Empirical Research Paper

Exploring Asymmetries in Self-Concept Change After Discrepant Feedback

Personality and Social Psychology Bulletin
1–14

DOI: 10.1177/01461672241232738
journals.sagepub.com/home/pspb
Sage

Franziska Brotzeller[1] **and Mario Gollwitzer**[1]

Abstract

Receiving self-relevant feedback that is discrepant from one's self-concept can lead to self-concept change. However, it is currently unclear whether positive or negative feedback has a larger effect on self-concept change. Across four studies (total $N = 1{,}438$), we demonstrate that intentions for self-concept change (Study 1) as well as actual self-concept change (Studies 2, 3, and 4) are larger (a) for larger discrepancies between self-concept and feedback and (b) for negative compared to positive discrepancies. Exploring these effects further in Study 4, we find no evidence that the opportunity for improvement influences whether self-concept change is positively or negatively biased. In sum, the present research provides consistent evidence for a negativity bias in self-concept change, investigates a theoretical explanation, and discusses alternative explanatory approaches.

Keywords

self-concept, self-concept change, performance feedback, negativity bias

Received May 30, 2023; revision accepted January 24, 2024

In everyday life, people often receive feedback about their traits, abilities, or physical appearance. Such feedback can be provided both in a formal setting (e.g., when receiving performance feedback at work) or in an informal setting (e.g., when being complimented on one's cooking skills at a private dinner party). In both settings, self-relevant feedback can shape people's self-concept, defined as a person's perception of themselves (Bem, 1972; Shavelson et al., 1976). However, the extent to which discrepant external feedback leads to changes in such self-perceptions (i.e., self-concept change) varies considerably: In some cases, receiving external feedback leads to self-concept change in accordance with the feedback, while, in other cases, even highly discrepant feedback does not lead to self-concept change. When receiving the feedback that one is a very good cook, for example, one might accept the feedback and adapt one's self-concept accordingly, or one might attribute the successful dish to a very detailed recipe and stick to the belief that one is a mediocre cook. Although a considerable amount of research has examined self-concept change after self-relevant feedback, many unanswered questions remain.

Research in the areas of social and clinical psychology suggests that the extent to which people change their self-concept in accordance with the feedback they received depends, among other factors, on (a) characteristics of the source of the feedback (e.g., the expertise of the person giving the feedback), (b) characteristics of the receiver (e.g., their self-esteem), and (c) features of the feedback itself (e.g., the discrepancy between feedback and self-perceptions; see Binderman et al., 1972; Kernis & Goldman, 2003; Shrauger & Schoeneman, 1979). In the present research, we focus on the latter of the three features and examine the *size* of the discrepancy between one's self-view and the feedback one receives as well as the *direction* of the discrepancy—that is, whether the feedback is positive and suggests an upward adjustment of one's self-concept (e.g., "I obviously cook better than I thought I would") or whether it is negative and suggests a downward adjustment (e.g., "I obviously cook worse than I thought I would"). As we will discuss in more detail in the following, research on the effects of discrepancy size on self-concept change has produced largely consistent findings: Larger discrepancies lead to more change. By contrast, research on the effects of discrepancy direction on self-concept change provides a stunningly inconsistent picture. Some studies have found larger self-concept change after positive than negative feedback, while others have observed exactly the opposite pattern. With the present research, we aim to contribute to this literature by (1) systematically examining the effect of positive and negative feedback on self-concept change under different conditions and by (2)

[1]Ludwig-Maximilians-Universität München, Germany

Corresponding Author:
Franziska Brotzeller, Department of Psychology, Ludwig-Maximilians-Universität München, Leopoldstraße 13, 80802 Munich, Germany.
Email: f.brotzeller@psy.lmu.de

testing a theoretical explanation for the inconsistent previous findings. In the process, we also provide further evidence on the effect of the size of discrepancy on self-concept change.

Discrepant Feedback and Self-Concept Change

Before reviewing the relevant literature on self-concept change, it is helpful to clarify which terms and conceptualizations have been used to describe (changes in) people's self-concept in different research areas and how we define these constructs here. Early research in the educational context uses the term self-concept to describe a person's self-perceptions regarding specific or more global self-relevant dimensions (Marsh & Shavelson, 1985; Shavelson et al., 1976). This term has been adopted by more recent research on changes in self-perceptions (Elder et al., 2022; Korn et al., 2012). The management literature often refers to (updating of) self-beliefs when examining how new self-relevant information impacts people's self-perceptions (Eil & Rao, 2011; Ertac, 2011; Möbius et al., 2022), while clinical research often speaks of (updating of) expectations (Kube et al., 2019, 2022). Our own definition of self-concept change builds upon the definition by Shavelson et al. (1976) and attempts to be even more precise: We argue that self-concept change has occurred whenever a person's perception of themselves on a specific self-relevant dimension at a given time point differs from a previous self-perception on the same dimension.

In addition, we should also clarify what we mean by feedback, given that this term is central for the research presented here: Feedback means any kind of external information that a person receives on a self-relevant dimension (e.g., on a trait or an ability). Importantly, this feedback must be perceived as diagnostically relevant for this specific dimension: If Anna thinks she is a mediocre cook, and Peter says "what a great dish" while tasting the dinner she prepared, then Peter's feedback is more diagnostically relevant for Anna's self-concept (regarding her cooking skills) than if Peter had remarked that he had already eaten something similar the other day. While giving and decoding feedback can entail misunderstandings, the extent to which feedback is quantitatively discrepant from one's self-concept on a specific dimension is often unambiguous, especially when both one's self-concept and the feedback are quantifiable (e.g., for performance expectations and tests).

Studies investigating the effect of the size of the discrepancy between self-concept and feedback consistently demonstrate that larger discrepancies lead to more self-concept change (i.e., larger differences between previous and current self-perceptions) than smaller discrepancies, except for extreme and likely implausible discrepancies (Bergin, 1962; Binderman et al., 1972; Kube et al., 2022). Regarding the direction of the discrepancy, however, the empirical findings are less conclusive. Several studies show that positive and negative feedback lead to different amounts of self-concept change. Interestingly, it is unclear which of the two types of feedback leads to larger self-concept change: The majority of studies find larger self-concept change after positive than after negative feedback, indicating a positivity bias in the processing of self-relevant information (Eil & Rao, 2011; Elder et al., 2022; Korn et al., 2012; Möbius et al., 2022). The term *positivity bias* hereby is not meant to imply that such processing of self-relevant information is irrational; we merely use it to describe cases in which positive feedback produces more self-concept change than negative feedback. Notably, two recent studies demonstrate larger self-concept change after negative than after positive feedback (Ertac, 2011; Müller-Pinzler et al., 2019)—a pattern that rather suggests a negativity bias.

To investigate self-concept change, all studies mentioned above-assessed participants' self-perceptions before and after presenting them with (discrepant) feedback. The two studies that have found a negativity bias produced discrepant feedback as follows: Ertac (2011) presented participants with performance feedback on several rounds of math and verbal problems. Müller-Pinzler et al. (2019) asked participants to estimate the properties of different objects over several rounds and presented them with fake feedback on their performance. Comparing these two studies to the ones that have found a positivity bias and identifying meaningful differences is difficult as the studies differ from each other in many aspects (e.g., different types of feedback were given regarding various aspects of the self-concept; Eil & Rao, 2011; Elder et al., 2022; Ertac, 2011; Müller-Pinzler et al., 2019). That said, it is worth mentioning that the two studies that found a negativity bias both (1) assessed self-concept change in the intellectual ability domain (i.e., estimation, verbal, and math abilities), (2) confronted participants with performance feedback over multiple rounds, and (3) measured self-concept in a situation-specific way as they assessed participants' performance expectations for each upcoming round. Yet, the question whether a negativity bias also occurs in other contexts has remained unresolved so far. In the present research, we investigate under which conditions positively or negatively biased self-concept change occurs. In particular, we are interested in whether a negativity bias also occurs (1) on other aspects of the self-concept rather than performance expectations, (2) when presenting participants with feedback only once, and (3) when examining more generalized rather than situation-specific self-perceptions. In doing so, we also examine the effect of the size of the discrepancy on self-concept change. While we are mainly interested in the main effects of size and direction of discrepancy separately, we also explore whether they interact in producing self-concept change. Prior studies have largely neglected the possible interaction effects of these variables. However, it is plausible, for example, that the direction of discrepancy is only relevant for large discrepancies and less impactful for small discrepancies.

