI 52.9–82.4), Specificity 90.7% (83.9–95.3)) with an AUC of 0.90. The AQ-10 within the
<table>
<thead>
<tr>
<th>Measure</th>
<th>t</th>
<th>p</th>
<th>Cohen’s d</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>Measure</th>
<th>t</th>
<th>p</th>
<th>Cohen’s d</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQ-10</td>
<td>10.07</td>
<td>&lt;0.001</td>
<td>1.81</td>
<td>1.40</td>
<td>2.21</td>
<td>ISQ total</td>
<td>3.47</td>
<td>&lt;0.001</td>
<td>0.62</td>
<td>0.26</td>
<td>0.98</td>
</tr>
<tr>
<td>RAADS-14 total</td>
<td>9.62</td>
<td>&lt;0.001</td>
<td>1.73</td>
<td>1.33</td>
<td>2.13</td>
<td>IUS total</td>
<td>3.44</td>
<td>&lt;0.001</td>
<td>0.62</td>
<td>0.26</td>
<td>0.98</td>
</tr>
<tr>
<td>RAADS-14 mentalising deficit</td>
<td>9.04</td>
<td>&lt;0.001</td>
<td>1.62</td>
<td>1.23</td>
<td>2.02</td>
<td>GSQ gustatory hypo</td>
<td>3.43</td>
<td>&lt;0.001</td>
<td>0.62</td>
<td>0.26</td>
<td>0.98</td>
</tr>
<tr>
<td>RBQ-2A total</td>
<td>8.95</td>
<td>&lt;0.001</td>
<td>1.61</td>
<td>1.21</td>
<td>2.00</td>
<td>SWEAA purchase of food</td>
<td>3.33</td>
<td>0.00</td>
<td>0.60</td>
<td>0.24</td>
<td>0.96</td>
</tr>
<tr>
<td>RAADS-14 sensory reactivity</td>
<td>8.59</td>
<td>&lt;0.001</td>
<td>1.54</td>
<td>1.15</td>
<td>1.93</td>
<td>GSQ vestibular hypo</td>
<td>3.14</td>
<td>0.00</td>
<td>0.56</td>
<td>0.21</td>
<td>0.92</td>
</tr>
<tr>
<td>GSQ auditory total</td>
<td>8.51</td>
<td>&lt;0.001</td>
<td>1.53</td>
<td>1.14</td>
<td>1.92</td>
<td>GSQ tactile hypo</td>
<td>3.01</td>
<td>0.00</td>
<td>0.54</td>
<td>0.18</td>
<td>0.90</td>
</tr>
<tr>
<td>CAT-Q compensation</td>
<td>7.71</td>
<td>&lt;0.001</td>
<td>1.39</td>
<td>1.00</td>
<td>1.77</td>
<td>SWEAA mealtime surroundings</td>
<td>2.86</td>
<td>0.01</td>
<td>0.51</td>
<td>0.16</td>
<td>0.87</td>
</tr>
<tr>
<td>GSQ auditory hyper</td>
<td>7.36</td>
<td>&lt;0.001</td>
<td>1.32</td>
<td>0.94</td>
<td>1.70</td>
<td>SBS total</td>
<td>2.57</td>
<td>0.01</td>
<td>0.46</td>
<td>0.11</td>
<td>0.82</td>
</tr>
<tr>
<td>RAADS-14 social anxiety</td>
<td>7.05</td>
<td>&lt;0.001</td>
<td>1.27</td>
<td>0.89</td>
<td>1.64</td>
<td>SWEAA simultaneous capacity</td>
<td>2.56</td>
<td>0.01</td>
<td>0.46</td>
<td>0.10</td>
<td>0.82</td>
</tr>
<tr>
<td>RBQ-2A repetitive motor behaviours</td>
<td>7.03</td>
<td>&lt;0.001</td>
<td>1.26</td>
<td>0.88</td>
<td>1.64</td>
<td>SWEAA pica</td>
<td>2.50</td>
<td>0.01</td>
<td>0.45</td>
<td>0.09</td>
<td>0.80</td>
</tr>
<tr>
<td>RBQ-2A insistence sameness</td>
<td>7.03</td>
<td>&lt;0.001</td>
<td>1.26</td>
<td>0.88</td>
<td>1.64</td>
<td>IUS inhibitory anxiety</td>
<td>2.42</td>
<td>0.02</td>
<td>0.43</td>
<td>0.08</td>
<td>0.79</td>
</tr>
<tr>
<td>GSQ auditory hypo</td>
<td>6.82</td>
<td>&lt;0.001</td>
<td>1.23</td>
<td>0.85</td>
<td>1.60</td>
<td>SPIN total</td>
<td>2.38</td>
<td>0.02</td>
<td>0.43</td>
<td>0.07</td>
<td>0.78</td>
</tr>
<tr>
<td>TAS difficulty describing feelings</td>
<td>6.75</td>
<td>&lt;0.001</td>
<td>1.21</td>
<td>0.84</td>
<td>1.59</td>
<td>SWEAA social situation</td>
<td>2.28</td>
<td>0.02</td>
<td>0.41</td>
<td>0.06</td>
<td>0.77</td>
</tr>
<tr>
<td>GSQ total</td>
<td>6.75</td>
<td>&lt;0.001</td>
<td>1.21</td>
<td>0.83</td>
<td>1.59</td>
<td>SWEAA eating behaviour</td>
<td>1.65</td>
<td>0.10</td>
