ORIGINAL ARTICLE

The Relationship of Aversive and Appetitive Appearance-Related Comparisons with Depression, Well-Being, and Self-esteem: A Response Surface Analysis

Pascal Schlechter[1] **· Sarah Katenhusen**[2] **· Nexhmedin Morina**[2]

Abstract

Purpose Aversive appearance-related comparisons (i.e., threatening one's own motives) show stronger associations with depression, psychological well-being, and self-esteem than appetitive comparisons (i.e., consonant with or challenging one's motives). However, the relevance of their congruent (i.e., equal) and incongruent (i.e., unequal) presence remains unknown.
Methods By using response surface analysis, we investigated differential associations of congruent high levels of aversive and appetitive comparisons with depression, well-being, and self-esteem relative to incongruent high levels of aversive (or appetitive) comparisons. Participants ($N = 1112$) responded to measures of depression, psychological well-being, self-esteem, and the Comparison Standards Scale for Appearance. The latter assesses aversive and appetitive social, temporal, counterfactual, criteria-based, and dimensional comparisons regarding their frequency, discrepancy to the standard, and affective impact.
Results Results confirmed our preregistered hypotheses. First, higher levels of congruent frequency, discrepancy, or affective impact were associated with higher depression, and lower well-being and self-esteem. Second, a greater predominance of aversive over appetitive comparisons was associated with higher depression, and lower well-being and self-esteem. Third, a predominance of appetitive over aversive comparison was associated with lower depression, and higher well-being and self-esteem.
Conclusions The distinct patterns of the (in-)congruence of aversive and appetitive comparisons have important research and clinical implications.

Keywords Comparison theory · Appearance-related comparisons · Depression · Well-being · Response surface analysis

Introduction

Appearance represents an important self-attribute in the lives of many individuals (Grogan, 2006; Lawler & Nixon, 2011; Quittkat et al., 2019). Body dissatisfaction is associated with depression and low psychological well-being and self-esteem (Barnes et al., 2020; Quittkat et al., 2019; Stice et al., 2000). Individuals frequently evaluate their current appearance with different comparison types. Most research in appearance-related comparisons has focused on *social comparison* (McCarthy & Morina, 2020), the process of comparing one's current appearance with someone else's (Corcoran et al., 2011; Crusius et al., 2022; Festinger, 1954; Mussweiler, 2014; Unkelbach et al., 2023). This research has focused on the overall social comparison tendency or on differential effects of upward (i.e., better looking) versus downward (i.e., less good looking) appearance-related social comparisons on outcomes like depression, well-being, or self-esteem (Buunk & Gibbons, 2006; Laker & Waller, 2020; Schaefer & Thompson, 2018). Current findings suggest that upward social comparisons in particular are centrally involved in appearance evaluations, body image disturbance, and eating pathology (Hill & Nolan, 2021; Laker & Waller, 2020; Myers & Crowther, 2009). With respect to the impact of the frequency of downward social comparisons on mental health, mixed results have been reported, ranging from positive (Steers et al., 2014), no effects (Butzer & Kuiper, 2006) to negative effects (Lup et al., 2015).

✉ Pascal Schlechter
ps798@medschl.cam.ac.uk

[1] Department of Psychiatry, University of Cambridge, Cambridge, UK

[2] Institute of Psychology, University of Münster, Münster, Germany

Comparison Types and Their Motivational Significance

Beyond social comparison, it is important to consider the role of other appearance-based comparison types in self-evaluation (Morina, 2021). One's current appearance is also often compared with a recollection of how one has looked at a certain time in the past or with an image of how one might look in the future (*temporal comparisons*; Albert, 1977), with that of a hypothetical self that might have occurred but did not actually occur (*counterfactual comparisons*; Kahneman & Miller, 1986; Woltin & Epstude, 2023), against aspirations or certain norms (*criteria-based comparisons*; Higgins, 1996; Lewin, 1951), or with some other personal attribute (*dimensional comparisons*; Möller & Marsh, 2013). The direction of these comparisons can be upward (e.g., perceiving myself as less good looking than my co-worker), lateral (e.g., perceiving oneself as adequately looking for one's age), and downward (e.g., thinking that if I had not regularly engaged in sports, my appearance would now be worse). According to the general comparative-processing model (gComp), comparing to these five types of standard can be perceived as aversive or appetitive, depending on whether the comparison outcome is processed as threatening or corresponding to the comparer's needs and goals (Morina, 2021). Factor analyses confirmed that these five comparison types share important parallels and can be conceptually summarized into the two factors aversive and appetitive comparisons (Morina et al., 2023). In particular, *upward* social, past temporal, counterfactual, and criteria-based comparisons, and *downward* prospective temporal comparisons constitute aversive comparison (Morina et al., 2023). *Downward* social, past temporal, counterfactual, criteria-based, and dimensional comparisons, and *upward* prospective temporal and dimensional comparisons constitute appetitive comparison. Therefore, conceptualizing aversive and appetitive comparison as distinct factors is particularly useful for the examination of appearance-related comparison as they encompass different comparison standards that form a comprehensive construal of one's appearance.

The Comparison Process and Mental Health Outcomes

Aversive and appetitive comparisons are operationalized as a process consisting of (and not limited to) the comparison standard selection (e.g., social or counterfactual), the evaluation of (dis-)similarities between the target and the standard that produces the comparison outcome (i.e.,

the perceived discrepancy), and the engendered affective impact (Morina, 2021). A recent examination of aversive and appetitive appearance-related comparisons revealed that aversive comparisons show stronger associations with depression, well-being and self-esteem than appetitive comparisons (Morina et al., 2023). In particular, a higher comparison frequency and perceived discrepancy between the target and the standard after engaging in aversive comparisons were strongly associated with depression, well-being, and self-esteem (Morina et al., 2023), whereas perceived discrepancy after engaging in appetitive comparisons was not significantly associated with depression and weakly positively associated with self-esteem and psychological well-being. Furthermore, higher levels of negative affective impact upon engagement in aversive and appetitive appearance-related comparisons were also associated with higher depressive symptoms, and lower psychological well-being and self-esteem (Morina et al., 2023).

The Proportion of Engaging in Aversive Relative to Appetitive Comparisons

However, the proportion of engaging in aversive relative to appetitive comparisons differs between individuals and their congruent (i.e., equal) and incongruent (i.e., unequal) presence may be differentially associated with mental health and other self-attributes. While such associations have not yet been examined, theoretically, a general tendency to engage equally in aversive and appetitive comparisons may be differently related to mental health and other outcomes than the dominance of one comparison type over the other (i.e., predominance of aversive but not appetitive comparisons and vice versa). Furthermore, this may differ for the *frequency*, perceived *discrepancy* and engendered *affective impact* after engaging in appearance-related comparisons. In this study, we therefore employed polynomial regression models and Response Surface Analysis (RSA; Shanock et al., 2010), to examine the association of congruent and incongruent combinations of aversive and appetitive comparisons with depression, psychological well-being, and self-esteem. The proportion of engaging in both aversive and appetitive comparisons bears relevance for mental health outcomes (Buunk & Gibbons, 2006). For instance, a general high *frequency* of both aversive and appetitive comparisons may reflect a general comparison tendency that is associated with higher depression, as well as lower psychological well-being and self-esteem (Buunk & Gibbons, 2006; Laker & Waller, 2020; Schaefer & Thompson, 2018). When individuals engage frequently in both aversive and appetitive comparison, aversive comparison may have a strong impact by informing the comparer that threat or harm to

their motives has taken place (Bowen et al., 2018). Moreover, against the backdrop of frequent aversive comparisons, appetitive comparisons may be deliberately employed to counteract negative outcomes of aversive comparisons. For example, a comparer may first appraise a colleague as better looking than themselves and feel bad as a result. Thereupon they might decide to engage in downward social comparison to ameliorate the negative engendered affective impact of the initial comparison. In gComp this represents a tertiary comparison, which serves the self-enhancement motive and aims at adjusting the consequences of a prior comparison, and differs from primary or secondary comparisons, with the latter two rather serving self-assessment or self-improvement motives (Morina, 2021). This may also explain why large *discrepancy* between the target and the standard after engaging in both aversive and appetitive comparisons may be associated with higher depression, lower well-being, and lower self-esteem. Large discrepancy after engaging in aversive comparisons indicates an unattainable end goal ("I will never look like my colleague") and may negatively influence any attempt to move towards the desired outcome (i.e., looking better). Large discrepancy after tertiary appetitive comparisons, on the other hand, mainly aims at maintaining a positive self-view for the sake of improving current affect. Again, this would suggest that in the case of large congruent discrepancy of aversive and appetitive comparisons, aversive comparisons are the main underlying factor that negatively influences behavior and mental health. Concerning the *engendered affective impact* after engaging in comparisons, different patterns may emerge for high levels of negative affective impact after engaging in aversive comparison and simultaneously high levels of positive affective impact after engaging in appetitive comparison. While the emotion context insensitivity theory (Bylsma, 2021) posits that depressed individuals show lower affective reactivity to stimuli in general, a study found an increased affective reaction to comparisons in depressed individuals (Giordano et al., 2000). Still other studies yielded no relationship of affect intensity with depression (Thompson et al., 2011). Thus, strong affective reaction to both aversive and appetitive comparisons may be predominated by the negative affective impact after engaging in aversive comparisons due to a greater salience of negative emotions compared with positive emotions (Bowen et al., 2018). Alternatively, high levels of positive affective impact may also counteract the emotional impact in general.

