

Psychological and Biological Validation of a Novel Digital Social Peer Evaluation Experiment (digi-SPEE)

Claudia Menne-Lothmann^{*†1}; Jeroen Decoster^{‡1,2}; Ruud van Winkel^{#1,2}; Dina Collip^{#1}, Bart P. F. Rutten¹; Philippe Delespaul¹; Marc De Hert²; Catherine Derom^{3,4}; Evert Thiery⁵; Nele Jacobs^{1,6}; Jim van Os^{1,7}; Marieke Wichers^{1,8}

¹Department of Psychiatry and Psychology, Maastricht University Medical Centre, Maastricht, The Netherlands

²University Psychiatric Centre KU Leuven, Leuven, Belgium

³Centre of Human Genetics, University Hospitals Leuven, Leuven, Belgium

⁴Department of Human Genetics, KU Leuven, Belgium

⁵Department of Neurology, Ghent University Hospital, Ghent University, Ghent, Belgium

⁶Faculty of Psychology, Open University of the Netherlands, Heerlen, The Netherlands

⁷Department of Psychosis Studies, Institute of Psychiatry, King's College London, King's Health Partners, London, United Kingdom

⁸Department of Psychiatry, University Medical Centre of Groningen, Groningen University, Groningen, The Netherlands

^{*}contributed equally; [#] contributed equally

ABSTRACT

Introduction: Negative social evaluation is associated with psychopathology. Given the frequency of evaluation through increasingly prevalent virtual social networks, increased understanding of the effects of this social evaluation is urgently required.

Methods: A new digital social peer evaluation experiment (digi-SPEE) was developed to mimic everyday online social interactions between peers. Participants received mildly negative feedback on their appearance, intelligence, and congeniality.

Two hundred and forty-one young people [58.9% female, aged 18.9 years (15 to 34)] from an ongoing novel general population twin study participated in this study. Positive affect (PA), negative affect (NA), implicit self-esteem, and cortisol were assessed before and after exposure to the social evaluation experiment.

Results: The social evaluation experiment decreased PA ($B=-5.25$, $p<.001$) and implicit self-esteem ($B=-.19$; $p<.001$), whereas it increased NA ($B=5.99$; $p<.001$) and cortisol levels ($B=.07$; $p<.001$). Females (PA: $B=-7.62$; $p<.001$; NA: $B=8.28$; $p<.001$) and participants with higher levels of general psychological distress (PA: $B=-.04$, $p=.035$; NA: $B=.06$; $p=.028$) showed stronger affective responses. Stressor-induced cortisol increase was stronger in adolescents under the age of 18 than in participants 18 years and older ($B=-.06$, $p=.002$).

Conclusion: The digi-SPEE represents a social evaluation stressor that elicits biological and implicit and explicit mental changes that are relevant to mechanisms of psychopathology.

Keywords: Social evaluation, psychological stress, psychopathology, experimental design, cortisol, risk

INTRODUCTION

Belonging and acceptance by peers are basic human needs. At the same time, humans are vigilant to social evaluation. Currently, social interactions are occurring increasingly in online social networks, particularly among young people. These networks offer the possibility of disclosing personal information and evaluating others, including their likes/dislikes, hobbies, and personal musings via "wall posts" and "status updates." The need to belong and self-presentation are the two main reasons for participating in online social networks; research indicates that these desires are partially met (1). For example, time spent in online social networks can enhance explicit self-esteem when observing one's own profile, may foster healthy identity development in adolescents, and helps individuals with low self-esteem to build social capital (2). Thus, it may not be surprising that individuals disclose a considerable amount of sensitive personal information in their social network profiles (3). However, this may also place these individuals at risk for negative social evaluation (3). For example, there is emerging evidence that bullying over the internet has even more disastrous consequences than its face-to-face counterpart, possibly because it occurs at a greater psychological distance and is more anonymous (4). Therefore, a greater understanding of the effects of the most typical features of online social networks, namely being evaluated by peers purely on the basis of personal characteristics, is urgently required.

Laboratory studies have reliably demonstrated that social rejection/exclusion stressors, which usually involve rejection of the participant from joining any kind of group, induce significant stress and yield affective, self-esteem, and cortisol reactions (5,6). Other laboratory studies have demonstrated that performance evaluation stressors, which usually involve performance evaluations of the participant by a panel or peers, increase the stress hormone cortisol (7). Research has indicated that stress responses are different for social rejection/exclusion and



Correspondence Address: Jim van Os, Department of Psychiatry and Psychology, Maastricht University Medical Centre Postbus 616 (DRT 12), 6200 MD Maastricht, The Netherlands E-mail: j.vanos@maastrichtuniversity.nl

Received: 23.08.2016 • Accepted: 07.10.2016

©Copyright 2017 by Turkish Association of Neuropsychiatry - Available online at www.noropsikiyatriarsivi.com

performance evaluation stressors (8). Online social interactions include selective elements of both stressors, i.e., feedback from peers as social rejection/exclusion stressors, and social evaluations as performance evaluation stressors. Therefore, there is good reason to assume that a slightly negative evaluation of personal characteristics by peers imposes significant stress.

Moreover, exposure to social stressors has been associated with psychiatric symptoms (9,10,11,12); sensitivity to such stressors may be enhanced in individuals at risk for psychiatric symptoms (13,14). Adolescents represent the group that most frequently interacts in online social networks; they place increased importance on peer relationships (14) and are particularly vulnerable to developing mood symptoms (4,14,15,16). This adds urgency to tools that can accurately and sensitively assess sensitivity to negative digital social evaluations of personal characteristics by peers.