A Psychological Explanation for Asymmetries in Self-Concept Change: Self-Enhancement and Self-Improvement

There are different theoretical approaches to explaining the positivity and negativity bias in self-concept change and the contradictory findings that have resulted from previous research. One such approach focuses on two processes that shape how people perceive and integrate feedback into the self-concept: self-enhancement and self-improvement. Both self-enhancement and self-improvement assume that people are motivated to maintain a positive view of themselves even (or particularly) in the face of disconfirming feedback (Taylor & Brown, 1988). While *self-enhancement* describes biases in processing and interpreting information in a self-serving fashion (Heine & Hamamura, 2007), *self-improvement* describes biases aimed at reducing discrepancies between an "is-state" and a desirable "ought-state" (Kurman, 2006).

When a person receiving feedback is motivated to self-enhance, they should focus on positive and dismiss negative information as the latter is perceived as threatening one's positive self-view. Therefore, a self-enhancement motive should lead to positively biased self-concept change. When a person receiving feedback is motivated to self-improve, however, negative feedback is more informative than positive feedback because the former highlights opportunities for improvement. In other words, such a person should be negatively biased in changing their self-concept. This is consistent with theoretical accounts of a general negativity bias in human perception, behavior, and decision-making. Such accounts argue that learning from negative stimuli is more adaptive than learning from positive stimuli (Norris, 2021; Vaish et al., 2008): Avoiding negative consequences in the future is often more vitally important than approaching positive consequences. A negativity bias in learning from self-relevant feedback might serve a similar purpose as focusing on negative feedback promotes learning from one's shortcomings and might therefore be advantageous in the long run.

While self-improvement is triggered in particular when a person perceives that they can overcome is-ought discrepancies (e.g., by practicing or rehearsing), self-enhancement should be triggered when a person perceives it as impossible to improve on the self-concept aspect in question (Müller-Pinzler et al., 2019). When the aspect of the self-concept is perceived as fixed and unimprovable, negative feedback does not have an informational value toward improving oneself, but is, instead, particularly threatening to one's positive self-view (Dunning, 1995; Dweck et al., 1995; Levy & Dweck, 1998). In such cases, the only possibility of maintaining one's positive self-view is to self-enhance. Perceiving little opportunity for improvement should therefore trigger self-enhancement and produce positively biased self-concept change. Supporting this theorizing, a positivity bias—reflecting a self-enhancement process—has been empirically demonstrated on those self-concept aspects that are most likely to be perceived as fixed and unchangeable by most people (e.g., intelligence or beauty, see Eil & Rao, 2011; Möbius et al., 2022) or if the study was designed such that participants likely saw little opportunity for improvement (e.g., one-shot feedback from third parties; see Elder et al., 2022; Korn et al., 2012). These findings are also consistent with other studies on belief updating after feedback (Lefebvre et al., 2017).

By contrast, when the self-concept aspect in question is perceived as improvable ("malleable"), negative feedback is more informative for self-improvement purposes than positive feedback (Strube, 2012). Perceiving a high opportunity for improvement should trigger self-improvement motives and, thus, make a negativity bias (regarding the effect of feedback on self-concept change) more likely. This may explain the effects that Ertac (2011) and Müller-Pinzler et al. (2019) reported: Participants in these studies may have perceived the respective aspects of their self-concept (i.e., estimation, verbal, and math skills) as improvable and may have seen an opportunity to improve on the respective ability due to the repeated feedback over multiple rounds, which rendered negative feedback more informative than positive feedback.

Besides this explanation that focuses on how two motivational processes might shape feedback integration, there are other explanations for the contradictory findings on asymmetric self-concept change. One such explanation might be the diagnosticity of positive and negative self-relevant information (i.e., the informational value of positive and negative feedback for one's self-knowledge). If positive information is perceived as more diagnostic than negative information under certain conditions, this might lead to positively biased self-concept change, while, under other conditions, negative information might be perceived as more diagnostic, causing negatively biased self-concept change. Research on person perception has demonstrated that positive compared to negative information on another person is perceived as more or less diagnostic under different conditions (Unkelbach et al., 2020). If this were the case for self-relevant information as well, it might explain the contradictory findings on self-concept change.

To sum up, the opportunity for improvement in conjunction with motives for self-enhancement and self-improvement may be a plausible explanation for the contradictory findings on self-concept change after negative vs. positive feedback. Yet, this explanation has not been systematically examined so far. Therefore, the present research investigates the role of the opportunity for improvement in asymmetric self-concept change.

The Present Studies

The present research aims to contribute to the literature on self-concept change after self-relevant feedback by (1)

investigating under which conditions a positivity or negativity bias occurs as well as by (2) testing whether the opportunity for improvement causes positively or negatively biased self-concept change. More specifically, a low perceived opportunity for improvement should lead to positively biased self-concept change, while a high perceived opportunity for improvement should lead to negatively biased self-concept change—if self-enhancement and self-improvement play a role here.

In four studies, we investigated the effects of size of discrepancy (SoD) and direction of discrepancy (DoD) on the intention for and on actual self-concept change. Study 1 investigated intentions for self-concept change after self-relevant feedback using an autobiographic recall design. In Study 2, we examined actual self-concept change after participants had received (manipulated) self-discrepant feedback on a specific aspect of their self-concept. Study 3 was designed to replicate and extend Study 2 by examining a different aspect of participants' self-concepts in a more ecologically valid fashion. Finally, in Study 4, we investigated whether the perceived opportunity to improve is decisive in whether negative or positive self-relevant feedback is associated with a larger self-concept change.

All details regarding manipulations, measures, and exclusions for all four studies as well as the data and the R code necessary to replicate all primary analyses are available online at https://osf.io/yadqw/.

Study 1

Study 1 investigated the relationship between SoD and DoD and the intention for self-concept change in an exploratory manner. Using an autobiographical recall design, we asked participants to remember the last time they had received self-relevant feedback and assessed the characteristics of and participants' reactions to the feedback. This enabled us to examine feedback across a variety of contexts, formats, and aspects of the self-concept.[1]

Method

Sample. Participants were recruited through university and other mailing lists in exchange for raffled vouchers (2 vouchers worth 50 Euros). As for all following studies, the only eligibility criterion was an age of at least 18 years. The study was online for 4 weeks, and our sampling strategy was to collect as many data as possible during this period. A total of 360 individuals completed the survey, of which $n = 12$ were excluded because, when asked, they indicated that their data should not be used. The final sample thus comprised $N = 348$ participants ($M_{age} = 38.70$ years, $SD_{age} = 16.67$ years; 250 female, 94 male, four "other"). For our main analyses, we did not consider participants who indicated that their feedback was neither positive nor negative (see below) as they were not relevant to our research question. This left us with a sample of $n = 239$ participants for these analyses. With this sample size, we could detect an effect of $sr^2 = .04$ for the two predictor variables of interest (i.e., SoD and DoD) according to a sensitivity analysis conducted in G*Power ($\alpha = .05$, $1 - \beta = .80$, total sample size = 239, number of tested predictors = 2; Faul et al., 2007).

