Check for updates

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 48 (2024) e70017
© 2024 The Author(s). *Cognitive Science* published by Wiley Periodicals LLC on behalf of Cognitive Science Society (CSS).
ISSN: 1551-6709 online
DOI: 10.1111/cogs.70017

# Beyond the Positivity Bias: The Processing and Integration of Self-Relevant Feedback Is Driven by Its Alignment With Pre-Existing Self-Views

Josué García-Arch,[a,b,c] Solenn Friedrich,[a] Xiongbo Wu,[d] David Cucurell,[a,b,c] Lluís Fuentemilla[a,b,c]

[a]*Department of Cognition, Development and Education Psychology, University of Barcelona*
[b]*Institute of Neuroscience (UBNeuro), University of Barcelona*
[c]*Bellvitge Institute for Biomedical Research, Hospitalet de Llobregat*
[d]*Department of Psychology, Ludwig-Maximilians-Universität München*

Received 8 March 2024; received in revised form 25 October 2024; accepted 28 October 2024

**Abstract**

Our self-concept is constantly faced with self-relevant information. Prevailing research suggests that information's valence plays a central role in shaping our self-views. However, the need for stability within the self-concept structure and the inherent alignment of positive feedback with the pre-existing self-views of healthy individuals might mask valence and congruence effects. In this study ($N = 30$, undergraduates), we orthogonalized feedback valence and self-congruence effects to examine the behavioral and electrophysiological signatures of self-relevant feedback processing and self-concept updating. We found that participants had a preference for integrating self-congruent and dismissing self-incongruent feedback, regardless of its valence. Consistently, electroencephalography results revealed that feedback congruence, but not feedback valence, is rapidly detected during early processing stages. Our findings diverge from the accepted notion that self-concept updating is based on the selective incorporation of positive information. These findings offer novel insights into self-concept dynamics, with implications for the understanding of psychopathological conditions.

*Keywords:* Self-concept; Belief updating; Social feedback processing; Self-concept clarity; Positivity bias

Correspondence should be sent to Josué García-Arch, Department of Cognition, Development and Education Psychology, University of Barcelona, PG de la Vall d'Hebron, 171, 08035 Barcelona, Spain. E-mail: j.garcia.arch@ub.edu, josue.g.arch@gmail.com

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

Individuals hold beliefs about their abilities and attributes that aid in understanding themselves and their environments (Epstein, 1973; Mokady & Reggev, 2022). How these beliefs are formed and updated is a topic that has received a lot of attention in recent years. The dominant perspective in this field suggests that when updating self-relevant beliefs, positive and negative information is differently weighted, contributing to the formation of positively biased self-representations (Korn, Prehn, Park, Walter, & Heekeren, 2012; Sharot & Garrett, 2016). While these principles apply to diverse self-relevant beliefs, further considerations are essential to understand self-concept updating. The self-concept is considered a cognitive schema comprising diverse self-representations, including beliefs about our personality traits (Campbell, 1990; Martinelli, Sperduti, & Piolino, 2013). These self-representations are embedded in a highly organized autobiographical knowledge system that protects the self-concept against stability disruptions (Conway, 2005). Consistently, there is evidence that individuals are motivated to seek self-congruent information, regardless of its valence (Swann & Brooks, 2012). This raises questions about the capacity of positive feedback to prompt belief updating independently of its compatibility with pre-existing self-knowledge. Moreover, the inherent positive bias in the self-concept of healthy individuals (Taylor et al., 1988) obscures the distinction between positive and self-congruent information (García-Arch, March Sabio Albert, & Lluis Fuentemilla, 2023), which might influence the interpretation of findings from previous behavioral and neuroimaging studies. Understanding how individuals form and update self-representations is crucial, since they play a central role in psychological functioning and well-being (Korn, La Rosée, Heekeren, & Roepke, 2016; Mokady & Reggev, 2022; Swann, Tafarodi, Wenzlaff, & Swann, 1992). Therefore, unraveling the distinct influences of feedback valence and feedback congruence on self-concept updating requires further inquiry.

Behavioral and neuroimaging studies suggest that desirable and undesirable information is processed and used differently to update self-relevant beliefs, resulting in valence-dependent learning asymmetries (Sharot & Garrett, 2016). Evidence suggests that positive information is readily integrated into our beliefs, while negative information is dismissed (Sharot, Korn, & Dolan, 2011). The pervasiveness of this phenomenon has led to the assumption that it reflects a fundamental property of learning (Sharot & Garrett, 2016). Recently, these principles have extended to the domain of self-concept updating (Korn et al., 2012, 2014, 2016), consistent with the notion that individuals are motivated to build a positively biased self-view (Hepper, Gramzow, & Sedikides, 2010). These studies have also shown differential behavioral and neural responses to positive and negative feedback, aligning with a valence-based belief updating bias. Importantly, the propensity toward a valence-dependent updating of self-representations may carry important implications for well-being (Korn et al., 2016; Sharot & Garrett, 2016).

To understand how self-representations might be updated, it is important to consider several important features of the self-concept. Although the self-concept evolves during the lifespan, it also exhibits a pronounced tendency toward stability and coherence (Conway, 2005; Nowak, Vallacher, Tesser, & Borkowski, 2000). Self-beliefs, as those related to our

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

personality traits, are well-grounded semantic representations supported by a wide range of autobiographical evidence, which provides certainty and stability to the self-concept (Conway, 2005; Martinelli et al., 2013). We are highly sensitive to information that matches our self-views. Behavioral and neuroimaging studies indicate that we are especially tuned to discern self-related from non-self-related information (Northoff et al., 2006). Information that aligns with our self-perceptions undergoes preferential processing, whereas identity-discrepant information is swiftly identified at the early stages of processing, and subsequently minimized or distorted (Abendroth, Nauroth, Richter, & Gollwitzer, 2022; Conway, 2005; Nowak et al., 2000). There is also evidence that individuals are motivated to seek self-congruent feedback and protect from self-discrepant evaluations. For example, when facing self-incongruent feedback, individuals experience negative emotional responses, and employ different strategies to mitigate its impact (Swann & Brooks, 2012). Consistently, novel theoretical models suggest that information that matches our self-views might trigger rewarding experiences (Mokady & Reggev, 2022). These findings underscore the pervasive human endeavor to reinforce the certainty and stability of the self-concept. This pursuit aligns with research indicating that a confident and stable self-concept is crucial for daily functioning, bolstering psychological continuity and well-being (Campbell, Assanand, & Di Paula, 2003; Jiang, Wang, Poon, Gaer, & Wang, 2023; Nowak et al., 2000).

