Meta-analytic prospective associations between self-esteem and eating disorders appear to be spurious: a reanalysis and comment on Krauss et al. (2023)

Kimmo Sorjonen*, Ata Ghaderi and Bo Melin

Karolinska Institutet (KI), Solna, Sweden

In a recent meta-analysis, Krauss et al. found support for a reciprocal model of low self-esteem and eating disorders where, in a vicious circle, low self-esteem makes people more vulnerable to developing eating disorders and eating disorders, in turn, scars individuals' self-esteem. However, in the present reanalyses of the same meta-analytic data, we found that the prospective effects between self-esteem and eating disorders are likely spurious, meaning they do not reflect a true causal effect, but rather correlations with residuals and regression to the mean. Consequently, the claims by Krauss et al. can be challenged. To avoid statistical artifacts, we recommend researchers to fit, as we did in the present study, complementary models to their data in order to evaluate if prospective effects may be genuinely increasing or decreasing or if they appear to be spurious.

Introduction

Self-esteem is often defined as our attitudes toward ourselves, where high and low self-esteem would mean a positive and a negative attitude, respectively (Kenrick et al., 2007). It has been proposed that low self-esteem increases the likelihood for eating pathology (Colmsee et al., 2021). It has also been hypothesized that eating pathology may, conversely, have a detrimental impact on self-esteem (Silverstone and Salsali, 2003).

Krauss et al. (2023) extracted zero-order correlations between self-esteem and eating disorders measured at two occasions from studies included in their meta-analysis and estimated the effect of prior self-esteem on subsequent eating disorders when adjusting for prior eating disorders and vice versa. As both meta-analytic adjusted cross-lagged effects were statistically significant ($\beta = -0.08$ and $\beta = -0.09$, respectively), Krauss et al. concluded that their findings supported a reciprocal-relations model of low self-esteem and eating disorders where, in a vicious circle, low self-esteem makes people more vulnerable to developing eating disorders and eating disorders, in turn, deteriorates individuals' self-esteem. However, it is well established that adjusted cross-lagged effects may be spurious due to correlations with residuals and regression to the mean rather than due to genuine increasing or decreasing effects (Glymour et al., 2005; Eriksson and Häggström, 2014; Castro-Schilo and Grimm, 2018; Sorjonen et al., 2019; Lucas, 2023).

For example, self-rated self-esteem and eating pathology are negatively correlated (Krauss et al., 2023). Therefore, we should expect that among individuals with the same initial degree of eating pathology, those with a higher initial self-esteem score have a lower true degree of

eating pathology and have, consequently, received a more positive residual in the measurement of eating pathology compared with those with the same initial measured eating pathology score but lower initial self-esteem. However, residuals tend to regress toward a mean value of zero between measurements. This means that we should expect a more negative, but spurious, change in self-rated eating pathology to a subsequent measurement among those with high initial self-esteem compared with those with the same initial eating pathology score but lower initial self-esteem. This combination of correlations with residuals and regression toward the mean might explain observed negative cross-lagged effects of initial self-esteem on subsequent eating pathology when adjusting for initial eating pathology and vice versa.

The objective of the present study was to reanalyze the meta-analytic data used by Krauss et al. (2023) in order to evaluate if the identified adjusted cross-lagged associations suggested genuine decreasing effects between self-esteem and eating disorders or if the associations may have been spurious.
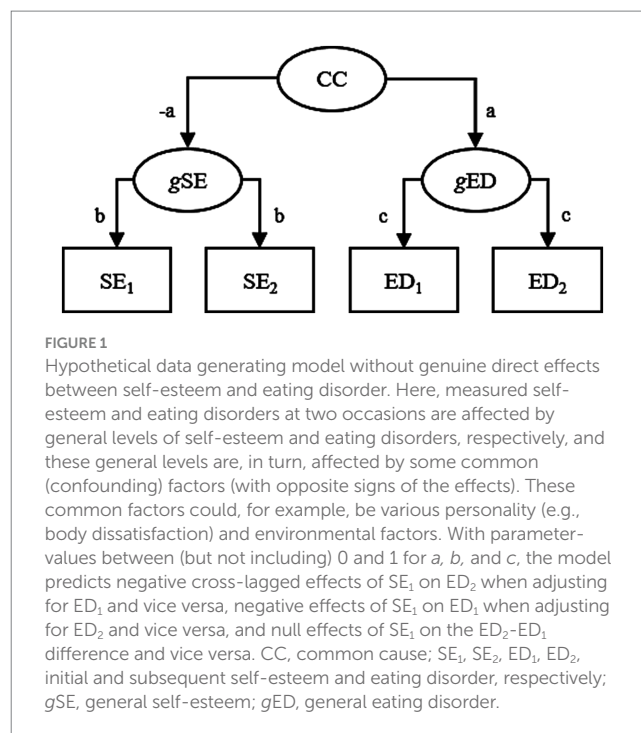
Method

Meta-analytic data

We refer to Krauss et al. (2023) for more comprehensive information on selection of studies, study populations, etc. In short, Krauss et al. extracted information from 44 sources (41 journal articles and three dissertations, published between 2001 and 2018) on 48 independent samples (ranging in size between 20 and 2,601, $M = 399.7$, total $N = 19{,}187$). In total, the data provided between 127 and 146 correlations between self-esteem and eating pathology measured at two occasions. Mean age at the initial measurement ranged between 6.5 years and 47.6 years ($M = 19.3$ years). Mean proportion of female participants was 79% (range from 0 to 100%). The two most commonly used measures of self-esteem and eating disorders were Rosenberg Self-Esteem Scale, Harter's Self-Perception Profile, Eating Disorder Examination, and Eating Disorder Inventory, respectively. Additionally, an Italian tool for measuring eating behavior regulation, the Tempest Self-Regulation Questionnaire for Eating, has been validated in the literature, showing adequate psychometric properties and gender invariance (Diotaiuti et al., 2022). Krauss et al. have made their data and analytic script available at https://osf.io/t6nge/.