Therefore, we developed a novel digital social peer evaluation experiment (digi-SPEE). To realistically mimic online social network interactions, it is important to "deliver self-relevant, salient, and believable social feedback" (14). We therefore considered that social evaluation as experienced in online social networks includes a greater degree of psychological distance than face-to-face contact, is generally based on personal traits as plotted in the individual's online profile, and includes feedback by peers. The experiment exposes participants to subtle negative digital social peer evaluation of fundamental personal characteristics (appearance, intelligence, and stance in life).

The aim of the current study is to validate this experiment as a negative social evaluation stressor in three ways. First, the experiment was designed to elicit psychological and biological responses following exposure to subtle negative social evaluation. As psychological changes are not always directly observable or measurable when questioned explicitly, both explicit and implicit measures are required. As an explicit measure of sensitivity to negative social evaluation, we used subjective ratings of affect (11,12,17,18,19). To obtain implicit responses, we measured self-esteem implicitly using an implicit association task (20). Also, cortisol was included as a biological response measure.

Secondly, we aimed to demonstrate that the current experiment is sensitive to individual differences to social stress sensitivity. Previous research has indicated that women are more sensitive to peer-induced stress than men (11). Also, young people, i.e., adolescents, are assumed to be particularly sensitive to social stress situations (14,21). Finally, individuals with psychopathological problems have been indicated to show increased reactions to social stressors (22,23,24). Therefore, this study examines whether the current experiment is sensitive to these demographic differences.

Finally, we aimed to demonstrate that responses to the current experiment are specifically due to social evaluation exposure (i.e., the idea of being evaluated by others) and not to negative stimuli per se, as employed in the experiment.

METHODS

Participants

Participants were recruited in the context of a novel general adolescent/young adult sampling frame from a population twin register (East Flanders Prospective Twin Survey, EFPTS), which prospectively registers multiple births from 1964 onwards (25). To oversample adolescent participants, twins between 15 and 18 years of age were sent letters inviting them to participate in an ongoing longitudinal research project called "TwinsCan,"

which began in April 2010. Additionally, all twins and their siblings between 15 and 34 years of age were eligible to participate and could register via the twin register newsletter. Two hundred and forty-one subjects (58.9% female; 18.8 years) participated (see Table 1 for sample demographics). Instruments for assessing cortisol and self-esteem (SC-IAT, see below) were only available starting from August and September 2010, respectively; as a result, 202 participants (59.4% female; 18.7 years) underwent cortisol analyses and 187 participants (53.5% female; 18.2 years) underwent self-esteem analyses. Additionally, one participant was excluded due to very high (and physiologically implausible) cortisol values (>40 nmol/L). For another individual, post-social evaluation data were missing for all indices due to technical recording failure (see Table 2 for N, means, SD for all outcome measures). All siblings of the twins were invited to participate. The undeceived subsample consisted of a convenient sample of 25 siblings of the twins (76% female; 20.5 years) who agreed to participate (Table 1). Approval was obtained from the local ethics committee (Commissie Medische Ethiek van de Universitaire ziekenhuisen KULeuven, Nr. B32220107766). Participants provided written informed consent before being included in the study. In cases where the participants were younger than 18 years, their parents provided additional written informed consent.

The Digital Social Peer Evaluation Experiment (Digi-SPEE)

Participants were told that the general aim of the experiment was to investigate why people like or dislike each other based on short profiles and videos. To this end, participants were asked to provide a short written profile and to record a short video (≤ 1 min) introducing themselves, including information about their age, profession or education, preferred music, leisure activities, relationship status, and stance in life. For the same purpose, participants were asked to rate, on a 7-point rating scale (where higher scores were more positive), the profiles and videos of five other study participants regarding appearance and intelligence and congenial-

Table 1. Demographics of both samples and differences between the samples

	Deceived sample		Undeceived sample		P
	n	Mean (SD)	n	Mean (SD)	
Demographics					
% female	241	58.92	25	76.00	.10 ¹
Age (years) ²	241	18.76 (4.54)	25	20.52 (4.77)	.07 ¹
SCL-90-R	239	41.47 (35.25)	24	35.83 (28.52)	.45 ¹
% white ethnicity	239	99.17	25	100	.64 ¹
Work situation					.24 ³
% household	241	0	25	1	
% studies	241	80.91	25	68.00	
% regular work	241	1.66	25	8.00	
Highest level of completed education					.24 ³
% lower education	241	0	25	0	
% secondary education	241	55.19	25	44.00	
% bachelor's degree	241	24.90	25	32.00	
% master's degree	241	18.26	25	16.00	

n: number of subjects; SD: standard deviation

SCL-90-R: Symptom Checklist-90-Revised (higher score=more psychological distress during the last week)

¹based on independent sample t-tests

²median=18, range 15-34 years

³based on chi-square tests

ity; they were also asked to provide short reasons for their ratings. Although participants were told that they would see videos of other study participants, in fact, they were presented with prepared videos from five volunteer peers from the same population (female and male individuals between 15 and 34 years from East Flanders, all white). Depending on the ages and genders of the participants, they were confronted with different sham-peers. These peers were always age (+/-1 year)- and gender (three of the same gender and two of the opposite gender as participants)-matched to each participant. These peers did not in any way participate in the study themselves; their videos were recorded before the study began. Also, in fact, the participants' profiles and videos were never shown to anyone else; therefore, privacy was maintained throughout the procedure. Twin couples were usually tested simultaneously but separately by two trained research assistants to prevent them from discussing the experiment before social evaluation had occurred. If siblings volunteered as well, they were usually tested immediately after the twins, also to prevent them from talking to one another. The experiment consisted of two sessions which were both held in the participants' homes and were conducted by the same experimenter.