Together, evidence suggests that individuals are motivated to maintain both a positively biased and stable self-concept. However, this dual motivation poses a conceptual challenge in the study of how self-representations are updated. As the self-concept becomes positively biased, positive and self-congruent information converge (García-Arch et al., 2023). This convergence is not trivial, as the distinct behavioral and neural responses elicited by positive and negative feedback might be also explained by variations in its alignment with the existing self-concept. Similarly, different degrees of overlap between feedback valence and self-congruence might produce divergent results across studies and populations. Hence, to unravel the behavioral and neural responses underlying self-relevant belief updating, feedback valence and self-congruence need to be experimentally orthogonalized. Similar concerns have been expressed from different research lines (Mokady & Reggev, 2022; Swann Jr. & Brooks, 2012).

Here, we explored the possibility that in healthy individuals, where a positive bias in the self-concept is already present (Taylor et al., 1988), the tendency toward self-concept stabilization might be as pronounced as, or even surpass, the drive toward incorporating positive evaluations. In contrast, individuals might prioritize the incorporation of positive inputs to enhance the positivity of their self-images, regardless of the congruence of the information with their current self-views. However, while valence-based belief updating may contribute to building a positively biased self-image, indiscriminate incorporation of positive feedback could undermine self-concept certainty and stability, which are crucial for psychological well-being (Campbell et al., 2003). Note that in healthy individuals, an enhanced focus toward self-concept stability would add certainty to the current self-view at no cost for its overall positivity. Moreover, self-congruent feedback might convey to individuals that their self-views are accurate, which can trigger positive feelings (Mokady & Reggev, 2022), especially in healthy populations. In contrast, an exclusive drive toward incorporating positive feedback

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in individuals with already well-established and positive self-concepts might involve trying to integrate information that does not match their existing self-views and autobiographical evidence. While receiving incongruent positive self-evaluations might not be problematic in itself, or it might even induce positive feelings, attempting to incorporate such feedback may have a cost for the coherence and certainty of the self-concept (Conway, 2005).

If this notion is true, feedback that conflicts the existing self-concept should be swiftly identified, which could help avoiding the contamination of self-representations by subjectively inaccurate information (Abendroth et al., 2022). Employing neuroimaging techniques such as the electroencephalography (EEG), with its excellent temporal resolution, can offer critical insights into these processes. The capability of EEG to rapidly distinguish the electrophysiological signatures associated with the processing of feedback valence and congruence may offer novel insights into the dynamics of self-relevant feedback processing.

Here, we required healthy participants to engage in a belief updating task while recording scalp electrophysiological (EEG) activity. Participants evaluated themselves before and after receiving self-relevant social feedback from their peers. We employed a well-known belief updating paradigm (Elder, Davis, & Hughes, 2022; Korn et al., 2016) with a recent procedure that allows to control the effect of the initial positive bias in participants self-concept. This procedure allowed us to examine the differential behavioral and electrophysiological signatures associated with the effects of feedback valence and feedback congruence on feedback processing and self-concept updating.

## 2. Methods

### *2.1. Participants*

Prior to the study, we conducted a power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the required sample size. Following previous literature with similar experimental design (Korn et al., 2012, 2014, 2016), we assumed a partial eta squared of .1 with a conservative correlation between measures of .5. Power analysis revealed that for an acceptable power of .8, 20 participants would be required. To accommodate typical methodological challenges in EEG research, such as data quality issues arising from artifacts like participant movement and blinking, we recruited an initial sample larger than calculated by our power analysis. These common artifacts often lead to a significant portion of data being unusable. Moreover, our study design included attention checks to ensure participant engagement. Participants who failed these checks were excluded from the analysis. Anticipating potential data loss and noncompliance, based on established norms and our prior experience, we decided to recruit 35 participants (22 females), all of them students from the University of Barcelona. Participants received €10 per hour for participation. Informed consent and consent to publish were obtained from participants following procedures approved by the Ethics Committee of the University of Barcelona. The study was reviewed and approved by the University of Barcelona Bioethics Commission (IRB00003099). Four participants were excluded because of extensive artifacts in the recorded electroencephalogram (EEG).

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

One participant was excluded due to failing all the attention checks implemented in the experimental task (see details in the next section). The final sample ($N = 30$, 19 females) consisted of native Spanish speakers; all were right-handed, had normal or corrected-to-normal vision, and had no previous or current neurological or psychiatric disorders. On average, the participants were 22.43 years old ($SD = 2.17$).

### 2.2. Procedure

Participants took part in a two-session experiment separated by 3 days. The first session was online and administered via Qualtrics (www.qualtrics.com). The first session aimed to create a situation in which participants believed they would receive social feedback during the second session. The second session consisted of performing the experimental tasks while EEG was recorded.