Measures. Participants responded to several measures, presented in the following order equivalent for all participants.

SoD. We measured the SoD of the discrepancy with one item ("Please remember how you evaluated yourself on the trait, ability, etc. prior to receiving the feedback. How much did the feedback deviate from your self-evaluation?"). Participants indicated their response on a scale from 1 = *not at all* to 7 = *very much* ($M = 3.05$, $SD = 1.66$).

DoD. We measured the DoD with one item ("Was the feedback more positive or negative than you would have rated yourself?"). Participants indicated whether the feedback was *more positive, more negative*, or *neither more positive nor more negative* by selecting one of these three response options (options chosen with frequency of 50%, 19%, and 31%, respectively).

Intentions for Self-Concept Change. We measured participants' intentions for self-concept change using three items based on a scale previously used by Henss and Pinquart (2022) to assess coping with violated expectations, "Based on the feedback, I have reconsidered or will reconsider my self-evaluation regarding the trait, ability, etc.," "The feedback has made me question whether my self-evaluation is correct," "The feedback had no impact on my self-evaluation" (reverse-coded); $\alpha = .74$. The items were rated on a scale from 1 = *strongly disagree* to 6 = *strongly agree* ($M = 3.12$, $SD = 1.34$).[2]

Results

As for all the following studies, we used *R* to conduct our analyses (R Core Team, 2018). To test whether SoD and/or DoD (as well as the interaction between the two) were significantly related to intentions for self-concept change, we conducted a regression analysis with SoD and DoD as well as their interaction as predictors of mean intentions for self-concept change. The DoD was effect-coded (i.e., "negative" = −1, "positive" = 1), and SoD was standardized to facilitate the interpretation of the regression coefficients. Overall, the model explained a significant amount of variance in intentions for self-concept change, $F(3, 235) = 24.59$, $p < .001$, $R^2 = .24$, 95% CI [.14, .32]. SoD was significantly related to intentions for self-concept change, $B = 0.22$, $t(235) = 2.63$, $p = .009$, 95% CI for B [0.06, 0.39], $sr^2 =$

.02, indicating that, across negative and positive discrepancies, larger discrepancies were associated with larger intentions for self-concept change. In addition, DoD was significantly related to intentions for self-concept change, $B = -0.47$, $t(235) = -5.19$, $p < .001$, 95% CI for B [−0.64, −0.29], $sr^2 = .09$, indicating that negative discrepancies were associated with larger intentions for self-concept change than positive discrepancies. Furthermore, a significant interaction emerged, $B = 0.37$, $t(235) = 4.37$, $p < .001$, 95% CI for B [0.20, 0.54], $sr^2 = .06$. Subsequent simple slopes analyses using the reghelper package (Hughes & Beiner, 2021) revealed that for positive discrepancies, SoD was positively related to self-concept change intentions, $B = 0.60$, $t(235) = 6.37$, $p < .001$, whereas no such relation emerged for negative discrepancies, $B = -0.15$, $t(235) = -1.04$, $p = .298$.[3]

Discussion

While our findings are in line with previous studies demonstrating a positive relationship between the SoD and self-concept change (Bergin, 1962; Binderman et al., 1972), we also find larger intentions for self-concept change after negatively than after positively discrepant feedback. This is especially noteworthy as we did not use a design in which participants were asked to repeatedly update performance expectations after receiving performance feedback. Instead, we asked participants about the most recent self-relevant feedback they had received without limitations regarding the content, context, or format of the feedback. Our findings thus provide the first evidence that a negativity bias might not be limited to the specific circumstances investigated by Ertac (2011) and Müller-Pinzler et al. (2019).

Even though Study 1 provides first evidence for a negativity bias in self-concept change after feedback, there are three limitations that need to be mentioned and discussed. First, we asked participants to recall the most recent feedback they had received, regardless of whether it was positive or negative. Different processing of positive and negative information or other recall errors might have biased our results. Second, we assessed intentions for self-concept change instead of actual self-concept change as the dependent variable. This was done as we wanted to examine a broad range of self-concept aspects, and it would not have been possible to assess participants' actual self-concepts on all possible aspects. However, we cannot be certain that intentions for self-concept change reflect patterns of actual self-concept change. Third, participants might have received their most recent feedback in a context similar to the ones in which a negativity bias was previously found. This seems unlikely, as it would mean that the majority of participants had most recently received repeated feedback on a performance task. However, we cannot be fully certain that this was not the case. Study 2 was designed to address these limitations.

Study 2

Study 1 provided evidence for a negativity bias in intentions for self-concept change. In Study 2, we aimed to examine actual self-concept change by presenting participants with feedback on a specific aspect of their self-concept and assessing subsequent self-concept change. To do so, we manipulated the SoD and DoD. Furthermore, we experimentally varied whether participants received feedback once or multiple times, resulting in a 2 × 2 × 2 design. The study was designed to investigate the effects of SoD and DoD on self-concept change under controlled conditions. Furthermore, we wanted to examine biases in self-concept change depending on whether participants receive feedback repeatedly compared to just once.

A preregistration detailing the study design, pre-planned stopping rule, and exclusion criteria is available at https://aspredicted.org/te8gc.pdf.[4]

Method

Sample. Participants were recruited through university and other mailing lists, social media, and flyers distributed on campus of a German university and could participate in a raffle for vouchers or receive course credit in return for their participation. Data were collected until the date specified in our preregistered stopping rule. In total, 627 participants completed the study. As preregistered, we excluded participants based on several exclusion criteria to ensure high data quality.[5] After following the preregistered exclusion criteria, the final sample included data from $N = 373$ participants ($M_{age} = 29.33$ years, $SD_{age} = 12.50$ years; 290 female, 77 male, five "other," one did not respond).[6] We conducted a sensitivity analysis using G*Power ($\alpha = .05$, $1 - \beta = .80$, total sample size = 373, number of tested predictors = 3; Faul et al., 2007) and discovered that we were able to detect an effect of $sr^2 = .03$ of any of the three predictors of interest in our main analysis (i.e., SoD and DoD, interaction DoD × frequency of feedback).

Procedure. Participants learned that the study would be about their spatio-visual ability, received an explanation of what this ability encompassed, and were told that it was normally distributed across the population. They were then asked to rate their ability (see below) and were immediately shown their self-perception score, which was created by converting participants' mean self-perception into a percentage. An exemplary feedback read: "On a scale of 0% (very low ability) to 100% (very high ability) your self-rated ability for spatio-visual thinking is at: 50%." At this point, participants were randomly assigned to two frequency of feedback conditions (once, three times). Depending on their condition, participants were asked to work on either one or three subsequent tasks measuring spatio-visual thinking. In these tasks, which were adapted from the Wiener Matrizen-Test 2 (WMT-2; Formann et al., 2011) and two subtests of

the Wilde Intelligenztest-2 (WIT-2; Kersting et al., 2008), participants were asked to mentally manipulate objects to find solutions to given questions (e.g., participants had to mentally fold sides of a cube). Each task contained 18 to 20 subtasks, and participants were asked to complete as many of them as they could within the given timeframe (2 minutes per task). Participants who worked on only one task were randomly assigned one of the three tasks. After completing a task, participants received (false) feedback about their performance in the test (i.e., their percentage of correctly solved subtasks; they received feedback either once or three times; see above). Specifically, participants were randomly assigned to one of two SoD conditions (small, large) and one of two DoD conditions (negative, positive). In case of a small SoD, participants received feedback that deviated from their self-perception score by around 5%; a large size of discrepancy referred to a deviation of around 20%.[7] For those assigned to the negative (positive) DoD condition, this number was subtracted from (added to) their self-perception score. An exemplary feedback read: "You correctly solved 70% of the task. As a reminder: Your self-perception was 50%." Afterward, participants were once again asked for their self-perception regarding their ability for spatio-visual thinking. At the end of the study, participants were debriefed and had the opportunity to learn their actual task score(s).