Predominantly engaging in aversive appearance-related comparisons along with low levels of appetitive comparison *frequency* should be related to higher depression, and lower well-being and self-esteem (Schaefer & Thompson, 2018; Zimmer-Gembeck et al., 2021). In such combinations, individuals have to deal with aversive comparison outcomes perceived as a threat to their core motives (Morina, 2021;

Sedikides & Strube, 1997). This way, individuals engage in frequent negative evaluations of their appearance. In this case, a comparer may constantly engage in comparison to physically more attractive celebrities or counterfactual scenarios in which they would be more attractive, if they had undergone fitness training. Moreover, depressed individuals engage more often in aversive comparison, which may reinforce the relationship between this incongruent comparison tendency and mental health outcomes (Appel et al., 2015). For instance, while being depressed, individuals may engage in selective social comparison, wherein they compare themselves primarily to individuals with better physical appearance, or selectively retrieve memories of their own physical features that were more appealing in the past. Perceiving high *discrepancy* after engaging in aversive comparison, but not after engaging in appetitive comparison is likely to signal threat to significant motives and thus also be related to higher depression, lower well-being, and lower self-esteem (Schaefer & Thompson, 2018; Zimmer-Gembeck et al., 2021). Last, perceiving high negative *affective reactions* after engaging in aversive appearance-related comparisons in conjunction with limited affective reactions after engaging in appetitive comparisons should also be related to higher depression and lower psychological well-being and self-esteem (Schaefer & Thompson, 2018; Zimmer-Gembeck et al., 2021). This may reflect dysfunctional emotion regulation strategies and a dominance of negative affective reactions to comparisons. On the other hand, a preponderance of appetitive appearance-related comparisons over aversive comparisons may lead to beneficial psychological outcomes. In these scenarios, individuals frequently conclude that they are more attractive than others or certain norms. Comparing predominantly to appetitive standards, perceiving a greater discrepancy and experiencing a more positive affective impact may thus improve one's well-being or self-esteem (Morina, 2021). However, this may rather apply to appetitive comparisons employed to satisfy self-assessment and self-improvement motives, which in gComp are defined as primary and secondary comparisons (Morina, 2021). As indicated above, appetitive comparisons may also serve to protect the self and may thus point to fragile self-esteem and a dysfunctional strategy. It has been posited that individuals with low self-esteem mainly engage in appetitive comparison as a means to protect the self (Wills, 1981). In these cases, individuals may selectively engage in appetitive comparison with limited utility as they are cognizant of their deliberate tendency to compare themselves to inferior individuals or standards.

The Present Study

In the present study, we examined the ratios of aversive and appetitive comparisons to discern the differential impact of the proportion of aversive and appetitive comparisons

Table 1 Overview of the preregistered hypotheses

Construct	Hypothesis
Depression	
Frequency-Congruency	H1.1a: Congruent high frequency of aversive and appetitive comparison is associated with higher depression.
Frequency-Incongruency	H1.1b: Greater preponderance of frequency of aversive comparisons over appetitive comparisons is associated with higher depression.
Discrepancy-Congruency	H1.2a: Congruent high discrepancy between the target and the standard is associated with higher depression.
Discrepancy-Incongruency	H1.2b: Greater preponderance of discrepancy between the target and the standard after engaging in aversive comparisons over discrepancy after appetitive comparisons is associated with higher depression.
Affect-Congruency*	H1.3a: Congruent high affective impact after both engaging in aversive and appetitive comparisons and its effects on depression will be analyzed exploratorily.
Affect-Incongruency*	H1.3b: Greater preponderance of engendered negative affective impact after engaging in aversive comparisons over positive affective impact after engaging in appetitive comparisons is associated with higher depression.
Psychological well-being	
Frequency-Congruency	H2.1a: Congruent high frequency of aversive and appetitive comparisons is associated with lower well-being.
Frequency-Incongruency	H2.1b: Greater preponderance of the frequency of aversive comparisons over appetitive comparisons is associated with lower well-being.
Discrepancy-Congruency	H2.2a: Congruent high discrepancy between the target and the standard is associated with lower psychological well-being.
Discrepancy-Incongruency	H2.2b: Greater preponderance of discrepancy between the target and the standard after engaging in aversive comparisons over a perceived discrepancy after appetitive comparisons is associated with lower well-being.
Affect-Congruency*	H2.3a: Congruent high affective impact after both engaging in aversive and appetitive comparisons and its effects on psychological well-being will be analyzed exploratorily.
Affect-Incongruency*	H2.3b: Greater preponderance of engendered negative affective impact after engaging in aversive comparisons over positive affective impact after engaging in appetitive comparisons is associated with lower psychological well-being.
Self-esteem	
Frequency-Congruency	H3.1a: Congruent high frequency of aversive and appetitive comparisons is associated with lower self-esteem.
Frequency-Incongruency	H3.1b: Greater preponderance of the frequency of aversive comparisons over appetitive comparisons is associated with lower self-esteem.
Discrepancy-Congruency	H3.2a: Congruent high discrepancy between the target and the standard after engaging in aversive and appetitive comparisons is associated with lower self-esteem.
Discrepancy-Incongruency	H3.2b: Greater preponderance of discrepancy between the target and the standard after engaging in aversive comparisons over discrepancy after appetitive comparisons is associated with lower self-esteem.
Affect-Congruency*	H3.3a: Congruent high affective impact after both engaging in aversive and appetitive comparisons and its association with self-esteem will be analyzed exploratorily.
Affect-Incongruency*	H3.3b: Greater preponderance of engendered negative affective impact after engaging in aversive comparisons over positive affective impact after engaging in appetitive comparisons is associated with lower self-esteem.

*These hypotheses deviate from the preregistration because of the way affective impact was conceptualized

in terms of their congruence (aversive = appetitive) and incongruence (aversive > appetitive or appetitive > aversive) in relation to depression, psychological well-being, and self-esteem. These outcomes present mental health variables with high public health and clinical relevance (Liu et al., 2020; Orth et al., 2018; Trudel-Fitzgerald et al., 2019). To disentangle nuanced effects resulting from the different combinations of aversive and appetitive comparisons, we used polynomial regression models and RSA (Edwards, 2002; Shanock et al., 2010) to test our preregistered hypotheses, which are outlined in Table 1. This approach enables granulated analyses of the effects of predictor combinations on our

mental health outcomes beyond other statistical approaches such as difference scores (Schönbrodt, 2016). We expected two general patterns. First, we hypothesized that congruent high comparison frequency and discrepancy reflect a general comparison tendency that is associated with higher depressive symptoms, and lower psychological well-being and self-esteem (Laker & Waller, 2020; McCarthy & Morina, 2020; Schaefer & Thompson, 2018). Congruent affective impact was analyzed exploratorily. Second, when looking at the incongruent relations, we expected that predominantly aversive comparison frequency, discrepancy, and affective impact along with low levels of appetitive comparison

frequency, discrepancy, and affective impact would affect depression, well-being and self-esteem negatively (Schaefer & Thompson, 2018; Zimmer-Gembeck et al., 2021, see Table 1). The discordant relationship of a preponderance of the frequency, discrepancy, and affective impact of appetitive appearance-related comparisons over aversive comparisons was tested exploratorily for all hypotheses [1].