Statistical analyses and predictions

Following Krauss et al. (2023), we used Equation 1 (Cohen et al., 2003) to estimate the effect of initial self-esteem on subsequent eating pathology while adjusting for initial eating pathology, and vice versa, in the included studies. Here, both a hypothesis of genuine reciprocal decreasing effects and a hypothesis of spurious effects (if data were, for example, generated as in Figure 1) predicted negative effects (Table 1, rows 1 and 4). Additionally, we used Equation 1 to estimate the effect of initial self-esteem on initial eating pathology while adjusting for subsequent eating pathology and vice versa. Here, a hypothesis of genuine decreasing reciprocal effects predicted positive effects. This positive effect would indicate that a high initial score on

FIGURE 1
Hypothetical data generating model without genuine direct effects between self-esteem and eating disorder. Here, measured self-esteem and eating disorders at two occasions are affected by general levels of self-esteem and eating disorders, respectively, and these general levels are, in turn, affected by some common (confounding) factors (with opposite signs of the effects). These common factors could, for example, be various personality (e.g., body dissatisfaction) and environmental factors. With parameter-values between (but not including) 0 and 1 for a, b, and c, the model predicts negative cross-lagged effects of SE_1 on ED_2 when adjusting for ED_1 and vice versa, negative effects of SE_1 on ED_1 when adjusting for ED_2 and vice versa, and null effects of SE_1 on the ED_2-ED_1 difference and vice versa. CC, common cause; SE_1, SE_2, ED_1, ED_2, initial and subsequent self-esteem and eating disorder, respectively; gSE, general self-esteem; gED, general eating disorder.

self-esteem had counteracted a high initial value on eating pathology and allowed individuals to reach the same subsequent value on eating pathology as individuals with a lower initial value on eating pathology but also with a lower initial value on self-esteem and vice versa. Contrarily, a hypothesis of spurious effects predicted these effects to be negative (Table 1, rows 2 and 5). Moreover, we used Equation 2 (Guilford, 1965) to estimate the effect of initial self-esteem on the subsequent – initial eating pathology difference and vice versa. Here, a hypothesis of genuine reciprocal decreasing effects predicted negative effects. On the other hand, a hypothesis of spurious effects predicted either effects close to zero (if concurrent and cross-lagged correlations were approximately equal in size) or positive effects (if concurrent correlations were more negative than cross-lagged correlations) (Table 1, rows 3 and 6).

$$E \mid \beta_{X1,Y2.Y1} \mid = \frac{r_{X1,Y2} - r_{X1,Y1} r_{Y1,Y2}}{1 - r_{X1,Y1}^2} \tag{1}$$

$$E \mid \beta_{X1,Y2-Y1} \mid = \frac{r_{X1,Y2} - r_{X1,Y1}}{\sqrt{2(1 - r_{Y1,Y2})}} \tag{2}$$

We conducted a multilevel random effects meta-analysis for all six regression effects in Table 1, as well as for each of the six zero-order correlations between self-esteem and eating disorders measured at two occasions. Multilevel analyses adjusted for non-independence due to the fact that some of the effect sizes (NE between 127 and 146) were estimated/clustered in the same sample (K between 37 and 48). In a first step, pooled effects were estimated across all effects within the same sample, and in a second step an aggregated effect was estimated across all samples. This way, artificially reduced heterogeneity in effect sizes and increased risk for false positive findings were counteracted

(Harrer et al., 2022). A random meta-analytic effect, with a 95% confidence interval, was estimated across the independent effect sizes (i.e., independent samples). Analyses were conducted on Fisher's z-transformed standardized regression effects and correlations, but these were inverted back to non-transformed effects and correlations for presentations. Analyses were conducted with R 4.3.1 statistical software (R Core Team, 2024) using the metafor (Viechtbauer, 2010) and osfr (Wolen et al., 2020) packages. Data, a list of included studies, analytic script, forest plots, and supplementary results are available at the Open Science Framework at https://osf.io/g8dkc/. Similarly as the original meta-analysis by Krauss et al. (2023), the present reanalysis of publicly available, anonymous, and aggregated data did not require ethical approval or consent from participants and it was not pre-registered.

Results

Meta-analytic estimates of associations between self-esteem and eating disorders are presented in Table 2. With one exception,

TABLE 1 Predicted sign of effects between self-esteem and eating disorders according to a hypothesis of genuine reciprocal decreasing effects and a hypothesis of spurious effects.

Effect	Genuine	Spurious
1. β(SE1,ED2.ED1)	Negative	Negative
2. β(SE1,ED1.ED2)	Positive	Negative
3. β(SE1,ED2-ED1)	Negative	Zero or positive
4. β(ED1,SE2.SE1)	Negative	Negative
5. β(ED1,SE1.SE2)	Positive	Negative
6. β(ED1,SE2-SE1)	Negative	Zero or positive

SE, self-esteem; ED, eating disorders; 1, time 1; 2, time 2; the variables are given in the order predictor, outcome, and covariate.

estimated associations (aggregated within samples) exhibited statistically non-significant and mostly low heterogeneity, as estimated by Cochran's Q and I^2, which estimates percentage of variation across effects attributable to heterogeneity rather than random variance. This means that associations across the included samples can, with one exception, be assumed to have come from the same distribution.

