

Does Repeated Testing improve the Validity of Self-Reported Emotional Eating through a process of Meaning Making?

Tatjana Van Strien^{1*}, Laura HH Winkens² and Hanna Konttinen³

¹*Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands*

²*Consumption and Healthy Lifestyles Chair Group, Wageningen University & Research, Wageningen, The Netherlands*

³*Faculty of Social Sciences, University of Helsinki, Helsinki, Finland*

Article Info

*Corresponding author:

Tatjana Van Strien

Professor
Radboud University
Behavioural Science Institute
Nijmegen
The Netherlands
E-mail: t.vanstrien@psych.ru.nl

Received: May 29, 2019

Accepted: July 23, 2019

Published: August 1, 2019

Citation: van Strien T, Winkens LHH, Konttinen H. Does Repeated Testing improve the Validity of Self-Reported Emotional Eating through a process of Meaning Making? *Int J Obes Nutr Sci.* 2019; 1(1): 11-21.
doi: 10.18689/ijons-1000103

Copyright: © 2019 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Published by Madridge Publishers

Abstract

In two experimental studies in women, we investigated whether repeated testing improved the predictive validity of self-reported emotional eating (EE) for distress-induced food intake. We also tested whether there is support for a process of meaning making where pre-test and re-test EE are indirectly related through a serial causal chain of alexithymia and poor introspective awareness (IA). In study 1 (n=80), self-reported alexithymia and IA were measured before retesting EE. In study 2 (n=128), alexithymia and IA were measured after re-testing EE. In support of a process of meaning making, in both studies there was a significant serial chain of pre-test EE to re-test EE through alexithymia and IA. Further, re-test EE predicted somewhat more variance in distress-induced food intake than pre-test EE, though the difference was not significant. In conclusion, repeated testing may help respondents get a better understanding of a measure, thereby improving the validity of that measure.

Keywords: Emotional Eating; Re-Testing; Meaning Making; Validity; Alexithymia; Poor Introspective Awareness.

Introduction

There is increasing evidence that emotional eating, defined as eating in response to negative emotions [1] acts as a mediator between depression and weight gain [2-4]. As depression and obesity are common conditions with severe medical consequences and high cost for society, a reduction in high emotional eating (HEE) is an important treatment target for both obesity and depression [2].

A pre-condition for this approach is valid assessment of emotional eating. Assessing emotional eating by means of self-report requires respondents to have a good awareness of specific emotions. It can be questioned whether respondents with HEE meet this requirement, because difficulty in identifying and describing specific emotions (alexithymia) count among the principal causes of HEE [5]. Repeated assessment improved the reliability and validity of test scores on measures for general anxiety and employment tests [6,7]. The present study addresses the question of whether repeated testing similarly improves the predictive validity of self-reported emotional eating for food intake, and possible underlying mechanisms of such improvement.

Test-retest and Meaning Making

Scale scores on tests, such as those on anxiety, tend to show a mean shift towards greater improvement -less anxiety- from test to re-test [8,9]. Knowles et al. [9] explained this drift in scores with the phenomenon of meaning making: as respondents had more experience with a test, they were "... better able to discern its meaning and to use that meaning to interpret that item" [9]. An improved understanding of the meaning of a measure was associated with higher reliability of later answers [6]. Schubert and Fiske [10] and Ferrando [11] also found evidence of higher reliabilities of re-test scores. Van Iddekinge and Arnold [7] made similar observations for the retaking of employment tests. In their review, re-test scores were associated with more improvement, higher reliability, and better validity in relation to prediction of academic and job performance.

Re-testing may also improve the validity of self-endorsed emotional eating (EE); eating in response to distress is an atypical response. The typical response is not eating because emotional distress is associated with physiological responses that suppress feelings of hunger [12]. It occurs in people with high degrees of poor introspective awareness of hunger and satiety and alexithymia – difficulty in identifying and describing emotions [13-16]. When assessing EE by means of self-report, people with a limited awareness of their emotional and visceral states judge their own awareness of their emotions and their desire to eat in response to these emotions. For people with HEE, this may be a mission impossible [17]. If, as suggested by Knowles et al. [9], repeated measurement facilitates respondents' process of meaning making, exposure to items on eating in response to negative emotions at a pre-test may make respondents more aware of their poor ability to identify and describe emotions; their high degree of alexithymia, and their poor introspective awareness. This in turn, may result in a better awareness of their actual tendency towards eating in response to negative emotions with as possible outcome a more valid endorsement on items on emotional eating at the re-test.

In the following two studies we tested

- Differences between test and re-test scores on EE
- Alexithymia and poor introspective awareness as possible underlying mechanisms between pre-test EE and re-testEE
- The validity for food intake of test versus re-test scores of EE.

In the first study, an investigation of 'health and physiology', we conducted between the pre-test and the re-test an experiment which tested food intake after a control task and a stress task (Trier Social Stress Task; TSST) [18], using a within subjects design. In the second study, 'an investigation on the influence of the senses on mood and behaviour', we conducted between the pre-test and the re-test an experiment where participants were randomly assigned to a fabric-feeling, or a chocolate taste test condition. This study had a between subjects design and no stress

induction was used. Both studies included only women, because emotional eating has a higher prevalence in women [19] and because there are indications that the underlying mechanisms of emotional eating may be different in women than in men [20].

Study 1

Rationale and aim

Bekker, van de Meerendonk & Mollerus [21] tested EE twice, namely several weeks before and just after a mood induction -failure vs no failure on a quiz. There were 52 female participants, and the study had a between subjects design. It was found that the participants in the negative affect condition had significantly higher scores on EE at the re-test than at the pre-test. This finding is contradictory to earlier evidence [7,9], that re-test scores tend to show a shift towards greater improvement – less anxiety, better employability. However, a difference is that these test re-test studies did not manipulate mood. Evers et al. [17] suggested that the prediction of EE may be more accurate when assessing EE under 'hot' states, (when being emotional), instead of 'cold' states (when not being emotional) 'because people predict the influence of past or future hot states more accurately when they are in a corresponding state' [17]. This would suggest that re-testing EE after a mood manipulation would have higher validity for food intake than when re-testing EE without manipulating mood. However, whether the higher re-test scores in the study by Bekker et al. [21] had higher validity for food intake than the pre-test scores was not tested in that study, and the same held true for possible explanations for the higher re-test scores.

In study 1, we conducted between the pre-test and the re-test an experiment where all participants were offered food after a control and a stress task (Trier Social Stress Task: TSST). Using a within-subjects design all participants performed the control task and the stress task on two consecutive days. The re-test always took place immediately after the stress task and the food intake, so EE was assessed right after a 'hot' state. During the experiment, the respondent also filled out scales on alexithymia – on the first day, before the control condition and the food intake – and on poor introspective awareness – on the second day, before the stress condition and the food intake-(see procedure).