Measures

Self-Concept Change. Participants' self-perceptions were measured at two occasions with the same five items on a 9-point scale from 1 = *strongly disagree* to 9 = *strongly agree* (e.g., "I have no difficulty at all in imagining shapes and objects in my mind's eye"; $\alpha_{t1} = .82$, $M_{t1} = 5.65$, $SD_{t1} = 1.29$; $\alpha_{t2} = .90$, $M_{t2} = 5.34$, $SD_{t2} = 1.65$). To create the self-concept change score, we subtracted self-perceptions at t1 from self-perceptions at t2. As we were interested in the absolute amount of change, we created absolute values of this score, resulting in the final absolute self-concept change score used in all analyses ($M = 0.76$, $SD = 0.71$).

Perceived SoD. We assessed participants' perceived SoD with one item ("How large was the difference between your self-perception and your percentage of correctly solved tasks on spatio-visual reasoning?") on a scale from 1 = *small* to 9 = *large*. This item was used as a manipulation check for SoD ($M = 4.04$, $SD = 2.17$).

Perceived DoD. We measured whether participants had correctly perceived the DoD with one item ("On average, was your score in the tasks better or worse than your previously submitted self-perception? [Did you end up solving fewer or more tasks correctly?]"). Participants indicated whether they had performed *better* or *worse* in the tasks than they had indicated in their self-perception (options chosen with a frequency of 51% and 49%, respectively).

Perceived Frequency of Feedback. We assessed whether participants had correctly perceived the number of times they had received task feedback throughout the study with one item ("How many times did you in total receive feedback related to your performance in the spatial-visual reasoning tasks?"). Participants indicated whether they had received task feedback *once, three times*, or *not at all* (options chosen with a frequency of 46%, 49%, and 5%, respectively).[8]

Results

First, we tested whether our experimental manipulations were successful. A Welch two-sample *t*-test showed that participants in the large discrepancy condition perceived larger discrepancies ($M = 5.41$, $SD = 1.81$) than those in the small discrepancy condition ($M = 2.53$, $SD = 1.40$), $t(345.94) = -16.91$, $p < .001$, Cohen's $d = -1.77$, 95% CI [−2.02, −1.53] (effsize package; Torchiano, 2020). Furthermore, for both the negative and the positive DoD condition, 98% of participants correctly indicated their respective condition on our one-item measure. For the frequency of feedback conditions, 89% of participants who had received feedback once and 94% of participants who had received feedback three times correctly indicated so in response to the respective item.

For our main analysis, we conducted a regression analysis with SoD, DoD, frequency, and all possible interaction terms as predictors of absolute self-concept change (i.e., the absolute difference between self-perceptions at t2 and t1; see Method section for further information). All predictors were effect-coded (i.e., for SoD, "small" = −1 and "large" = 1; for DoD, "negative" = −1 and "positive" = 1; for frequency of feedback, "once" = −1 and "three times" = 1). The results are summarized in Table 1. Two effects turned out to be statistically significant: First, larger discrepancies led to more self-concept change than smaller discrepancies, $B = 0.18$, $p < .001$. Second, negative discrepancies led to more self-concept change than positive discrepancies, $B = -0.15$, $p < .001$. No other effects were significant.

Discussion

Our findings in Study 2 are consistent with our findings from Study 1 in that large SoDs lead to more self-concept change than small ones. In addition, we show that negative discrepancies are associated with more self-concept change than positive discrepancies regardless of the frequency of feedback, providing evidence that a negativity bias in self-concept change does not merely occur in contexts of repeated performance feedback. However, Study 2 only investigates one specific aspect of the self-concept. Therefore, we expand our research to another aspect of the self-concept in Study 3.

Table 1. Regression Analysis Summary for Size and Direction of Discrepancy as well as Frequency of Feedback Predicting Absolute Self-Concept Change in Study 2.

Predictor	*B*	*SE B*	*t*	*p*	95% CI for *B* [LL, UL]	sr^2
Size of Discrepancy (SoD)	0.18	0.03	5.15	<.001	[0.11, 0.25]	.06
Direction of Discrepancy (DoD)	−0.15	0.03	−4.43	<.001	[−0.22, –0.09]	.05
Frequency of Feedback (FoF)	0.05	0.03	1.46	.144	[−0.02, 0.12]	.01
SoD × DoD	−0.03	0.03	−0.78	.435	[−0.10, 0.04]	.00
SoD × FoF	0.04	0.03	1.15	.251	[−0.03, 0.11]	.00
DoD × FoF	−0.03	0.03	−0.93	.356	[−0.10, 0.04]	.00
SoD × DoD × FoF	0.01	0.03	0.34	.734	[−0.06, 0.08]	.00

Note. $R^2 = .12$ ($N = 373$, $p < .001$). SoD: small = −1, large = 1. DoD: negative = −1, positive = 1. FoF: once = −1, three times = 1. *B* represents unstandardized regression weights. LL and UL indicate the lower and upper limits of a confidence interval, respectively. sr^2 represents the squared semipartial correlation.

Study 3

In Study 3, we expand our findings from Study 2 by examining self-concept change after self-relevant feedback regarding a different aspect of the self-concept. To increase ecological validity, we presented participants with their actual instead of artificially created feedback, reflecting feedback in naturally occurring situations. The study was designed to further test the effects of SoD and DoD on self-concept change.

A preregistration for the study design, pre-planned stopping rule, and exclusion criteria can be found at https://aspredicted.org/4hr79.pdf.[9]

Method

Sample. The recruiting channels, sampling strategy, and participation rewards for Study 3 were equivalent to those of Study 2. Data were collected until the date preregistered in our stopping rule. When data collection was stopped, 463 individuals had participated in the complete study. As per our preregistered exclusion criteria, we excluded $n = 87$ participants, resulting in a final sample of $N = 376$ participants ($M_{age} = 35.29$ years, $SD_{age} = 14.42$ years; 271 female, 100 male, five "other"). Beyond the preregistered exclusions, equivalent to Study 1, we did not consider participants with neither positive nor negative discrepancies in our main analyses, resulting in the further exclusion of $n = 3$ participants and a sample of $n = 373$ for these analyses. A sensitivity analysis conducted with G*Power ($\alpha = .05$, $1 - \beta = .80$, total sample size = 373, number of tested predictors = 2; Faul et al., 2007) suggested that an effect of $sr^2 = .03$ can be detected with this sample size.

Procedure. Participants were told that this study would be about their emotion-recognition skills—the ability to correctly identify an emotion experienced by a target person based on this person's eye area. The study procedure was similar to that of Study 2 in that participants first indicated their self-perceived emotion-recognition skills, which were then feedbacked to them in a percentage format. Afterward, participants completed a short version of the Reading the Mind in the Eyes Test (Bölte, 2005). In this test, participants were repeatedly presented with photographs of human eye areas and tasked with choosing the correct out of four possible emotions felt by the person in the photograph. After completing the task, participants received feedback on their task score, the percentage of correctly chosen emotions in the task.[10] Contrary to Study 2, the feedbacks reflected participants' actual task scores.[11] Afterwards, participants were once again asked about their self-perceived emotion-recognition skills. At the end of the study, participants had the opportunity to complete and receive feedback on the full Reading the Mind in the Eyes Test (Bölte, 2005).