Concurrent and cross-lagged meta-analytic correlations were negative (Table 2, rows 1–2 and 5–6, respectively) while auto-correlations were positive (Table 2, rows 3–4). As for regression effects, initial self-esteem had a negative effect on subsequent eating pathology when adjusting for initial eating pathology (Table 2, row 7 and Figure 2A) and vice versa (Table 2, row 10 and Figure 2D). Initial self-esteem also had a negative effect on initial eating pathology when adjusting for subsequent eating pathology (Table 2, row 8 and Figure 2B) and vice versa (Table 2, row 11 and Figure 2E). On the other hand, initial self-esteem had a positive effect on the subsequent-initial eating pathology difference (Table 2, row 9 and Figure 2C) and vice versa (Table 2, row 12 and Figure 2F). The regression effects in Table 2, illustrated in Figure 2, indicate that initial self-esteem had a decreasing effect on eating pathology only when conditioning on initial eating pathology, and vice versa. The other effects indicate, contrarily and paradoxically, increasing effects. The negative effect of initial self-esteem on initial eating pathology when adjusting for subsequent eating pathology and vice versa (Table 2, rows 8 and 11 and Figures 2B,E) suggested that low, rather than high, initial self-esteem had counteracted high initial eating pathology and allowed individuals to attain the same subsequent degree of eating pathology as individuals with lower initial eating pathology but with higher initial self-esteem, and vice versa. The meta-analytic regressions effects agreed better with a hypothesis of spurious prospective effects than with a hypothesis of genuine decreasing prospective effects (compare effects on rows 7–12 in Table 2 with predictions in Table 1).

While a hypothesis of genuine decreasing effects predicted the six effects on rows 7–12 in Table 2, and across Figures 2A–F, to be (1)

TABLE 2 Meta-analytic correlations and regression effects between self-esteem and eating disorders measured at two occasions.

Association	K	NE	N	Estimate (95% CI)	Q (df)	I^2 (95% CI)
1. r(SE1,ED1)	48	146	19,187	−0.339 (−0.378; −0.298)	49.2 (47)	0 (0; 41.7)
2. r(SE2,ED2)	37	127	15,737	−0.398 (−0.442; −0.352)	39.9 (36)	0 (0; 50.4)
3. r(SE1,SE2)	39	133	15,878	0.572 (0.512; 0.626)	38.7 (38)	0 (0; 41.6)
4. r(ED1,ED2)	46	140	19,046	0.587 (0.540; 0.630)	44.8 (45)	0 (0; 36.8)
5. r(SE1,ED2)	46	140	19,046	−0.266 (−0.300; −0.232)	43.3 (45)	0 (0; 34.5)
6. r(ED1,SE2)	39	133	15,878	−0.273 (−0.312; −0.233)	35.8 (38)	0 (0; 35.7)
7. β(SE1,ED2.ED1)	46	140	19,046	−0.079 (−0.096; −0.062)	47.6 (45)	0 (0; 48.9)
8. β(SE1,ED1.ED2)	46	140	19,046	−0.202 (−0.236; −0.167)	44.8 (45)	0 (0; 37.8)
9. β(SE1,ED2-ED1)	46	140	19,046	0.083 (0.054; 0.112)	46.1 (45)	0 (0; 40.8)
10. β(ED1,SE2.SE1)	39	133	15,878	−0.088 (−0.102; −0.075)	56.8 (38)*	0 (0; 84.0)
11. β(ED1,SE1.SE2)	39	133	15,878	−0.202 (−0.243; −0.161)	48.3 (38)	0 (0; 60.7)
12. β(ED1,SE2-SE1)	39	133	15,878	0.068 (0.043; 0.094)	44.5 (38)	0 (0; 57.9)