Session 1

During the first session, the participants entered their profiles into a pre-programmed computerized form and recorded their videos with a webcam. The participants also evaluated the profiles and videos of three of the five sham participants. The participants were told that their profiles and videos would be shown to the same five participants for evaluation between the sessions and that they could learn how they had been evaluated by their peers during the second session.

Session 2

The second session was also held at the participants' homes several days later; the session started approximately an hour before the social evaluation exposure occurred and was conducted by the same researcher individuals had seen the previous week to prevent elevated cortisol levels at the start of the experiment due to experienced novelty. The participants were first asked to rate the remaining two volunteers to increase the impression of a genuine ongoing interaction. Next, the participants watched their own video again to increase self-awareness. Subsequently, the actual social evaluation experiment took place. Participants learned how they had been rated by their sham peers. In fact, all participants received exactly the same evaluation, which had been previously generated by the experimenter. First, the participants would see two vertical bars, one with the heading "your evaluation" and one with the heading "average evaluation for all individuals within the study." Both bars would start to fill up simultaneously; the bar with the highest fill indicated a more positive evaluation and, therefore, a less negative evaluation. The bar of "your evaluation" would not fill further than approximately halfway, whereas the average bar filled to approximately 80%. This result was very similar for the three rated characteristics (intelligence, appearance, and congeniality). Following this, the participants were presented with the individual evaluations by the five sham peers, accompanied by a short written reason for their evaluation. Of the fifteen evaluations for appearance, intelligence, and congeniality, seven were neutral or positive, such as "seems friendly" (congeniality) or "seems to know a lot" (intelligence), and eight were mildly negative, such as "strange nose" (appearance). The reasons for the evaluations were provided in a very general manner (i.e., "funny nose") so that they would be applicable to every individual. To prevent participants from refusing to look at their evaluations in the first place, they were required to indicate their levels of agreement with the evaluation of each of the five sham peers by sorting them from 1 (agree most) to 5 (agree least). This evaluation constituted the social evaluation exposure.

The participants were debriefed on the true nature of the procedure immediately after finishing the experiment with the accompanying affective and implicit self-esteem measurements. Thus, the participants were under the impression that other participants truly evaluated them for less than 30 minutes (Figure 1).

Design

A within-subject (pre-post stressor) design was employed to assess the effects of the social evaluation experiment. Positive affect (PA), negative affect (NA), implicit self-esteem, and cortisol were assessed both before and (for cortisol: multiple times) after exposure to the social evaluation (Figure 1). In addition, the subsample of 25 siblings underwent the same procedure; however, they were fully debriefed and therefore undecieved about the nature of the evaluation before it was presented. These participants were thus exposed to the same stimuli but knew that the evaluations had been invented by the experimenter and that no other peers had actually seen their videos or evaluated their profiles.

Measures

Subjective report of positive and negative affects

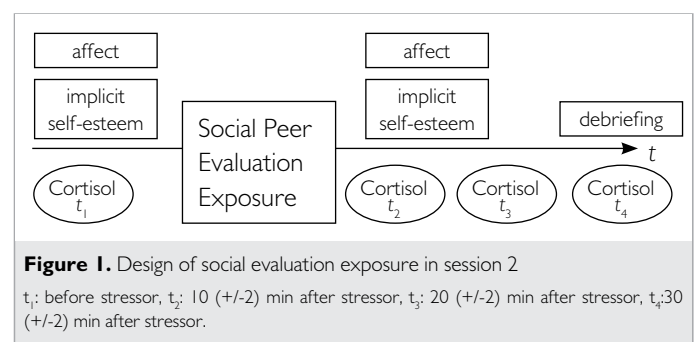
The positive and negative affect scales schedule (PANAS) was assessed using visual analogue scales (VAS) (26). The PANAS consists of 10 mo-

Table 2. Outcome measures for both samples and differences between samples

	Deceived sample		Undeceived sample		P
	n	Mean (SD)	n	Mean (SD)	
Outcome measure					
PA _{pre}	241	50.82 (13.75)	25	52.96 (13.71)	.46
PA _{post}	240	45.60 (16.67)	25	50.45 (12.92)	.093
NA _{pre}	241	10.85 (9.72)	25	9.85 (8.82)	.60
NA _{post}	240	16.76 (15.33)	25	12.53 (13.69)	.16
Self-esteem _{pre}	187	0.36 (0.35)	25	0.39 (0.35)	.69
Self-esteem _{post}	186	0.17 (0.28)	23	0.28 (.27)	.076
Cortisol t ₁	202	1.78 (0.61)	24	1.86 (0.58)	.50
Cortisol t ₂	200	1.75 (0.62)	24	1.74 (0.57)	.95
Cortisol t ₃	201	1.89 (0.67)	23	1.72 (0.56)	.24
Cortisol t ₄	201	1.90 (0.68)	23	1.78 (0.61)	.35

n: number of subjects; SD: standard deviation

Cortisol is depicted as log-transformed values. Pre: before social evaluation; post: after social evaluation; PA: positive affect (higher score=more positive affect); NA: negative affect (higher score=more negative affect); t₁:before social evaluation; t₂:10 min after social evaluation; t₃:20 min after social evaluation; t₄:30 min after social evaluation



mentary positive and 10 momentary negative affect items. For each affect characteristic (i.e., irritable, distressed, enthusiastic) the participants were presented with a horizontal line 105 mm in length with the label "not" on its left end and "very much" on its right end; the participants were required to indicate by mouse click the degree to which they were experiencing this affective state at that moment. The closer this position was to the right side of the line, the stronger the affective state in question was experienced. The standard procedure for analyzing the PANAS was followed (26). That is, the ratings for positive and negative affect items were averaged per person and per assessment to form measurements of positive affect (PA) and negative affect (NA), respectively. As affect was assessed on a visual analogue scale, values were expressed in mm; higher scores indicated higher levels of PA or NA, respectively.