Measures

Self-Concept Change. Participants' self-perceptions were measured at two occasions with the same four items on a 9-point scale from 1 = *strongly disagree* to 9 = *strongly agree* (e.g., "It's very easy for me to read a person's emotions from their eyes"; $\alpha_{t1} = .91$, $M_{t1} = 4.94$, $SD_{t1} = 1.62$; $\alpha_{t2} = .94$, $M_{t2} = 4.80$, $SD_{t2} = 1.61$). Absolute self-concept change scores were calculated following the same approach as in Study 2 ($M = 0.89$, $SD = 0.94$).

Perceived SoD. The perceived SoD between self-perceptions and feedback was measured with one item ("How large did you perceive the difference between the feedback on your self-perception and the photo task to be?") on a 9-point rating scale from 1 = *very small* to 9 = *very large* ($M = 5.17$, $SD = 2.30$).

Results

In preparation for our analyses, we created scores for SoD and DoD. Since their percentage scores for self-perceived ability and their actual scores in the emotion-recognition task were scaled identically, we subtracted participants'

self-perception percentage scores at t1 from their task scores. The absolute values of this variable served as the SoD. The variable was also used to create the DoD variable: Negative scores were coded as *negative* and positive scores as *positive* DoDs. Cases with no discrepancy between task sore and self-perception at t1 were coded as *neither positive nor negative* discrepancies.

First, we checked whether participants' perceptions of the SoD were consistent with the actual SoD between self-perceptions and feedback. The correlation between the perceived and the actual SoD was $r(374) = .73$, $p < .001$.

For our main analysis, we fitted a linear regression model with SoD and DoD as well as the interaction between the two as predictors of absolute self-concept change. Again, DoD was effect-coded (i.e., "negative" = −1, "positive" = 1). SoD was standardized on the sample mean and standard deviation. Overall, the model explained a significant amount of variance in self-concept change, $F(3, 369) = 6.90$, $p < .001$, $R^2 = .05$, 95% CI [.01, .10]. SoD was positively related to self-concept change, $B = 0.27$, $t(369) = 3.86$, $p < .001$, 95% CI for B [0.13, 0.41], $sr^2 = .04$. Moreover, DoD was significantly related to self-concept change, $B = -0.16$, $t(369) = -2.65$, $p = .009$, 95% CI for B [−0.28, −0.04], $sr^2 = .02$, such that negative discrepancies were associated with larger self-concept change than positive discrepancies. The SoD x DoD interaction effect was not significant, $B = -0.08$, $t(369) = -1.17$, $p = .245$, 95% CI for B [−0.22, 0.06], $sr^2 = .00$.

Discussion

Just as in the two previous studies, we find that (a) larger SoDs and (b) negative discrepancies are associated with more self-concept change. The latter finding lends further support to the notion of a negativity bias in self-concept change. Study 4 was designed to explore a potential mechanism underlying asymmetric self-concept change.

Study 4

Study 4 aimed to explore whether the perceived opportunity for improvement is the psychological mechanism underlying the negativity bias we have observed so far. To do so, we aimed to manipulate participants' subjective expectation that they can (vs. cannot) improve on the ability in question (in other words, whether they see an opportunity to improve or not). More specifically, half of the participants were led to believe that it is possible to improve on the ability in question (i.e., emotion-recognition skills), while the other half learned that emotion-recognition skills cannot be improved via rehearsal etc. With this manipulation and the result pattern it produces, we aimed at indirectly inferring whether self-enhancement and self-improvement play a role here: If the negativity bias was indeed due to self-improvement processes, then such a bias should occur in the high, but not in the low opportunity for improvement condition. In the latter condition, negative feedback should be uninformative and even threatening for the self. Here, self-enhancement processes should lead to a positivity bias. We therefore hypothesized and preregistered an interaction effect of DoD × opportunity for improvement on self-concept change. Furthermore, we hypothesized and preregistered a main effect of SoD on self-concept change.

A preregistration detailing the study design, pre-planned stopping rule, exclusion criteria, and planned analyses is available at https://aspredicted.org/rr4r7.pdf.

Method

Sample. Participants were recruited through university and other mailing lists in return for raffled vouchers. Data collection was stopped according to the preregistered stopping rule, with 548 individuals having completed the survey. Applying the preregistered exclusion criteria, we excluded $n = 89$ participants. The final sample thus comprised $N = 459$ participants ($M_{age} = 32.55$ years, $SD_{age} = 14.02$ years; 332 female, 119 male, 8 "other"). As preregistered, we did not consider participants whose feedback was neither positive nor negative for all analyses that included the DoD. This led to the exclusion of $n = 6$ participants and left us with a sample of $n = 453$ participants for these analyses. With this sample size, we could detect an effect of $sr^2 = .02$ for the two hypothesized effects (i.e., SoD and interaction DoD x opportunity for improvement) according to a sensitivity analysis conducted in G*Power ($\alpha = .05$, $1 - \beta = .80$, total sample size = 453, number of tested predictors = 2; Faul et al., 2007).

Procedure. The design of Study 4 was very similar to Study 3, with one important main difference[12]: immediately after giving participants feedback about their task performance (i.e., their percentage score from the short Reading the Mind in the Eyes Test adapted from Bölte, 2005; see Study 3), they were randomly assigned to one of two experimental conditions (opportunity for improvement: low, high). In the *high opportunity for improvement* condition, participants read the following information:

> "You have just received feedback on your ability to recognize emotions from people's eyes. Research in this area shows: People can change their ability to do this. People who train the ability perform better on subsequent tests of the ability than they did before."

In the *low opportunity for improvement* condition, they read the following information:

Table 2. Regression Analysis Summary for Size and Direction of Discrepancy as well as Opportunity for Improvement Predicting Absolute Self-Concept Change in Study 4.

Predictor	*B*	*SE B*	*t*	*p*	95% CI for *B* [LL, UL]	sr^2
Size of Discrepancy (SoD)	0.38	0.05	7.12	<.001	[0.28, 0.49]	.09
Direction of Discrepancy (DoD)	−0.38	0.05	−8.07	<.001	[−0.47, –0.28]	.12
Opportunity for Improvement (OfI)	0.03	0.05	0.65	.517	[−0.06, 0.12]	.00
SoD × DoD	−0.31	0.05	−5.68	<.001	[−0.41, –0.20]	.06
SoD × OfI	−0.00	0.05	−0.04	.972	[−0.11, 0.10]	.00
DoD × OfI	−0.04	0.05	−0.76	.449	[−0.13, 0.06]	.00
SoD × DoD × OfI	0.02	0.05	0.30	.762	[−0.09, 0.12]	.00

Note. $R^2 = .17$ ($N = 453$, $p < .001$). SoD values are scaled. DoD: negative = −1, positive = 1. OfI: low = −1, high = 1. *B* represents unstandardized regression weights. LL and UL indicate the lower and upper limits of a confidence interval, respectively. sr^2 represents the squared semipartial correlation.

> "You have just received feedback on your ability to recognize emotions from people's eyes. Research in this area shows: People can hardly change their ability to do this. People who train the ability do not perform better on subsequent tests of the ability than they did before."

Study 4 further differed from Study 3 in that we conducted a manipulation check regarding the opportunity for improvement manipulation after measuring participants' self-perceptions at t2.[13] Afterward, like in Study 3, participants were debriefed and could complete and receive feedback on the full Reading the Mind in the Eyes Test (Bölte, 2005).