Methods

Openness and Transparency

All data and the R code to reproduce the current results are provided in the open science framework (https://osf.io/6us2n/?view_only=2f16cc32a93e4ba7b0fd2997d40b0d6c). We also provide the survey material that has been used. The present study is part of a larger project examining cognitive and social variables in relation to mental health outcomes. One manuscript of this project has been published on the development of the Comparison Standards Scale for Appearance (Morina et al., 2023), which we use in the present contribution to assess appearance-related comparisons. The initial project was not preregistered, however the current study was preregistered (https://osf.io/xam4j/) following suggestions for preregistrations of secondary data analyses (Akker et al., 2021). Hypotheses concerning the affective impact after engaging in comparison differ from the preregistration as explained below and indicated by an * (see Table 1).

Participants and Procedure

We conducted secondary data analysis of a study that recruited $N = 1121$ participants (Morina et al., 2023) from online panel provider Prolific Researcher (Palan & Schitter, 2018). The survey was open to all panel members who had indicated to be fluent in English and were older than 17 years. The study was approved by the Ethics Committee of the University of Münster. Because of missing data, we had to exclude nine participants, resulting in a final number of $N = 1112$ participants that were on average 28.7 ($SD = 9.7$) years old. Table 2 depicts the sociodemographic details of the entire sample and for women ($n = 479$) and men ($n = 621$). Overall, 12 participants did not identify as women or men.

Measures

The Comparison Standards Scale for Appearance (CSS-A) was used to assess engagement in appearance-related upward and downward comparisons via social, temporal, counter-factual, criteria-based, and dimensional standards (Morina et al., 2023). The scale comprises three subscales. It starts with (a) 16 obligatory items on the frequency of comparisons in the past three weeks using a six-point Likert scale (0 = *not at all* to 5 = *very often*). Then, there are two elective subscales comprising (b) 16 items addressing the perceived discrepancy between the target and the comparison standard on a six-point Likert scale (0 = *not at all* to 5 = *much better/worse*), and (c) 16 items addressing the engendered affective impact on a bipolar seven-point Likert scale for affective impact (−3 = *much worse* to +3 = *much better*). Participants only responded to the sub-items (b) and (c) of the respective item when they reported to have engaged in this comparison type. For example, the upward social comparison item first asks about the frequency: "*Over the past three weeks when considering your appearance, how often have you compared with others in your close circles who look better than you?*". If participants indicate more than "0 – not at all", they are asked "*How much better have you considered their appearance to be?*" (i.e., discrepancy assessment) and "*On average during the past three weeks, how did the comparison make you feel?*" (i.e., affect assessment). Factor analysis indicated that *upward* social, past temporal, counterfactual, and criteria-based comparisons, and *downward* prospective temporal comparisons build one factor of *aversive* comparisons. *Downward* social, past temporal, counterfactual, criteria-based, and dimensional comparisons, and *upward* prospective temporal and dimensional comparisons can be defined as *appetitive* comparisons. However, deviating from the preregistration, data on the affective impact scale were adjusted to enhance interpretation. Note that the affective impact variable asks whether the comparer felt any negative or positive affective change following a specific comparison on a scale of -3 to +3, with −3 and +3 best represented as equivalent on strength of affect. Since some participants reported positive (or negative) affective change following aversive (or appetitive) comparisons, we needed to adjust the data generated with the affective impact subscale given that negative scores cannot be treated as being lower than positive scores and vice versa. As a consequence, negative affect scores following appetitive comparisons and positive affect following aversive comparisons were set to zero because they deviated in an unexpected direction and would confound the evaluation of the strength of the comparison affective impact. Accordingly, the hypotheses concerning engendered affective impact after engaging in comparisons differ from the preregistration (marked with an *). Internal consistencies were α = 0.73 for aversive comparison

---

[1] We may find a linear effect or a curvilinear effect for the line of incongruence.

Table 2 Sociodemographic variables

Variable	Entire sample		Women		Men	
	n	%	n	%	n	%
Education						
Less than high school degree	13	1.2	3	0.6	9	1.4
High school degree or equivalent	267	24.0	106	22.1	158	25.4
Some college but no degree	227	20.4	94	19.6	129	20.8
Associate degree	51	4.6	16	3.3	34	5.5
Bachelor's degree	354	31.8	165	34.4	187	30.1
Graduate degree	200	18.0	95	19.8	104	16.7
Marital Status						
Married/Domestic Partnership	343	30.8	150	31.3	191	30.8
Widowed	2	0.2	2	0.4	0	0.0
Divorced	10	0.9	6	1.3	4	0.6
Separated	12	1.1	7	1.5	5	0.8
Single/never married	745	67.0	314	65.6	421	67.8

frequency, α = 0.62 for aversive comparison discrepancy, and α = 0.81 for aversive comparison affective impact and α = 0.73 for appetitive comparison frequency, α = 0.70 for appetitive comparison discrepancy, and α = 0.71 for appetitive comparison affective impact.

The Patient Health Questionnaire (PHQ-8; Kroenke et al., 2009) was used for the assessment of depressive symptoms over the last two weeks. Depressive symptom severity (e.g., "*Feeling tired or having little energy*") was assessed on a 4-point scale (0 = *not at all* to 3 = *nearly every day*). Cronbach's alpha was 0.88 in the present study.

Psychological well-being was assessed with the eighteen-item Scale for Psychological Well-being (SPWB; Ryff & Keyes, 1995), covering six areas of psychological well-being: autonomy, self-acceptance, environmental mastery, personal growth, positive relations with others, and purpose in life. Items (e.g., "*In general, I feel I am in charge of the situation in which I live*") were assessed on a 6-point scale (1 = *strongly disagree* to 6 = *strongly agree*). In the present analyses, Cronbach's alpha was 0.85.

Self-esteem was assessed with the Rosenberg Self-Esteem Scale (RSES, Rosenberg, 1965) consisting of ten items. Items (e.g., "*On the whole, I am satisfied with myself*") are rated on a 4-point Likert scale (0 = *strongly disagree* to 3 = *strongly agree*). Cronbach's alpha in the current study was 0.91.

Analysis Procedure

Tested Models

Analyses were performed in R version 4.2.1 (R. Core Team, 2021). RSA analyses were conducted with the RSA package (Schönbrodt & Humberg, 2021). To test the effect of the proportion of aversive and appetitive comparisons on our outcome variables, we used a polynomial regression model including the two types of comparisons (i.e., aversive comparisons, X, and appetitive comparisons, Y), their squared terms, and their interaction to predict an outcome measure (Z) (e.g., depression), where ϵ denotes an error term:

$$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + \epsilon$$

Predictors were grand-mean centered before they were included in the analyses, while we used sum scores for the outcome variables. In total, we computed nine different models to test our hypotheses. First, we ran one model in which we investigated the *frequency* of aversive and appetitive comparisons as predictor variables X and Y with depression as criterion variable Z (Model 1). Second, a similar model was run with the perceived *discrepancy* after engaging in aversive (X) and appetitive (Y) comparisons as predictor variables and depression as criterion variable (Z) (Model 2).

Third, we ran a model with the engendered *affective impact* after engaging in aversive (X) and appetitive (Y) comparisons as predictor variables and depression as criterion variable (Z) (Model 3). These three models were tested again with the same predictor variables but different criterion variables, namely psychological well-being (Models 4–6) and self-esteem (Models 7–9), leading to nine models.

Response surface methodology was used to visualize the models in a three-dimensional coordinate system. To understand the effects of the proportion of aversive and appetitive comparisons better, the line of congruence (LOC), and the line of incongruence (LIOC) can be interpreted. The LOC reflects predicted outcome scores (Z) for levels of aversive comparisons (X) that are perfectly congruent to the levels of appetitive comparisons (Y) ($X = Y$). The relationship along the LOC can be understood with the following response surface parameters derived from the polynomial regression model: $a1 = b_1 + b_2$ and $a2 = b_3 + b_4 + b_5$. When $a2$ is not significant, a significant positive $a1$ indicates a linear effect of the LOC in that depression is higher when aversive and appetitive comparisons are congruent at a higher level than when they are congruent at a lower level. A significant negative $a2$ value signifies a curvilinear slope of the LOC, suggesting higher depression when aversive and appetitive comparison ratings correspond at medium levels rather than at extreme levels. When $a1$ is simultaneously significant, it signifies the position of the extreme value, when it is not significant the extreme value is at (0/0). The LOIC is orthogonal to the LOC and describes depression severity for ratings that are equal in magnitude but opposite in sign ($X = -Y$). Again, derived response surface parameters aid interpretation, $a3 = b_1 - b_2$ and $a4 = b_3 - b_4 + b_5$. When $a4$ is not significant, a significant positive a3 indicates that depression is higher when the discrepancy is such that X (aversive comparisons) is higher than Y (appetitive comparisons). A significant a4 indicates a curvilinear slope.[2] When $a3$ is simultaneously significant, it signifies the position of the extreme value, when it is not significant the extreme value is at (0/0). Note that we only derived linear hypotheses based on prior literature and tested curvilinear effects exploratorily.

