Acute hunger modifies responses on the Three Factor Eating Questionnaire

Hunger and Disinhibition, but not Restraint, scales.

Martin R. Yeomans¹²* and Keri McCrickerd³⁴

School of Psychology¹ and Sussex Neuroscience², University of Sussex, Falmer,

Brighton BN1 9QH, UK.

³ Clinical Nutrition Research Centre, Centre for Translational Medicine, Yong Loo

Lin School of Medicine, Singapore

⁴ Singapore Institute for Clinical Sciences, Agency for Science, Technology and

Research (A*STAR), Singapore

Running head: Hunger modifies responses on TFEQ

Address for correspondence:

Prof Martin R Yeomans

School of Psychology,

University of Sussex,

Brighton, BN1 9QH, UK

Email: martin@sussex.ac.uk

Tel: 01273 678617

Fax: 01273 678058
Abstract

It is widely assumed that responses on the Three Factor Eating Questionnaire (TFEQ) represent long-term (trait) attitudes to eating behaviour. However, the questionnaire requires agreement with a number of food related statements, and it is possible that some are easier to agree with when assessed hungry than sated. To test this potential state-dependency, participants completed a 100 mm visual analogue scale rating of their current hunger at the time they completed the TFEQ. Data were collected from two cohorts: Cohort 1 (507 women and 119 men) completed both measures on paper, while the hunger rating was computerised in Cohort 2 (179 women). Regression analysis revealed significant effects of rated hunger on scores on the hunger (TFEQ-H) and disinhibition (TFEQ-D) but not restraint (TFEQ-R) subscales, with higher TFEQ-H and TFEQ-D scores when participants were more hungry. In addition, 61 women and two men from Cohort 1 completed the measures on two separate occasions. Here, scores on TFEQ-H were higher on days when these participants were hungrier, but no differences in TFEQ-D or TFEQ-R were found. Overall these data suggest TFEQ-H could be interpreted as an indirect measure of current hunger, that scores on TFEQ-D are partly moderated by hunger but TFEQ-R is a more trait-like measure of restraint.

Key words:

Restraint, disinhibition, hunger
Since its publication, the Three Factor Eating Questionnaire (Stunkard & Messick, 1985) has become one of the most widely used psychometric tools identifying individual differences in eating attitudes in studies of human ingestive behaviour. As its name indicates, the questionnaire was devised to measure three aspects of human eating, defined by the authors as restraint (TFEQ-R), disinhibition (TFEQ-D) and hunger (TFEQ-H). Originally, these scales were designed to measure long-term attitudes to eating and consequently scores are often considered as long-term or trait measures (Barkeling, King, Näslund, & Blundell, 2007; Bryant, King, & Blundell, 2008; Finlayson, Cecil, Higgs, Hill, & Hetherington, 2012; Gallant, et al., 2013; Gallant, et al., 2010; Lattimore, Fisher, & Malinowski, 2011). However, this implies that answers to items on the TFEQ reflect long-term influences on each individual’s eating and are thus insensitive to the acute appetitive state of the participant. To our knowledge this assumption has never been formally tested.

At present there are three widely used measures of restrained eating: TFEQ-R, the restraint scale from the Dutch Eating Behaviour Questionnaire (DEBQ: Van Strien, Frijters, Bergers, & Defares, 1986) and the Revised Restraint Scale (RRS: Polivy, Herman, & Howard, 1988), which itself was developed from the original attempt to measure habitual restrained eating (Herman & Mack, 1975). Restraint is often initiated as a response to weight gain and all three measures are based on responses to items that measure the tendency to cognitively control eating behaviours and restrict intake due to concerns with current body weight.
The external focus restrained eaters makes it less likely that scores on restraint scales will vary with acute hunger state.

However, both TFEQ-D and TFEQ-H measure attitudes and responses to food. The TFEQ-D scale has been described variously as a measure of trait disinhibition (Lattimore, et al., 2011; Neale, Mazzeo, & Bulik, 2003), uncontrolled eating (Keskitalo, et al., 2008; Yeomans, Leitch, & Mobini, 2008) or opportunistic eating (Bryant, et al., 2008; Finlayson, et al., 2012), reflecting a recognition that the name disinhibition is confounded with the disinhibition effect arising from the breakdown of dietary restraint. Indeed, a shortened version of the TFEQ combined items from the original TFEQ-D and TFEQ-H scales into measure of uncontrolled and emotional eating (Karlsson, Persson, Sjostrom, & Sullivan, 2000). High scores on TFEQ-D have been associated with higher body-weight both on its own (French, Mitchell, Finlayson, Blundell, & Jeffery, 2014; Hays & Roberts, 2008; Lawson, et al., 1995; Tepper & Ullrich, 2002), and in combination with scores on TFEQ-R (Williamson, et al., 1995). In all cases, TFEQ-D is interpreted as a longer term dispositional measure of self-reported tendency to struggle to control eating. However, the statements underlying TFEQ-D typically relate to situations that challenge the ability to resist eating, either by the presence of desirable food or by emotional states that may promote eating. Thus it is plausible that some participants might be more likely to respond positively to these questions when more hungry than when sated.