Measures

Self-Concept Change. Participants' self-perceptions were measured at two occasions using the same items as in Study 3 ($\alpha_{t1} = .94$, $M_{t1} = 5.10$, $SD_{t1} = 1.72$; $\alpha_{t2} = .94$, $M_{t2} = 4.84$, $SD_{t2} = 1.58$). Self-concept change scores were computed in the same fashion as in Study 3 ($M = 0.89$, $SD = 0.93$).

Perceived SoD . The perceived SoD was assessed using the same item as in Study 3 ($M = 5.33$, $SD = 2.32$).

Perceived Opportunity for Improvement. To check whether our opportunity for improvement manipulation was successful, we assessed the perceived opportunity for improvement with eight items adapted from De Castella and Byrne's (2015) revised scale for measuring implicit theories of intelligence (e.g., "I believe that I can significantly improve my ability to recognize emotions based on the eye area"; $\alpha = .94$). Participants indicated their agreement on a scale from 1 = *strongly disagree* to 9 = *strongly agree* ($M = 5.18$, $SD = 1.76$).

Results

Mirroring our approach from Study 3, we created scores for SoD and DoD. Like in Study 3, we first checked whether participants' perceptions of the SoD were consistent with the actual SoD between self-perceptions and feedback. The correlation between the perceived and the actual SoD was $r(457) = .75$, $p < .001$. Then, we checked whether our opportunity for improvement manipulation was successful by conducting a Welch two-sample *t*-test. The *t*-test revealed that the perceived opportunity for improvement was significantly higher in the high ($M = 5.94$, $SD = 1.52$) than in the low opportunity for improvement condition ($M = 4.45$, $SD = 1.67$), $t(455.7) = 10.01$, $p < .001$, Cohen's $d = 0.93$, 95% CI [0.74, 1.13] (effsize package; Torchiano, 2020).

To test our hypotheses, as preregistered, we conducted a regression analysis with SoD (absolute values; standardized on sample mean and standard deviation), DoD (negative = −1 and positive = 1), and opportunity for improvement (low = −1, high = 1) as well as all interaction terms as predictors of absolute self-concept change. Results are displayed in Table 2. As expected, SoD was significantly related to self-concept change, $B = 0.38$, $p < .001$. Furthermore, we found a significant main effect of DoD, $B = -0.38$, $p < .001$. Unexpectedly, the interaction effect between DoD and opportunity for improvement was not significant, $B = -0.04$, $p = .449$.

The SoD × DoD interaction effect was significant, $B = -0.31$, $p < .001$. We conducted simple slope analyses for a model including SoD and DoD as well as their interaction as predictors of self-concept change using the reghelper package (Hughes & Beiner, 2021) to further examine this interaction effect. For negative discrepancies, the relationship between SoD and self-concept change was significant, $B = 0.68$, $t(449) = 7.49$, $p < .001$. This was not the case for positive discrepancies, $B = 0.08$, $t(449) = 1.67$, $p = .095$.

Discussion

In Study 4, we replicate our findings from Study 3, showing larger self-concept changes for larger discrepancies and after negative compared to positive feedback. Moreover, we find no evidence that the opportunity for improvement leads to biases in self-concept change.

General Discussion

In the present research, we examined the effects of SoD and DoD on self-concept change and explored an explanation for asymmetric self-concept change after positive and negative feedback. We found that larger discrepancies led to more self-concept change than smaller discrepancies, which is in line with previous findings (Bergin, 1962; Binderman et al., 1972). Regarding the DoD, we found that negative feedback had a stronger impact on (intended or actual) self-concept change compared to positive feedback. This finding was consistent across different aspects of the self-concept (a variety of self-concept aspects in Study 1, spatio-visual thinking in Study 2, and emotion recognition abilities in Studies 3 and 4) and across paradigms (autobiographical recall in Study 1, manipulated task feedback of different frequencies in Study 2, and feedback reflecting naturally occurring task performance in Studies 3 and 4). As findings on the interaction effect between SoD and DoD were inconsistent, we do not interpret them here.

In addition, we aimed at testing whether two motives—self-enhancement and self-improvement—can explain our pattern of results by manipulating the opportunity for improvement regarding the trait in question in Study 4. The rationale behind this was that, if self-improvement and/or self-enhancement motives actually played a role, then a negativity bias should be more likely to occur in the "high opportunity for improvement" condition, whereas a positivity bias should be more likely to occur in the "low opportunity for improvement" condition. However, our results do not support this explanation as we find a negativity bias regardless of a low or high opportunity for improvement in Study 4.

Our manipulation check indicates that the manipulation successfully impacted the perceived opportunity for improvement as participants actually perceived a higher opportunity for improvement in the high than in the low opportunity for improvement condition. Notably, the experimental manipulation aimed at *indirectly* shaping participants' self-improvement or self-enhancement motives (i.e., self-improvement should only play a role in the high opportunity for improvement, self-enhancement in the low opportunity for improvement condition). That said, it is important to note that we did not *directly* manipulate self-enhancement and self-improvement motives in our study. Therefore, while our research does not suggest that these motives can explain contradictory findings on positively and negatively biased self-concept change, the specific role that they play here remains to be scrutinized more directly by future research.

Theoretical Implications

Our findings provide several contributions to the existing research on self-concept change after discrepant feedback. First, our findings challenge the assumption of a robust positivity bias in self-concept change. Previous studies by Ertac (2011) and Müller-Pinzler et al. (2019) have already demonstrated a negativity bias specifically for updates in (intellectual) performance expectations over multiple feedback rounds. The present research contributes to this literature by showing that a negativity bias occurs (1) for self-concept aspects outside the intellectual domain, (2) even after single instances of feedback, and (3) regarding generalized instead of situation-specific self-perceptions. The latter finding is consistent with theoretical accounts on the hierarchical structure of the self-concept as feedback regarding a specific task should not only impact one's situation-specific self-concept but also be indicative of more general aspects of the self-concept (Shavelson et al., 1976). In sum, the present research broadens the current knowledge on the conditions under which a negativity bias in self-concept change can emerge.

Second, the present research questions whether self-enhancement and self-improvement motives can explain when a positivity or negativity bias emerges in self-concept change. According to theorizing by Müller-Pinzler et al. (2019), these motives should be triggered by a high or low opportunity for improvement, respectively, and cause negatively or positively biased self-concept change. However, our results do not support the assumption that the opportunity for improvement is decisive in whether a positivity or negativity bias emerges. We find that negative information is overweighed compared to positive information regardless of a low or high opportunity for improvement (Study 4). As we did not directly manipulate motives for self-enhancement and self-improvement, we cannot make strong claims about the role that these motives play in biased self-concept change. Nonetheless, if the motives did play a role, we should have observed an interaction effect between the DoD and the opportunity for improvement manipulation. We found no such effect despite our study being sufficiently powered.

Besides motives for self-enhancement and self-improvement, a myriad of other motivational processes might be involved—although we have focused on the most prominently discussed motives and know of no other equally plausible candidates. Moreover, other factors such as the diagnosticity of positive and negative self-relevant feedback might be relevant in producing negatively or positively biased self-concept change. Research has shown that negative information is generally more impactful than positive information in person perception and in forming impressions of others (Unkelbach et al., 2020). One explanation for this effect is based on the differing properties of positive and negative information in our environment. More specifically, negative information is less frequent but more diverse, extreme, intense, and surprising than positive information (Leising et al., 2012; Unkelbach et al., 2020). This might lead to negative information being overweighed when forming and updating impressions of other people. Transferring these findings to the domain of self-concept change, it is

possible that negative information is perceived as more diagnostic in learning about oneself than positive information. While negative feedback does not necessarily contain information on how to improve, it conveys that there is room for or even the necessity to improve. Coupled with negative information being rarer and more unexpected, this might increase attention to and elaboration on negative feedback and induce a negativity bias. Still, the question arises as to why several other studies have found a positivity bias. In the literature on person perception, it has been shown that under specific conditions, positive information can be more diagnostic than negative information (e.g., for specific domains of traits; Unkelbach et al., 2020). It remains an avenue for future research to investigate the diagnosticity of self-relevant information and the conditions under which positive versus negative self-relevant information is more diagnostic.