#### 2.2.1. First session

This session consisted of an online survey. At the beginning of the survey, participants encountered three embedded audio recordings containing personality descriptions. They were informed that these recordings belonged to anonymous participants contributing to the same experiment within the next 72 h and had already completed the online survey. Participants' task was to evaluate the speakers' personalities using a provided list of adjectives. To ensure the authenticity of voice samples, recordings were made by independent collaborators who were initially unaware of the aim of the study. Each recording lasted approximately 8 min (ranging from 7.45 to 8.29 min). After completing their contributions, collaborators were briefed on the study's purpose and provided informed consent for data use. The recordings were presented in random order to the participants. After listening to each personality description, participants evaluated the speaker by choosing applicable adjectives from a predetermined list. Subsequently, they were instructed to record themselves following detailed guidelines and using the presented recordings as examples. These guidelines incorporated 12 randomly chosen items from each of the six HEXACO personality factors (https://hexaco.org/), such as "I feel reasonably satisfied with myself overall" and "I rarely express my opinion in social meetings." Participants were required to speak for at least 30–45 s of each statement, expressing their level of agreement and providing contextual examples or anecdotes. Upon completion, they attached their recordings to the online questionnaire.

Next, participants were instructed to evaluate themselves using a list of 150 adjectives (75 positive, see *Stimuli*). The process was designed to control the initial positive bias in participants' self-concept and orthogonalize feedback valence and feedback self-congruence effects. Participants used a drag-and-drop interface to categorize each adjective as "Yes (Me)" or "No (not Me)." They were also instructed to classify adjectives that were unfamiliar to them in an auxiliary box. Adjectives were listed in random order within blocks of positive and negative adjectives, which were also randomized. Participants were instructed to make a minimum of 28 positive and 28 negative decisions, ensuring that negative decisions represented a realistic percentage among the total sample of adjectives (∼18%) (García-Arch et al., 2023). Once this data was obtained, we conducted a nonproportional stratified

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

random sampling on participants' positive and negative decisions. That is, the same number of positive and negative decisions were randomly drawn from their respective populations. This strategy allowed us to orthogonalize feedback valence and feedback self-congruence in session 2.

Finally, participants were requested to complete the Beck Depression Inventory (BDI-II). BDI-II score was used as an exclusion criterion. Following previous research (Garcia-Arch, Barberia, Rodríguez-Ferreiro, & Fuentemilla, 2022; Kappes & Sharot, 2019), participants who scored >19 in the BDI were excluded from the data analysis. In the current experiment, none of the participants met this criterion.

### 2.2.2. Second session

In this session, participants performed a belief-updating task similar to those previously used to study the impact of positive and negative feedback on participant's self-representations (Elder et al., 2022; Korn et al., 2012, 2014, 2016). The task consisted of three blocks: self-evaluation, social evaluative feedback, and re-evaluation phase (Fig. 1). In the first block, participants were presented with their own judgments from the first session. Each judgment was displayed on the screen in the format "You think you are [adjective]" or "You think you are not [adjective]," with each adjective (e.g., "Sociable") presented one at a time in random order. Participants were instructed to rate their confidence in each self-assessment using a 0 to 100 slider scale (10 s), where 0 represented no confidence at all and 100 represented perfect confidence. Participants were instructed to confirm their selection by pressing the space bar within a 10-s interval. The second block introduced social evaluative feedback, purportedly from three other participants who had listened to the participant's voice clip describing their personality. Participants were led to believe that the feedback represented the most frequent judgment among the three evaluators. Each trial began with the question "Do others think you are [adjective]?", that was on the screen for 3 s, followed by a fixation cross displayed for a jittered duration of 300, 400, or 500 ms. The evaluators' decision ("Yes" or "No") was then shown for 1.5 s. An inter-trial interval of 1.5 s including a jittered fixation cross on the screen separated at the start of the next trial. Feedback on each adjective was presented three times across three separate blocks, interspersed with rest periods, to ensure a sufficient number of observations per condition for robust statistical analysis of EEG data. The feedback was manipulated such that in 25% of cases, participants received positive feedback that matched their self-evaluations (positive + self-congruent), in 25% of cases, the feedback was positive but did not match their self-evaluations (positive + self-incongruent), in another 25% of cases, they received negative feedback that matched their self-evaluations (negative + self-congruent), and in the remaining 25%, the negative feedback did not match their self-evaluations (negative + self-incongruent). We employed categorical feedback to ensure no ambiguities in the perception of its alignment with participants' decisions or its valence. In addition to the main trials, the feedback block also included catch trials to ensure participant engagement and attentiveness. These catch trials followed the same format as the main trials, with the prompt, "Do others think you are [catch]?", however, in these cases, "[catch]" was replaced with nonadjective words (e.g., "Whistle"). Participants were instructed to identify