<td>0.30</td>
<td>−0.06</td>
<td>0.65</td>
</tr>
<tr>
<td>GSQ total hyper</td>
<td>6.71</td>
<td>&lt;0.001</td>
<td>1.21</td>
<td>0.83</td>
<td>1.58</td>
<td>GSQ olfactory hypo</td>
<td>1.54</td>
<td>0.13</td>
<td>0.28</td>
<td>−0.08</td>
<td>0.63</td>
</tr>
<tr>
<td>TAS difficulty identifying feelings</td>
<td>6.56</td>
<td>&lt;0.001</td>
<td>1.18</td>
<td>0.80</td>
<td>1.55</td>
<td>HADS-A</td>
<td>1.28</td>
<td>0.20</td>
<td>0.23</td>
<td>−0.12</td>
<td>0.58</td>
</tr>
<tr>
<td>GSQ vestibular hypo</td>
<td>6.55</td>
<td>&lt;0.001</td>
<td>1.18</td>
<td>0.80</td>
<td>1.55</td>
<td>CAT-Q masking</td>
<td>1.08</td>
<td>0.28</td>
<td>0.19</td>
<td>−0.16</td>
<td>0.55</td>
</tr>
<tr>
<td>GSQ visual total</td>
<td>6.51</td>
<td>&lt;0.001</td>
<td>1.17</td>
<td>0.79</td>
<td>1.54</td>
<td>SWEAA hunger satiety</td>
<td>0.92</td>
<td>0.36</td>
<td>0.17</td>
<td>−0.19</td>
<td>0.52</td>
</tr>
<tr>
<td>TAS externally oriented thinking</td>
<td>6.43</td>
<td>&lt;0.001</td>
<td>1.16</td>
<td>0.78</td>
<td>1.53</td>
<td>HADS-D</td>
<td>−0.14</td>
<td>0.89</td>
<td>−0.03</td>
<td>−0.38</td>
<td>0.33</td>
</tr>
<tr>
<td>CAT-Q total</td>
<td>6.25</td>
<td>&lt;0.001</td>
<td>1.12</td>
<td>0.75</td>
<td>1.50</td>
<td>PEP-S capturing others attention</td>
<td>−0.17</td>
<td>0.87</td>
<td>−0.03</td>
<td>−0.38</td>
<td>0.32</td>
</tr>
<tr>
<td>GSQ visual hyper</td>
<td>6.18</td>
<td>&lt;0.001</td>
<td>1.11</td>
<td>0.74</td>
<td>1.48</td>
<td>SWEAA other behaviour disturbed eating</td>
<td>−0.67</td>
<td>0.51</td>
<td>−0.12</td>
<td>−0.47</td>
<td>0.23</td>
</tr>
<tr>
<td>GSQ total hypo</td>
<td>6.06</td>
<td>&lt;0.001</td>
<td>1.09</td>
<td>0.72</td>
<td>1.46</td>
<td>PEP-S healthy weight eating</td>
<td>−0.94</td>
<td>0.35</td>
<td>−0.17</td>
<td>−0.52</td>
<td>0.18</td>
</tr>
<tr>
<td>GSQ proprioception total</td>
<td>5.85</td>
<td>&lt;0.001</td>
<td>1.05</td>
<td>0.68</td>
<td>1.42</td>
<td>SATAQ internalisation athlete</td>
<td>−1.15</td>
<td>0.25</td>
<td>−0.21</td>
<td>−0.56</td>
<td>0.15</td>
</tr>
<tr>
<td>GSQ vestibular total</td>
<td>5.64</td>
<td>&lt;0.001</td>
<td>1.01</td>
<td>0.64</td>
<td>1.38</td>
<td>SCS total</td>
<td>−1.18</td>
<td>0.24</td>
<td>−0.21</td>
<td>−0.56</td>
<td>0.14</td>
</tr>
<tr>
<td>CAT-Q assimilation</td>
<td>5.54</td>
<td>&lt;0.001</td>
<td>1.00</td>
<td>0.63</td>
<td>1.36</td>
<td>PEP-S weight loss food control thinness</td>
<td>−1.23</td>
<td>0.22</td>
<td>−0.22</td>
<td>−0.57</td>
<td>0.13</td>
</tr>
<tr>
<td>GSQ proprioception hyper</td>
<td>5.34</td>
<td>&lt;0.001</td>
<td>0.96</td>
<td>0.59</td>
<td>1.33</td>
<td>PEP-S total</td>
<td>−1.39</td>
<td>0.17</td>
<td>−0.25</td>
<td>−0.60</td>
<td>0.10</td>
</tr>
<tr>
<td>GSQ tactile hyper</td>
<td>5.26</td>
<td>&lt;0.001</td>
<td>0.95</td>
<td>0.58</td>
<td>1.31</td>
<td>BFNE total</td>
<td>−1.57</td>
<td>0.12</td>
<td>−0.28</td>
<td>−0.63</td>
<td>0.07</td>
</tr>
<tr>
<td>GSQ visual hypo</td>
<td>5.21</td>
<td>&lt;0.001</td>
<td>0.94</td>
<td>0.57</td>
<td>1.30</td>
<td>PEP-S outperforming others social</td>
<td>−1.65</td>
<td>0.10</td>
<td>−0.30</td>
<td>−0.65</td>
<td>0.06</td>
</tr>
</tbody>
</table>
regression model correctly classified 29 out of 42 individuals within the autistic group and 107 out of 118 in the non-autistic group.