Limitations and Future Research

The present research is subject to several limitations and raises issues to be addressed in future studies. One limitation is that some features of our research might limit the generalizability of our findings. While we broadly assessed real life feedback across a variety of contexts, formats, and aspects of the self-concept in Study 1, we examined intentions for instead of actual self-concept change for feasibility reasons. Aiming for a highly ecologically valid but standardized examination of actual self-concept change in Studies 3 and 4, we presented participants with real task feedback. The paradigm used in these studies is similar to how certain types of feedback for online-self-assessments or in an educational or work context are produced. However, it does not reflect the full range of feedback people receive in everyday life. To produce more generalizable insights, future research should investigate the effects of self-relevant feedback on actual self-concept change in natural settings, for example, using field experiments.

Furthermore, future research should systematically investigate the interplay of feedback, person, and study design characteristics in producing self-concept change, especially when aiming to investigate conditions under which a positivity or negativity bias emerges. While we systematically examined the role of certain aspects, such as the frequency of feedback, we used a similar paradigm and kept other context factors constant in Studies 2, 3, and 4. Conducting additional research systematically examining the impact of other factors such as content, source, and format of the feedback would be a further step in understanding self-concept change after self-relevant feedback. Moreover, we did not investigate the effects of personal characteristics even though there is evidence that they play an important role in reactions to feedback. People with depressive symptoms, for example, have been shown to be less optimistically biased in belief updating than those without such symptoms (Korn et al., 2014; Kube et al., 2019). Similar effects have been shown in the domain of social anxiety (Koban et al., 2017) and dispositional risk aversion (Niv et al., 2012). Furthermore, previous studies as well as exploratory analyses for the current studies (see Online Supplemental Materials) have demonstrated gender differences in self-concept change after discrepant feedback (Ertac, 2011; Möbius et al., 2022), some of them showing, for example, that women are less optimistic or even more pessimistic in updating their beliefs after feedback (Study 4 of the current research; Ertac, 2011). A person's confidence regarding their prior (i.e., initial self-concept) has been identified as a relevant determinant for their reaction to the feedback, as well (Ertac, 2011). Yet, in Study 2, we randomly assigned participants to feedback of different sizes and directions; at least in this particular study, the negativity bias we found does not seem to be produced by interindividual differences. Nevertheless, examining the interplay between feedback, study design, and personal characteristics could produce valuable insights into the mechanisms underlying differential reactions to feedback.

Another issue raised by our findings is the question of whether motives for self-enhancement and self-improvement are the driving factor behind biased self-concept change. While this has been presumed in previous research (Müller-Pinzler et al., 2019), the present research does not support such theorizing. Future research should further investigate their role, either by more directly assessing the motives themselves or by improving indirect approaches such as through opportunity for improvement. The latter could be done by manipulating the opportunity for improvement via different traits that are perceived as more or less malleable instead of using the same trait. Furthermore, instead of realizing a "low" opportunity for improvement condition (as we did in Study 4), it might be necessary to induce the perception that there is no opportunity for improvement at all. In addition, other explanations such as the diagnosticity of self-relevant information should be examined more directly. From our perspective, it would first be interesting to further examine whether negative information is perceived as more diagnostic than positive information in the context of self-relevant feedback. Second, identifying factors that inverse diagnosticity is key to explaining the contradictory findings on positivity and negativity biases in self-concept change.

Conclusion

Self-relevant feedback provides people with external information about themselves and impacts their self-concepts. However, feedback is not always integrated into the self-concept in a way that results in an accurate representation of the world. Consistent with previous research, we find that larger discrepancies are associated with more self-concept change. Contrary to several previous studies, however, we find that negative feedback is overweighed in comparison to positive feedback, resulting in negatively biased self-concept change. Aiming to explain when self-concept change after

feedback is negatively versus positively biased, we find no evidence that the opportunity for improvement causes biased self-concept change.

Acknowledgments

The authors thank Veronika Fischer, Alina Kolb, Emma Runge, Vivian Swoboda, and Leonie Thürmer for their help in creating study materials and collecting data for Study 4.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

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

ORCID iDs

Franziska Brotzeller https://orcid.org/0009-0002-3069-1503
Mario Gollwitzer https://orcid.org/0000-0003-4310-4793

Supplemental Material

See Online Supplemental Materials are available at https://osf.io/yadqw/.

Notes

1. Due to the exploratory nature of this study, we did not preregister it.
2. In addition to the measures described in detail, participants also completed some other measures. A complete list of all measured variables for this and all following studies can be accessed at https://osf.io/yadqw/. As the additional measures are not central to the present research, they will not be further discussed.
3. We conducted additional analyses exploring gender differences for this and all following studies. The results of these analyses can be found at https://osf.io/yadqw/ (Appendix A for Study 1, B3 for Study 2, C2 for Study 3, and D3 for Study 4).
4. We deviated from the preregistered analyses for H1 to be consistent with the analyses used for the other studies in this paper. The preregistered analysis, however, produces the same pattern and significance of results as the analysis used in the present research. We did not test the preregistered H2–H4 because they were part of another project examining the role of self-concept clarity in self-concept change, which is not relevant to the present research.
5. To do so, we assessed several attention check items as well as a *use me*-item in this and all following studies.
6. Most of the $n = 254$ participants were excluded due to the following two criteria: First, to ensure that participants could be randomly assigned to the experimental feedback conditions, participants with too low or high initial self-perceptions had to be excluded ($n = 126$). Second, a further $n = 112$ participants were excluded because they did not pass both of our two attention checks. Conducting the main analysis without excluding participants who failed the attention checks does not change the pattern or significance of our results.
7. To avoid the impression that feedback was systematically manipulated, small discrepancies deviated by either 4%, 5%, or 6% from participants' self-perception scores, while large discrepancies deviated by 19%, 20%, or 21%.
8. Among the additionally measured constructs, which can be accessed at https://osf.io/yadqw/, was a measure for participants reactions to the feedback using a 13-item scale based on Henss and Pinquart's (2022) scale for coping with expectation violations. This scale included four items measuring participants' intentions for self-concept change. As this and all following studies focus on actual self-concept change as the dependent variable, we do not report the results for intentions to change here. Instead, they can be found at https://osf.io/yadqw/ for this and all following studies (Appendix B4 for Study 2, C3 for Study 3, and D4 for Study 4). When conducting the main analyses using intentions for self-concept change as the dependent variable, patterns and significances of results in this and all following studies are largely identical to the results for actual self-concept change.
9. We deviate from the preregistered analyses as they were part of a research project examining the role of reflection in self-concept change and are not applicable to the research questions investigated in the present research.
10. Feedback for Study 3 was equivalent to those in Study 2 in content and very similar in wording. An exemplary feedback for Study 3 can be found in the study materials at https://osf.io/yadqw/.
11. As this study was originally geared toward examining a different research question, the following manipulation was included at this point in the study: Participants were randomly assigned to one of two conditions, in which they were either asked to reflect on the feedback they had received or to work on a distractor task aimed at inducing cognitive load and inhibiting reflection. This manipulation is not central to the present research and is therefore not further discussed. Entering the reflection conditions into our main analyses does not change the pattern or significance of our results.
12. The manipulation of inducing or hindering reflection from Study 3 was omitted in Study 4.
13. We employed one additional measure assessing upward or downward comparison at this point in the study: Participants were given the opportunity to compare themselves to one of two other alleged participants, one of whom had performed worse and one of whom had performed better than the participant. This measure served as an additional exploratory measure of motives for self-enhancement and self-improvement, as previous research has shown that upward comparisons are found when motives for self-improvement are present, while downward comparisons are used to self-enhance (Buunk & Gibbons, 2007). The full description of this measure as well as the exploratory analyses conducted with it can be accessed at https://osf.io/yadqw/ (see Appendices D5 and D6 for the analyses).