The least researched of the three TFEQ subscales is TFEQ-H, originally seen as a measure of susceptibility to hunger cues: i.e. a reflection of long-term individual
differences in responses to hunger rather than a state measure of hunger per se.

High scores on TFEQ-H have again been associated with higher body-weight (Dykes, Brunner, Martikainen, & Wardle, 2004; French, et al., 2014), which runs counter to the original idea that high scores on TFEQ-H might relate to greater interoceptive awareness and consequently lower susceptibility to overeating.

However, as with TFEQ-D responses on hunger items could again be enhanced by actual state hunger: for example, it seemed plausible that hungry individuals might more readily agree with the statement “When I see a real delicacy, I often get so hungry that I have to eat right away”. Indeed, some papers interpret TFEQ-H as a measure of perceived or state hunger (de Castro & Lilenfeld, 2005; Rutters, Nieuwenhuizen, Lemmens, Born, & Westerterp - Plantenga, 2009; Williamson, et al., 1995), implying this is more an acute than trait measure, although others describe TFEQ-H in trait terms (Barkeling, et al., 2007; Bond, McDowell, & Wilkinson, 2001; e.g. Gendall, Joyce, Sullivan, & Bulik, 1998; Provencher, et al., 2005). This uncertainty in the nature of the TFEQ-H measure can be clarified by examining the acute sensitivity of TFEQ-H to actual hunger when the TFEQ was completed.

The key question in the present paper thus relates to the sensitivity of responses to the TFEQ to acute appetitive state. The majority of the scores on the original 51-item TFEQ derive from simple “True/False” statements which describe various eating situations. Although the intention was to use this simple questionnaire format to promote long-term responses, it is possible that the degree to which someone agreed with these statements depended on their acute hunger. The present study tested this idea by assessing how within and
between-person responses on all three TFEQ sub-scales varied as a function of hunger at the time of completion to explicitly test the extent to which TFEQ scores were sensitive to the self-reported hunger state at the time of testing.

Methods

Design
The study examined how scores on the three subscales of the TFEQ varied depending on the rated hunger of the participant at the time when the TFEQ was completed.

Participants
Participants were 805 male and female volunteers who completed the TFEQ and at the same time rated their hunger. Most data were obtained as part of a standardised recruitment process between 2006 and 2008: additional data came from a subset of specific studies conducted between 2012 and 2014, where participants again completed the TFEQ alongside a rating of current hunger. All data collection was approved by the University of Sussex Science and Technology Cross-Schools Research Ethics Committee (C-REC) and was conducted in line with the British Psychological Society code of conduct, ethical principles and guidelines. As this was an analysis of data originally collected for other purposes, we confirm all participants consented to the use of their responses in future research.
**Procedure**

Data were collected from two cohorts of participants. Cohort 1 was collected between 2006 and 2008 and comprised of staff and students at the University of Sussex who completed a standardised recruitment questionnaires part of their voluntary admission to the Sussex Ingestive Behaviour Unit (SIBU) participant pool. This combined the standard 51-item TFEQ, a series of questions about drinking habits, food aversions and allergies, and finally a single rating of hunger presented as a 100mm visual analogue scale (VAS) below the instruction “Please put a mark on the line to show how hungry you are right now, paying attention to the descriptions at the end of the line”, with the end-anchors “Not at all Hungry”, coded as zero, and “Extremely hungry”, coded as 100. The VAS rating was on a separate page from all TFEQ measures. In total 626 completed questionnaires were available for analysis in Cohort 1 (507 women and 119 men). Each person’s age at the time of completion was recorded (mean age: males = 23.4 ± 5.9, range 18-62 years; females = 21.6 ± 4.4, range 18-61 years).

Of the participants in Cohort 1, 63 (59 women and 4 men) completed the questionnaire more than once (separated by between 1 and 15 months). This provided the opportunity to test how within-participant differences in rated hunger altered the way they completed the TFEQ.