Implicit measurement of self-esteem

In the single category implicit association task (SC-IAT) for self-esteem (27,28), individuals were asked to categorize 12 personalized self-words (e.g., their first name or last name) either with positive or with negative words in a reaction time task. The faster a person can categorize self-words with positive words relative to negative words, the higher that person's implicit self-esteem (i.e., $RT_{\text{self+neg words}} - RT_{\text{self+pos words}}$ in ms). In total, there are two blocks. In one block, subjects are requested to sort self-words in the same category as positive words. In the other block, subjects are requested to sort self-words together with negative words. Each block contains a total of 81 stimuli [30 self-items, 30 words in the opposite direction (i.e., negative in the positive block and positive in the negative block), and 21 in the direction of the block]. The data were prepared according to recommendations for analyses from previous literature on the SC-IAT (for more detail, see reference 28). As the SC-IAT is a reaction time task, the values are expressed in ms, with higher scores indicating higher implicit self-esteem.

Cortisol measurements

Cortisol was measured in saliva samples which were collected immediately before the computer paradigm started during session 2 and after exposure to the social rejection stress (at 10 (+/-2) min, 20 (+/-2) min, and 30 (+/-2) min) (Fig. 1). These time points were based on suggestions in the literature and on several reported experiments on cortisol assessment (29,30,31). Participants were restrained from eating and smoking for at least one hour and from drinking water at least 40 minutes before the first cortisol measure. The subjects collected saliva samples with cotton swabs (Salivette; 32) and placed the swabs in salivette tubes. The interviewer recorded the exact collection time. Uncentrifuged samples were maintained at -20°C until analysis. Saliva samples were frozen and stored at -20°C until analysis. After thawing, the salivettes were centrifuged at 3,000 rpm for 5 min, which resulted in a clear supernatant of low viscosity. Salivary concentrations were measured using a commercially available chemiluminescence immunoassay with high sensitivity (33). The intra- and interassay coefficients for cortisol were below 8%.

The cortisol values, in nmol/L, were then log-transformed to reduce the skewness of the distribution. We aimed to assess the individuals in the afternoon to avoid the influence of time of day on the cortisol measurements. However, it was not feasible to avoid variations in timing. The average time of the baseline cortisol assessment was 3:21 pm (median 3:45 pm; standard deviation 2.75 hours; range 10:45 am–9:50 pm). The average time between the baseline measurement and the final cortisol measurement was 36 minutes (range 23–46 min). The sampling times of the measurements were added as a covariate to control for diurnal changes in cortisol. For this purpose, both the variables time and time² were included in the model, as described previously in similar work (34). The addition of higher order polynomial terms did not improve the model fit.

General psychological distress

The current level of general psychological distress was measured using the Symptom Checklist-90-Revised (SCL-90-R), a 90-item self-assessment questionnaire regarding the presence of several psychiatric symptoms within the past week, including several subscales (35). It has previously been suggested that the overall scale is a good indicator for general psychological distress and transdiagnostic psychopathology (36). For the current sample, the scale had high internal consistency ($\alpha=.97$). Higher total scores indicate greater levels of general psychological distress.

Deception

After being debriefed, all individuals responded to three questions about how much they believed the evaluations were real. These scores were combined into one deception score, with lower scores indicating stronger belief in the authenticity of the evaluation.

Analyses

Multilevel regression models were employed to investigate whether affect (PA, NA), implicit self-esteem, and salivary cortisol levels changed from before to after exposure to the social rejection stressor. The current data had a hierarchical structure because multiple observations (assessments of affect, cortisol, and implicit self-esteem pre- and post-social evaluation exposure) were clustered within subjects (level 1), who were related as they were members of twin pairs or were siblings of twins (level 2).

Impact of the Digi-SPEE on Affect, Implicit Self-Esteem, and Cortisol

We first investigated whether the time of assessment (pre- or post-social evaluation exposure) predicted the level of PA, NA, or implicit self-esteem with multilevel linear regression, using the XTMIXED command in STATA (StataCorp, 2009; College Station, TX, USA). PA, NA, and implicit self-esteem were each regressed on time of assessment. For the cortisol levels, the same analyses were conducted; however, this analysis included four repeated measures per individual and was controlled for contraception and alcohol use, as described previously in similar work (34). Cortisol increase was thus expressed as the regression slope (i.e., the change) of cortisol concentration in nmol/L over time.

Differential Impact of the Digi-SPEE Depending on Age, Gender, or Level of General Psychological Distress

Age, gender, and level of general psychological distress were added to the multilevel linear regression models, each in a separate analysis, as moderators of the time effects. Interaction effects between the moderators and time (pre- and post-social evaluation exposure) on the outcome measures (PA, NA, implicit self-esteem, cortisol) were investigated and followed up by simple effects if applicable.

Manipulation of Perceived Social Evaluation

Manipulation of perceived social evaluation was assessed through two strategies. Deception scores were explored by calculating the percentage of individuals that had no doubts, some doubts, or many doubts as to whether the experiment was real or knew that the experiment was not real. Associations between deception score and SCL-90-R score, age, and gender were assessed. The second strategy was the inclusion of an undeceived subgroup. We examined whether the experiment had similar impacts on affect, implicit self-esteem, and cortisol for undeceived and deceived participants. For this purpose, unpaired sample t-tests for the pre-post evaluation changes in PA, NA, and self-esteem or interactions with the above time effects (only for cortisol) were employed.