In regard to the question

- Differences between the pre-test and re-test scores, we left it an open empirical question whether the re-test scores would be significantly lower, as could be expected from the earlier test-re-test studies, or significantly higher, as in the one study by Bekker et al. [21].
- Alexithymia and poor introspective awareness as possible underlying mechanisms between pre-test EE and re-test EE, we expected that the pre-test and re-test scores of emotional eating would be indirectly related through the following serial chain of meaning making. Exposure to items on emotional eating at the pre-test may have made the respondents with HEE more aware of their poor

ability to identify and describe emotions: their degree of alexithymia, with as possible outcome a higher endorsement on subsequent questions on alexithymia and poor introspective awareness. This, in turn, may have resulted in a better awareness of their tendency toward emotional eating: EE at the pre-test → alexithymia → poor introspective awareness (also including awareness of hunger and satiety) → EE at the re-test.

- The validity for food intake of the pre-test versus re-test scores of EE, we expected that the re-test scale scores on EE would have a higher predictive validity for distress induced food intake than the pre-test scale scores.

Method Study 1

Design

Participants were recruited from a pool of female students taking introductory psychology or pedagogy courses who had completed the emotional eating scale at our university online portal. Results on the first 47 respectively 60 participants of the present sample, have been reported earlier [22-24]. The data for the additional participants in the present study were collected between October 2012 and May 2013.

Using a within-subjects design, female students who were preselected on the basis of extremely high or low scores on the DEBQ (Dutch Eating Behaviour Questionnaire; [25]) scale for emotional eating were subjected to a control task and a stress task (Trier Social Stress Task: TSST) on two consecutive days. The TSST involves speaking in front of a jury, coupled with an arithmetic challenge. Since some participants may perceive the stress condition as very stressful, we deliberately always started with the control condition and did not counter balance the order of the two conditions. We were afraid that we would lose too many subjects if we started with the stress condition because they would refuse to come back the following day for the control condition. We also were afraid that the control condition would suffer from carry over effects if we started with the stress condition [24]. As already noted, this means for the present study that all participants filled out the re-test of emotional eating immediately after the stressor and the subsequent food intake.

Of the additional women that participated in the present study, 17 women did not fulfill the requirement of having extreme values on the pre-test of emotional eating, because we had increasing difficulties in finding participants with extremely low values on emotional eating – extremely high values were not much of a problem. Nevertheless, with over 75% of our sample having extreme values on emotional eating we well met the advice of Whisman & McClelland [26] to oversample participants with extreme scores to enhance the power of the study (p. 118). Following Preacher [27], to preserve “the individual differences within each extreme” (p.2), we kept the data on emotional eating in the present study in their original, continuous form, instead of using the earlier dichotomy of low versus high emotional eating.

The study protocol was approved by the ethical board of the Faculty of Social Sciences of the Radboud University Nijmegen (ECG 29042010). Before participating, the participants filled out informed consent forms on both the control and the stress day.

Participants

Study 1 had 85 female participants, but we obtained complete information on pre-test and re-test emotional eating from 82 women (27 LEE, 36 HEE and 19 women with intermediate EE scores). The cut-off points for LEE and HEE were <1.82 and >3.25. They correspond with the 20th percent lowest and 80th percent highest scores on the DEBQ scale for emotional eating in the Dutch norm group of females. Their mean age was 23.01 (SD=2.25) (range 20–31) and their mean BMI (body mass index, weight/height²) was 21.09 (SD=2.50) (range 16.30–31.53; BMI<17.50: n=4; BMI>25.00: n=6). We had 80 participants with complete information on the main other variables of the present study.

Procedure

The sessions were scheduled on consecutive week days between 11 am and 15 pm. The participants filled out the scale for alexithymia at the control day, before the control condition. In this condition the participants had to rate various fabrics (e.g. fur and silk) on various attributes (e.g. softness and warmth) for 15 minutes. After this, they were led to a separate room to fill out questionnaires at a table which also held a glass of water and four bowls filled with, respectively, white grapes, pieces of carrot, M&M's and pieces of butter cake. Participants were invited to help themselves to the water and the food with the words: “Please help yourself to the water and the food. You have earned it”.

The scale for poor introspective awareness was administered to the subjects on the stress day, before the stress condition. In this stress condition the participants were subjected to a modified version of the TSST [18], which consisted of preparing (5 min) and delivering (5 min) a videotaped speech, followed by a serial subtraction task (5 min). The speech and subtraction task were presented in front of a two person jury who sat behind a table and wore white doctor's coats. To enhance the stress, the participants had to stand in stocking feet on a Wii© balance board, in front of the jury, and to prolong the period of stress, the participants had to wait for the jury's judgment of their performance. After this judgment, the participants were led to the separate room to fill out a further set of questionnaires. Once again, participants were invited to help themselves to the water and the food on the table in the same words as the previous day. After 20 minutes, the experimenter returned to administrate the emotional eating scale of the DEBQ. The final task for the experimenter was to measure the weight and height of the participant, and debrief, thank and compensate the participants with course credits. It should be noted that the experimenter was kept blind to the emotional eating status of the participants and that none of the participants was aware that their food intake was being measured.

Measures

Emotional eating was measured with the scale for emotional eating of the Dutch Eating Behaviour Questionnaire (DEBQ) [25]. This scale has 13 items (e.g., "Do you have a desire to eat when you are irritated?") and all items have to be rated on a 5-point scale with response categories that range from 1 'never' to 5 'very often'. The DEBQ has been rated as 'up to the mark' or 'good' by the Dutch Committee on Tests and Testing (COTAN) on all EFPA (European Federation of Psychologists' Association) criteria (e.g. norms, reliability (internal consistency, test-re-test) and validity (dimensional validity, construct validity and criterion validity) [28]. The COTAN rated the scale for emotional eating to have good validity for distress induced food intake [22].

The alexithymia aspects Difficulty identifying feelings and Difficulty describing feelings were measured with the Toronto Alexithymia Scale-20, the TAS-20 [29,30]. The subscale difficulty identifying feelings has seven items (e.g. "I have feelings that I can't quite identify"). The subscale difficulty describing feelings has five items (e.g. "It is difficult for me to find the right words for my feelings"); response categories range from 1 "never" to 5 "always".

Poor introspective awareness was measured with a subscale of the revised Eating Disorder Inventory (EDI-II) [31]. The scale for poor interceptive awareness has 10 items (e.g. "I get confused as to whether or not I am hungry"). Response categories ranged from 1 "never" to 6 "always". In contrast to the EDI manual [31], in which a transformation of responses into a four-point scale is advocated, the present study utilized untransformed responses, as scale transformation was found to reduce the validity of the EDI among a non-clinical population [32].