---

[2] Further, the first principal axis (FPA) describes the ridge of the surface. The FPA is the line on the surface with maximal upward (or minimal downward) curvature (Edwards, 2002). One can test whether the ridge equals the LOC with $a5 = b_3 - b_5$ to determine which combinations of aversive and appetitive comparisons lead to predicted depressive symptoms on the FPA. If the FPA equals the LOC, the combinations of aversive and appetitive comparisons that lead to predicted depressive symptoms on the ridge are exactly the congruent combinations.

Table 3 Descriptive statistics and Pearson correlation coefficients of all variables

Variable	2.	3.	4.	5.	6.	7.	8.	9.	*M*	*SD*
1. Av. comp. frequency	0.28***	0.83***	0.24***	0.70***	− 0.01	0.48***	− 0.39***	− 0.48***	2.15	1.00
2. App. comp. frequency	–	0.29***	0.84***	− 0.12***	0.54***	0.04	0.06*	0.12***	1.58	0.80
3. Av. comp. discrepancy		–	0.34***	0.52***	0.07*	0.35***	− 0.32***	− 0.37***	2.09	0.85
4. App. comp. discrepancy			–	− 0.17***	0.49***	0.01	0.05	0.13***	1.44	0.73
5. Av. comp. affect				–	− 0.17***	0.52***	− 0.37***	− 0.53***	0.73	0.66
6. App. comp. affect					–	− 0.11***	0.24***	0.24***	0.37	0.40
7. Depressive symptoms						–	− 0.53***	− 0.63***	16.53	5.89
8. Psychological well-being							–	0.78***	85.77	16.32
9. Self-esteem								–	27.43	6.44

Av. aversive, *app.* appetitive, *comp.* comparisons

Prerequisites for the RSA

For RSA, the sample sizes should be around three times higher as required for simple regression models (Humberg et al., 2019). Despite not conducting an a priori power analyses for our secondary data analysis, our sample size appears sufficiently large for stable parameter estimation. To conduct RSA, scales measuring the predictors need to be commensurable (Humberg et al., 2019), which is achieved in the present analysis because appetitive and aversive comparisons were assessed with the same instrument, the CSS-A. Moreover, for a valid interpretation of RSA parameters, it is important to establish that there are enough data pairs for each combination of (in)congruent predictors. There needs to be enough data for cases where aversive comparison frequency exceeds appetitive comparison frequency and vice versa, and where the frequency of both is equal. The same accounts for the other predictor combinations (i.e., discrepancy and affective impact). In addition, following recommendations by Rodrigues (2021), we excluded outliers that were below the 0.25 percentile and above the 0.75 percentile. However, models including outliers were tested as sensitivity analyses. For the final model, we tested the assumptions of the independence of residuals with the Durbin-Watson-Test, deviations from the normal distribution with the Shapiro-Wilk-Test, deviations from the assumption of homoscedasticity with the Breusch-Pagan-Test, and multicollinearity with the variance inflation factor (VIF). The former three tests should be non-significant, while the VIF should be below 5 to satisfy these conditions. Confidence intervals and p-values for all parameters were estimated with bootstrap resampling procedures with 10,000 iterations (Humberg & Grund, 2021).

Model Selection

The full polynomial model is computationally complex, and thus prone to overfitting. Therefore, we tested the full polynomial model against other models that have been suggested in literature (Schönbrodt, 2016). These models are nested within the full model but have fewer degrees of freedom. Different models may be suited to explain the data and need to be tested against each other based on different criteria. First, we evaluated the (corrected) Akaike Information Criterion (AIC), which balances model fit and complexity, where lower values indicate better fit when taking model complexity into account. In addition, we calculated the Akaike weight, which is the probability that the model is the best fitting model based on the AIC of all tested models. The difference (Δ) of the AIC of two models can also be used for model selection. When the difference is below 2 models are equivalent, an ΔAIC > 10 indicates significant differences between the models (Schönbrodt, 2016). The evidence ratio was also used, which is the ratio of the Akaike weight in comparison to the best fitting model, indicating how much more likely the best model is compared to the respective model. A further consideration was the explained variance of the model (R^2). Moreover, the Comparative Fit Index (CFI) should be over 0.95 to indicate good model fit.

Results

Descriptive statistics and Pearson correlation coefficients of all scales that have been used can be found in Table 3. For all models, there were sufficient data for each predictor combination (Supplemental Table S1). Depending on the predictors, between 6 and 45 outliers were excluded, while the variables depression, well-being, and self-esteem had no outliers (Supplemental Table S1). In six out of the nine tested models, the full polynomial model was the best fitting model according to our criteria. In the remaining cases, it was among the best models and could not be rejected compared to the other models based on model selection criteria. Therefore, we continued with the full

Fig. 1 Response Surface Plots for aversive and appetitive comparisons as predictor variables and depressive symptoms as outcome variable (Models 1–3). Predictor variables were grand-mean centered. Affective impact indicates positive affective impact after engaging in appetitive comparisons and negative affective impact after engaging in aversive comparisons. Models are based on the full polynomial model $Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + \varepsilon$
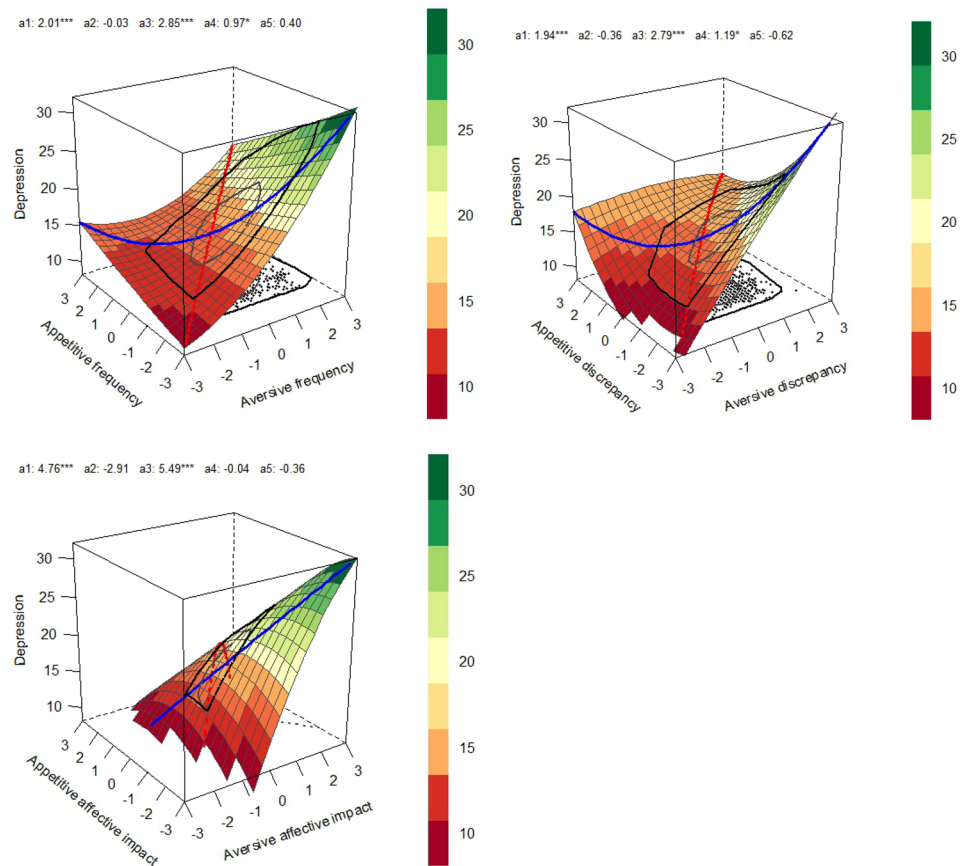
polynomial model for all nine models to enhance comparability of the results. The full results of the model comparisons can be found in Supplemental Table S2. All models met the condition of independence of residuals and did not exhibit multicollinearity (Supplemental Table S3). In some models, the tests indicated a violation of normal distribution and homoscedasticity. Given the large sample size and that confidence intervals and p-values were estimated with bootstrap resampling procedures, these violations were not considered consequential (Humberg & Grund, 2021; Pek et al., 2018).