RESULTS

Impact of the Digi-SPEE on Affect, Implicit Self-Esteem, and Cortisol

Exposure to the digi-SPEE resulted in a significant decrease in PA (N=241, B=-5.25, p<.001) and a significant increase in NA (N=241, B=5.99, p<.001) (Figure 2). The mean PA after the stressor was approximately 5 mm (on the VAS) lower than before the stressor and the mean NA after the stressor was approximately 6 mm higher than before the stressor. Second, implicit self-esteem decreased significantly from before to after exposure to the social evaluation (N=187, B=-.19; p<.001); this was indicated by the fact that the difference between positive and negative self-evaluation, as assessed in the SC-IAT, decreased by 0.19 ms (Figure 3). Finally, cortisol levels increased significantly from before to after the exposure (N=143, B=.07; p<.001) (Figure 4). The observation that cortisol concentrations did not differ between T1 and T2 (b=-.01, p=.55) suggests that the cortisol levels were not increased at the baseline. The changes in PA and NA were negatively correlated with each other (r=-.51, p<.001), the changes in PA and self-esteem were positively correlated with each other (r=.15, p=.003), and the changes in NA and self-esteem were negatively correlated with each other (r=-.15, p=.003). Changes in cortisol were also significantly associated in the expected directions with changes in psychological measures (PA: B=-.002, p=.008; NA: B=.001, p=.015; implicit self-esteem: B=-.04, p=.015).

Moderation of the Impact of the Digi-SPEE by Age, Gender, or Current General Psychological Distress

Changes in affect were significantly moderated by gender (N=241) and current general psychological distress (N=239) (Table 3). Females and participants with higher levels of psychological distress showed larger decreases in PA and larger increases in NA (simple effects for moderation by gender: PA: males: B=-2.62, p=.012; females: B=-7.09, p<.001; NA: males: B=2.73, p=.10; females: B=8.28, p<.001). Cortisol responses were significantly moderated by age (N=186), indicating that younger participants showed stronger increases in cortisol following exposure to the social evaluation (Table 3). To further examine this age effect, we dichotomized age in post hoc analyses; we found that when the participants were 18 years or older (N=137), the increase in cortisol was significantly less than that of participants younger than 18 years (N=104) (B=-.06, p=.002).

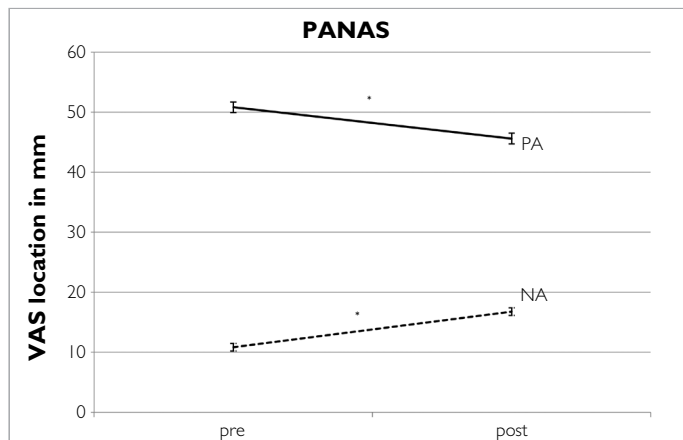


Figure 2. Change in affect
PANAS: positive and negative affect scales; VAS: visual analogue scales; PA: positive affect; NA: negative affect; Pre: before social evaluation; Post: after social evaluation; * p<.001.

Manipulation of Perceived Social Evaluation

The majority of individuals believed that the experiment was real (67.52% had no doubts at all about the authenticity of the experiment, 27.35% had few doubts, 5.13% had some doubts, and 0% had serious doubts or knew that the experiment was not real). Furthermore, the level of deception was not significantly associated with general psychological distress (SCL-90-R; β=-.09, p=.19) or age (-0.06, p=.35) but was significantly associated with gender (mean_{males}=2.89, mean_{females}=2.29, difference=.60,

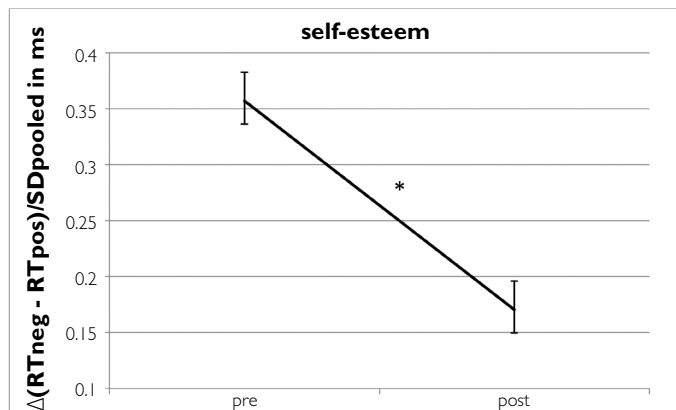


Figure 3. Change in implicit self-esteem
Pre: before social evaluation; Post: after social evaluation; * p<.001.