K, number of samples; NE, number of effects; N, total sample size; Q, Cochran's Q; I^2, percentage of variation due to heterogeneity rather than randomness; SE, self-esteem; ED, eating disorders; 1, time 1; 2, time 2; the variables are given in the order predictor, outcome, and covariate; A hypothesis of genuine decreasing effects predicted the effects on rows 7–12 to be (7) negative, (8) positive, (9) negative, (10) negative, (11) positive, and (12) negative, respectively (see Table 1). As seen, the observed effect on rows 7–12 did not agree with these predictions. However, the observed effects did agree with predictions by a hypothesis of spurious effects of (7) negative, (8) negative, (9) zero or positive, (10) negative, (11) negative, and (12) zero or positive effects, respectively (see Table 1); *$p < 0.05$.
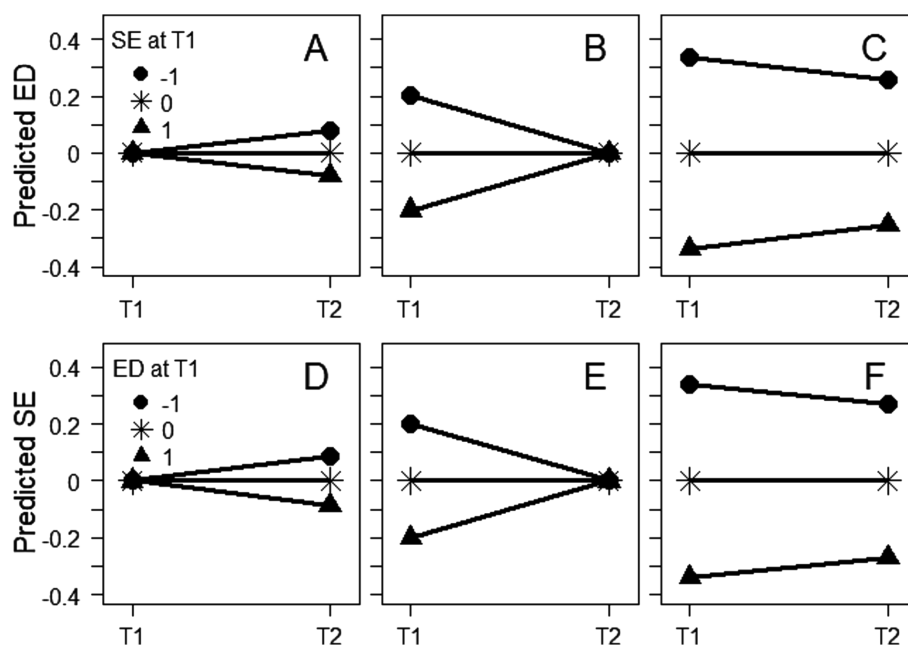
FIGURE 2
Predicted eating disorders **(A–C)** and self-esteem **(D–F)** at time 1 and time 2, when conditioning on mean initial **(A,D)** and subsequent **(B,E)** score on the outcome variable and when not conditioning **(C,F)**. Separately for individuals with low ($Z = -1$), mean, and high ($Z = 1$) initial self-esteem **(A–C)** and eating disorder **(D–F)**, respectively. The depicted effects of initial self-esteem **(A–C)** and initial eating disorders **(D–F)** across panels **(A–F)** were **(A)** negative, **(B)** negative, **(C)** positive, **(D)** negative, **(E)** negative, and **(F)** positive, respectively. These effects did not agree with predictions by a hypothesis of genuine decreasing effects (see Table 1) but they did agree with predictions by a hypothesis of spurious effects (see Table 1). SE, self-esteem; ED, eating disorder; T1, time 1; T2, time 2.

negative, (2) positive, (3) negative, (4) negative, (5) positive, and (6) negative, respectively (see Table 1), the observed effects were (1) negative, (2) negative, (3) positive, (4) negative, (5) negative, and (6) positive, respectively (see rows 7–12 in Table 2 and Figures 2A–F). Consequently, the predictions by a hypothesis of genuine decreasing effects (see Table 1) did not agree with the empirical findings (see rows 7–12 in Table 2 and Figures 2A–F). On the other hand, a hypothesis of spurious effects predicted the six effects on rows 7–12 in Table 2, and across Figures 2A–F, to be (1) negative, (2) negative, (3) zero or positive, (4) negative, (5) negative, and (6) zero or positive, respectively (see Table 1). These predictions by a hypothesis of spurious effects agreed with the observed effects (see rows 7–12 in Table 2 and Figures 2A–F).

The results above are for total eating pathology, i.e., for an aggregation of various eating disorder related variables. Krauss et al. (2023) also analyzed meta-analytic data on specific eating disorder related variables, namely restrained eating, binge eating, bulimia, eating concerns, body dissatisfaction, and drive for thinness. In reanalyses we found mainly a similar pattern of effects for these specific variables as for total eating pathology, i.e., (1) A negative effect of initial self-esteem on subsequent eating pathology when adjusting for initial eating pathology and vice versa (indicating decreasing effects); (2) A negative effect of initial self-esteem on initial eating pathology when adjusting for subsequent eating pathology and vice versa (indicating increasing effects); (3) A positive effect of initial self-esteem on the subsequent-initial eating pathology difference and vice versa (indicating increasing effects) (see Supplementary Tables S1–S6 in results at https://osf. io/g8dkc/).

Discussion

The objective of the present study was to reanalyze the meta-analytic data used by Krauss et al. (2023) in order to evaluate if the identified adjusted cross-lagged associations suggested genuine decreasing effects between self-esteem and eating disorders or if the associations may have been spurious. The disparate findings of simultaneous and paradoxical decreasing and increasing effects indicated that prospective effects between initial and subsequent self-esteem and eating pathology were spurious, presumably due to correlations with residuals and regression to the mean. Among individuals with the same initial eating disorder score, those with a higher initial self-esteem score may be suspected to have received a too high eating disorder score, i.e., a positive residual, or those with a lower self-esteem score may have received a too low eating disorder score, i.e., a negative residual. However, as residuals tend to regress toward a mean value of zero between measurements, individuals with a higher initial self-esteem score should be expected to experience a more negative, but spurious, change in the eating disorder score between measurements compared with individuals with the same initial eating disorder score but with a lower initial self-esteem score. This might explain the negative effect of initial self-esteem on subsequent eating disorders when adjusting for initial eating disorders. The effect of initial eating disorders on subsequent self-esteem when adjusting for initial self-esteem could be given a similar explanation based on correlations with residuals and regression to the mean.

In a taxonomy of treatment effects (Sorjonen et al., 2024a), the present results do not agree with scenarios with a true decreasing effect of a treatment (e.g., initial self-esteem) on an outcome (e.g.,

TABLE 3 Meta-analytic cross-lagged panel analyses, and their conclusions, that we have reanalyzed and challenged.