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

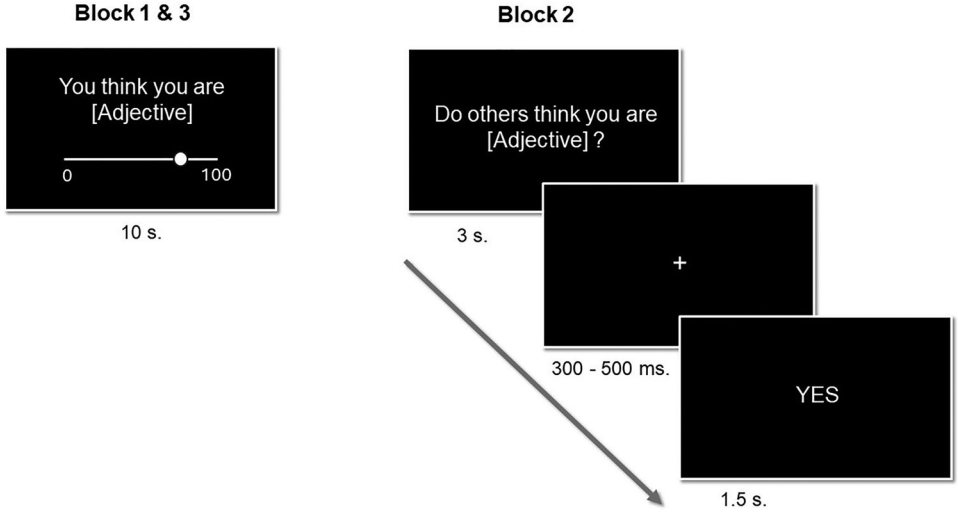


Fig. 1. Overview of the Experimental task. The task is divided into three main blocks. Self-Assessment Rating (Block 1): Participants are presented with statements about their self-judgments from a prior session, formatted as "You think you are [adjective]" or "You think you are not [adjective]." Each adjective is shown individually in a random sequence. Participants rate their confidence in these self-assessments on a 0 to 100 scale, where 0 indicates no confidence and 100 indicates complete confidence. Confirmation of each rating is done via space bar press. Social Evaluative Feedback (Block 2): Participants receive feedback, purportedly from three peers, on whether others perceive them as described by the adjectives. Feedback is presented in a structured sequence, beginning with a query ("Do others think you are [adjective]?"), followed by a variable-duration fixation cross, the evaluators' decision ("Yes" or "No"), and another fixation cross before proceeding to the next trial. Feedback is systematically manipulated to include positive and negative evaluations, both congruent and incongruent with the participant's self-assessment. Catch trials with nonadjective prompts are included to monitor engagement and attentiveness. Post-Feedback Reassessment (Block 3): Following the feedback phase, participants revisit the initial confidence rating.

them by pressing the space bar. After the social evaluative feedback phase, the experiment returned to the initial confidence judgment task (Block 3).

Following the completion of their second session, participants were debriefed. They were informed that the feedback they received was generated pseudo-randomly, and that nobody had actually evaluated their voice clips. They were also informed that the voice recordings they evaluated were made by external collaborators. Additionally, a set of final questions was posed to evaluate any confusion about the stimuli, the task, or the setup. No problems were reported.

### *2.3. Stimuli*

Following previous studies (Elder et al., 2022; García-Arch et al., 2023; Korn et al., 2012, 2014, 2016), we chose personality adjectives to study self-concept updating (i.e., trait words such as "Sociable," "Organized," etc.). For the current study, we randomly selected 75 positive (e.g., "Honest") and 75 negative adjectives (e.g., "Anxious") from classifications

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

employed in previous studies, which come from widely studied lists of personality descriptors (Anderson, 1968) (see, osf.io/x98mu/?view_only = ddf54d1c650942488f97f17f88c0c7d8).

### 2.4. Main behavioral measures

The target dependent variable for behavioral analysis was *update scores*. These scores represent the change in participants' beliefs (i.e., confidence ratings in this study) in the direction suggested by the feedback. That is, post – pre confidence ratings for (positive and negative) congruent feedback and pre – post confidence ratings for (positive and negative) incongruent feedback, representing a measure of "feedback acceptance" (Korn et al., 2012). Note that this measure reflects feedback-consistent variations in the certainty with which self-concepts are held, rather than direct changes in self-evaluation (Pelham & Swann, 1989).

All analyses included two binary categorical variables representing the experimental conditions: feedback valence (positive/negative) and feedback self-congruence (self-congruent/self-incongruent). Feedback valence represented whether participants received positive or negative evaluations, while feedback self-congruence was defined by whether those evaluations matched or not participants' decisions. A control measure was included to control for how much space within the scale participants had available for updating (*Update Space*).

### 2.5. EEG recording and preprocessing

EEG was recorded in a Faraday cage. Participants were seated in front of the screen at a distance of approximately 57 cm from the center of the screen. The EEG recording was conducted with a 64-channel system at a sampling rate of 250 Hz, using an actiChamp amplifier (Brain Products) and Ag/AgCl electrodes mounted in an electrocap (ANT neuro) located at 60 standard positions (Fp1, Fp2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCz, FC2, FC4, FC6, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2) and the left and right mastoids. One electrode (FT9) was excluded due to technical problems. Eye movements were monitored with an electrode placed at the infraorbital ridge of the right eye. Electrode impedances were kept below 10 kΩ during the recording. FCz served as an online reference. The signal was re-referenced offline to the linked mastoids and bad channels were interpolated (spherical interpolation). A high-pass filter at 0.1 Hz and a low-pass filter at 30 Hz were implemented offline. The continuous EEG data was then epoched into 1 s segments. Each epoch spanned a time window from −100 ms pre-stimulus to 900 ms post-stimulus and a pre-stimulus interval of 100 ms was used as the baseline for absolute baseline correction. Trials exceeding ± 100 μV in EEG and/or EOG channels were automatically rejected offline. Trials containing noise not detected through the amplitude threshold approach were also rejected manually. Preprocessing and statistical analysis of EEG data were conducted in *MATLAB* (Version R2021a) in conjunction with *EEGLAB* (Version 2022.0, Delorme & Makeig, 2004) and *Fieldtrip* (Oostenveld, Fries, Maris, & Schoffelen, 2011).

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

### 2.6. EEG data analysis

The EEG analysis was designed to explore the electrophysiological signatures of feedback valence and self-congruence without the constraints of parametric assumptions and the specificity required for Event-Related Potentials (ERP) analysis in terms of time windows and spatial locations. The literature on ERP and social feedback processing is still expanding, with substantial variability in the selection of time windows, spatial locations, hypotheses tested, and effects obtained (Peters et al., 2024). Moreover, to our knowledge, this is the first EEG study to experimentally orthogonalize feedback valence and self-congruence as distinct experimental conditions, which presents challenges in reliably testing specific hypotheses based on well-established ERP correlates. Consequently, to explore the electrophysiological signatures for self-congruent, self-incongruent, positive, and negative feedback, we employed a nonparametric cluster-based permutation test, which provides a way to formulate a null hypothesis (identical probability distribution in the different experimental conditions) without prior assumptions of possible effects (Maris & Oostenveld, 2007).