3.3.2 | Improving the model

Additional variables were then included using the manual multistep procedure described above in ‘Analysis’. The first measure to provide a significant, albeit modest, addition to the model was the RAADS‐14 which increased the model’s explanatory power to 58.0% and correctly classified 83.8% of cases (Sensitivity 71.4% (55.4–84.3), Specificity 88.1% (80.9–93.4)) with an AUC of 0.92. The addition of the GSQ Auditory subscale only increased the specificity with one additional participant being correctly identified as having autism. The addition of the GSQ Auditory subscale meant the RAADS‐14 became non-significant and was therefore dropped from the model. Further iterations were constructed until no more significant variables increased the accuracy of the model. The final model consisted of the AQ‐10 along with the three subscales: GSQ Auditory, CAT‐Q Compensation and TAS Externally Oriented Thinking (EOT). This model increased the explanatory ability to 65.8% and correctly classified 86.9% of cases (Sensitivity 76.2% (60.6–88.0), Specificity 92.4% (86.0–96.5)) with an AUC of 0.94. The AQ‐10, inclusion of RAADS‐14 and the final model are depicted in Table 2 below.

3.3.3 | Secondary analysis

There were 21 (18%) participants within the RED group that, despite self-reporting an absence of autism, scored above cut-off on both autism screening measures, AQ‐10 and the RAADS‐14 and above cut-off on the restricted
and repetitive behaviours scale the RBQ-2A. Removing them from the analysis and re-running the model increases the explanatory power to 86.1% and correctly classified 93.5% of the cases (Sensitivity 85.7% (71.5–94.6), Specificity 96.9% (91.2–99.4)). This marginally outperforms the AQ-10 with the same participants removed, by correctly identifying one more case in the non-autism group.

4 | DISCUSSION

This study aimed to identify a brief questionnaire-based screening procedure for autism within a sample of women with a RED. We intended to identify a screening method that can be tested subsequently in an independent sample. When comparing the questionnaire responses of the autistic and non-autistic groups, the largest significant differences were unsurprisingly the three autism measures of diagnostic features of autism, indicating their efficacy in discriminating between autistic and non-autistic women in a RED sample. However, none of them were included in our final screening model, suggesting it is unlikely to be helpful to screen for autism in an ED group with more than one autism screening measure. However, including questionnaire subscales on auditory sensitivity (GSQ), social compensation (CAT-Q) and externally orientated thinking (TAS) significantly improved the model’s ability to discriminate between the two groups. This suggests the possibility that including these questionnaire subscales could increase the accuracy of the screening process. Including these subscales result in a screening process involving 33 questions, including the 10 from the AQ-10, and correctly identified five more individuals out of 160 (3.1%) compared to using the AQ-10 alone. This means 3.1% of women with a RED could go on to receive a more appropriate referral and clinical treatment pathway as a direct result of more accurate screening.