Challenged study	Challenged presented prospective effect	Challenging reanalysis
Dapp et al. (2023)	Reciprocal between general and domain-specific self-esteem (positive effects)	Sorjonen and Melin (2024b)
Giletta et al. (2021)	Initial peer behavior on subsequent target youth behavior (positive effect)	Sorjonen et al. (2023c)
Harris and Orth (2020)	Reciprocal between self-esteem and quality of social relations (positive effects)	Sorjonen et al. (2023a)
Krauss and Orth (2022)	Reciprocal between self-esteem and work experiences (positive and negative effects)	Sorjonen et al. (2023b)
Prieto-Fidalgo et al. (2022)	Initial mindfulness on subsequent anxiety and depressive symptoms (negative effects)	Sorjonen and Melin (2024a)
Wang et al. (2021)	Reciprocal between social support and posttraumatic stress disorder (negative effects)	Sorjonen and Melin (2023)
Wu et al. (2021)	Reciprocal between academic self-concept and academic achievement (positive effects)	Sorjonen et al. (2024b)

eating disorders). However, the results do agree with a situation where data are generated as in Figure 1, without any direct effects between self-esteem and eating disorders, if we assume an influence by state factors resulting in stronger (more negative) concurrent compared with cross-lagged correlation which, in turn, would result in a positive association between initial self-esteem and subsequent change in eating disorders and vice versa (Equation 2), which is what we found in the present analyses. Actually, in the taxonomy of Sorjonen et al. (2024a), the present results agree with some scenarios where self-esteem would have an increasing, rather than a decreasing, effect on eating disorders and vice versa. However, we prefer the more cautious conclusion that the associations presumably were spurious. Consequently, Krauss et al.s' suggestion of a vicious circle, where low self-esteem makes people more vulnerable to developing eating disorders and eating disorders, in turn, scars individuals' self-esteem, may be challenged.

The present findings carry some clinical relevance as they warn against assuming causal effects between low self-esteem and eating pathology. This suggests that measures to improve self-esteem in order to prevent eating pathology, and vice versa, may not be the best use of limited resources. However, our conclusion that self-esteem might not have a decreasing effect on eating pathology is, admittedly, probably of limited practical clinical relevance in itself. But, if challenges of potentially clinically relevant but unsubstantiated claims (e.g., "Wearing an amulet of gold alleviates depression.") are not allowed, with the argument that the challenges are not clinically relevant ("What is the practical clinical relevance of knowing that wearing an amulet of gold probably does not alleviate depression?"), we risk creating a situation with dire consequences.

The present study is part of a series where we have reanalyzed meta-analyses where researchers have extracted zero-order correlations between constructs measured at two occasions from included studies and estimated adjusted cross-lagged regression effects with Equation 1 (Table 3). Based on our reanalyses, we have concluded that prospective effects were spurious due to correlations with residuals and regression to the mean and, consequently, challenged suggestions of genuine increasing or decreasing prospective effects by authors of the original meta-analyses. This means that the methodological point made in the present paper has been made before, also by others (e.g., Glymour et al., 2005; Eriksson and Häggström, 2014; Castro-Schilo and Grimm, 2018; Lucas, 2023), and that the degree of methodological novelty is limited. However, as researchers continue to overinterpret findings from cross-lagged panel models, despite repeated mentions of biases and flaws in such models, we believe the point is worth repeating. We do not think that it would

be tenable to argue that because bias and flaws in adjusted cross-lagged effects, meta-analytically aggregated or not, has been pointed out before, researchers should be allowed to use the method and overinterpret and publish findings without being challenged. Hopefully, we might help spread knowledge that adjusted cross-lagged effects often do not allow causal inference any more than zero-order correlations do.

When claiming, as we do here, that other researchers probably have overinterpreted findings due to limitations in used statistical methods, it is common to be asked what statistical methods should have been used instead. This suggests that it is common to believe that inference about true causal effects is possible if using the right statistical method, even with correlational, i.e., non-experimental, data. The random-intercept cross-lagged panel model (RI-CLPM) is an extension, and supposed improvement, of the traditional cross-lagged panel model (Hamaker et al., 2015; Mulder and Hamaker, 2021). Compared with the traditional model, RI-CLPM adjusts for individuals' trait-like levels on the constructs and estimates auto-regressive and cross-lagged effects between within-individual residuals not accounted for by the trait-like levels. Consequently, effects are estimated within, rather than between, individuals, which presumably allows more confident inference of causality (Usami et al., 2019; Raymaekers et al., 2020; Bailey et al., 2024). However, there are indications that RI-CLPM can be susceptible to bias (Lüdtke and Robitzsch, 2021; Lucas, 2023; Sorjonen et al., 2023d; Murayama and Gfrörer, 2024), meaning that statistically significant effects in RI-CLPM should not be viewed as irrefutable evidence of causality. For example, Sorjonen et al. (2024c) showed that RI-CLPM is susceptible to spurious findings when scores on the two constructs are affected by common auto-correlated state factors, something that might be quite common in psychological research. It should be noted that we could not conduct analyses with RI-CLPM in the present study, as it requires data from at least three waves of measurement. In the present reanalyses we reanalyzed meta-analytic data used by Krauss et al. (2023), which only contained data from two waves of measurement.

Broadly speaking, we do not share the optimism that strong causal inference is possible in analyses of correlational data (see Hammerton and Munafò, 2021; Munafò and Davey Smith, 2018 for a similar opinion). We believe that causal inference is more a question about how data were collected than how they were analyzed. If data were collected with a randomized controlled trial (RCT) with high quality (e.g., adequate sample size and randomization, independent assessments, reliable and valid measures, etc.), a statistically significant

treatment condition by time interaction-effect may suggest a causal effect of the treatment. Greater emphasis on experimental designs, such as longitudinal studies with frequent sampling or controlled intervention experiments, could provide more reliable data to investigate causal relationships between self-esteem and eating disorders.