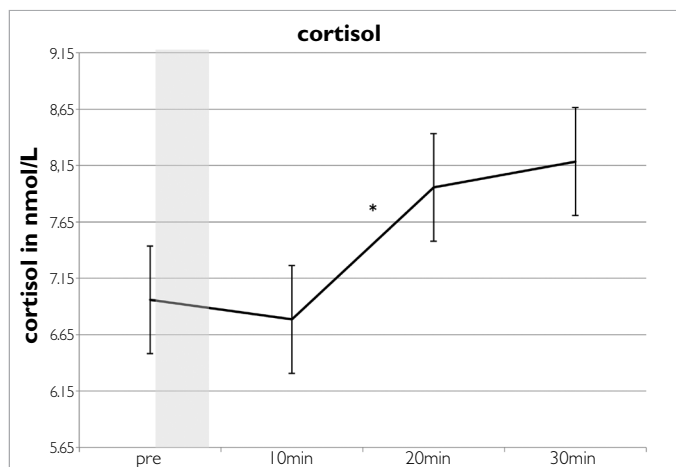


Figure 4. Change in cortisol
Pre: before social evaluation; *p<.001.

Table 3. Regression betas with p-values for the interaction between assessment time (pre and post-social evaluation) and potential moderators

	Outcome			
	PA	NA	Self-esteem	Cortisol
time x gender ¹	- 4.47** (p=.002)	5.55** (p=.001)	-0.07 (p=.28)	-0.03 (p=.16)
time x age	0.02 (p=.90)	0.00 (p=.99)	0.00 (p=.61)	-0.01** (p<.001)
time x SCL-90-R total score	-0.04* (p=.035)	0.06* (p=.018)	0.00 (p=.37)	0.00 (p=.38)

PA: positive affect; NA: negative affect; time: pre- and post-social evaluation for PA, NA, and self-esteem and pre- and 10 min, 20 min, and 30 min post-evaluation for cortisol; SCL-90-R=Symptom Checklist-90-Revised. 1 male: 0, female: 1

$p < .001$). Post hoc correction for deception in the analyses of the moderating effect of gender on the PA- and NA-measures did not change the results.

Second, in the undeceived subgroup, neither affect (PA: $B = -2.52$, $p = .14$; NA: $B = 2.68$, $p = .19$), implicit self-esteem ($B = -.10$, $p = .21$), nor cortisol ($B = -.02$, $p = .38$) significantly changed. None of these changes were significantly different when compared to the deceived sample, although the effects occurred in the expected direction, i.e., smaller than in the deceived sample (see also Table 2 for differences at separate assessment times). Controlling these analyses for age and gender (as this subgroup was slightly older and consisted of more females than the deceived sample) did not change the results.

DISCUSSION

Validation of the Digi-Spee

The current study introduced a new digital social peer evaluation experiment, the digi-SPEE that can easily be administered in individuals' homes. The current results demonstrate that the digi-SPEE elicits stress responses in measures of subjective affect, implicit self-esteem, and cortisol; these responses were particularly pronounced in groups expected to be more sensitive to social stress. The fully-computerized real-life interaction format opens a door for innovative research in the area of social peer evaluation experiences relating to emerging symptoms of mental disorders.

More specifically, the results showed that (i) the experiment induced all the expected changes in affect, implicit self-esteem, and cortisol. Furthermore, (ii) groups hypothesized to be more sensitive to social stress—women (11), young adolescents (14,21), and individuals with higher current levels of psychological distress (22,23,24)—showed stronger responses to the current social peer evaluation exposure. This suggests that the experiment is sufficiently sensitive to identify differences between differentially exposed populations. Finally, (iii) the results cautiously suggest that not only exposure to unspecific experiment characteristics but the experience of being evaluated by peers per se accounted for the experiment-induced stress responses. These findings support that the digi-SPEE may prove to be a valid tool to induce the experience of social peer evaluation.

Group Differences in Peer Evaluation-Induced Stress Responses

In accordance with previous studies (9,10,11), we found here that individuals who were female, of adolescent age, or who had more current psychological distress showed particularly pronounced effects. The results indicated that females demonstrate an increased affective sensitivity to digital social peer evaluation exposure. These findings are in line with previous research reporting that females find unpleasant stimuli to be more unpleasant than males (37,38). However, other research has showed that there are no differences between women and men for mood ratings in response to a social rejection stressor (11). Alternatively, the gender effect may be interpreted as a cultural report bias (i.e., in Western cultures, it is acceptable for women to express emotions, whereas this is less acceptable for men). Also, symptoms of mental distress were associated with a more pronounced affective reaction to the current experiment, which concurs with previous findings. For example, it has previously been reported that individuals with mood symptoms respond with a particularly pronounced increase in NA to situations of (slight) disapproval (23). Furthermore, the current results showed that participants younger than 18 years were particularly biologically reactive to the experiment, as indicated by their amplified cortisol re-

sponses. Adolescence is marked by an increased attunement to socially evaluative information and therefore involves increased vigilance to only slightly evaluative contexts (14), such as daily online social interactions. Additionally, the hypothalamic-pituitary-adrenal (HPA) axis is particularly sensitive during adolescence (24). As the HPA axis is responsible for translating social and other stressful experiences to cortisol release (22), the current findings were expected. Although we cannot conclude from the current findings whether the attunement to social evaluations, the increased HPA-axis reactivity, or both are responsible for this effect, it is valid to conclude that younger individuals react to social evaluative situations with a greater release of stress hormones. We also expected to find that cortisol levels were moderated by levels of psychological distress; however, this did not emerge from the current data. It is known that increased levels of mood symptoms are associated with both increased and decreased cortisol responses; this could be the reason for the observed heterogeneity in cortisol reactions. The direction of this effect has been suggested to depend on trait differences and the chronicity of depressive symptoms (39,22). Thus, overall, the different outcome measures were not equally affected by gender, age, or level of psychological distress. However, the current study may demonstrate that reactivity to the digi-SPEE could vary between individuals at varying levels of expected sensitivity to social evaluation.