Mood

On both days, mood was measured upon arrival and at three more time points: immediately after the task, after the feedback and during the food intake. The Positive and Negative Affect Schedule (PANAS) [33], measured on a 5-point scale ('not at all' to 'extremely'), the degree to which participants experienced 10 positive and 10 negative moods. Scale-scores were obtained by calculating the mean of the items comprising a scale.

Food intake

Before and after participants ate, the bowls with grapes, carrots, M&M's and butter cakes were weighed with a professional balance (model 200, Kern[®]). We then translated weight into calories for each food type, and summed the caloric intake over the food. Before carrying out statistical analyses, the food intake data were scrutinized for outliers, defined as >3 SD above or below the mean for each assessment.

Analytic plan

All analyses were carried out using SPSS version 23.0 [34]. The data were first inspected for Skewness, Kurtosis and

outliers (>3 SD above and below the mean). Though regression analyses are very robust for violations of normality [35], particularly so when using bootstrapping: "bootstrapping does not impose the assumption of normality of the sampling distribution" [36], when necessary steps were taken to normalise the data. With repeated measures GLM we conducted manipulation checks using the values on mood in the two conditions and the various time points. Greenhouse-Geisser corrections were applied where appropriate.

We calculated the Pearson correlations, means, and standard deviations of the variables. The PROCESS macro of SPSS, developed by Hayes [35], was applied to test serial mediation (Model 6) as well as single mediation (Model 4). In this approach, effects are assessed with bias corrected bootstrap confidence intervals that are significant when the upper and lower bound of the bias corrected 95% confidence intervals (CI) does not contain zero. We used bootstrapping with 5,000 samples. The effects on distress induced food intake of pre-test vs post-test emotional were assessed with regression analyses. Distress induced food intake was measured by regressing the food intake (Kcal) at the stress day on the food intake at the control day, a positive score meaning more food intake after stress than after control. Since we had one-sided hypotheses in regard to the effects of EE on distress induced food intake (higher EE is associated with higher food intake under stress compared to control), we could test these effects at 90% CI ($p < 0.10$).

Results Study 1

Manipulation check

Table 1 shows the values on negative mood in the control and the stress condition upon arrival (T1), immediately after the task (T2), after the feedback in the stress condition (T3), and during the food intake (T4). In both conditions the values on negative mood were significantly affected by time (control condition: $F(2,352, 183,430)=6,403, p=.001, \eta_p^2=.08$; stress condition: $F(1,865, 147,368)=60,167, p<.001, \eta_p^2=.43$). In the control condition, negative mood showed slow improvement; here, the linear model reached the highest significance ($F(1,78)=10,882, p<.001, \eta_p^2=.12$). In the stress condition, negative mood showed a sharp peak immediately after the stressor but markedly improved during the food intake; here, the quadratic model reached the highest significance ($F(1,79)=86,314, p<.001, \eta_p^2=.52$). As could be expected, there were significantly higher values on negative mood in the stress than in the control condition on all time points except T1. The overall moderator effect of the stress condition on the mood values over time was significant ($F(3,75)=26,420, p<.001, \eta_p^2=.51$).

Next, the data were scrutinized for Skewness, Kurtosis, and outliers. There was one outlying value for food intake (kcal) in the stress condition of 894.47. This value was made less extreme by replacing the outlying value with the value of 3 SD above the mean (kcal=800.96). After this, no problems were observed.

Table 1. Means and standard deviations of the negative mood values in the control and the stress condition in study 1 (n=80).

Negative mood	Control condition		Stress condition	
	Mean	SD	Mean	SD
T1	1.39	0.56	1.28	0.50
T2	1.36	0.53	1.84	0.75
T3	1.32	0.54	1.81	0.77
T4	1.31	0.47	1.42	0.61

Notes: SD: Standard Deviation; T1: Upon arrival; T2: Immediately after the task; T3: After the feedback (in the stress condition); T4: During the food intake.

Descriptive

Table 2 shows the means and standard deviations of all variables in the study, in addition to the correlations between the variables. Emotional eating at the pre-test and at the re-test were highly interrelated (r=.61). As could be expected

from earlier test-re-test studies, but contradictory to the results by Bekker et al. [21], the re-test scores on emotional eating were significantly lower than the pre-test scores (T(79)=-2.22, p=.03, 95% CI [-0.43, -0.02]). Further, in line with our expectations, the re-test emotional eating was somewhat more strongly associated with food intake after stress (r=0.22, p=0.053) than the pre-test emotional eating (r=0.13, p=0.236). Both emotional eating at the pre-test and emotional eating at the re-test were significantly associated with alexithymia identifying feelings, but not alexithymia describing feelings. Although both emotional eating at the pre-test and at the re-test were significantly associated with poor introspective awareness, the association with poor introspective awareness was somewhat higher for emotional eating at the re-test. Alexithymia identifying feelings and describing feelings were significantly associated, which finding was also in line with expectations.

Table 2. Means, standard deviations and correlations among the variables in study 1 (n=80).

	Pre-test EE	Re-test EE	kcal control	kcal stress	Alex identifying	Alex describing	Poor introspective awareness
Pre-test EE							
Re-test EE	0.61**						
kcal control	-0.02	0.08					
kcal stress	0.13	0.22 (p=.053)	0.51				
Alex identifying	0.25*	0.34**	0.07	0.09			
Alex describing	-0.05	-0.04	0.06	0.16	0.43**		
Poor introspective awareness	0.23*	0.44**	0.05	0.15	0.69**	0.43**	
mean	2.79	2.57	157.75	188.39	2.19	2.44	2.07
SD	1.10	0.93	144.67	204.63	0.74	0.93	0.62

Notes: EE: Emotional Eating; Alex: Alexithymia.

*p<.05; **p<.01 (2-tailed)

Serial mediation

We next used the serial mediation model with the three mediators to test whether alexithymia difficulty identifying feelings, alexithymia difficulty describing feelings and poor introspective awareness mediated the association between pre-test EE and re-test EE. The total indirect effect (c-c') was significant (B=0.09, 95% CI=0.02, 0.18). The serial chain between pre-test and re-test emotional eating through alexithymia identifying feelings, alexithymia describing feelings and poor introspective awareness was significant (B=0.007, 95% CI=0.0006, 0.02), but also the serial chain through alexithymia identifying feelings and poor introspective awareness was significant (B=0.04, 95% CI=0.01, 0.11). The indirect effect of pre-test EE on re-test EE through the three mediators in serial was statistically different from the indirect effect of pre-test EE on re-test EE through the two mediators in serial (95% CI=0.01, 0.10). This indicated that that the pre-test EE had a larger effect on the re-test EE through the two mediators (alexithymia identifying feelings and poor introspective awareness) in serial than through the three mediators (alexithymia identifying feelings, alexithymia describing feelings and poor introspective awareness) in serial. When we tested the serial mediation model with the two mediators, the serial chain through alexithymia identifying feelings and poor introspective awareness remained

significant at 95% CI (B=0.04, 95% CI=0.01, 0.11) (Table 3, study 1). The full model, containing pre-test emotional eating, and the two mediators was significant (F(3,76)=22.39, p<.001) and explained 47% of the variance in re-test emotional eating. See figure 1 for the B (95% CI) associated with the various paths in the model. Finally, results from two separate single mediation models suggested that both alexithymia identifying feelings and poor introspective awareness acted as individual mediators between pre-test and re-test emotional eating at 95% CI (Table 3, Study 1).

