

Authenticity is more than self-enhancement: behavioural and neurophysiological evidence

Chengli Huang^{1,2,*}, Emily K. Penney², Constantine Sedikides², Nicholas J. Kelley²

¹School of Mental Health, Wenzhou Medical University, Wenzhou, 325035, China

²Centre for Research on Self and Identity, University of Southampton, Southampton, SO17 1BJ, United Kingdom

*Corresponding author. School of Mental Health, Wenzhou Medical University, Wenzhou, 325035, China. E-mail: C.Huang@wmu.edu.cn.

Abstract

Negative self-descriptive information can be threatening to the self. This may depend, however, on the self-representation for which the information is relevant. We focused on two self-presentations, the authentic self and the presented self. In particular, we examined how the authentic and presented selves are influenced by emotional self-descriptiveness. Participants (N = 147) completed a self-referent emotional Stroop task while electroencephalogram (EEG) was recorded. They viewed in coloured text positive or negative traits exemplifying the authentic self ('I am genuinely honest'), the presented self ('I am outwardly honest'), or control ('It is clearly honest'). Colour naming latency was slower to negative (vs. positive) traits for the presented self and control. Colour naming latency was faster to negative (vs. positive) traits for the authentic self. Event-related potentials indicated that at both early (P1) and later (P3) stages of attentional processing, the authentic self exhibited comparable amplitudes to negative and positive traits. However, P1 was larger for negative, and P3 was larger for positive, traits for the presented self. Taken together, the findings highlight that the presented self is more pursuant of positivity, whereas the authentic self is more tolerant of negativity.

Keywords: authentic self; presented self; self-enhancement; self-consistency; self-accuracy; emotional Stroop effect

The self-concept is multifaceted (Marsh 1990, Sande 1990). A key facet involves the internal-external distinction. William James (1890) pioneered this distinction in terms of the spiritual (internal) and material or social (external) self. Contemporary theorists have framed the distinction as one between the private and public self (Baumeister 1986, Fenigstein 2009). Individuals are motivated to perceive the external expressions of the self (henceforth 'the presented self') in a positive light, extolling their strengths and under-emphasizing their weaknesses (Baumeister 1982, Leary 1995). Whereas some researchers suggest that self-enhancement motivation extends to aspects of the internal self (henceforth 'the authentic self'), others propose that the authentic self is driven by self-consistency or self-accuracy motivation (Sedikides and Schlegel 2024). We test these theoretical views by examining how negative (vs. positive) self-relevant information captures attention, both behaviourally and neurophysiologically, when the presented and authentic selves are salient.

The authentic self

Authenticity is 'the subjective perception that one is being the true, unvarnished "me"' (Sedikides et al. 2019, p. 73). Despite the construct's long history, dating back to Aristotelian thinking (Tredennick and Thomson 1976), its meaning has been a matter of controversy (Baumeister 2019, Hicks et al. 2019). A traditional view conceptualized authenticity through the lens of *self-accuracy*, the motivation to form an accurate image of the self or unbiasedly process self-relevant information (Kernis and Goldman 2006). In

support of this view, individuals high on authenticity are less defensive when faced with evidence that their behaviour is short of their ideals (Lahey et al. 2008). However, self-accuracy is difficult to attain or measure (Vazire and Wilson 2012). Moreover, the more unbiased individuals believe themselves to be, the more likely they are to report possessing more positive than negative characteristics, casting doubt on the veracity of their self-beliefs (Gillath et al. 2010). Another view conceptualizes authenticity through the lens of *self-consistency*, the motivation to maintain coherence among one's cognitions, emotions, and behaviours (Kernis and Goldman 2006, Wood et al. 2008). In support of this view, inconsistency between one's gender identity (female) and experimentally allocated self-presentation (masculine) reduces authenticity (Dormanen et al. 2020). However, individuals appraise their socially desirable behaviours as authentic, whether or not they align with their self-concept (Sheldon et al. 1997), and consider their positive behaviours as more authentic than negative ones (Jongman-Sereno and Leary 2016). These findings challenge the self-consistency view but are readily explained by the self-enhancement view.