Data for Cohort 2 was collected later (2012-2014) from a further 179 female participants who consented to be part of research studies in the SIBU. On this occasion participants made computerised ratings of hunger alongside fullness, thirst and desire to eat. Participants were asked “How Hungry do you feel right
now?” and responded on a similar 100mm VAS as in Cohort 1, with the end- 
anchors “not at all hungry” (0) and “Extremely hungry” (100). This time, 
however, the rating was completed on a computer. All participants completed 
the TFEQ (on paper) straight after rating their appetite. Finally, Body Mass Index 
(BMI: calculated from height and weight measurements) and age at the time of 
testing was recorded (mean age = 20.6 ± 3.0, range 18-38 years; mean BMI 
kg/m² = 23.2 ± 3.6, range = 17-37).

Data analysis

The key question for this study was the extent to which ratings on the three sub- 
scales of the TFEQ depended on a person’s reported hunger at the time when the 
TFEQ was completed. Principle analyses regressed Rated Hunger against each 
TFEQ factor in separate regression models. Cohort (1 vs 2), Gender (male vs. 
female), Age (years) were entered as control variables in the first step of each 
model and their interaction with Hunger ratings were tested in a second step. 
BMI was not included as these data were only available for the smaller Cohort 2. 
Since 29 participants did not provide their age, analysis was on the 776 for 
whom we had complete data. Where a participant had completed the 
questionnaire twice, only data from their first questionnaire was included in 
these analyses. In the regression analysis, the % variance accounted for by each 
model (and specifically the variables of interest: rated hunger and the control 
variables gender, age and cohort) can be taken as the effect size and these data 
are reported throughout the results.
The second set of analysis examined data for those 63 participants from Cohort 1 who had completed the questionnaire twice. Here, individual hunger ratings were used to nominally classify participants as More Hungry (i.e. the session when their hunger score was highest) and Less Hungry (the session when their rating was lower). The three TFEQ scores were then contrasted between the More and Less Hungry conditions using 2-way ANOVA and report partial eta values as measures of effect size.

**Results**

Across all participants, responses on the three TFEQ sub-scales were somewhat related. TFEQ-R was significantly positively correlated with TFEQ-D ($r = 0.321$, $p < 0.001$) but not with TFEQ-H ($r = 0.008$, $p = 0.817$). The strongest relationship was noted for TFEQ-D and TFEQ-H subscales ($r = 0.462$, $p < 0.001$).

**Relationship between TFEQ scores and rated hunger**

Regression analyses found significant effects of hunger at time of completion on scores on TFEQ-H and TFEQ-D, but not TFEQ-R sub-scales (see Table 1). For TFEQ-R, there was no significant effect of rated hunger or age. Gender significantly influenced TFEQ-R, resulting from higher average TFEQ-R scores for women ($8.0 \pm 0.2$) than men ($4.7 \pm 0.3$), which would be expected. There was also a significant effect of test cohort on TFEQ-R, with scores tending to be higher in Cohort 2 ($9.2 \pm 0.4$) than Cohort 1 ($7.1 \pm 0.2$). Since Cohort 2 only used women, it is possible that the main effect of cohort is simply a reflection of this gender difference. Overall these variables accounted for 7% of the variance in TFEQ-R. Rated hunger did not interact with age, gender or cohort to affect
TFEQ-R scores, and these interaction terms did not significantly improve the regression model ($R^2$ change = 0.2%, $p = 0.702$) and were removed from the analysis.

There was a significant effect of hunger rating on TFEQ-D, with lower TFEQ-D associated with less hunger (Table 1). As with TFEQ-R scores, there was no effect of age on TFEQ-D but there was a significant effect of gender and a marginal effect of cohort: women tended to have higher scores on average (6.9 ± 0.1) than men (5.2 ± 0.3), and scores tended to be higher for cohort 2 (7.5 ± 0.2) than cohort 1 (6.4 ± 0.1). These variables accounted for 6% of the variance in TFEQ-D. Rated hunger did not interact with age, gender or cohort to affect TFEQ-R scores, and these interaction terms did not significantly improve the regression model ($R^2$ change = 0.1%, $p = 0.906$) and were removed from the analysis.

Finally, there was a significant effect of rated hunger on TFEQ-H, larger than the effect seen for TFEQ-D (Table 1), where increased hunger at the time of completing the TFEQ was associated with higher TFEQ-H. This model accounted for 13% of the variance in TFEQ-H. Rated hunger did not interact with age, gender or cohort to affect TFEQ-H scores, and these interaction terms did not significantly improve the regression model ($R^2$ change < 0.01%, $p = 0.474$) and were excluded from the analysis.