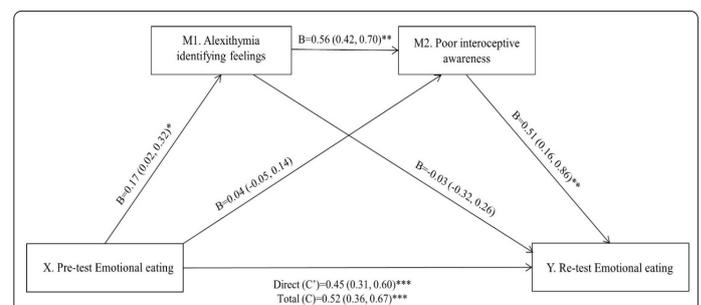


Figure 1. Serial multiple mediation of alexithymia difficulty identifying feelings and poor interoceptive awareness in the association between pre-test emotional eating and re-test emotional eating (n=80). Unstandardised beta coefficients (with bias-corrected and accelerated bootstrap 95% confidence intervals) are shown on the arrows. *p<.05, **p<.01, and ***p<.001.

Table 3. Indirect effects of the pre-test EE on the re-test EE through the two serial mediators (alexithymia identifying feelings and poor introspective awareness) or through the single mediators (alexithymia identifying feelings or poor introspective awareness).

	Study 1 (80 females)			Study 2 (128 females) ^a		
	unstandardized beta (B)	Confidence Interval (CI)	Model R ²	unstandardized beta (B)	Confidence Interval (CI)	Model R ²
Serial multiple mediation model						
Indirect effect via alexithymia identifying feelings and poor introspective awareness	0.04	95% CI=0.01, 0.11	0.47	0.01	90% CI=0.0007, 0.03	0.62
Single mediation model						
Indirect effect via alexithymia identifying feelings	0.04	95% CI=0.005, 0.12	0.41	0.02 ^b	95% CI=0.0001, 0.08	0.61
Single mediation model						
Indirect effect via poor introspective awareness	0.06	95% CI=0.01, 0.13	0.47	0.03 ^c	95% CI=0.0003, 0.09	0.62

^a63 females in the chocolate tasting condition and 65 females in the fabric feeling condition.

^bCondition moderated this mediation effect at 90% CI (index of moderated mediation: B=0.03; 90% Boot CI=0.0015, 0.09).

The mediation effect of alexithymia identifying feelings was higher in the chocolate tasting condition (B=0.04; 0.008, 0.10) than in the fabric feeling condition (B=0.01; -0.003, 0.06).

^cCondition did not moderate this mediation effect at 95% CI or at 90% CI (index of moderated mediation: B=-0.02; 90% Boot CI=-0.07, 0.02).

Pre-test and re-test emotional eating and prediction of stress induced food intake

In line with the Pearson correlation coefficients, pre-test EE did not significantly predict distress induced food intake at 90% CI (B=0.15, 90% CI=-0.02, 0.32). Only re-test emotional eating significantly predicted distress induced food intake at 90% CI (B=0.22, 90% CI =0.03, 0.49). Pre-test EE explained 2,7% of the variance in distress induced food intake, and re-test EE 4,5%. In a hierarchical regression analysis we tested the difference in explained variance in distress induced food intake by entering pre-test EE in step 1, and re-test EE in step 2. The increase in explained variance of re-test EE was not significant (R² change=0.017; F change (1,77)=1.388, p=.242).

Discussion Study 1

In study 1, we tested whether re-test scores on EE were significantly different from pre-test scores and whether pre-test EE was indirectly associated with re-test emotional eating through the serial chain of alexithymia and poor introspective awareness. We further examined whether re-test EE scores are associated with a better predictive validity for distress induced food intake than pre-test EE scores.

Results indicated that the re-test scores on EE were significantly lower than the pre-test scores. Further, there was support for a serial causal chain of pre-test to re-test EE through alexithymia difficulty identifying feelings, alexithymia difficulty describing feelings and poor introspective awareness. There was, however, also support for a serial

causal chain through alexithymia difficulty identifying feelings and poor introspective awareness, with a larger effect of pre-test EE on re-test EE through the two mediators than through the three mediators. Finally though only re-test EE significantly predicted distress induced food intake, there was no significant difference between pre-test and re-test EE in the amount of variance they explained in distress induced food intake.

Our finding that the re-test scores on EE were significantly lower than those on the pre-test is in correspondence with the robust evidence that scale scores on tests, such as those on anxiety, tend to show a mean shift towards greater improvement-less anxiety – at the retest [8,9]. Our finding is, however, contradictory to the higher re-test scores on DEBQ EE found by Bekker et al. [21] in their sample of 52 female college students. Those higher re-test scores were found in the negative affect condition, where EE was assessed right after a negative mood induction (failure on a quiz). Moreover all participants had been offered food-bowls with sweets and salty biscuits- during the quiz, therefore, EE was re-tested when the participants seemed to be in a ‘hot’ state [21]. In the present study 1, EE was also re-tested when the participants were in a ‘hot’ state (after a stress induction and after food intake), but a difference is that a major part of the sample in study 1 had been preselected on their extreme scores on emotional eating, whereas no such pre selection was used by Bekker et al. [21]. The lower re-test scores in our study may therefore be explained by regression toward the mean, of which scores at the extremes are more vulnerable [27]. Alternatively, the additional administration of tests on alexithymia and poor interceptive awareness between pre-test and re-test in our study may have had an additional facilitating effect on the respondent’s process of meaning making. In support of the meaning making process, suggested by Knowles et al. [9], we found that pre-test and re-test scores of EE were causally associated through alexithymia identifying feelings and poor introspective awareness.

A possible drawback of assessment of emotional eating at the very end of the experiment, after the stress and the food intake, is that the participants’ scores on emotional eating may have been affected by the amount of food they had just consumed. Indeed, see table 1, only re-test scores were associated with the food intake (kcal) after stress. So the question which arises reads: can the difference between pre-test and re-test scores also be explained by the amount of food consumed after stress? To investigate this possibility we also ran a *post hoc* mediation analysis with food intake (kcal) as mediator after stress. The indirect effect through food intake after stress between pre-test and re-test emotional eating was not significant at 95% CI (B=0.02, 95% CI=-0.01, 0.07) and also not at 90% CI (90% CI=-0.002, 0.06). Therefore, food intake after the stress induction and just before the filling out of the EE items at the re-rest did not explain the difference between the scores on EE at the pre-test and those at the post test.