The same confidence about causal inference as in RCTs can probably never be achieved by analyses of correlational, i.e., non-experimental, data. However, with correlational data researchers could, as we did in the present study, analyze models that predict different signs of effects/associations depending on if they are genuinely increasing/decreasing or spurious, i.e., use triangulation (Munafò and Davey Smith, 2018; Hammerton and Munafò, 2021). If findings converge, inference about causality is supported (although never finally proven). If, on the other hand and as in the present study, findings diverge, conclusions about causality would appear premature.

All this said, we still agree with Grosz et al. (2020) that it may be better if researchers make explicit causal claims even when analyzing correlational data, rather than implicit ones in the form of policy recommendations etc., because explicit claims can be challenged. However, this would require a research climate where challenges are welcomed. Unfortunately, today's scientific atmosphere, including publication and citations, appears to favor statistically significant findings and positive claims (Song et al., 2009; Ioannidis et al., 2014; Antonakis, 2017; Duyx et al., 2017), which means that challenging studies advising caution risk being rejected and ignored. We predict that Krauss et al. (2023) conclusion, that self-esteem and eating disorders reciprocally affect each other, will receive many more citations than our present warning that their findings may have been spurious due to correlations with residuals and regression to the mean.

Limitations

The present reanalysis suffers from some of the same limitations as the original meta-analysis by Krauss et al. (2023). For example, samples in the included studies were predominantly from Western cultures, of White ethnicity, and of medium to high socioeconomic status. Therefore, it is unclear if and to what degree our main conclusion, that prospective effects between self-esteem and eating disorders appear to be spurious due to correlations with residuals and regression to the mean, applies to other cultural, ethnic, and socioeconomic contexts. A review of literature on eating disorders in different cultural groups showed mixed findings and concluded that it is still unclear if the presentation of eating disorders differs across cultures (Soh et al., 2006).

Used instruments, timing of measurements, the study sample, etc. in the included studies may not always have been optimal. However, it is important to bear in mind that such factors were constant across the analyzed models and cannot, consequently, explain why the models suggested simultaneous and paradoxical decreasing and increasing prospective effects between self-esteem and eating disorders. Therefore, such possible methodological deficiencies should not be a threat against our main conclusion that meta-analytic prospective effects between self-esteem and eating disorders were spurious. We do not think that it would be tenable to argue that due to possible methodological deficiencies in the included studies we should assume genuine causal decreasing reciprocal effects between self-esteem and eating disorders. On the contrary, possible methodological deficiencies should make us even more cautious to assume causality.

We do not claim to have proven, once and for all, that self-esteem and eating pathology have no genuine decreasing effects on each other. The present findings agreed with predictions from a data generating model without any genuine direct effects between self-esteem and eating pathology, like the one in Figure 1, and did not agree with predictions of a hypothesis of genuine decreasing effects. Nevertheless, it is possible that data were generated by some alternative model. However, we do claim that the meta-analytic data analyzed by Krauss et al. (2023) could have been generated by a model without any genuine direct effects between self-esteem and eating pathology, like the one in Figure 1, and that it, therefore, is premature to assume a true decreasing effect between these constructs. We do not think that it would be tenable to argue that because we cannot be absolutely certain that effects between self-esteem and eating disorders are spurious, we should assume that they are genuinely decreasing. It should also be pointed out that, as mentioned above, the meta-analytic data analyzed by Krauss et al., and reanalyzed by us here, actually cannot rule out increasing prospective effects between self-esteem and eating disorders (Sorjonen et al., 2024a).

The present paper was not meant as a comprehensive review of theories, causes, and consequences of self-esteem and eating disorders. Rather, and more specifically, the objective was to reanalyze the meta-analytic data used by Krauss et al. (2023) in order to evaluate if the identified adjusted cross-lagged associations suggested genuine decreasing effects between self-esteem and eating disorders or if the associations may have been spurious due to correlations with residuals and regression to the mean. Readers interested in reviews of research on self-esteem are recommended to read, for example, Leary and MacDonald (2003), Orth and Robins (2014), and Pyszczynski et al. (2004). For reviews of eating disorders, we recommend Schaumberg et al. (2017) and Treasure et al. (2020).

Conclusion

The present reanalyses of meta-analytic data used by Krauss et al. (2023) indicated that reciprocal prospective decreasing effects between self-esteem and eating disorders appear to have been spurious, presumably due to correlations with residuals and regression to the mean. Consequently, we suggest that the conclusion by Krauss et al. regarding a reciprocal-relations model of low self-esteem and eating disorders is not supported. It is important for researchers not to overinterpret findings when analyzing correlational, i.e., non-experimental, data. Neither should results from meta-analyses be taken at face value, as they may be nothing more than aggregations of biased findings. Our results suggest caution in interpreting the prospective effects between self-esteem and eating disorders. However, future studies with more robust methodological approaches may better clarify these dynamics.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://osf.io/g8dkc/.

Author contributions

KS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. AG: Conceptualization, Investigation, Supervision, Validation, Writing – review & editing. BM: Conceptualization, Investigation, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antonakis, J. (2017). On doing better science: from thrill of discovery to policy implications. *Leadersh. Q.* 28, 5–21. doi: 10.1016/j.leaqua.2017.01.006

Bailey, D. H., Hübner, N., Zitzmann, S., Hecht, M., and Murayama, K. (2024). Illusory traits: wrong but sometimes useful. *Psychol. Rev.* doi: 10.1037/rev0000522

Castro-Schilo, L., and Grimm, K. J. (2018). Using residualized change versus difference scores for longitudinal research. *J. Soc. Pers. Relat.* 35, 32–58. doi: 10.1177/0265407517718387

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences. *Third* Edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Colmsee, I.-S. O., Hank, P., and Bošnjak, M. (2021). Low self-esteem as a risk factor for eating disorders: a meta-analysis. *Z. Psychol.* 229, 48–69. doi: 10.1027/2151-2604/a000433