Importantly, in the current sample, general psychological distress and age did not concur with higher levels of deception; also, controlling for deception did not change gender effects, suggesting that none of the moderator effects are attributable to increased perception of the authenticity of the experiment.

The Digi-SPEE and Future Psychopathology

Based on the current cross-sectional findings (i.e., increased reactivity in females, adolescents, and individuals with current psychological distress), we cautiously propose that the current experiment will demonstrate predictive validity for the development of mental symptoms and may contribute to elucidating genetic and environmental causes in reactivity to peer evaluation as important risk factors for psychopathology. For example, there is an increasing number of literature studies on clinical high risk samples of individuals with soft expression of psychotic symptoms who may make transitions to full-blown psychotic disorders. It is of interest to examine to what degree the use of the digi-SPEE, adding information on sensitivity to social peer evaluation, would add to the algorithm used to predict transition, in combination with information about actual social peer exposure over time. This becomes particularly important when considering the increased possibility of exposure to peer evaluation through online social network interactions in young people (15). Being able to identify individuals at increased risk to develop future symptoms of psychopathology through their sensitivity to the current experiment will ultimately enable research into resilience factors. For example, potential resilience factors that may modify the impact of social peer evaluation sensitivity on long-term outcome are warm parenting and interpretation biases in ambiguous situations (40). Understanding the risk and resilience mechanisms operating on the association between the experience of social peer evaluation and the incidence of psychopathology, particularly in adolescence, will enable the development of target-oriented prevention.

Methodological Considerations

The current findings should be viewed in light of certain methodological considerations. First, as this study particularly aimed to include participants between 15 and 18 years of age relative to older participants, the findings may not be generalized to older samples. Second, the undeceived

subgroup was considerably smaller than the deceived group. This was because reactivity to the social evaluation experiment was the variable of interest, and the number of undeceived individuals was maintained at a reasonable minimum. However, the non-significant result within the undeceived sample is unlikely the result of low power, as the effect sizes were also much smaller in this sample. Third, negative social evaluative situations may give rise to a mixture of experiences, such as feeling evaluated but also feeling rejected or excluded. The current study cannot isolate these separate effects. Future studies with this specific aim can use the current experiment for this purpose, combined with response questionnaires on evaluation, rejection, and exclusion. Finally, the current analyses regarding cortisol were controlled for age, gender, alcohol use, and use of contraception. However, some other factors may have influenced cortisol levels, such as phase of the menstrual cycle, body mass index, socio-economic status, or smoking behavior. These potential influences were not incorporated into the current analyses. To the extent that the above factors influence the within-person change in cortisol response from pre- to post-exposure, this may have affected the results, likely by inducing noise. Additionally, it should be acknowledged that there was variation in the time of day at which cortisol was sampled. However, the time between the baseline and final sampling was short (average 36 minutes), and the time of day was included as a covariate in the analyses.

Ethics Committee Approval: Ethics committee approval was received for this study from the ethics committee of Maastricht University Hospital.

Informed Consent: Written informed consent was obtained from patients who participated in this study.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept - M.C.W., R.V.W., J.V.O.; Design - M.C.W., R.V.W.; Supervision - M.C.W., R.V.W., J.V.O., M.D.H., B.P.F.R.; Resource - J.V.O., C.D., E.T., B.P.F.; Materials - M.C.W., R.V.W., J.V.O., C.D., E.T., M.D.H., N.J.; Data Collection and/or Processing - C.M.L., J.D., D.C., E.T., C.D., N.J.; Analysis and/or Interpretation - C.M.L., J.D., M.C.W., R.V.W., J.V.O.; Literature Search - C.M.L., J.D.; Writing - C.M.L., J.D., D.C.; Critical Reviews - C.M.L., J.D., M.C.W., R.V.W., J.V.O., M.D.H., N.J., E.T., C.D., D.C. B.P.F.R.

Acknowledgements: The authors thank Jill Ielegems, Katrien Lyssens, Davinia Verhoeven, and Debora op't Eijnde for data collection and all twins who volunteered for participation.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: This study was supported by the European Community's 7th Framework Program under grant agreement No. HEALTH-F2-2009-241909 (Project EU-GEL).