Study 2

Rationale and aim

In study 1, EE was re-tested when all participants were in a 'hot' state (after a stress induction, so when they were emotional) and all participants had been offered food. For study 2, it would be of interest to determine whether similar results are obtained when EE is re-tested when the participants are in a 'cold' state, that is, when *no* stress induction took place, and when part of the participants did not receive food. A further characteristic of study 1 was that 75% of the participants had been pre-selected for their extremely low vs high EE scores; therefore, the lower EE scores at the re-test could also (partly) be explained by regression toward the mean, of which scores at the extreme are more vulnerable. Will re-test scores of EE also be lower than pre-test scores of EE when the participants had not been pre-selected for their extreme scores on EE? A final characteristic of study 1 was that the scales for alexithymia and poor introspective awareness had been filled out by the participants between the pre-test and the re-test, thereby, perhaps, having a facilitating role in the participants' meaning making process. Will there be support for a similar chain of mediation when the scales for alexithymia and poor introspective awareness are presented to the participants after the re-test, instead of between the pre-test and the re-test?

In Study 2, between the pre-test and the re-test an experiment was conducted where the participants were randomly assigned to a fabric-feeling- or a chocolate taste test condition, using a between subjects design. Mood was not artificially manipulated by a mood induction procedure. A further characteristic of study 2 was that scales for alexithymia and poor introspective awareness had been filled out by the participants after the re-test. In sum, in study 2 i) no mood manipulation was used, ii) participants were not pre-selected for their extreme EE scores, iii) scales for alexithymia and poor introspective awareness were administered after the re-test, and iiiii) part of the participants did not receive food.

Participants

For an experiment on the influence of 'the senses on mood and behaviour', participants were recruited from a pool of female students taking introductory psychology or pedagogy courses after completion of a scale on EE in our research participant portal (pre-test EE). The study had been approved by the IRB and after the students had signed an informed consent form, they were randomly assigned to a fabric-feeling, or a chocolate taste test condition. After removal of the data of four male students and of one female student with an outlying value ($> \text{mean} + 3 \text{ SD}$) on chocolate intake (namely, 257.4 grams), the number of participants was 128 (63 in the chocolate taste test condition). Their mean age was 20.57 years ($\text{SD}=1.93$) and their mean BMI (Body Mass Index; $\text{weight}/\text{height}^2$) was 21.50 ($\text{SD}=2.32$).

Procedure

Participants were instructed to refrain from food intake for at least 2 hours before the experiment. Experimental sessions were scheduled from 9.30 until 13.30 hour. After the participants had filled out a questionnaire on mood -not of relevance for the present study- they were taken to a separate room where they were invited to sit at a table. In the fabric feeling condition the participants had to rate six different fabrics (wool, fur, felt, linen and cotton) on various attributes (e.g., softness, pleasantness, warmth). In the chocolate taste test condition the table had a glass of water, three bowls of chocolate of the brand Delicata©, white (0% cocoa), milk (36% cocoa) or dark chocolate (58% cocoa), and three rating forms. The participants had to taste and rate each type of chocolate on various aspects after which they could eat as much from the chocolate as they wanted. After 10 minutes they were taken to their original room to fill out a further set of questionnaires also containing a scale on emotional eating (the re-test) and scales on poor introspective awareness and alexithymia (in this order). Though no mood manipulation was used, the participation in an experiment may in and out of itself provoke the diffuse and out of control sort of anxiety that elicits distress induced food intake in HEE [37-39]. The final task for the experimenter was to measure the weight and height of the participant, debrief, thank and compensate the participants with course credits. It should be noted that none of the participants was aware that their food intake was being measured.

Measures

Emotional eating was measured with a brief version of the emotional eating scale of the DEBQ [25], namely with 6 of the 13 items [1,40]. See further study 1.

For a description on the scales of poor introspective awareness and alexithymia, see Study 1.

Chocolate consumption

Before and after the participants ate, the individual bowls with chocolate (white, milk and dark) were weighed with a professional balance. The difference in weights was the intake of a specific type of chocolate. For total intake of chocolate (grams), we summed the intake of the specific types of chocolate.

In the analyses on the total sample, condition (fabric feeling vs chocolate tasting), was treated as possible confounder.

Analytic plan

The analytic plan of study 2 was similar to the one of study 1. A difference was that we only tested the serial mediation model with the two mediators alexithymia identifying feelings and poor introspective awareness (to test the robustness of the results of study 1). A further difference was that we used the single mediation models to test whether condition moderated the indirect effects of alexithymia identifying feelings and poor introspective awareness (moderated mediation). Significance of moderated mediation

was tested with Hayes' index of moderated mediation [41]. The effects on food intake of pre-test vs post-test EE in the chocolate tasting condition was tested with correlation and regression analyses. Because we expected that HEE would be associated with more food intake, which hypothesis is unidirectional, we tested these effects at 90% CI ($p < .10$).

Results Study 2

Descriptive

Table 4 shows the means and standard deviations of all variables in study 2, in addition to the correlations between

Table 4. Means, standard deviations and correlations among the variables in study 2 (n=128).

	Pre-test EE	Re-test EE	Alex identifying	Alex describing	Poor interoceptive awareness	Intake chocolate (grams)
Pre-test EE						
Re-test EE	0.77**					
Alex identifying	0.15	0.26**				
Alex describing	0.17	0.17	0.59**			
Poor introspective awareness	0.17	0.30**	0.44**	0.25**		
Intake chocolate (grams)	0.18	0.24*	0.04	0.08	-0.32**	
Mean	2.40	2.35	1.97	2.22	28.52	29.06
SD	0.74	0.85	0.66	0.67	3.87	13.11

Note: EE: Emotional Eating; Alex: Alexithymia; SD: Standard Deviation.

* $p < .05$; ** $p < .01$.

Serial mediation

We next tested the serial multiple mediation of alexithymia difficulty identifying feelings and poor introspective awareness in the association between pre-test EE and re-test EE (Table 3, study 2). The total indirect effect (c-c') was significant at 95% CI ($B=0.04$, 95% CI=0.01, 0.09). The serial chain between pre-test and re-test emotional eating through alexithymia identifying feelings and poor introspective awareness was only significant at 90% CI ($B=0.01$, 90% CI=0.001, 0.03). The full model, containing pre-test emotional eating, the two mediators and the confounder (condition) was significant ($F(4,123)=50,865$, $p < .001$) and explained 62% of the variance in re-test emotional eating. See figure 2 for the B (90% CI) associated with the various paths in the model.