According to the self-enhancement view, the authentic self encompasses predominantly positive characteristics. Consequently, individuals process self-relevant information so as to accentuate their strengths and downplay their weaknesses (Guenther and Sedikides, in press). In support of this view, individuals evaluate their true self as moral and positive (Strohming et al. 2017), and label the times when they behaved in accordance with a positive (vs. negative) trait as authentic (Bailey and Iyengar 2023). Similarly, the more positively individuals assess a change in their lives, the more

Received: 10 December 2024; Revised: 18 August 2025; Accepted: 1 October 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

likely they are to think that this change was fuelled by authenticity (Bench et al. 2015). Finally, laboratory experiments, individual difference studies, and daily diary studies point to reciprocal positive associations between authenticity and self-enhancement (Guenther et al. 2024).

The presented self

The presented self is a mental representation as integral to one's self-concept as the authentic self. Stakes are high for the presented self, given that it can facilitate or undermine cooperation, reputation, respect, status, as well as access to social networks, professional opportunities (e.g. jobs, promotions, and housing), and personal resources (e.g. friends and partners; Vonasch et al. 2018, Dores Cruz et al. 2021). Hence, self-presentation promotes a sanitized portrait of the individual, exaggerating, if not glorifying strengths, while minimizing, if not concealing, weaknesses (Baumeister 1982, Hancock and Toma 2009).

Humanistic theories and person-centred therapies propose that individuals often modify or distort their social behaviour to conform to perceived 'conditions of worth' imposed by their surrounding environment (Rogers 1951, Tunnel 1984). For instance, smiling at an unfunny joke to fit in the social gathering, or feigning enthusiasm for the company's mission during a job interview, might be motivated by the desire to evade negative evaluations from others. Empirical research supports this notion, as individuals often engage in strategic self-presentation, denying negative traits and drawing attention to positive ones (Roth et al. 1986, Lee et al. 1999). Furthermore, positive self-relevant contexts can elicit a smaller N400 than positive other-relevant contexts, suggesting that individuals are more likely to expect positive information when referring to themselves (Fields and Kuperberg 2015).

Current investigation

Emotional Stroop task

To test the three views of authenticity, we recorded behavioural and neurophysiological (event-related potential or ERP) responses to emotionally charged self-evaluations in a modified Stroop task, the Emotional Stroop Task (Mathews and MacLeod 1985, Watts et al. 1986). Here, participants identify the ink colour of emotionally evocative (or neutral) words (Fig. 1). Both the traditional Stroop and Emotional Stroop tasks demonstrate that the meaning of the text captures attention, even when it is irrelevant to the task. The classic Stroop effect arises from a direct semantic conflict (or

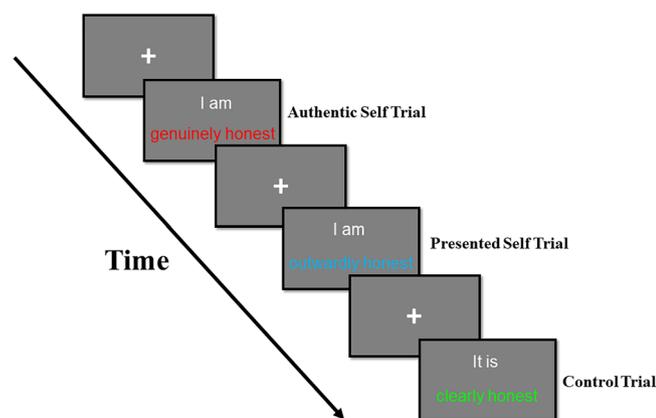


Figure 1. Trial event diagram.

incongruence) between word meaning and ink colour (e.g. the word 'red' printed in green vs. red ink). In contrast, the emotional Stroop effect delves into the discrepancy between emotionally salient versus neutral words (e.g. 'death' vs. 'door' printed in red), without semantic conflict central to the classic Stroop effect. This fundamental distinction positions the emotional Stroop task as a unique paradigm: where the traditional Stroop task reveals competition between automatic reading and controlled colour-naming processes, the emotional Stroop task reflects a generalized slowing due to threat detection mechanisms (Algom et al. 2004). Consequently, the emotional Stroop effect refers to the slowdown in responding to the ink colour of negative (vs. positive or neutral) words (Williams et al. 1996, Bar-Haim et al. 2007, Phaf and Kan 2007), resulting from the automatic attention allocation to threatening stimuli at the expense of concurrent task demands (Öhman 1993, Öhman et al. 2001). Moreover, such negativity occurs rapidly and automatically in emotion information processing (Pratto and John 1991, Anderson et al. 2003, Zald 2003) and is stronger than positivity (Baumeister et al. 2001). For example, the motivation to avoid negative self-definitions typically outweighs the drive to pursue positive self-definitions (Sedikides 2012, Sedikides et al. 2016). Leveraging the emotional Stroop effect's sensitivity to negativity, we used this task to assess the resistance of the authentic and presented self to negativity.