This data-driven analytical strategy was used to identify clusters of significant points in the spatiotemporal 2D matrix (time and electrodes). This method addresses the multiple-comparison problem by employing a nonparametric statistical testing strategy. The procedure is based on a cluster-level randomization testing to control for the family-wise error rate. Statistics for each time point were calculated, identifying spatiotemporal points with statistical values exceeding a predefined threshold ($p < .05$, two-tailed). Next, these points were grouped into clusters based on their adjacency along the x and y axes within the 2D matrix. The observed cluster-level statistics were computed by taking the sum of all values from the contrast statistics within a cluster. Condition labels were then permuted 1000 times (Monte Carlo randomization) to approximate the null hypothesis, and the maximum cluster statistic was chosen to construct a distribution of the cluster-level statistics under the null hypothesis. The significance of the nonparametric statistical test was determined by the proportion of randomized test statistics that exceeded the observed cluster-level statistics. In this analysis, we included the main effects of feedback self-congruence and feedback valence, as well as their two-way $2 \times 2$ interaction.

## 3. Results

### 3.1. Behavioral analysis

Of primary interest, we examined whether participants incorporated more self-congruent than self-incongruent feedback in their self-representations as well as more positive negative feedback. To analyze the data, we employed both repeated measures ANOVA and linear mixed-effects models (MLMs). While acknowledging the superior analytical flexibility and accuracy of MLMs, particularly in handling individual differences and nested data structures, we include ANOVA results to maintain continuity and to facilitate comparative analyses with existing literature in the field.

First, we conducted a repeated measures analysis of variance (rmANOVA) with average update scores as the dependent variable and feedback self-congruence, feedback valence, and

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

their interaction as within-subjects effects. The results of this analysis revealed that participants tended to adjust significantly more their confidence ratings in a feedback-consistent direction in response to self-congruent than in response to self-incongruent feedback ($F(1, 29) = 5.22$, $p = .029$, $\eta_p^2 = .15$). No significant effects were found for feedback valence ($F(1, 29) = 2.43$, $p = .129$, $\eta_p^2 = .07$) and feedback self-congruence × feedback valence interaction ($F(1, 29) = .11$, $p = .743$, $\eta_p^2 = .003$). Next, we aimed to test whether the observed differences in update scores between self-congruent and self-incongruent feedback could be attributed to participants integrating self-congruent feedback (indicated by update scores above zero) and dismissing self-incongruent feedback (reflected by update scores at or below zero), among other possible patterns. Post hoc analysis (one-sample *t*-test) revealed confirmed that participants tended to integrate self-congruent feedback ($M = 3.15$, $SE = .71$, 95% $CI[1.71, 4.59]$ $t(29) = 4.45$, $p < .001$, $d = .81$) and dismiss self-incongruent feedback ($M = .33$, $SE = .82$, 95% $CI[-1.34, 2.01]$ $t(29) = .41$, $p = .687$, $d = .07$) (Fig. 2b). To further investigate the lack of preferential integration of positive feedback indicated by the nonsignificant effect of feedback valence, we conducted a Bayes Factor analysis. In line with our findings, the results of this analysis indicated strong evidence against an enhanced integration of positive (vs. negative) feedback (BF = 12.032).

Next, we sought to carry out a more detailed analysis using linear mixed-effects models. This modeling technique allows to account for individual differences in parameter estimates, include within-subjects covariates (such as update space), compute proper post hoc tests with all the information included in the model, and incorporate additional random effects in the covariance structure of the tested model (Barr, Levy, Scheepers, & Tily, 2013; Brown, 2021). We constructed alternative models that varied in their inclusion of fixed effects for feedback self-congruence and feedback valence (each one separate, both main effects, and both with interaction) as well as different combinations of random slopes (see Table S1). All models included partially crossed random effects between adjectives and participants' IDs and update space as a covariate. Model selection was conducted using the Bayesian Information Criteria (BIC), which penalizes model complexity. *p*-Values were determined by Satterthwaite's approximation of degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2017). Maximal random effects structures were kept when supported by the data and model convergence (Barr et al., 2013).

Consistent with the rmANOVA results, the winning model (Marginal R2 = .294, Conditional R2 = .409) included feedback self-congruence as a fixed effect, as well as its random slope. The results of this analysis showed that participants tended to incorporate more self-congruent than self-incongruent feedback ($\beta_{\text{Self-congruent}} = 18.01$, $SE = 1.91$, 95% CI[14.24, 21.81], $t(34.329) = 9.44$, $p < .001$). All models and their associated BICs are reported in Table S1 (Supplementary Materials).

### 3.2. EEG results

To investigate the electrophysiological signatures associated with feedback self-congruence and feedback valence, we conducted a cluster-based permutation test on the EEG data recorded during the feedback phase (Fig. 1).