We proceeded with testing whether alexithymia identifying feelings and poor introspective awareness would act as individual mediators between pre-test and re-test emotional eating and whether condition (fabric feeling vs chocolate tasting) would moderate possible mediation effects (Table 3, study 2). Alexithymia identifying feelings acted as mediator between pre-test and re-test emotional eating at 95% CI ($B=.02$, 95% CI=0.0001; 0.08), and condition moderated this mediation effect at 90% CI (index of moderated mediation: $B=0.03$, 90% CI=0.002; 0.09). The mediation effect of alexithymia identifying feelings was higher in the chocolate tasting condition ($B=0.04$; 90% CI=0.008, 0.10) than in the fabric feeling condition ($B=0.01$; 90% CI= -0.003, 0.06). Poor introspective awareness also acted as mediator between pre-test and re-test emotional eating at 95% CI ($B=.03$, 95% CI=0.003, 0.09), but there was no moderator effect of condition at 95% or 90% CI on the mediation effect of poor introspective awareness.

the variables. Emotional eating at the pre-test and at the re-test were highly interrelated ($r=.77$). As could be expected from earlier test-re-test studies, but contradictory to the results by Bekker et al. [21], the re-test EE scores were somewhat lower than the pre-test EE scores, though the difference was not statistically significant ($T(127)=1.064$, $p=0.290$, 95% CI [-0.05; 0.01]). Further, only re-test EE was significantly associated with intake of chocolate. Finally, only EE at the re-test was significantly associated with alexithymia identifying feelings and poor introspective awareness (but not with alexithymia describing feelings).

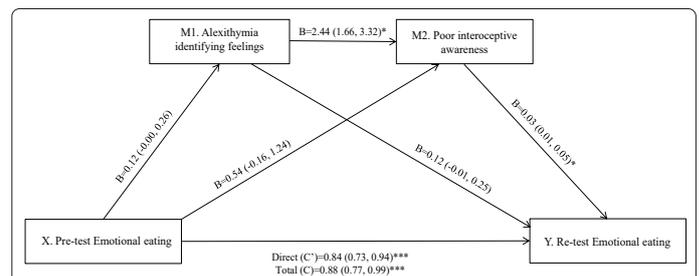


Figure 2. Serial multiple mediation of alexithymia difficulty identifying feelings and poor introspective awareness in the association between pre-test EE and re-test EE. Unstandardized beta coefficients (with bias-corrected and accelerated bootstrap 90% confidence intervals) are shown on the arrows (n=128). * $p < .05$, ** $p < .01$, and *** $p < .001$.

Pre-test EE vs re-test EE: difference in explained variance in food intake

Though only re-test emotional eating significantly predicted higher intake of chocolate (re-test EE: $B=3.622$, 90% CI=0.45, 6.80; pre-test EE: $B=3.258$, 90% CI=-0.46, 6.97), the increase in explained variance of re-test EE over and above the explained variance of pre-test EE was, however, not significant (R^2 change=0.022; F change (1,60)=1.420, $p=.238$).

Discussion Study 2

Results of study 2 all went in the same direction as those of study 1. As in study 1, the re-test scores on EE were lower than the pre-test scores, though the difference was this time not significant. The serial causal chain of pre-test EE to re-test EE through alexithymia difficulty identifying feelings and poor introspective awareness of study 1 could be replicated in study 2 at 90% CI. Though only re-test EE predicted higher

intake of chocolate, it did not explain more variance in intake of chocolate than pre-test EE. This result is also similar to the one of study 1.

As in study 1, we tested with a *post hoc* analysis the possible mediation effect of the amount of chocolate eaten between pre-test and re-test EE in the chocolate taste-test condition. Also here, this seemed not the case, as there was no indirect effect of intake of chocolate at 95% CI or at 90% CI ($B=0.02$, 90% CI= -0.004 , 0.08).

By using a between subjects design with a fabric feeling and a chocolate tasting condition, we also obtained some new results. When testing the individual mediation effects of alexithymia identifying feelings and poor introspective awareness both effects were significant at 95% CI. Though there was no moderator effect of condition on the mediation effect of poor introspective awareness, condition acted as moderator (at 90% CI) in the mediation effect of alexithymia difficulty identifying feelings. Interestingly, the mediation effect for alexithymia difficulty identifying feelings between pre-test and re-test EE was stronger in the chocolate tasting condition than in the fabric feeling condition.

General Discussion

In two studies, we tested i) whether there are differences between pre-test and re-test scores of EE, ii) whether there is support for a process of meaning making where pre-test and re-test EE are indirectly related through a serial causal chain of alexithymia and poor introspective awareness, and iii) whether re-test EE has a higher predictive validity for food intake than pre-test EE. In both studies, we found that the re-test scores on EE were lower than the pre-test scores, though the difference was only significant in study 1. In both studies we found that there was a serial causal chain of pre-test EE to re-test EE through alexithymia difficulty identifying feelings and poor introspective awareness, though this serial mediation was in study 2 only significant at 90% CI. Further, though re-test EE predicted in both studies somewhat more variance in distress induced food intake than pre-test EE, the difference in explained variance in food intake between pre-test and re-test EE was not significant.

The fact that most results of study 1 went in the same direction in study 2, that is when using no mood manipulation, no extreme scores on EE for 75% of the participants, and when assessing alexithymia and poor introspective awareness after the re-test of EE, instead of before the re-test EE indicates that the obtained results of study 1 are robust for experimental design and assessment timing.

The significant lower scores on EE at the re-test than at the pre-test of study 1, though in line with most earlier studies [7,9] are in contrast with the results by Bekker et al. [21]. In both studies re-test EE was assessed when the participants were in a 'hot' state, that is when they were emotional and additionally had eaten food. A difference was that in the study by Bekker et al. [21] the participants had not been pre-selected for their extreme scores on EE. A possible explanation

for the lower scores on EE in study 1 could be regression towards the mean, of which extreme scores are more vulnerable [27]. The fact that the re-test scores on EE in study 2, where there was no pre-selection on extreme EE scores, were only somewhat lower, but not significantly lower than those on pre-test EE would suggest support for regression towards the mean as possible explanation for the significant lower re-test scores in Study 1.

The lower scores at a re-test have also been interpreted as 'artefact' and outcome of improved awareness of the low social desirability of the items of a test [42]. High endorsement of items on emotional eating is indeed associated with low social desirability [43]. However, according to Nunnally [44] "adjustment and self-desirability (or self-esteem) are much the same thing only a poorly adjusted person would be so unfamiliar with social expectations as not to know how to 'fake good' on a self-inventory" (pp. 480-481). In the same line, McCrae and Costa [45] convincingly showed that correction for social desirability decreased rather than increased the validity of self-reports in relation to the external criterion of spouse ratings on various personality traits. In support; in our present two studies, the predictive validity of the re-test scores on EE was even somewhat higher than the ones of the pre-test scores, though not significantly higher. Further, the finding that in both study 1 and study 2, re-test EE had stronger associations with alexithymia difficulty identifying feelings and poor introspective awareness than pre-test EE indicates that the re-test scores also had a better construct validity than the pre-test scores.