REFERENCES

- Nadkarni A, Hoffmann S. Why do people use facebook? Personality and Individual Differences. 2012; 52:243-249.
- Steinfeld C, Ellison NB, Lampe C. Social capital, self-esteem, and use of on-line social networks: a longitudinal analysis. Journal of Applied Developmental Psychology 2008; 29:434-445.
- Nosko A, Wood E, Molema S. All about me: disclosure in online social networking profiles: the case of FACEBOOK. Computers in Human Behavior 2010; 26:406-418.
- Donnerstein E. Internet bullying. Pediatr Clin North Am 2012; 59:623-633.
- Blackhart GC, Nelson BC, Knowles ML, Baumeister RF. Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: a meta-analytic review of 192 studies on social exclusion. Pers Soc Psychol Rev 2009; 13:269-309.
- Blackhart GC, Eckel LA, Tice DM. Salivary cortisol in response to acute social rejection and acceptance by peers. Biol Psychol 2007; 75:267-276.
- Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol Bull 2004; 130:355-391.
- Stroud LR, Foster E, Papandonatos GD, Handwerker K, Granger DA, Kivlighan KT, Niaura R. Stress response in the adolescent transition: Performance versus peer rejection stressors. Dev Psychopathol 2009; 21:47-68.
- Björkqvist K. Social defeat as a stressor in humans. Physiol Behav 2001; 73:435-442.
- Selten JP, Cantor-Graae E. Social defeat: risk factor for schizophrenia? Br J Psychiatry 2005; 187:101-102.
- Strout LR, Salovey P, Epel ES. Sex differences in stress responses: social rejection versus achievement stress. Biological Psychiatry 2002; 52:318-327.
- Williams K, Cheing C, Choi W. Cyberostracism: effects of being ignored over the internet. J Pers Soc Psychol 2000;79:748-762.
- Maner JK, Miller SL, Schmidt NB, Eckel LA. The endocrinology of exclusion: rejection elicits motivationally tuned changes in progesterone. Psychol Sci 2010; 21:581-528.
- Sommerville LH. The teenage brain: sensitivity to social evaluation. Psychol Sci 2013; 22:121-127.
- Pempek TA, Yermolayeva YA, Calvert SL. College students' social networking experiences. Journal of Applied Developmental Psychology 2009; 30:227-238.
- Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. J Abnorm Psychol 1993; 102:133-144.
- Blackhart G, Eckel LA, Tice DM. Salivary cortisol in response to acute social rejection and acceptance by peers. Biol Psychol 2007;75:267-276.
- Buckley KE, Winkel RE, Leary MR. Reactions to acceptance and rejection: effects of level and sequence of relational evaluation. Journal of Experimental Social Psychology 2004;40:14-28.
- Nesdale D, Lambert A. Effects of experimentally manipulated peer rejection in children's negative affect, self-esteem, and maladaptive social behavior. International Journal of Behavioral Development 2007; 31:115-122.
- Bosson JK, Swann WB, Pennebaker JW. Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? J Pers Soc Psychol 2000; 79:631-643.
- Collins JK, Thomas NT. Age and susceptibility to same sex peer pressure. British Journal of Educational Psychology 1972; 42:83-85.
- Booij SH, Bouma EMC, de Jonge P, Ormel J, Oldehinkel AJ. Chronicity of depressive problems and the cortisol response to psychosocial stress in adolescents: The TRAILS study. Psychoneuroendocrinology 2012.
- Cuellar AK, Johnson SL. Depressive symptoms and affective reactivity to maternal praise and criticism. Journal of Social and Clinical Psychology 2009; 28:1173-1194.
- Hankin BJ, Badanes LS, Abela JRZ, Watamura SE. Hypothalamic-pituitary-adrenal axis dysregulation in dysphoric children and adolescents: cortisol reactivity to psychosocial stress from preschool through middle adolescence. Biol Psychiatry 2010; 68:484-490.
- Derom C, Thiery E, Peeters H, Vlietinck R, Defoort P, Fijns JP. The East Flanders prospective twin survey (EFPTS). An actual perception. Twin Res Hum Genet 2013; 16:58-63.
- 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; 54:1063-1070.
- Karpinsky A, Steinman RB. The single category implicit association test as a measure of implicit social cognition. J Pers Soc Psychol 2006; 91:16-32.
- Greenwald AG, Farnham SD. Using the implicit association test to measure self-esteem and self-concept. J Pers Soc Psychol 2000; 79:1022-1038.
- Brooks KP, Robles TF. Recent depressive and anxious symptoms predict cortisol responses to stress in men. Psychoneuroendocrinology 2009; 34:1041-1049.
- Nicolson NA. Measurement of cortisol. In: Luecken LJ, Gallo LC, editors. Handbook of Physiological Research Methods in Health Psychology. Thousand Oaks, CA.; Sage Publications; 2007. p. 37-74.
- Young EA, Nolen-Hoeksema S. Effect of ruminations on the saliva cortisol response to a social stressor. Psychoneuroendocrinology 2001; 26:319-329.

32. Sarstedt. Salivette. Hygienic saliva collection for diagnostics and monitoring.
33. IBL. Cortisol Luminescence immunoassay - Luminescence immunoassay for the in-vitro diagnostic quantitative determination of cortisol un human saliva and serum. In: GmbH II, editor. Hamburg 2009.
34. Collip D, Nicolson NA, Lardinois M, Lataster T, van Os J, Myin-Germeys I. Daily cortisol, stress reactivity and psychotic experiences in individuals at above average genetic risk for psychosis. *Psychol Med* 2011; 41:2305-2315.
35. Arrindell WA, Ettema J, M. SCL-90. Handleiding bij een multidimensionele psychopathologie-indicator: Swets & Zeitlinger; 1986.
36. Bufka LF, Camp N. Brief measures for screening and measuring mental health outcome. In: Antony MM, Barlow DH, editors. *Handbook of Assessment and Treatment Planning for Planning for Psychological Disorders*. 2. New York, London: The Guilford Press; 2010. p. 62-94.
37. Bianchin M, Angrilli A. Gender differences in emotional responses: a psychophysiological study. *Physiol Behav* 2012; 105:925-932.
38. Codispoti M, Surcinelli P, Baldaro B. Watching emotional movies: affective reactions and gender differences. *Int J Psychophysiol* 2008; 69:90-95.
39. Tops M, Riese H, Oldenhinkel AJ, Rijdsdijk FV, Ormel J. Rejection sensitivity relates to hypocortisolism and depressed mood state in young women. *Psychoneuroendocrinology* 2008; 33:551-559.
40. MacLeod C, Cohen L. Anxiety and the interpretation of ambiguity: a text comprehension study. *J Abnorm Psychol* 1993; 102:238-247.