The presented self is predominantly shaped by self-enhancement motivation, which promotes linking the self-concept to positive than negative information (Baumeister 1982, Leary and Kowalski 1990, Sedikides and Gregg 2008). Hence, negative self-relevant information will disproportionately affect self-representation. Specifically, when maintaining a positively biased self-presentation, exposure to negative self-descriptors creates greater cognitive-affective dissonance than in the authentic self, thereby amplifying their salience and disruptive potential. As such, we hypothesized that in an emotional Stroop task, interference of negative self-evaluations would be more pronounced for the presented self (an amplified emotional Stroop effect). The case for the authentic self is equivocal. On the one hand, building on the self-consistency/self-accuracy views, individuals would integrate rather than avoid negative self-aspects, and such balanced processing of all self-relevant information would decrease their attentional salience to negativity and thus minimize defensive responses to negative self-representation. If so, the interference of negative self-evaluations would be less pronounced for the authentic than presented self (an attenuated emotional Stroop effect relative to the presented self, but amplified compared to the control condition). On the other hand, self-enhancement motivation might drive the authentic self, much like the presented self. If so, the interference of negative self-evaluations would be on par with the effect observed for the presented self (an amplified emotional Stroop effect).

Event-related potential assessment

We also recorded electroencephalography (EEG) activity during the emotional Stroop task. Prior ERP research has delineated distinct temporal stages of emotional information processing: the P1, which reflects initial threat detection; the N170 and early posterior negativity, which reflect emotional and non-emotional discrimination; and the later components (the P3 and late positive component), which distinguish between positive and negative words (Luo et al. 2010, Zhang et al. 2014). This temporal framework provides a neurocognitive basis for examining emotional processing dynamics in our self-referential Emotional Stroop task. Taken together, we focused on the ERPs at early-stage selective attention (P1; Batty

and Taylor 2003, Pourtois et al. 2005, Zhang et al. 2014), attention allocation to emotionally evocative stimuli (N170; Williams et al. 2006, Zhang et al. 2014, Cai et al. 2016), and early (P2; Fields and Kuperberg 2012, Fan et al. 2016, San Martín et al. 2016) as well as late (P3; Gray et al. 2004, Tacikowski and Nowicka 2010) stages of self-relevant information processing. We describe the pertinent components below.

P1. The P1 is an early sensory-evoked component, emanating from parieto-occipital regions as early as 60 ms post-stimulus (Luck 2014). It reflects the selective amplification of sensory information (Hillyard et al. 1998) and is sensitive to emotional stimuli (Mueller et al. 2013, Schindler and Bublatzky 2020), with larger P1 amplitudes evoked by negative compared to neutral stimuli (e.g. faces, words; Batty and Taylor 2003, Van Hooff et al. 2008, Luo et al. 2010) indicating that the P1 can differentiate between threatening and non-threatening information (Zhang et al. 2014, Huang et al. 2025). The P1 emerges during the initial phase, representing an evolutionary adaptation for threat vigilance (Luo et al. 2010, Zhang et al. 2014). Such early emotional response might signify rapid extraction of emotion-related information and function, at least in part, independently of subsequent emotional processes (N170; Vuilleumier and Pourtois 2007).

N170. The N170 is a negative deflection that typically peaks approximately 170 ms after stimulus onset over lateral occipito-temporal regions (Luck 2014). The N170 reflects early rapid attention to visual stimuli (Rossion and Jacques 2012), with larger N170 amplitudes representing the allocation of more attentional resources (Cai et al. 2016). Moreover, the N170 is modulated by emotional stimuli, with a substantially enhanced amplitude for emotional relative to neutral stimuli (Williams et al. 2006, Luo et al. 2010, Zhang et al. 2014, Schindler et al. 2023), especially for negative (vs. positive) ones (Montalan et al. 2008, Cai et al. 2016).