In both study 1 and study 2 we found that pre-test and re-test EE were causally associated through alexithymia identifying feelings and poor introspective awareness. The serial chain of mediation we found in the present two studies suggests a process of meaning making where exposure to items on eating to negative emotions at the pre-test made respondents more aware of their poor ability to identify emotions, resulting in a higher endorsement on questions on the alexithymia aspect difficulty identifying feelings, in turn, resulting in a higher endorsement on questions on poor introspective awareness (including awareness of hunger and satiety), in turn, resulting in a better awareness of their tendency toward emotional eating at the re-test.

The finding that the serial chain of mediation was significant at 95% CI in study 1, and at 90% CI in study 2, would suggest that the filling out of items on alexithymia and poor introspective awareness before re-testing EE, may have been an additional facilitating factor in the process of meaning making of items on EE. In study 2, the scales on alexithymia and poor introspective awareness had been administered to the participants after the re-test of EE. A further difference between the two studies was that in study 1, the full 13-item version of the DEBQ-EE scale was used, whereas in study 2 only the brief, 6-item version of the EE scale was used. Participants in study 2 had therefore less opportunity to get experienced with the test, which may have had an additional negative effect on their process of meaning making.

The process of meaning making may possibly also be facilitated by the intake of food. In study 2, where food intake was manipulated by subjecting participants to a fabric feeling vs chocolate tasting condition, the mediation effect of alexithymia difficulty identifying feelings between pre-test and re-test EE was stronger in the chocolate tasting condition than in the fabric feeling condition. It does not, however, seem plausible that the difference between EE scores at pre-test vs re-test is explained by the intake of food just before the re-test. When we tested in a *post hoc* analysis the possible mediation effect of food consumption between pre-test and re-test EE, we did not find support for such mediation in either study 1 or study 2.

Implications for research

In experiments with food intake, emotional eating is usually administered at the very end of the experiment. The reason is, that this lowers the chance that participants get aware of the purpose of the study and that outcome variables such as the food intake are affected by this awareness [46]. Earlier it was shown that participants are less food when they suspected that their food intake was being monitored [47]. Awareness of the purpose of an experiment may therefore lower the validity of the experimental outcome. An entirely different matter is an improved understanding of the meaning of a measure, because an adequate understanding of the questions and the construct that they measure is a sine qua non for the predictive validity of that measure. This is supported by results of the present study. They suggest that additionally testing emotional eating at a pre-test may facilitate the participants' meaning making process, thereby enhancing both the predictive and construct validity of the re-test at the end of the study.

Implications for therapy

Improving the process of meaning making may also be helpful for the therapeutic process. Therefore, repeated assessment of constructs that are central to therapy may help patients to get a better understanding of that construct.

Strengths

The present study has various strengths. It consists of two studies that tackle the research questions with different experimental designs and timings of the assessments.

Limitations

Emotional eating is closely associated with binge eating and depressive feelings [2,15,46]. Therefore, it is highly probable that participants of study 1 with HEE had other symptomatology such as depressive feelings. This is a limitation of that study, which should deserve attention in future studies with more participants. An additional limitation is that both studies were conducted in predominantly normal weight females; hence the study needs replication in participants with overweight and in men. A further limitation is that we cannot rule out the possibility that social desirability or acquiescence may have affected scores on alexithymia, poor introspective awareness, negative affect and EE, but see McCrae & Costa [45].

Conclusion

Re-test emotional eating explained more variance in distress induced food intake than pre-test scores, but the difference in explained variance was not significant. In support of the contention that re-testing facilitates the process of meaning making, pre-test emotional eating was indirectly associated with re-test emotional eating through alexithymia and poor introspective awareness. Repeated testing may help respondents get a better understanding of the underlying construct of a measure, thereby improving the construct validity of that measure.

References

- van Strien T, Donker MH, Ouwens MA. Is desire to eat in response to positive emotions an 'obese' eating style: Is kummerspeck for some people a misnomer? *Appetite*. 2016; 100: 225-235. doi: 10.1016/j.appet.2016.02.035
- van Strien T, Konttinen H, Homberg JR, Engels RC, Winkens LH. Emotional eating as a mediator between depression and weight gain. *Appetite*. 2016; 100: 216-224. doi: 10.1016/j.appet.2016.02.034
- Vittengl JR. Mediation of the bidirectional relations between obesity and depression among women. *Psychiatry Research*. *Psychiatry Res*. 2018; 264: 254-259. doi: 10.1016/j.psychres.2018.03.023
- Konttinen H, van Strien T, Männistö S, Jousilahti P, Haukka A. Depression, emotional eating and long-term weight changes: a population-based prospective study. *Int J Behav Nutr Phys Act*. 2019; 16(1): 28. doi: 10.1186/s12966-019-0791-8
- van Strien T. Causes of emotional eating and matched treatment of obesity. *Curr Diab Rep*. 2018; 18(6): 35. doi: 10.1007/s11892-018-1000-x
- Knowles ES, Byers B. Reliability shifts in measurement reactivity: Driven by content engagement or self-engagement? *J Pers Soc Psychol*. 1996; 70(5): 1080-1090.
- Van Iddekinge CH, Arnold JD. Retaking employment test: What we know and what we still need to know. *Annual Review of Organizational Psychology and Organizational Behavior*. 2017; 4: 445-471. doi: 10.1146/annurev-orgpsych-032516-113349
- Windle C. Test-retest effect on personality questionnaires. *Educational and Psychological Measurement*. 1954; 14: 617-633. doi: 10.1177/001316445401400404
- Knowles ES, Coker MC, Scott RA, Cook DA, Neville JW. Measurement-induced improvement in anxiety: mean shifts with repeated assessment. *J Pers Soc Psychol*. 1996; 71(2): 352-363.
- Schubert DS, Fiske DW. Increase of item response consistency by prior item response. *Educ Psychol Meas*. 1973; 33(1): 113-121. doi: 10.1177/001316447303300112
- Ferrando PJ. Analyzing retest increases in reliability: A covariance structure modeling approach. *Struct Equ Modeling*. 2003; 10(2): 222-237. doi: 10.1207/S15328007SEM1002_4
- Gold PW, Chrousos GP. Organization of the stress system and its dysregulation in melancholic and atypical depression: high vs low CRH/NE states. *Mol Psychiatry*. 2002; 7(3): 254-275.
- Larsen JK, van Strien T, Eisinga R, Engels RC. Gender differences in the association between alexithymia and emotional eating in obese individuals. *J Psychosom Res*. 2006; 60(3): 237-243. doi: 10.1016/j.jpsychores.2005.07.006
- Pinaquy S, Chabrol H, Simon C, Louvet JP, Barber P. Emotional eating, alexithymia and binge eating disorder in obese women. *Obes Res*. 2003; 11(2): 195-201. doi: 10.1038/oby.2003.31
- van Strien T, Engels RC, Van Leeuwen J, Snoek HM. The Stice model of overeating: Tests in clinical and non-clinical samples. *Appetite*. 2005; 45(3): 205-213. doi: 10.1016/j.appet.2005.08.004