Appetitive Appearance-Related Comparisons and Depression

### Frequency and Depression (Model 1)

The response surfaces for aversive and appetitive comparison frequency, discrepancy and affective impact models are shown in Fig. 1 (Models 1–3). Supplemental Table S4 shows all regression- and response surface parameters for all nine full polynomial models. Congruent comparison frequency of aversive and appetitive appearance-related comparisons was positively linearly related with depression (a1 = 2.01,

$p < .001$; a2 = 0.03, $p = .901$). This indicates that congruent high frequencies of aversive and appetitive comparisons were associated with higher depression compared with lower levels of congruency. For the LOIC, a positive curvilinear relationship emerged between the incongruent frequency of comparisons and depression, whose extreme value was shifted towards the appetitive-predominant range (a3 = 2.85, $p < .001$; a4 = 0.97, $p = .016$). More aversive comparison frequency predominating over appetitive comparison frequency was associated with higher depression. Note that the increase on the other side of the extreme value should not be interpreted due to too few data points (interpretation permitted up to the black mark, Fig. 1). Hypothesis 1.1a and 1.1b were thus confirmed.

Discrepancy and Depression (Model 2)

As expected, congruent discrepancy perceived after aversive and appetitive comparisons displayed a positive linear relationship with depression (a1 = 1.94, $p < .001$; a2 = − 0.36, $p = .238$). Specifically, congruent higher discrepancy was associated with higher depression. A positive curvilinear relationship emerged between the incongruent discrepancy comparison combinations and depression (a3 = 2.79,

Fig. 2 Response Surface Plots for aversive and appetitive comparisons as predictor variables and psychological well-being as outcome variable (Models 4–6). Predictor variables were grand-mean centered. Affective impact indicates positive affective impact after engaging in appetitive comparisons and negative affective impact after engaging in aversive comparisons. Models are based on the full polynomial model: $Z = b_0 + b_1 X + b_2 Y + b_3 X^2 + b_4 XY + b_5 Y^2 + \varepsilon$
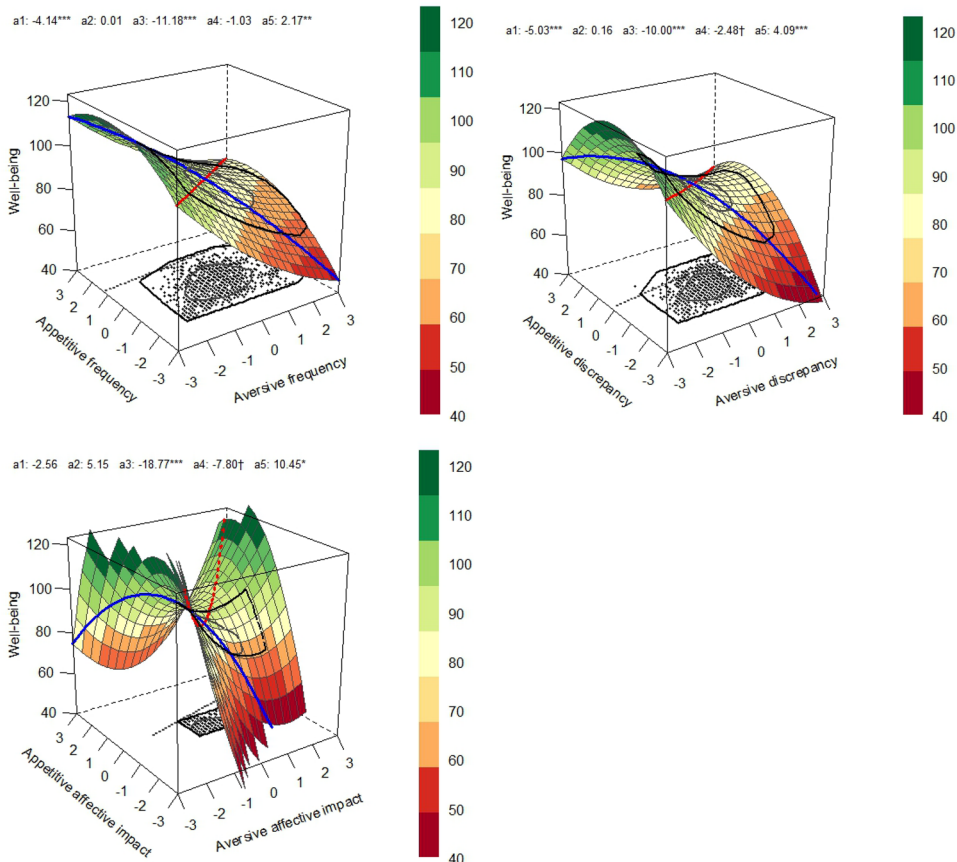
$p = < 0.001$; a4 = 1.19, $p = .026$). The extreme value was again shifted to the appetitive-predominant area. In the interpretable range, a stronger preponderance of perceived aversive discrepancy over perceived appetitive discrepancy was associated with higher depression. Hypotheses 1.2a and 1.2b were thus confirmed. A greater preponderance of appetitive discrepancy over aversive discrepancy was associated with lower depression (Fig. 1; Table s4).

Affective Impact and Depression (Model 3)

Comparative affective impact after engaging in both types of comparisons displayed a positive linear relationship with depression (a1 = 4.76, $p < .001$; a2 = −2.91, $p = .106$). Congruent high affective impact was associated with higher depression. As hypothesized in Hypothesis 1.3b, incongruent affective impact was positively linearly related to depression (a3 = 5.49, $p < .001$; a4 = −0.04, $p = .998$). A greater preponderance of negative affective impact following aversive comparisons over positive affective impact following appetitive comparisons was associated with higher depression. A stronger predominance of positive affective impact after appetitive comparisons over negative affective
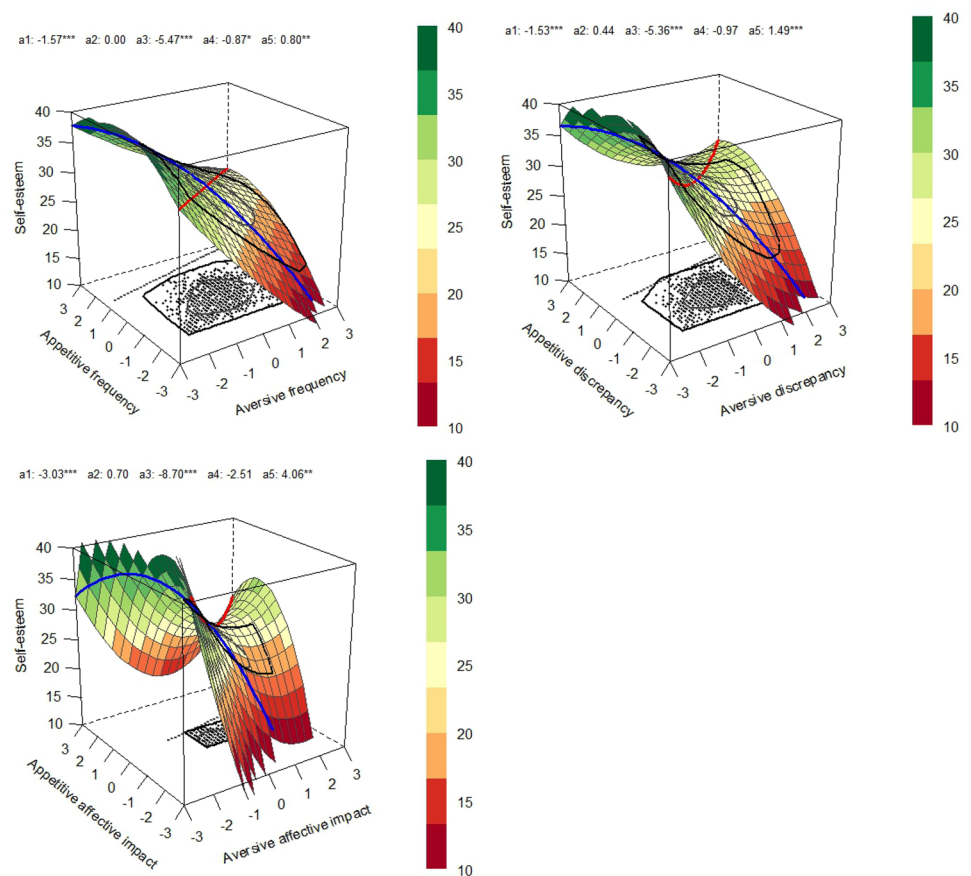
impact after aversive comparisons was associated with lower depression (Fig. 1; Table s4).