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

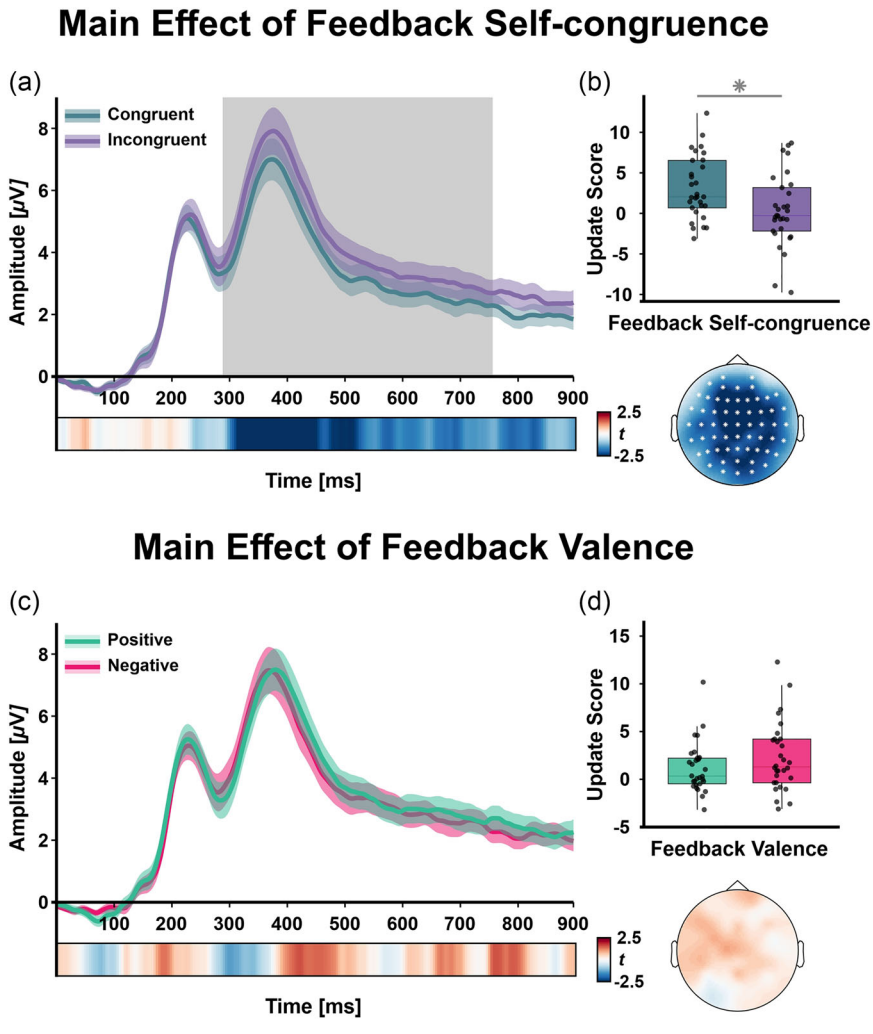


Fig. 2. Behavioral and electrophysiological signatures of feedback self-congruence and feedback valence. Panel (a) presents ERP amplitudes in response to congruent (teal blue) and incongruent (purple) feedback over time, with shaded error bands indicating the standard error of the mean. The inset displays the scalp topography of the *t*-statistic for the main effect of feedback congruence. Panel (b) shows box plots of the main feedback self-congruence on update scores, jittered points represent participants' average. Panel (c) depicts ERP responses to positive (green) and negative (pink) feedback. Panel (d) shows box plots of the main feedback valence on update scores.

The analysis of the EEG data elicited at the feedback cue revealed a significant negative cluster distributed throughout the scalp electrodes between ~300 and ~750 ms from cue onset ($p = .003$, mean $t$ value $= -2.82$, $d = -.51$, peak $t$ value $= -5.32$, $d = -.97$), indicating that self-congruent feedback elicited lower ERP amplitudes than self-incongruent feedback (Fig. 2a). No significant clusters were found for the contrasts including feedback valence

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Fig. 2c) or feedback self-congruence × feedback valence interaction (all $p > .124$) (Fig. S1). These results were in line with those obtained in the behavioral analysis. To provide complementary evidence of the effects of our experimental conditions, we implemented a Bayes factor statistical analysis. A Bayes factor was calculated for each contrast between conditions for each point of the resulting 2D electrodes × time matrix. Results were in line with those obtained by cluster-based permutation test (Fig. S2).

#### *3.2.1. Control analysis*

In our task, feedback on each adjective was presented three times across three separate blocks. This strategy aimed to enhance the reliability of our electrophysiological data by increasing the number of trials, a common approach in EEG research to improve the signal-to-noise ratio. Considering the nature of our EEG findings, which potentially feature a P300 component known for its association with mismatch and expectancy effects, we pursued an exploratory analysis to examine the main effect of block and its potential interaction with the congruency effect. Results from a cluster-based permutation test revealed a statistically significant higher overall mean amplitude in the first block (mean $t$ value $= 5.4489$, $p < .001$), indicating enhanced initial sensitivity to the feedback. However, the decrease in amplitude across blocks did not differ significantly across experimental conditions ($p = .156$). This indicates that while a general habituation to the task occurred, it did not differ across different types of feedback, suggesting that the fundamental target effects were not compromised by the repetition of feedback across blocks.

## 4. Discussion

In this study, we examined the behavioral and neurophysiological responses to social feedback by systematically manipulating feedback valence and self-congruence. Our findings revealed a pronounced asymmetry in the responses to self-congruent and self-incongruent feedback, both at the behavioral and neurophysiological levels. We found that feedback self-congruence was detected at early stages of processing, and that self-congruent information was readily integrated, whereas self-incongruent information failed to influence individuals' certainty in their self-representations. Interestingly, feedback valence did not modulate either behavioral or neurophysiological responses. This finding challenges the widely accepted notion that there is a strong, universal bias toward positive feedback in the updating of self-representations (Korn et al., 2012, 2014). Our experimental orthogonalization of feedback self-congruence and feedback valence provides novel insights into the behavioral and neural signatures of self-relevant feedback processing and self-concept updating.

Our findings revealed a behavioral tendency to preferentially assimilate self-congruent over self-incongruent feedback. This is consistent with the notion that self-beliefs are embedded in a rich system of autobiographical information that necessitates mechanisms to stabilize self-representations and protect against conflicting information (Conway, 2005; Nowak et al., 2000). The preferential integration of self-congruent feedback may facilitate the

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

differentiation between self-descriptive and non-self-descriptive attributes, enhancing self-concept clarity (Campbell, 1990). Such clarity in self-concept is crucial for daily functioning, enabling accurate predictions about future behaviors, strategic planning of actions, selection of suitable social partners, and maintenance of psychological well-being (Campbell, 1990; Mokady & Reggev, 2022; Swann & Hill, 1982).