References

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (*Vol. 6*, pp. 1–62). Academic Press. https://doi.org/10.1016/S0065-2601(08)60024-6

Bergin, A. E. (1962). The effect of dissonant persuasive communications upon changes in a self-referring attitude1. *Journal of Personality*, *30*(3), 423–438. https://doi.org/10.1111/j.1467-6494.1962.tb02314.x

Binderman, R. M., Fretz, B. R., Scott, N. A., & Abrams, M. H. (1972). Effects of interpreter credibility and discrepancy level of results on responses to test results. *Journal of Counseling Psychology*, *19*, 399–403. https://doi.org/10.1037/h0033232

Bölte, S. (2005). *Reading the Mind in the Eyes Test für Erwachsene (dt. Fassung) von S. Baron-Cohen*. J.W. Goethe-Universität Frankfurt/M.

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, *102*(1), 3–21. https://doi.org/10.1016/j.obhdp.2006.09.007

De Castella, K., & Byrne, D. (2015). My intelligence may Be more malleable than yours: The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education*, *30*(3), 245–267. https://doi.org/10.1007/s10212-015-0244-y

Dunning, D. (1995). Trait importance and modifiability as factors influencing self-assessment and self-enhancement motives. *Personality and Social Psychology Bulletin*, *21*(12), 1297–1306. https://doi.org/10.1177/01461672952112007

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*(4), 267–285. https://doi.org/10.1207/s15327965pli0604_1

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114–138. https://doi.org/10.1257/mic.3.2.114

Elder, J., Davis, T., & Hughes, B. L. (2022). Learning about the self: Motives for coherence and positivity constrain learning from self-relevant social feedback. *Psychological Science*, *33*(4), 629–647. https://doi.org/10.1177/09567976211045934

Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior & Organization*, *80*(3), 532–545. https://doi.org/10.1016/j.jebo.2011.05.012

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Formann, A. K., Waldherr, K., & Piswanger, K. (2011). *Wiener Matrizen-Test 2 (WMT-2): Ein Rasch-skalierter sprachfreier Kurztest zu Erfassung der Intelligenz* [Viennese Matrices Test 2 (WMT-2): A Rasch-scaled culture-fair short intelligence test]. Beltz Test, Hogrefe.

Heine, S. J., & Hamamura, T. (2007). In search of east asian self-enhancement. *Personality and Social Psychology Review*, *11*(1), 4–27. https://doi.org/10.1177/1088868306294587

Henss, L., & Pinquart, M. (2022). Dispositional and situational predictors of coping with violated achievement expectations. *Quarterly Journal of Experimental Psychology*, *75*(6), 1121–1134. https://doi.org/10.1177/17470218211048108

Hughes, J., & Beiner, D. (2021). *Reghelper: Helper functions for regression analysis. R package version 1.1.0*. https://CRAN.R-project.org/package=reghelper

Kernis, M. H., & Goldman, B. M. (2003). Stability and variability in self-concept and self-esteem. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 106–127). The Guilford Press.

Kersting, M., Althoff, K., & Jäger, A. O. (2008). *Wilde-Intelligenz-Test 2: WIT-2*. Hogrefe.

Koban, L., Schneider, R., Ashar, Y. K., Andrews-Hanna, J. R., Landy, L., Moscovitch, D. A., Wager, T. D., & Arch, J. J. (2017). Social anxiety is characterized by biased learning about performance and the self. *Emotion*, *17*(8), 1144–1155. https://doi.org/10.1037/emo0000296

Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *Journal of Neuroscience*, *32*(47), 16832–16844. https://doi.org/10.1523/JNEUROSCI.3016-12.2012

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*(3), 579–592. https://doi.org/10.1017/S0033291713001074

Kube, T., Kirchner, L., Lemmer, G., & Glombiewski, J. A. (2022). How the discrepancy between prior expectations and new information influences expectation updating in depression—The greater, the better? *Clinical Psychological Science*, *10*(3), 430–449. https://doi.org/10.1177/21677026211024644

Kube, T., Rief, W., Gollwitzer, M., Gärtner, T., & Glombiewski, J. A. (2019). Why dysfunctional expectations in depression persist—Results from two experimental studies investigating cognitive immunization. *Psychological Medicine*, *49*(9), 1532–1544. https://doi.org/10.1017/S0033291718002106

Kurman, J. (2006). Self-enhancement, self-regulation and self-improvement following failures. *British Journal of Social Psychology*, *45*(2), 339–356. https://doi.org/10.1348/014466605X42912

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, *1*(4), Article 4. https://doi.org/10.1038/s41562-017-0067

Leising, D., Ostrovski, O., & Borkenau, P. (2012). Vocabulary for describing disliked persons is more differentiated than vocabulary for describing liked persons. *Journal of Research in Personality*, *46*(4), 393–396. https://doi.org/10.1016/j.jrp.2012.03.006

Levy, S. R., & Dweck, C. S. (1998). Trait- versus process-focused social judgment. *Social Cognition*, *16*(1), 151–172. https://doi.org/10.1521/soco.1998.16.1.151

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, *20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2022). Managing self-confidence: Theory and experimental evidence. *Management Science*, *68*(11), 7793–7817. https://doi.org/10.1287/mnsc.2021.4294

Müller-Pinzler, L., Czekalla, N., Mayer, A. V., Stolz, D. S., Gazzola, V., Keysers, C., Paulus, F. M., & Krach, S. (2019). Negativity-bias

in forming beliefs about own abilities. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-50821-w

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*, *32*(2), 551–562. https://doi.org/10.1523/JNEUROSCI.5498-10.2012

Norris, C. J. (2021). The negativity bias, revisited: Evidence from neuroscience measures and an individual differences approach. *Social Neuroscience*, *16*(1), 68–82. https://doi.org/10.1080/17470919.2019.1696225

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-Project.Org/

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407–441.

Shrauger, J. S., & Schoeneman, T. J. (1979). Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, *86*, 549–573. https://doi.org/10.1037/0033-2909.86.3.549

Strube, M. J. (2012). From "out there" to "in here": Implications of self-evaluation motives for self-knowledge. In S. Vazire & T. D. Wilson (Eds.), *Handbook of self-knowledge* (pp. 397–412). The Guilford Press.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210. https://doi.org/10.1037/0033-2909.103.2.193

Torchiano, M. (2020). *Effsize: A package for efficient effect size computation*. https://doi.org/10.5281/zenodo.1480624

Unkelbach, C., Alves, H., & Koch, A. (2020). Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. In B. Gawronski (Ed.), *Advances in experimental social psychology* (*Vol. 62*, pp. 115–187). Academic Press. https://doi.org/10.1016/bs.aesp.2020.04.005

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*, 383–403. https://doi.org/10.1037/0033-2909.134.3.383