Appearance-Related Comparisons and Psychological Well-Being

Frequency and Psychological Well-Being (Model 4)

The response surfaces for aversive and appetitive comparison frequency, discrepancy and affective impact are shown in Fig. 2 (Models 4–6, Supplemental Table s4 for the regression- and response surface parameters). Congruent high frequency of aversive and appetitive comparisons was related to lower well-being. This relationship was negatively linear (a1 = − 4.14, $p < .001$; a2 = 0.01, $p = .992$). The relationship between incongruent frequencies of comparisons was found to be negatively linear (a3 = − 11.18, $p < .001$; a4 = − 1.03, $p = .346$). A greater preponderance of the frequency of aversive comparisons over appetitive comparisons was associated with lower well-being. Hypotheses 2.1a and 2.1b were confirmed. A greater preponderance of the frequency of appetitive comparisons over aversive comparisons was associated with higher well-being.

Fig. 3 Response Surface Plots for aversive and appetitive comparisons as predictor variables and self-esteem as outcome variable (Models 7–9). Predictor variables were grand-mean centered. Affective impact indicates positive affective impact after engaging in appetitive comparisons and negative affective impact after engaging in aversive comparisons. Models are based on the full polynomial model:
$$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + \varepsilon$$

Discrepancy and Well-Being (Model 5)

A negative linear relationship emerged between the strength of perceived congruent discrepancy after aversive and appetitive comparisons and psychological well-being (a1 = −5.03, $p < .001$; a2 = 0.16, $p = .840$): Congruent high aversive and appetitive discrepancy was related to lower well-being. In addition, we found a negative linear relationship of incongruent perceived discrepancies after comparisons and psychological well-being (a3 = − 10.00, $p < .001$; a4 = − 2.48, $p = .119$). A greater preponderance of perceived discrepancy after aversive comparisons over perceived discrepancy after appetitive comparisons was associated with lower well-being. This confirms hypotheses 2.2a and 2.2b. Well-being was higher for individuals who perceived the discrepancy to be greater after appetitive comparisons than after aversive comparisons (Fig. 2; Table s4).

Affective Impact and Psychological Well-Being (Model 6)

Congruent comparative affective impact had no significant relationship with well-being (a1 = − 2.56, $p = .272$; a2 = 5.15, $p = .350$). A negative linear relationship occurred between incongruent affective impact and well-being (a3 = − 18.77, $p < .001$; a4 = − 7.80, $p = .093$). A stronger

predominance of negative affective impact following aversive comparisons over engendered positive affective impact following appetitive comparisons was associated with lower well-being. Hypothesis 2.3b was thus confirmed. A stronger predominance of positive affective impact after appetitive comparisons over negative affect after aversive comparisons was associated with higher well-being (Fig. 2; Table s4).

Appearance-Related Comparisons and Self-esteem

Frequency and Self-esteem (Model 7)

The response surfaces concerning self-esteem can be found in Fig. 3 (Models 7–9, Table s4 for regression- and response surface parameters). Congruent comparison frequency displayed a negative linear relationship with self-esteem (a1 = − 1.57, $p < .001$; a2 = 0.00, $p = .998$). Congruent high aversive and appetitive comparisons were associated with lower self-esteem. A negative curvilinear relationship with an extreme value in the appetitive-predominant range emerged between incongruent frequencies and self-esteem (a3 = − 5.47, $p < .001$; a4 = − 0.87, $p = .022$). In the interpretable range, a preponderance of frequency of aversive over appetitive comparisons was associated with lower self-esteem. Hypotheses 3.1a and 3.1b were thus confirmed. Greater

frequency of appetitive over aversive comparison frequency was associated with higher self-esteem.

Discrepancy and Self-esteem (Model 8)

Congruent discrepancy following aversive and appetitive comparisons was negatively linearly related to self-esteem (a1 = −1.53, $p < .001$; a2 = 0.44, $p = .197$). Congruent high discrepancy was associated with lower self-esteem. Incongruent discrepancy was also negatively linearly related to self-esteem (a3 = −5.36, $p < .001$; a4 = −0.97, $p = .125$). A greater preponderance of aversive discrepancy over appetitive discrepancy was associated with lower self-esteem. Hypotheses 3.2a and 3.2b were thus confirmed. A greater preponderance of perceived discrepancy after appetitive comparisons over perceived discrepancy after aversive comparisons was associated with higher self-esteem (Fig. 3; Table s4).

Affective Impact and Self-esteem (Model 9)

Congruent affective impact after engaging in comparisons was negatively linearly related to self-esteem (a1 = −3.03, $p < .001$; a2 = 0.70, $p = .749$). Congruent strong affective impact was related to lower self-esteem. Incongruent engendered affective impact was negatively linearly related to self-esteem (a3 = −8.70, $p < .001$; a4 = −2.51, $p = .102$). A stronger predominance of negative affective impact following aversive comparisons over positive affective impact following appetitive comparisons was associated with lower self-esteem. A stronger predominance of positive affective impact after appetitive comparisons over negative affective impact after aversive comparisons was associated with higher self-esteem (Fig. 3; Table s4). Hypothesis 3.3b was thus confirmed.

Sensitivity Analyses

As sensitivity analyses, we tested all models again including outliers and separately for men and women. The interpretation of the results remained essentially the same, and we present them in the supplemental material in part B (analyses with outliers) and part C (analyses separately for gender).

Discussion

We investigated the relationship of congruent and incongruent aversive and appetitive appearance-related comparisons with depression, psychological well-being, and self-esteem using polynomial regression models and RSA. Findings mainly confirmed our preregistered hypotheses. First, congruent high frequency, discrepancy, or engendered affective impact were associated with higher depression and lower psychological well-being and self-esteem. Second, a greater predominance of aversive over appetitive comparisons was associated with higher depression, and lower psychological well-being and self-esteem. Third, a stronger predominance of appetitive over aversive comparison was associated with lower depression and higher psychological well-being and self-esteem.

Congruent high *frequency* of aversive and appetitive comparisons was associated with higher depression, and lower psychological well-being and self-esteem. This aligns with social comparison literature indicating that comparison orientation is negatively associated with well-being (Buunk & Gibbons, 2006; Corcoran et al., 2011; Laker & Waller, 2020; Schaefer & Thompson, 2018; Unkelbach et al., 2023). Our findings yet suggest that aversive comparison may drive this pattern as its negative outcomes may be more salient than the outcomes of appetitive comparison (Bowen et al., 2018). Specifically, for single associations of all combinations, aversive comparison displayed higher correlations with the outcome variables than appetitive comparison, for which some single correlations with the outcomes were nonsignificant. At least for the analyses concerning depression, this is plausible since no regression parameter of the appetitive component was significant in these models. Future studies need to disentangle temporal effects as this pattern may reflect attempts to counteract negative outcomes of aversive comparisons by deliberately engaging in tertiary appetitive comparisons, which are conducted as an attempt to restore one's own self-esteem following prior aversive comparison (Morina, 2021). The association of mental health outcomes with congruent high aversive and appetitive *discrepancy* between the target and the standard may point to the instability of the self-concept as dissonant information has to be integrated. Notably, the finding that high *engendered affective impact* after engaging in both types of comparisons was associated with more mental health problems did not align with the emotion context insensitivity theory (Bylsma, 2021) that states that depressed individuals show lower affective reactivity to stimuli in general. In contrast, these results align with a study that found an increased affective reaction to comparisons in depressed individuals (Giordano et al., 2000). Again, strong affective reaction following both aversive and appetitive comparisons may be predominated by the negative engendered affective impact after engaging in aversive comparisons due to a greater salience of negative emotions compared with positive emotions (Bowen et al., 2018). Only for psychological well-being, there was no effect of a congruent engendered affective impact after engaging in appearance-related comparisons. In this model, most regression parameters for appetitive comparisons were significant, pointing to a potentially more important role of appetitive comparisons in psychological well-being.