Since we also had BMI data for Cohort 2 along with other appetite ratings, we explored further predictors of TFEQ subscales in more detail with this smaller
cohort. BMI was not significantly related to TFEQ-R scores, but BMI was
positively associated with TFEQ-D (Beta = 0.30, t(176) = 9.09, p<0.001) and
negatively associated with TFEQ-H (Beta = -0.14, t(176) = -2.28, p = 0.024).
Neither rated fullness or desire to eat were significantly related to any of the
TFEQ sub-scales.

Effects of relative hunger on TFEQ scores
For the 63 participants who completed the TFEQ on two separate occasions,
analysis of actual rated hunger in the nominal More or Less hungry conditions
confirmed the difference was clear and significant (More Hungry, 54±3: Less
Hungry, 18± 2: t(62) = 12.92, p<0.001). Average TFEQ scores were higher when
participants were More than Less hungry (F(1,62) = 30.52, p<0.001, ηp² = 0.33),
but this depended on TFEQ subscale (F(2,124) = 22.24, p<0.001, ηp²= 0.26). As
can be seen (Figure 1), while relative hunger had no significant effect on TFEQ-R
or TFEQ-D, TFEQ-H scores were significantly higher in the More than Less
hungry condition (t(62) = 7.93, p<0.001).
Discussion

The key question posed in this paper was the extent to which responses on the TFEQ are truly “trait” measures as has often been assumed, but which to our knowledge has not previously been formally tested. The answer is that the three sub-scales of the TFEQ differed in their dependence on appetitive state at the time of questionnaire completion. Scores on TFEQ-R were independent of rated hunger, consistent with restraint being a long term dieting-related attitude. Scores on TFEQ-H were clearly influenced by actual hunger, with higher TFEQ-H scores when tested more hungry. The surprising finding was the small but significant influence of hunger on TFEQ-D scores.

Rated hunger at the time when the TFEQ was completed was associated with scores on TFEQ-H: the more hungry an individual was, the more likely they were to agree with items associated with the TFEQ-H scale. Moreover, where people had completed the TFEQ twice, they scored higher on TFEQ on the day when they were hungry than when more sated. Thus rather than being a trait measure, these data strongly suggest that TFEQ-H is at least partially an expression of actual hunger at the time when the questionnaire is completed.

Since the majority of items contributing to the TFEQ-H score are simple True/False statements like “Since I am often hungry, I sometimes wish that while I am eating, an expert would tell me that I have had enough or that I can have something more to eat” (TFEQ item 8), it must be easier to agree with this if hungry than sated. This questions the outcomes of studies that have interpreted effects of TFEQ-H as trait effects of sensitivity to physiological hunger.
In contrast to TFEQ-H, TFEQ-R scores were clearly independent of hunger at the
time of completion, in line with the general interpretation of restraint as a
longer-term (trait-like) attitude to eating. What was unexpected was the small
but significant effect of rated hunger on TFEQ-D scores. Notably the within-
participant contrast of TFEQ-D scores between the same person tested hungry or
sated was also not significant. Given the small effect size of the influence of
hunger on TFEQ-D in the regression analysis, and the lack of effect when
contrasted between hungry and sated states, it is unlikely that hunger will have
significantly influenced the outcomes of studies which have tested effects of
TFEQ-D on other behaviours without controlling for acute hunger state.
However it does suggest it would be prudent to record actual hunger and covary
this in future studies exploring effects of TFEQ-D, and certainly to do so for
TFEQ-H. Indeed, for studies exploring differences between groups defined by
TFEQ scores but with smaller sized samples, the best approach would be to
standardise the level of hunger to minimise any potential effects of acute hunger
state.

As in previous studies, there was also evidence that scores on the sub-scales of
the TFEQ are inter-related. TFEQ-D and TFEQ-R were positively correlated here,
and similar positive correlations have been reported previously in some studies
both from the same study population (Brace & Yeomans, 2016) and elsewhere
(e.g. Contento, Zybert, & Williams, 2005; Dykes, et al., 2004; Van Strien, Cleven, &
Schippers, 2000), but notably other well-powered studies have not found this
relationship (e.g. Lawson, et al., 1995; Williamson, et al., 1995), and some even
reported lower disinhibition with increasing restraint (Westenhoefer, 1991). An
explanation for these differences may lie in the observation by Westenhoefer (1991) that Dietary Restraint could be sub-divided into Rigid Control, which was positively correlated with TFEQ-D, and Flexible Control, which was negatively correlated with TFEQ-D. Thus the extent to which restraint and disinhibition may vary between populations, reflecting the relative distribution of flexible and rigid diet control amongst respondents. Weight status may also be key to the relationship between TFEQ-R and TFEQ-D: in a previous study, a positive relationship was noted between these measures in participants with a BMI within what the author’s defined as a healthy range for that population (between 18-27 kg/m²), but this relationship reversed in overweight and obese individuals (Bellisle, et al., 2004). The current sample was largely normal weight (although we only had actual weight measures for Cohort 2), which fits with the relationship we found between TFEQ-R and TFEQ-D.