Consistent with our behavioral results, we found that self-congruent and self-incongruent feedback elicited distinct electrophysiological signatures suggesting a rapid discrimination between congruent and incongruent information compatible with a P300 waveform. Our findings are consistent with ERP literature suggesting that schema-incongruent information triggers rapid electrophysiological responses (Höltje, Lubahn, & Mecklinger, 2019; Richter, 2020). These responses are postulated to reflect a mismatch between incoming information and activated schemas, triggering error signals that result in the updating of mental representations. In contrast, our findings suggested that self-incongruent information did not update participants' self-representations. These differences might be explained by the special nature of the self-concept, which unlike other cognitive schemas is considered to be a highly integrated, emotionally charged structure supported by a lifetime of accumulated evidence (Campbell, 1990; Conway, 2005). These self-concept features promote psychological continuity and might shield self-representations from immediate updates in the face of self-incongruent information (Conway, 2005; Nowak et al., 2000). In line with these notions, recent research suggests that identity-discrepant inputs are detected at the early stages of processing and treated as "false" information (Abendroth et al., 2022) which suggests that the rapid detection of self-incongruent feedback helps protecting self-representations from being disrupted by subjectively inaccurate information.

The EEG correlates found in this study might be also in line with existing literature on the P300 ERP component. Specifically, the P300 has been found to reflect expectancy violations (Polich, 2007). Our findings suggest that P300 modulations related to feedback self-congruence may occur with relative independence of feedback valence. While this is in line with several studies, whether P300 is insensitive to outcome valence is still under debate, and current studies are trying to clarify it (Paul et al., 2022). This work might contribute to the debate by providing a novel experimental approach that might help to experimentally orthogonalize valence-based and expectancy effects. Our findings might also help understanding the ERP correlates associated with social feedback processing. For example, prior studies have examined the P300 waveform in the context of different types of social feedback (e.g., social acceptance vs. rejection), mainly related to valence conditions, yielding mixed findings (Peters et al., 2024). When possible and theoretically justified, the orthogonalization of valence and self-congruence effects together with the selection of nonparametric data-driven approaches to analyze EEG data could shed light on the underlying electrophysiological signatures of social feedback processing.

We did not find significant differences at either the behavioral or electrophysiological level in response to positive and negative feedback, nor did we observe significant interactions with self-congruence. The lack of asymmetry in the updating of self-representations' certainty, favoring neither positive nor negative feedback, confronts the notion that psychologically healthy adults exhibit a strong tendency to integrate self-relevant information in a

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

valence-dependent manner (Korn et al., 2012). Similarly, we observed no differential electrophysiological responses between positive and negative feedback, diverging from current works that suggest a specialized neural tuning for discerning feedback valence (Korn et al., 2012). However, the lack of significant differences between positive and negative feedback should not be taken as strong evidence against valence effects. Rather, they suggest that the role of valence might be more nuanced and potentially overshadowed by factors like self-congruence in healthy adult samples. Future studies should replicate our findings with larger sample sizes. Along these lines, future research should investigate potential moderators, such as individual differences in self-esteem, to better understand how these variables influence the processing of feedback valence and self-congruence.

We suggest that healthy individuals with a positively skewed self-view might have a stronger drive to maintain self-concept stability, which would be compromised if responses to social feedback primarily involved unselective integration of positive feedback. Note that reinforcing a positively biased self-concept with confirming evidence would further crystallize self-representations while maintaining its overall positivity. However, we do not dispute the existence of self-related positivity biases. Indeed, the ubiquity of those biases is in itself manifested in the need to control for the initial positive skew in individuals' self-concept to orthogonalize feedback valence and self-congruence. Future research should test narrower hypotheses that include the control of feedback self-congruence and employ larger effect sizes than those reported in prior research to explore its potential effects. Moreover, although individuals with a positive self-concept seem to prioritize self-concept stabilization, it is possible that this drive toward stability diminishes during pivotal life transitions that require self-concept updates (Conway, 2005). In such instances, a valence-dependent integration of new information might preserve individuals' well-being during adaptive changes. Additionally, we found important variability in responses to self-incongruent feedback, as indicated by wide confidence intervals in update scores. This opens the door to studying potential moderators of this phenomenon in future studies. One interesting candidate is self-esteem, as it is a construct that can moderate self-relevant feedback processing (Mokady & Reggev, 2022; Swann et al., 1992).

Our findings may have important implications. The experimental orthogonalization of feedback self-congruence and valence might help reinterpreting findings obtained in previous studies. Moreover, our approach could also improve our understanding of different psychopathological conditions. As a remarkable example, it has been suggested that patients suffering from borderline personality disorder (BPD) display a reduced tendency toward valence-dependent learning asymmetries (Korn et al., 2016). However, this population is also characterized by a more negative self-concept, which can mask congruence and valence effects. Notably, BPD patients are not only characterized by negative self-views, but also by unstable self-concepts (Kaufman & Meddaoui, 2021). Therefore, unraveling congruence and valence effects might help in understanding their neural and behavioral responses to self-relevant information. Finally, the insights extracted from our work could enhance novel approaches based on the modification of maladaptive schemas through schema-incongruent learning in clinical populations (Moscovitch, Moscovitch, & Sheldon, 2023), potentially opening the door to more effective interventions.

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

## 5. Limitations

Following previous works, we focused on the updating of beliefs about personality adjectives. However, the self-concept contains a multiplicity of self-representations such as social roles or group memberships. Future research should extend the current findings to different components of the self-concept.

## Acknowledgments

We thank CERCA Programme/Generalitat de Catalunya for institutional support.

## Funding

This work was supported by a grant from the Spanish Ministerio de Ciencia, Innovación y Universidades, which is part of Agencia Estatal de Investigación, project: PID2022-140426NB-I00 (funded by MCIN/AEI/10.13039/501100011033/ and FEDER a way to make Europe), to L.F.

## Data availability statement

Data materials and analysis code have been made publicly available in the Open Science Framework (OSF): osf.io/x98mu/?view_only=ddf54d1c650942488f97f17f88c0c7d8.

## Ethics approval statement

The study was reviewed and approved by the University of Barcelona's Bioethics Commission (IRB00003099).

## References

Abendroth, J., Nauroth, P., Richter, T., & Gollwitzer, M. (2022). Non-strategic detection of identity threatening information: Epistemic validation and identity defense may share a common cognitive basis. *PLoS ONE*, *17*, e0261535.