Considering the limitations of the AQ-10 as a screening tool within ED populations, it performed well within this RED sample, correctly classifying 85% of autistic women with only 10 questions. Our findings have some consistencies with the original validation of the AQ-10 with identical specificity rates of 91% (Allison et al., 2012). However, The AQ-10 in our RED sample had a 69% sensitivity rate, suggesting the presence of many false negatives, an almost 19% reduction versus the validation study. This suggests that in our RED sample, the AQ-10 was less sensitive when identifying autistic people than in a general population sample. Also, our finding of relatively low sensitivity and high specificity for the AQ-10 is in contrast to Ashwood et al. (2016), who found a high sensitivity rate of 77% but a very low specificity rate of 28%, due to many false positives. One possibility is that our finding of low sensitivity for the AQ-10 arose, in part, because our sample comprised only women; whereas the AQ-10 has been clinically validated on majority male samples (Wigham et al., 2019). The AQ-10 in our RED sample does a good job at identifying those that do not have autism with a 9% false positive rate but

<table>
<thead>
<tr>
<th>Variables</th>
<th>B</th>
<th>SE</th>
<th>z</th>
<th>Wald test</th>
<th>95% confidence interval Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wald</td>
<td>df</td>
<td>p</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>5.10</td>
<td>0.77</td>
<td>−6.65</td>
<td>44.20</td>
<td>0.01</td>
</tr>
<tr>
<td>AQ-10</td>
<td>0.68</td>
<td>0.11</td>
<td>6.32</td>
<td>39.99</td>
<td>1.97</td>
</tr>
<tr>
<td>AQ-10</td>
<td>0.40</td>
<td>0.13</td>
<td>2.99</td>
<td>8.91</td>
<td>1.49</td>
</tr>
<tr>
<td>RAADS-14</td>
<td>0.10</td>
<td>0.03</td>
<td>2.79</td>
<td>7.77</td>
<td>1.10</td>
</tr>
<tr>
<td>AQ-10</td>
<td>0.54</td>
<td>0.14</td>
<td>3.76</td>
<td>14.12</td>
<td>1.72</td>
</tr>
<tr>
<td>GSQ auditory thinking</td>
<td>0.42</td>
<td>0.12</td>
<td>3.38</td>
<td>11.44</td>
<td>1.52</td>
</tr>
<tr>
<td>CAT-Q compensation</td>
<td>0.07</td>
<td>0.03</td>
<td>2.50</td>
<td>6.27</td>
<td>1.07</td>
</tr>
<tr>
<td>TAS externally oriented thinking</td>
<td>−0.02</td>
<td>0.01</td>
<td>−2.71</td>
<td>7.35</td>
<td>0.98</td>
</tr>
</tbody>
</table>
misses 31% of individuals who are in fact autistic. However, overall accuracy is important when making clinical decisions for patient care, and we also need to consider the impact on the individual of scoring negative on a screening tool when they are in fact autistic and could therefore benefit from an appropriate referral and treatment adaptations.

Including the auditory sensitivity subscale from the GSQ, which includes both hyper- and hypo-sensitivity, is the biggest contributor to the improved model suggesting that it might be a more specific feature of autism that is less likely to be seen in those with a RED without autism. Auditory sensitivity is not considered to be characteristic of RED, but it is a recognised feature of autism and is commonly included in screening and assessment measures (American Psychiatric Association, 2013). Furthermore, the AQ-10 includes only one question that relates to auditory hypersensitivity, the remaining questions tap into domains such as social communication and cognitive differences that can be present in both autism and REDs (Westwood et al., 2017). The GSQ has been shown to be highly correlated with the full AQ, meaning that using the full versions of both measures would likely be unhelpful (Ujiie & Wakabayashi, 2015). However, within the GSQ validation study, the auditory subscale was one of the least correlated to the full AQ and therefore more likely to make a unique contribution to a screening questionnaire that already includes the shortened AQ (Sapey-Triomphe et al., 2018).