There was also a clear positive relationship between TFEQ-D and TFEQ-H scores here, and in previous studies (Dykes, et al., 2004; Stunkard & Messick, 1985). In contrast, TFEQ-R and TFEQ-H were unrelated here, and elsewhere (Dykes, et al., 2004). The inter-relationships between responses on TFEQ subscales have led some to question the overall validity of the TFEQ (Karlsson, et al., 2000) and shorter versions have been developed with the aim of making these subscales more distinct (de Lauzon, et al., 2004; Karlsson, et al., 2000). Whether these shortened versions are also sensitive to acute hunger state remains untested, but given that they include items from the original TFEQ-H and have more response categories our current data suggest they may be equally sensitive to acute hunger.
Because the data collection for Cohort 1 was as part of a broader participant recruitment, we did not have BMI data for that part of the dataset, but notably for the more detailed data BMI and TFEQ-H were positively correlated, as has been reported elsewhere (Bellisle, et al., 2004; Dykes, et al., 2004; Hays, et al., 2002; Provencher, Drapeau, Tremblay, Despres, & Lemieux, 2003). However, this does raise some limitations in how far the present data can be generalised: Cohort 1 was an opportunity sample of undergraduates in a female-dominated University, and a replication with a larger male population would be useful. Likewise, the data are mainly from younger, healthy adults, and it would be interesting to note whether similar relationships are seen in a more representative population and in overweight/obese populations. Finally, it is notable for the within-participants contrast that the average hunger on the “More hungry” day was only 54pt on the 100pt hunger VAS, suggesting even the “More hungry” condition did not represent a truly hungry state. This raises an interesting issue which cannot be answered by the present dataset as to whether acute studies conducted in a fasted state would generate larger differences in state TFEQ responses, and this should be looked at in future studies.

In conclusion, the present data clearly show that the TFEQ-H scale is influenced by hunger state rather, undermining the use of this scale to measure the trait influence of sensitivity to hunger state. We also identify a weak influence of hunger on TFEQ-D scores, suggesting that studies using that measure should ideally co-vary actual hunger to remove any influence of current hunger state on disinhibition scores.
References


questionnaire-R18 is able to distinguish among different eating patterns in a general population. *Journal of Nutrition, 134*, 2372-2380.


Table 1. Outputs from regression analysis of predictors of scores on the three TFEQ sub-scales. Values given are the unstandardised $b$ values [bootstrapped 95 % CIs] and their associated significance from the entire cohort (n = 776).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>TFEQ-R</th>
<th>p</th>
<th>TFEQ-D</th>
<th>p</th>
<th>TFEQ-H</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hunger</td>
<td>&lt;0.01 [-0.02, 0.01]</td>
<td>0.772</td>
<td>0.01 [0.004, 0.02]</td>
<td>0.005</td>
<td>0.04 [0.03, 0.05]</td>
<td>0.001</td>
</tr>
<tr>
<td>Age</td>
<td>&lt;0.01 [-0.08, 0.08]</td>
<td>0.930</td>
<td>-0.01 [-0.06, 0.04]</td>
<td>0.593</td>
<td>-0.02 [-0.07, 0.02]</td>
<td>0.342</td>
</tr>
<tr>
<td>Gender</td>
<td>-3.01 [-3.84, -2.15]</td>
<td>0.001</td>
<td>-1.66 [-2.30, -0.95]</td>
<td>0.001</td>
<td>0.10 [-0.50, 0.75]</td>
<td>0.750</td>
</tr>
<tr>
<td>Cohort</td>
<td>1.52 [0.46, 2.59]</td>
<td>0.006</td>
<td>0.53 [-0.02, 1.15]</td>
<td>0.070</td>
<td>0.01 [-0.58, 0.57]</td>
<td>0.978</td>
</tr>
</tbody>
</table>

Each regression model for each TFEQ sub-scale was significantly better than using the mean model to predict TFEQ R, D and H ($F(4,771) \geq 11.24, p<0.001$). The interaction terms between Hunger and age, gender and cohort added in the second step were not significant and these variables were not included in the final model.
Figure 1. Scores on the three sub-scales of the TFEQ by the same participants when completed More (open bars) or Less (closed bars) hungry. Data are mean ± SE, n=63.