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*, (3), 272–279.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, (3), 255–278.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*, (1), 1–19. https://doi.org/10.1177/2515245920960351

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, *59*, (3), 538–549.

Campbell, J. D., Assanand, S., & Di Paula, A. (2003). The structure of the self-concept and its relation to psychological adjustment. *Journal of Personality*, *71*, (1), 115–140.

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, (4), 594–628.
Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, (1), 9–21.
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.
Epstein, S. (1973). The self-concept revisited or a theory of a theory. *American Psychologist*, *28*, (5), 404–416.
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.
García-Arch, J., Albert, M. S., & Fuentemilla, L. (2023). Selective integration of social feedback promotes a stable and positively biased self-concept. *Psyarxiv*, https://doi.org/10.31234/osf.io/3yd6g
Garcia-Arch, J., Barberia, I., Rodríguez-Ferreiro, J., & Fuentemilla, L. (2022). Authority brings responsibility: Feedback from experts promotes an overweighting of health-related pseudoscientific beliefs. *International Journal of Environmental Research and Public Health*, *19*, (22), 15154.
Hepper, E. G., Gramzow, R. H., & Sedikides, C. (2010). Individual differences in self-enhancement and self-protection strategies: An integrative analysis. *Journal of Personality*, *78*, (2), 781–814.
Höltje, G., Lubahn, B., & Mecklinger, A. (2019). The congruent, the incongruent, and the unexpected: Event-related potentials unveil the processes involved in schematic encoding. *Neuropsychologia*, *131*, 285–293.
Jiang, T., Wang, T., Poon, K. T., Gaer, W., & Wang, X. (2023). Low self-concept clarity inhibits self-control: The mediating effect of global self-continuity. *Personality and Social Psychology Bulletin*, *49*, (11), 1587–1600.
Kappes, A., & Sharot, T. (2019). The automatic nature of motivated belief updating. *Behavioural Public Policy*, *3*, (1), 87–103.
Kaufman, E. A., & Meddaoui, B. (2021). Identity pathology and borderline personality disorder: An empirical overview. *Current Opinion in Psychology*, *37*, 82–88.
Korn, C. W., Fan, Y., Zhang, K., Wang, C., Han, S., & Heekeren, H. R. (2014). Cultural influences on social feedback processing of character traits. *Frontiers in Human Neuroscience*, *8*, (1), 192.
Korn, C. W., La Rosée, L., Heekeren, H. R., & Roepke, S. (2016). Social feedback processing in borderline personality disorder. *Psychological Medicine*, *46*, (3), 575–587.
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.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, (13), 1–26.
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, (1), 177–190.
Martinelli, P., Sperduti, M., & Piolino, P. (2013). Neural substrates of the self-memory system: New insights from a meta-analysis. *Human Brain Mapping*, *34*, (7), 1515–1529.
Mokady, A., & Reggev, N. (2022). The role of predictions, their confirmation, and reward in maintaining the self-concept. *Frontiers in Human Neuroscience*, *16*, 824085.
Moscovitch, D. A., Moscovitch, M., & Sheldon, S. (2023). Neurocognitive model of schema-congruent and -incongruent learning in clinical disorders: Application to social anxiety and beyond. *Perspectives on Psychological Science*, *18*, (6), 1412–1435.
Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain-A meta-analysis of imaging studies on the self. *NeuroImage*, *31*, (1), 440–457.
Nowak, A., Vallacher, R. R., Tesser, A., & Borkowski, W. (2000). Society of self: The emergence of collective properties in self-structure. *Psychological Review*, *107*, (1), 39–61.
Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869.
Pelham, B. W., & Swann, W. B. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology*, *57*, (4), 672.

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Peters, A., Helming, H., Bruchmann, M., Wiegandt, A., Straube, T., & Schindler, S. (2024). How and when social evaluative feedback is processed in the brain: A systematic review on ERP studies. *Cortex*, *173*, 187–207.

Paul, K., Angus, D. J., Bublatzky, F., Dieterich, R., Endrass, T., Greenwood, L. M., Hajcak, G., Jack, B. N., Korinth, S. P., Kroczek, L. O. H., Lucero, B., Mundorf, A., Nolden, S., Peterburs, J., Pfabigan, D. M., Schettino, A., Shing, Y. L., Turan, G., Molen, M. J. W. van der, Wieser, M. J., Willscheid, N., Mushtaq, F., Pavlov, Y., & Pourtois, G. (2022). Revisiting the electrophysiological correlates of valence and expectancy in reward processing–A multi-lab replication. https://doi.org/10.31234/osf.io/4uy2c

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, (10), 2128–2148.

Richter, F. R. (2020). *Prediction errors indexed by the P3 track the updating of complex long-term memory schemas*. https://doi.org/10.1101/805887

Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in Cognitive Sciences*, *20*, (1), 25–33.

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*, (11), 1475–1479.

Swann, Jr., W. B., & Brooks, M. (2012). Why threats trigger compensatory reactions: The need for coherence and quest for self-verification. *Social Cognition*, *30*, (6), 758–777.

Swann, W. B., & Hill, C. A. (1982). When our identities are mistaken: Reaffirming self-conceptions through social interaction. *Journal of Personality and Social Psychology*, *43*, (1), 59–66.

Swann, W. B., Tafarodi, R. W., Wenzlaff, R. M., & Swann, L. B. (1992). Depression and the search for negative evaluations: More evidence of the role of self-verification strivings. *Journal of Abnormal Psychology*, *101*, (2), 314–317.

Taylor, S. E., Brown, J. D., Cantor, N., Emery, E., Fiske, S., Green-wald, T., Hammen, C., Lehman, D., McClintock, C., Nisbett, D., Ross, L., & Swann, B. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, (2), 193–210.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplementary Material

15516709, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70017 by Readcube (Labtiva Inc.), Wiley Online Library on [22/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License