Mainly engaging in aversive (rather than appetitive) appearance-related comparisons was associated with higher depression and lower well-being and self-esteem. This aligns with previous research using simple regression models (Buunk & Gibbons, 2006; Laker & Waller, 2020; Schaefer & Thompson, 2018). Depressed individuals frequently engage in aversive social comparison (McCarthy & Morina, 2020), which may reinforce the believe that one is not doing well (Appel et al., 2015). Similarly, high *discrepancy* after engaging in aversive (rather than appetitive) comparison standards was also associated with higher depression and lower well-being and lower self-esteem. Potentially, individuals are frequently reminded that they are not as good looking as the standard at hand with a large discrepancy, while simultaneously realizing that they look only marginally better than an appetitive standard. Last, perceiving only *negative engendered affective impact* after engaging in aversive appearance-related comparisons, but no or little engendered affective impact after engaging in appetitive comparisons was associated with higher depression, lower psychological well-being, and lower self-esteem, also in line with previous work (Schaefer & Thompson, 2018; Zimmer-Gembeck et al., 2021). Altogether, frequent comparisons to aversive standards only, accompanied by strong discrepancy to the standard as well as strong negative affective impact, signal to the comparer that they are not attractive enough, which they seem to perceive as a significant threat to their self-concept. This perception, in turn, negatively influences their depression, well-being, and self-esteem. This further points to dysfunctional emotion regulation strategies.

More frequent appetitive comparisons over aversive comparisons were related to positive psychological outcomes. Additionally, perceiving oneself as much better looking than appetitive standards, while perceiving small discrepancy aversive standards may be psychologically beneficial. Noteworthy, small discrepancy to better off standards may also appear as attainable and therefore beneficial (Morina, 2021). Likewise, experiencing high positive affective impact following appetitive comparisons and low affective impact after engaging in aversive comparisons appears functional. These findings do not support the proposition that individuals with low self-esteem engage in appetitive comparisons to restore their self-esteem (Wills, 1981). However, the temporal sequence and underlying mechanisms need to be disentangled as these individuals may have learned strategies to use comparisons in a way that serves their self-esteem.

As preregistered, we tested the effects of gender exploratorily to discern potential gender differences on the effects of different combinations of aversive and appetitive comparisons on the mental health outcomes, given gender differences in body satisfaction (He et al., 2020) and appearance-related comparisons (Morina et al., 2023). Importantly, the results did not differ in their core interpretation, indicating that these patterns are similar across gender despite mean-level difference (Morina et al., 2023).

Clinical Implications and Future Research

Social comparisons have been suggested to be associated with mental health outcomes (Schaefer & Thompson, 2018; Unkelbach et al., 2023). Our findings yielded that multiple types of comparison are related to depression, well-being, and self-esteem. In particular, they suggest that high levels of aversive appearance-based comparisons are negatively related to psychological outcomes independent of appetitive comparisons. They further suggest that frequent appetitive comparisons are positively related to well-being only in the absence of aversive comparisons. Future studies need to examine underlying mechanisms of comparative behavior in general (i.e., beyond appearance) and their relevance to mental health and beyond. For instance, disentangling underlying motives for engaging in aversive and appetitive comparisons and associated attitudes towards engagement in comparative behavior may prove beneficial. Furthermore, we need to better understand the extent to which some appetitive comparisons are actively made to counteract negative outcomes of aversive comparisons. Increased knowledge of comparative behavior will inform about therapeutic tools to reduce the frequency and impact of dysfunctional comparisons.

Limitations

This study has noteworthy limitations. First, data were cross-sectional. Although the models imply directionality, the terms predictor and outcome variable need to be understood in the context of the model and cannot discern causal relationships. As discussed, different and reciprocal directionalities are plausible and need to be tested in subsequent studies. Second, despite the preregistration of our hypotheses and the plan for our secondary analyses, we deviated from some of the prespecified hypotheses because of the way affective impact was conceptualized. Third, the sample was drawn from an online panel provider and results may thus not be generalized to different populations. Investigations with representative and culturally diverse samples are warranted. Fourth, some of the internal consistencies of the CSS-A were rather low, which may be attributable to the complexity and breadth of the comparison constructs. Nonetheless, the CSS-A is a psychometrically sound scale to assess appearance-related comparisons (Morina et al., 2023). Fifth, while theoretical considerations concerning the underlying motives and the strategic use of comparisons informed our hypotheses (Morina, 2021), we were unable to conduct a formal test

of this idea using our data, as our data did not distinguish between primary, secondary, or tertiary comparisons. Last, the interpretation of RSA parameters is challenging and should only be approached with caution (Humberg et al., 2019).

Conclusion

Our study sheds light on the complexity of the comparison process in the domain of appearance. Higher levels of congruent frequency, perceived discrepancy, or engendered affective impact were associated with higher depression and lower psychological well-being and self-esteem. In addition, a greater predominance of aversive over appetitive comparisons was associated with higher depression, lower well-being, and lower self-esteem. This points to the importance of aversive appearance-related comparisons in the context of mental health. By the means of RSA, this study discerned relevant patterns underlying comparisons that may inform future studies and intervention efforts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10608-023-10369-6.

Author Contributions PS, SK: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Visualization; Roles/Writing—original draft; Writing—review & editing. NM: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Roles/Writing—original draft; Writing—review & editing.

Declarations

Conflict of Interest Pascal Schlechter, Sarah Katenhusen and Nexhmedin Morina declare that they have no conflict of interest.

Informed Consent Additional informed consent was obtained from any subjects for whom identifying information appears in this paper.

Research Involving Animal and/or Human Participants No animal studies were carried out by the authors for this article.

Openness and Transparency All data and the R code to reproduce the current results are provided in an anonymized form in the open science framework (https://osf.io/6us2n/?view_only=2f16cc32a93e4ba7b0fd2997d40b0d6c). We also provide the survey material that has been used. The present study is part of a larger project examining cognitive and social variables in relation to mental health outcomes. One manuscript of this project has been published on the development of the Comparison Standards Scale for Appearance (Morina et al., 2023), which we use in the present contribution to assess appearance-related comparisons. The initial project was not preregistered, but the current study was preregistered (https://osf.io/xam4j/) following suggestions for preregistrations of secondary data analyses (Akker et al., 2021). Hypotheses concerning the affective impact after engaging in comparison differ from the preregistration as explained in the main text and indicated by an *.

References

Akker, O. R., Weston, S., Campbell, L., Chopik, B., Damian, R., Davis-Kean, P., Hall, A., Kosie, J., Kruse, E., Olsen, J., Ritchie, S., Valentine, K. D., Veer, A., & Bakker, M. (2021). Preregistration of secondary data analysis: A template and tutorial. *Meta-Psychology*. https://doi.org/10.15626/MP.2020.2625

Albert, S. (1977). Temporal comparison theory. *Psychological Review*, *84*, 485–503. https://doi.org/10.1037/0033-295X.84.6.485.

Appel, H., Crusius, J., & Gerlach, A. L. (2015). Social comparison, envy, and depression on facebook: A study looking at the effects of high comparison standards on depressed individuals. *Journal of Social and Clinical Psychology, 34*(4), 277–289. https://doi.org/10.1521/jscp.2015.34.4.277

Barnes, M., Abhyankar, P., Dimova, E., & Best, C. (2020). Associations between body dissatisfaction and self-reported anxiety and depression in otherwise healthy men: A systematic review and meta-analysis. *PLoS ONE, 15*(2), e0229268. https://doi.org/10.1371/journal.pone.0229268

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: Negative emotional valence enhances recapitulation. *Psychonomic Bulletin & Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9.

Butzer, B., & Kuiper, N. A. (2006). Relationships between the frequency of social comparisons and self-concept clarity, intolerance of uncertainty, anxiety, and depression. *Personality and Individual Differences*, *41*(1), 167–176. https://doi.org/10.1016/j.paid.2005.12.017.

Buunk, A. P., & Gibbons, F. X. (2006). Social comparison orientation: A new perspective on those who do and those who don't compare with others. *Social comparison and social psychology: Understanding cognition, intergroup relations, and culture* (pp. 15–32). Cambridge University Press.

Bylsma, L. M. (2021). Emotion context insensitivity in depression: Toward an integrated and contextualized approach. *Psychophysiology*, *58*(2), e13715. https://doi.org/10.1111/psyp.13715.

Corcoran, K., Crusius, J., & Mussweiler, T. (2011). *Social comparison: Motives, standards, and mechanisms. D. Chadee (Ed.), Theories in social psychology*, 119–139.