Dapp, L. C., Krauss, S., and Orth, U. (2023). Testing the bottom-up and top-down models of self-esteem: a meta-analysis of longitudinal studies. *J. Pers. Soc. Psychol.* 124, 1111–1131. doi: 10.1037/pspp0000444

Diotaiuti, P., Girelli, L., Mancone, S., Valente, G., Bellizzi, F., Misiti, F., et al. (2022). Psychometric properties and measurement invariance across gender of the Italian version of the tempest self-regulation questionnaire for eating adapted for young adults. *Front. Psychol.* 13:941784. doi: 10.3389/fpsyg.2022.941784

Duyx, B., Urlings, M. J. E., Swaen, G. M. H., Bouter, L. M., and Zeegers, M. P. (2017). Scientific citations favor positive results: a systematic review and meta-analysis. *J. Clin. Epidemiol.* 88, 92–101. doi: 10.1016/j.jclinepi.2017.06.002

Eriksson, K., and Häggström, O. (2014). Lord's paradox in a continuous setting and a regression artifact in numerical cognition research. *PLoS One* 9:e95949. doi: 10.1371/journal.pone.0095949

Giletta, M., Choukas-Bradley, S., Maes, M., Linthicum, K. P., Card, N. A., and Prinstein, M. J. (2021). A meta-analysis of longitudinal peer influence effects in childhood and adolescence. *Psychol. Bull.* 147, 719–747. doi: 10.1037/bul0000329

Glymour, M. M., Weuve, J., Berkman, L. F., Kawachi, I., and Robins, J. M. (2005). When is baseline adjustment useful in analyses of change? An example with education and cognitive change. *Am. J. Epidemiol.* 162, 267–278. doi: 10.1093/aje/kwi187

Grosz, M. P., Rohrer, J. M., and Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspect. Psychol. Sci.* 15, 1243–1255. doi: 10.1177/1745691620921521

Guilford, J. P. (1965). Fundamental statistics in psychology and education. *4th* Edn. New York: McGraw-Hill.

Hamaker, E. L., Kuiper, R. M., and Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychol. Methods* 20, 102–116. doi: 10.1037/a0038889

Hammerton, G., and Munafò, M. R. (2021). Causal inference with observational data: the need for triangulation of evidence. *Psychol. Med.* 51, 563–578. doi: 10.1017/S0033291720005127

Harrer, M., Cuijpers, P., Furukawa, T. A., and Ebert, D. D. (2022). Doing meta-analysis with R: A hands-on guide. *First* Edn. Boca Raton London New York: CRC Press, Taylor & Francis Group.

Harris, M. A., and Orth, U. (2020). The link between self-esteem and social relationships: a meta-analysis of longitudinal studies. *J. Pers. Soc. Psychol.* 119, 1459–1477. doi: 10.1037/pspp0000265

Ioannidis, J. P. A., Munafò, M. R., Fusar-Poli, P., Nosek, B. A., and David, S. P. (2014). Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. *Trends Cogn. Sci.* 18, 235–241. doi: 10.1016/j.tics.2014.02.010

Kenrick, D. T., Neuberg, S. L., and Cialdini, R. B. (2007). Social psychology: Goals in interaction. *4th* Edn. Boston: Pearson.

Krauss, S., Dapp, L. C., and Orth, U. (2023). The link between low self-esteem and eating disorders: a meta-analysis of longitudinal studies. *Clin. Psychol. Sci.* 11, 1141–1158. doi: 10.1177/21677026221144255

Krauss, S., and Orth, U. (2022). Work experiences and self-esteem development: a meta-analysis of longitudinal studies. *Eur. J. Personal.* 36, 849–869. doi: 10.1177/08902070211027142

Leary, M. R., and MacDonald, G. (2003). "Individual differences in self-esteem: a review and theoretical integration" in Handbook of self and identity Eds. Leary, M. R, and Tangney, J. P. (New York, NY: The Guilford Press), 401–418.

Lucas, R. E. (2023). Why the cross-lagged panel model is almost never the right choice. *Adv. Methods Pract. Psychol. Sci.* 6:25152459231158378. doi: 10.1177/25152459231158378

Lüdtke, O., and Robitzsch, A. (2021). A critique of the random intercept cross-lagged panel model. *PsyArXiv.* doi: 10.31234/osf.io/6f85c

Mulder, J. D., and Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Struct. Equ. Model. Multidiscip. J.* 28, 638–648. doi: 10.1080/10705511.2020.1784738

Munafò, M. R., and Davey Smith, G. (2018). Robust research needs many lines of evidence. *Nature* 553, 399–401. doi: 10.1038/d41586-018-01023-3

Murayama, K., and Gfrörer, T. (2024). Thinking clearly about time-invariant confounders in cross-lagged panel models: a guide for choosing a statistical model from a causal inference perspective. *Psychol. Methods.* doi: 10.1037/met0000647

Orth, U., and Robins, R. W. (2014). The development of self-esteem. *Curr. Dir. Psychol. Sci.* 23, 381–387. doi: 10.1177/0963721414547414

Prieto-Fidalgo, Á., Gómez-Odriozola, J., Royuela-Colomer, E., Orue, I., Fernández-González, L., Oñate, L., et al. (2022). Predictive associations of dispositional mindfulness facets with anxiety and depression: a meta-analytic structural equation modeling approach. *Mindfulness* 13, 37–53. doi: 10.1007/s12671-021-01756-9

Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J., and Schimel, J. (2004). Why do people need self-esteem? A theoretical and empirical review. *Psychol. Bull.* 130, 435–468. doi: 10.1037/0033-2909.130.3.435

R Core Team (2024). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Raymaekers, K., Luyckx, K., and Moons, P. (2020). A guide to improve your causal inferences from observational data. *Eur. J. Cardiovasc. Nurs.* 19, 757–762. doi: 10.1177/1474515120957241

Schaumberg, K., Welch, E., Breithaupt, L., Hübel, C., Baker, J. H., Munn-Chernoff, M. A., et al. (2017). The science behind the academy for eating disorders' nine truths about eating disorders. *Euro Eating Disord. Rev.* 25, 432–450. doi: 10.1002/erv.2553

Silverstone, P. H., and Salsali, M. (2003). Low self-esteem and psychiatric patients: part I – the relationship between low self-esteem and psychiatric diagnosis. *Ann. Gen Hosp. Psychiatry* 2. doi: 10.1186/1475-2832-2-2

Soh, N. L., Touyz, S. W., and Surgenor, L. J. (2006). Eating and body image disturbances across cultures: a review. *Eur. Eat. Disord. Rev.* 14, 54–65. doi: 10.1002/erv.678

Song, F., Parekh-Bhurke, S., Hooper, L., Loke, Y. K., Ryder, J. J., Sutton, A. J., et al. (2009). Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. *BMC Med. Res. Methodol.* 9:79. doi: 10.1186/1471-2288-9-79

Sorjonen, K., Arshamian, A., Lager, E., Nilsonne, G., and Melin, B. (2024a). A taxonomy of treatment effects in data with two waves of measurement and a promotion of triangulation. doi: 10.31234/osf.io/q47s6

Sorjonen, K., Ingre, M., Melin, B., and Nilsonne, G. (2023a). Unmasking artifactual links: a reanalysis reveals no direct causal relationship between self-esteem and quality of social relations. *Heliyon* 9:e20397. doi: 10.1016/j.heliyon.2023.e20397

Sorjonen, K., Ingre, M., Melin, B., and Nilsonne, G. (2024b). Questioning the reciprocal effects model of academic self-concept and achievement: a reanalysis of a meta-analysis of longitudinal studies and a simulation. *SAGE Open* 14:21582440241292826. doi: 10.1177/21582440241292826

Sorjonen, K., and Melin, B. (2023). Prospective associations between social support and posttraumatic stress disorder may be spurious: a re-meta-analysis of longitudinal studies and a simulation. *J. Affect. Disord.* 340, 174–180. doi: 10.1016/j.jad.2023.08.014

Sorjonen, K., and Melin, B. (2024a). Prospective effects of mindfulness on anxiety and depressive symptoms may be spurious: simulated reanalysis of a meta-analytic cross-lagged panel analysis. *PLoS One* 19:e0302141. doi: 10.1371/journal.pone.0302141

Sorjonen, K., and Melin, B. (2024b). Spurious prospective effects between general and domain-specific self-esteem: a reanalysis of a meta-analysis of longitudinal studies. *PLoS One* 19:e0298158. doi: 10.1371/journal.pone.0298158

Sorjonen, K., Melin, B., and Ingre, M. (2019). Predicting the effect of a predictor when controlling for baseline. *Educ. Psychol. Meas.* 79, 688–698. doi: 10.1177/0013164418822112

Sorjonen, K., Melin, M., and Melin, B. (2023b). Spurious prospective effects between work experiences and self-esteem: a reanalysis of a meta-analysis of longitudinal studies. *Eur. J. Personal.* 38, 534–543. doi: 10.1177/08902070231197876

Sorjonen, K., Nilsonne, G., Ghaderi, A., and Melin, B. (2024c). Spurious effects in random-intercept cross-lagged panel models: results from simulations and reanalyses of data on self-esteem and problematic eating behaviors used by Beckers et al. (2023). *PsyArXiv*. doi: 10.31234/osf.io/hferw

Sorjonen, K., Nilsonne, G., and Melin, B. (2023c). Distorted meta-analytic findings on peer influence: a reanalysis. *Heliyon* 9:e21458. doi: 10.1016/j.heliyon.2023.e21458

Sorjonen, K., Nilsonne, G., Melin, B., and Ingre, M. (2023d). Uncertain inference in random intercept cross-lagged panel models: an example involving need for cognition and anxiety and depression symptoms. *Personal. Individ. Differ.* 201:111925. doi: 10.1016/j.paid.2022.111925

Treasure, J., Duarte, T. A., and Schmidt, U. (2020). Eating disorders. *Lancet* 395, 899–911. doi: 10.1016/S0140-6736(20)30059-3

Usami, S., Murayama, K., and Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychol. Methods* 24, 637–657. doi: 10.1037/met0000210

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Soft.* 36. doi: 10.18637/jss.v036.i03

Wang, Y., Chung, M. C., Wang, N., Yu, X., and Kenardy, J. (2021). Social support and posttraumatic stress disorder: a meta-analysis of longitudinal studies. *Clin. Psychol. Rev.* 85:101998. doi: 10.1016/j.cpr.2021.101998

Wolen, A., Hartgerink, C., Hafen, R., Richards, B., Soderberg, C., and York, T. (2020). Osfr: an R interface to the open science framework. *JOSS* 5:2071. doi: 10.21105/joss.02071

Wu, H., Guo, Y., Yang, Y., Zhao, L., and Guo, C. (2021). A meta-analysis of the longitudinal relationship between academic self-concept and academic achievement. *Educ. Psychol. Rev.* 33, 1749–1778. doi: 10.1007/s10648-021-09600-1