The second significant addition to the improved model was the compensation subscale from the CAT-Q, which is defined as strategies actively employed to compensate for difficulties in social situations, for example, mirroring body language or learning social cues from movies and books (Hull et al., 2019). This is another set of characteristics that are likely to be prevalent amongst autistic women, who commonly utilise camouflaging strategies to manage the challenges of being autistic in social environments that are generally designed by and for non-autistic people (Cook et al., 2021). Social camouflaging is one of the many reasons why some women go undetected until much later in life, in comparison to males who typically receive a diagnosis at a younger age (Ratto et al., 2018). Furthermore, women are likely to score differently to men on the gold standard diagnostic observation measure due to the measure focussing on social communication difficulties, which are more likely to be successfully masked by women (Ratto et al., 2018). Social camouflaging is less common in autistic males (Cook et al., 2021) and it is therefore uncertain whether this subscale would remain significant in a sample that included males.

The final subscale that significantly contributed to the improved model was the TAS Externally Orientated Thinking Scale, which is one aspect of alexithymia. Alexithymia is broadly described as a tendency to focus on concrete external events, rather than attend to one’s inner experience such as feelings and fantasies (Bagby et al., 1994). Alexithymia is a trait found in both REDs and autism, especially in terms of difficulties identifying and describing one’s own feelings (Nuske et al., 2013), which were captured in other subscales of the TAS. However, this study suggests that externally oriented thinking is an aspect of alexithymia more likely to be seen in autism. The TAS Externally Oriented Thinking Subscale was the weakest addition to the model and became a non-significant contributor once women that may have been autistic were removed from the RED only group. This suggests that if the groups were screened using a full autism assessment, the TAS externally orientated thinking is not likely to significantly contribute to the differentiating ability of the model.

4.1 | Limitations

Whilst participants in the non-autistic group were carefully screened for not having an autism diagnosis, nor a suspected diagnosis, there is a likelihood that some may actually be undiagnosed autistic women. Indeed, 18% of participants in the non-autistic group scored above clinical cut-off on both screening measures and a measure assessing a core diagnostic feature of autism. This will likely have caused us to underestimate the true validity of our model by depressing specificity and positive predictive value estimates (i.e., some findings that in our study are counted as false positives will be true positives). Future studies will need independent face-to-face autism assessment of all participants, including those in the non-autistic group, to investigate thoroughly our proposed algorithm.

Our findings indicate areas of assessment that could improve traditional autism screening measures within a RED population. However, it should not be used clinically until further validation is completed in an independent sample with participants who have undergone full autism diagnostic assessments and a reliable cutoff is ascertained. Furthermore, the sample consisted of only adult women and therefore we cannot be sure if our findings would be replicated in male populations or those under 18.

In conclusion, improvements in autism screening measures for individuals with REDs should look towards questions in the areas of auditory sensitivity, social compensation, and externally orientated thinking. These areas are more likely to be specific to autism and less likely to be influenced by non-autistic RED symptoms,
which means they may be helpful in differentiating between the two conditions. This can lead to more accurate referrals to the autism diagnostic pathway and the application of appropriate treatments. Importantly, an accurate autism screening measure can support clinicians in understanding the profile of autistic traits in their patients with REDs, so that treatment adaptations can be considered regardless of diagnostic status. Furthermore, identifying autistic characteristics that improve the identification of possible autism in women with a RED reduces the chance of overinflation from autistic-like traits that can be exaggerated in more severe illness states of REDs (Westwood & Tchanturia, 2017). Future research is needed to test the model we generated in a larger sample where all participants are given full autism diagnostic assessments to confirm group eligibility. Our model had stronger specificity (91%) than sensitivity (76%), which is not ideal for pre-assessment screening, where false negatives are more problematic than false positives. Therefore, as part of future work to test and develop the screening model we propose, it will be useful to investigate how the threshold for possible autism can be manipulated, to further reduce false negative rates.

ACKNOWLEDGEMENTS
The authors would like to thank all the participants who took part in the study and contributed to our understanding of both conditions. This study received no individual funding.

CONFLICT OF INTEREST
The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT
Data available upon request.

ORCID
James Adamson https://orcid.org/0000-0002-1153-9680
John Fox https://orcid.org/0000-0003-3039-8024
Will Mandy https://orcid.org/0000-0002-3564-5808

REFERENCES