16. van Strien T, Ouwens MA. Effects of distress, alexithymia and impulsivity on eating. *Eat Behav.* 2007; 8(2): 251-257. doi: 10.1016/j.eatbeh.2006.06.004
17. Evers C, de Ridder DT, Adriaanse MA. Assessing yourself as an emotional eater. Mission impossible? *Health Psychol.* 2009; 28(6): 717-725. doi: 10.1037/a0016700
18. Kirschbaum C, Pirke KM, Hellhammer DH. The Trier Social Stress Test—a tool for investigating psychosocial stress responses in a laboratory setting. *Neuropsychobiology.* 1993; 28(1-2): 76-81. doi: 10.1159/000119004
19. O'Connor DB, Jones F, Conner M, McMillan B, Ferguson E. Effects of daily hassles and eating style on eating behavior. *Health Psychol.* 2008; 27(1S): S20-S31. doi: 10.1037/0278-6133.27.1.S20
20. van Strien T, Levitan RD, Engels RC, Homberg JR. Season of birth, the dopamine D4 receptor gene and emotional eating in males and females. Evidence of a genetic plasticity factor? *Appetite.* 2015; 90: 51-57. doi: 10.1016/j.appet.2015.02.024
21. Bekker MH, van de Meerendonk C, Mollerus J. Effects of negative mood induction and impulsivity on self-perceived emotional eating. *Int J Eat Disord.* 2004; 36(4): 461-469. doi: 10.1002/eat.20041
22. van Strien T, Herman CP, Anschutz DJ, Engels RC, deWeerth C. Moderation of distress-induced eating by emotional eating scores. *Appetite.* 2012; 58(1): 277-284. doi: 10.1016/j.appet.2011.10.005
23. van Strien T, Roelofs K, de Weerth C. Cortisol reactivity and distress-induced emotional eating. *Psychoneuroendocrinology.* 2013; 38(5): 677-684. doi: 10.1016/j.psyneuen.2012.08.008
24. van Strien T, Ouwens MA, Engel C, de Weerth C. Hunger, inhibitory control and distress-induced emotional eating. *Appetite.* 2014; 79: 124-133. doi: 10.1016/j.appet.2014.04.020
25. van Strien T. Nederlandse Vragenlijst voor Eetgedrag (NVE). Handleiding. [Dutch Eating Behaviour Questionnaire. Manual. 80p]. *Hogrefe.* 2015.
26. Whisman MA, McClelland GH. Designing, testing and interpreting interactions and moderator effects in family research. *J Fam Psychol.* 2005; 19(1): 111-120. doi: 10.1037/0893-3200.19.1.111
27. Preacher KJ. Extreme groups designs. In: Cautin RL, Lileinfield SO (eds). *The encyclopedia of clinical psychology.* Hoboken, NJ: John Wiley & Sons, Inc. 2015; 2: 1189-1192.
28. COTAN. Beoordeling Nederlandse Vragenlijst voor eetgedrag, 2013. [Review Dutch Eating Behaviour Questionnaire (DEBQ)]. 2013.
29. Bagby R, Parker JDA, Taylor GJ. The twenty-item Toronto Alexithymia Scale: I Item selection and cross-validation of the factor structure. *J Psychosom Res.* 1994; 38(1): 23-32. doi: 10.1016/0022-3999(94)90005-1
30. Kooiman CG, Spinhoven P, Trijsburg RW. The assessment of alexithymia: a critical review of the literature and a psychometric study of the Toronto Alexithymia Scale-20. *J Psychosom Res.* 2002; 53(6): 1083-1090.
31. Garner DM. *Eating Disorder Inventory-2: professional manual.* Psychological Assessment Resources. 1991.
32. Schoemaker C, van Strien T, van der Staak C. Validation of the Eating Disorder Inventory in a non-clinical population using transformed and untransformed responses. *Int J Eat Disord.* 1994; 15(4): 387-393.
33. Watson D, Clark LA, Tellegen A. Development and validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *J Pers Soc Psychol.* 1988; 54(6): 1063-1070.
34. IBM Corporation. *IBM SPSS Statistics for Windows, Version 23.0.* Armonk, NY: IBM Corp. 2015.
35. Hayes ASF. *Introduction to mediation, moderation, and conditional Process analysis, A regression-based approach.* The Guilford Press. 2013.
36. Preacher KJ, Hayes A. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods.* 2008; 40(3): 879-891. doi: 10.3758/BRM.40.3.87
37. Polivy J, Herman CP. The effects of alcohol on eating behavior: disinhibition or sedation? *Addict Behav.* 1976; 1(2): 121-125. doi: 10.1016/0306-4603(76)90004-6
38. Kaplan S, Hobart JL. New technique for recording skin resistance. *Am J Med Electron.* 1965; 4: 117-120.
39. van Strien T, Ouwens MA. Counterregulation in female obese emotional eaters: Schachter, Goldman, and Gordon's (1968) test of psychosomatic theory revisited. *Eat Behav.* 2003; 3(4): 329-340.
40. Paans NPG, Bot M, Van Strien T, Brouwer IA, Visser M, Penninx BWJH. Eating styles in major depressive disorder: Results from a large-scale study. *J Psychiatr Res.* 2018; 97: 38-46. doi: 10.1016/j.jpsychires.2017.11.003
41. Hayes AF. An index and test of linear moderated mediation. *Multivariate Behav Res.* 2015; 50(1): 1-22. doi: 10.1080/00273171.2014.962683
42. Jorm AF, Duncan-Jones P, Scott R. An analysis of the re-test artefact in longitudinal studies of psychiatric symptoms and personality. *Psychol Med.* 1989; 19(2): 487-493.
43. van Strien T, Frijters JER, Roosen RG, Knuiman-Hijl WJ, Defares PB. Eating behavior, personality traits and body mass in women. *Addict Behav.* 1985; 10(4): 333-343.
44. Nunnally JC. *Psychometric theory.* McGraw-Hill, New York. 1967.
45. McCrae RR, Costa PT. Social desirability scales: More substance than style. *J Consult Clin Psychol.* 1983; 51(6): 882-888. doi: 10.1037/0022-006X.51.6.882
46. Orne MT. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist.* 1962; 17(11): 776-783. doi: 10.1037/h0043424
47. Robinson E, Kersbergen I, Brunstrom JM, Field M. I'm watching you. Awareness that food consumption is being monitored is a demand characteristic in eating-behaviour experiments. *Appetite.* 2014; 83: 19-25. doi: 10.1016/j.appet.2014.07.029