Crusius, J., Corcoran, K., & Mussweiler, T. (2022). Social Comparison: Theory, Research, and Applications. Theories in Social Psychology, 165.

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression and response surface methodology. *Advances in Measurement and Data Analysis, 1*, 350–400.

Festinger, L. (1954). A theory of social communication processes. *Human Relations*, *7*, 117–140.

Giordano, C., Wood, J. V., & Michela, J. L. (2000). Depressive personality styles, dysphoria, and social comparisons in everyday life. *Journal of Personality and Social Psychology*, *79*, 438–451. https://doi.org/10.1037/0022-3514.79.3.438.

Grogan, S. (2006). Body image and health: Contemporary perspectives. *Journal of health psychology*, *11*(4), 523–530. https://doi.org/10.1177/1359105306065013

He, J., Sun, S., Zickgraf, H. F., Lin, Z., & Fan, X. (2020). Meta-analysis of gender differences in body appreciation. *Body Image*, *33*, 90–100. https://doi.org/10.1016/j.bodyim.2020.02.011.

Higgins, E. T. (1996). Activation: Accessibility, and salience. *Social psychology: Handbook of basic principles*, 133–168.

Hill, E. M., & Nolan, M. T. (2021). Examining eating disorder-related social comparison orientation and body dissatisfaction in the relationship between fit ideal internalization and drunkorexia engagement. *Eating Behaviors*, *41*, 101480. https://doi.org/10.1016/j.eatbeh.2021.101480.

Humberg, S., & Grund, S. (2021). Response surface analysis with Missing Data. *Multivariate behavioral research*, *57*, 1–22. https://doi.org/10.1080/00273171.2021.1884522.

Humberg, S., Nestler, S., & Back, M. D. (2019). Response surface analysis in personality and social psychology: Checklist and Clarifications for the case of congruence hypotheses. *Social Psychological and Personality Science*, *10*(3), 409–419. https://doi.org/10.1177/1948550618757600.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153. https://doi.org/10.1037/0033-295X.93.2.136.

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*(1), 163–173. https://doi.org/10.1016/j.jad.2008.06.026.

Laker, V., & Waller, G. (2020). The development of a body comparison measure: The CoSS. *Eating and Weight Disorders Studies on Anorexia Bulimia and Obesity, 25*(4), 879–888. https://doi.org/10.1007/s40519-019-00698-5

Lawler, M., & Nixon, E. (2011). Body dissatisfaction among adolescent boys and girls: The effects of body mass, peer appearance culture and internalization of appearance ideals. *Journal of youth and adolescence*, *40*(1), 59–71.

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers (Edited by Dorwin Cartwright.)* (S. xx, 346). Harpers.

Liu, Q., He, H., Yang, J., Feng, X., Zhao, F., & Lyu, J. (2020). Changes in the global burden of depression from 1990 to 2017: Findings from the Global Burden of Disease study. *Journal of Psychiatric Research*, *126*, 134–140. https://doi.org/10.1016/j.jpsychires.2019.08.002.

Lup, K., Trub, L., & Rosenthal, L. (2015). Instagram #Instasad?: Exploring Associations among Instagram Use, depressive symptoms, negative Social Comparison, and Strangers followed. *Cyberpsychology Behavior and Social Networking*, *18*(5), 247–252. https://doi.org/10.1089/cyber.2014.0560.

McCarthy, P. A., & Morina, N. (2020). Exploring the association of social comparison with depression and anxiety: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *27*(5), 640–671. https://doi.org/10.1002/cpp.2452.

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*, 544–560. https://doi.org/10.1037/a0032459.

Morina, N. (2021). Comparisons inform me who I am: A General Comparative-Processing Model of Self-Perception. *Perspectives on Psychological Science*, *16*(6), 1281–1299. https://doi.org/10.1177/1745691620966788.

Morina, N., McCarthy, P., Meyer, T., & Schlechter, P. (2023). Measuring appearance-related comparisons: Validation of the comparison Standards Scale for Appearance. *Cognitive Therapy and Research*, *47*(1), 123–139. https://doi.org/10.1007/s10608-022-10337-6.

Mussweiler, T. (2014). Same or different?: How similarity versus dissimilarity focus shapes social information processing. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind (pp. 328–339). The Guilford Press.*

Myers, T. A., & Crowther, J. H. (2009). Social comparison as a predictor of body dissatisfaction: A meta-analytic review. *Journal of Abnormal Psychology*, *118*, 683–698. https://doi.org/10.1037/a0016763.

Orth, U., Erol, R. Y., & Luciano, E. C. (2018). Development of self-esteem from age 4 to 94 years: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *144*, 1045–1080. https://doi.org/10.1037/bul0000161.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004.

Pek, J., Wong, O., & Wong, A. C. M. (2018). How to address non-normality: A taxonomy of approaches, reviewed, and illustrated. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.02104

Quittkat, H. L., Hartmann, A. S., Düsing, R., Buhlmann, U., & Vocks, S. (2019). Body dissatisfaction, importance of appearance, and body appreciation in men and women over the lifespan. *Frontiers in Psychiatry*. https://doi.org/10.3389/fpsyt.2019.00864

R. Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rodrigues, A. C. (2021). Response surface analysis: A tutorial for examining linear and curvilinear effects. *Revista de Administração Contemporânea*. https://doi.org/10.1590/1982-7849rac2021200293.en

Rosenberg, M. (1965). Rosenberg self-esteem scale. *Journal of Religion and Health*. https://doi.org/10.1037/t01038-000

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69*, 719–727. https://doi.org/10.1037/0022-3514.69.4.719.

Schaefer, L. M., & Thompson, J. K. (2018). The development and validation of the physical appearance comparison Scale–3 (PACS-3). *Psychological Assessment*, *30*, 1330–1341. https://doi.org/10.1037/pas0000576.

Schönbrodt, F. (2016). Testing fit patterns with polynomial regression models. *OSF Preprints*. https://doi.org/10.31219/osf.io/ndggf.

Schönbrodt, F., & Humberg, S. (2021). *Package 'RSA'*. https://cran.r-project.org/web/packages/RSA/RSA.pdf

Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. P. Zanna (Ed.), *Advances in experimental social psychology (Bd. 29, S. 209–269)*. Academic Press.

Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful Approach for Examining Moderation and Overcoming Limitations of difference scores. *Journal of Business and Psychology*, *25*(4), 543–554. https://doi.org/10.1007/s10869-010-9183-4.

Steers, M. L. N., Wickham, R. E., & Acitelli, L. K. (2014). Seeing everyone else's highlight reels: How Facebook usage is linked to depressive symptoms. *Journal of Social and Clinical Psychology*, *33*(8), 701.

Stice, E., Hayward, C., Cameron, R. P., Killen, J. D., & Taylor, C. B. (2000). Body-image and eating disturbances predict onset of depression among female adolescents: A longitudinal study.

Journal of Abnormal Psychology, *109*, 438–444. https://doi.org/10.1037/0021-843X.109.3.438.

Thompson, R. J., Berenbaum, H., & Bredemeier, K. (2011). Cross-sectional and longitudinal relations between affective instability and depression. *Journal of Affective Disorders*, *130*(1), 53–59. https://doi.org/10.1016/j.jad.2010.09.021.

Trudel-Fitzgerald, C., Millstein, R. A., von Hippel, C., Howe, C. J., Tomasso, L. P., Wagner, G. R., & VanderWeele, T. J. (2019). Psychological well-being as part of the public health debate? Insight into dimensions, interventions, and policy. *Bmc Public Health*, *19*(1), 1712. https://doi.org/10.1186/s12889-019-8029-x.

Unkelbach, C., Alves, H., Baldwin, M., Crusius, J., Diel, K., Galinsky, A. D., Gast, A., Hofmann, W., Imhoff, R., & Genschow, O. (2023). Relativity in social cognition: Basic processes and novel applications of social comparisons. *European Review of Social Psychology, 1*, 1–54.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, *90*, 245–271. https://doi.org/10.1037/0033-2909.90.2.245.

Woltin, K. A., & Epstude, K. (2023). Should I have been more careful or less careless? The comparative nature of counterfactual thoughts alters judgments of their impact. *Cognition*, *235*, 105402.

Zimmer-Gembeck, M. J., Hawes, T., & Pariz, J. (2021). A closer look at appearance and social media: Measuring activity, self-presentation, and social comparison and their associations with emotional adjustment. *Psychology of Popular Media*, *10*, 74–86. https://doi.org/10.1037/ppm0000277.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.