P2. The P2 is a positive deflection spanning 150–250 ms over the anterior-central region (Luck 2014). It exhibits greater amplitudes in response to stimuli containing target features, indicating early selective attention towards task-relevant stimuli (Potts 2004, Potts et al. 2006). This amplification is pronounced when the targets are infrequent (Luck and Hillyard 1994, Glazer and Nusslock 2022). Furthermore, the P2 has been associated with emotional processing, suggesting that its role in modulating selective attention is influenced by emotional content (Kotz and Paulmann 2011, Hajcak et al. 2012). However, findings regarding the modulation of P2 by emotion are mixed. Although some studies reported increased P2 amplitudes with emotional stimuli compared to neutral ones, others found the opposite pattern (Schindler and Bublatzky 2020). Unlike the earlier emotion-detection stages represented by the P1 and N170, the P2 is associated with higher-order, deeper, and conscious emotional processing (Prete et al. 2015, 2018, Nie et al. 2020). Moreover, research has implicated the P2 in self-referential processing, but findings are inconsistent. Some studies obtained reduced P2 amplitudes for self-related stimuli (Keyes et al. 2010, Liu et al. 2019), whereas others reported increased amplitudes (Fields and Kuperberg 2012, Fan et al. 2016, San Martín et al. 2016) or null effects (Yang et al. 2014).

P3. The P3 is a maximal positive wave that typically peaks around 300 ms post-stimulus at the midline parietal region (Luck 2014). The P3 has attracted a lot of interest for its iconic increased positivity following the presentation of infrequent and surprising (low probability) stimuli (Pritchard 1981, Polich 2007). Although P2 and P3 are larger for infrequent stimuli, the P2 effect occurs only when the target is defined by simple stimuli, whereas the P3 effect can occur for complex stimuli (Song et al. 2005, Barkaszi et al. 2013,

Luck 2014). Furthermore, ERP studies of self-referential processing show that P3 is often associated with the discrimination of self from others: a larger P3 wave follows the presentation of self-related objects, words, names, and faces relative to the same stimuli of other persons (Gray et al. 2004, Miyakoshi et al. 2007). Thus, the amplitude of the P3 might reflect increased attention or deeper processing of self-relevant stimuli (Porter et al. 2021).

Hypotheses

We assessed behavioural (reaction times) and neurophysiological (P1, N170, P2, and P3) responses to positive and negative traits indicative of the authentic and presented selves in a modified Emotional Stroop Task. In terms of the presented self, we hypothesized an amplified emotional Stroop effect alongside an elevated P1 and N170 for negative versus positive traits. In terms of the authentic self, we offered competing hypotheses. First, according to the self-enhancement view, we would observe the same behavioural (amplified emotional Stroop effect) and neurophysiological (elevated P1 and N170) pattern as for the authentic self. However, according to the self-consistency/self-accuracy views, we would observe an attenuated emotional Stroop effect alongside an attenuated P1 and N170 for negative versus positive traits. We approached the ERPs for the P2 and P3 exploratorily, due to mixed findings regarding emotional and self-referential processing (P2) and lack of electrophysiological studies on authenticity (P3).

Method

Design and participants

We implemented a 3 (self: authentic self, presented self, control) × 2 (valence: positive traits, negative traits) within-subjects design. We used G*Power (Faul et al. 2009) assuming a small effect size (Cohen's $f = .10$), six measures (based on the 3 × 2 design), $\alpha = .05$, power ($1 - \beta$) = .80, and a moderate correlation among the repeated measures ($r = .50$). Based on these parameters, a minimum $N = 109$ was required. We decided to recruit participants throughout the academic year, testing 162 University of [MASKED] undergraduate psychology student volunteers. Based on a-priori criteria, we excluded 15 participants: five encountered EEG acquisition device failures (Tacikowski and Nowicka 2010), two evinced over 50% missing data after cutting the 1% slowest and 1% fastest correct trials (Ratcliff 1993), two manifested mean reaction time exceeding ± 3 SDs (Cai et al. 2016), and six had more than 50% of their trials rejected due to artefacts in the EEG data (Imbir et al. 2021). The final sample comprised 147 participants (114 women, 31 men, 2 non-binary) aged 18–46 years ($M = 19.56$, $SD = 2.87$). We did not collect ethnicity information, but over 90% of the university's undergraduate students are White. Sensitivity analyses (G*Power; Faul et al. 2009) indicated 80% power to detect effects as small as Cohen's $f = .086$ ($\eta^2 = .007$).

Stimuli and procedure

To generate stimulus materials, we relied on Anderson's (1968) list, a compendium of personality traits rated for likableness and meaningfulness. We selected 60 positive and 60 negative traits. The likableness of the selected positive traits ($M = 4.66$, $SD = 0.45$) was higher than the likableness of the selected negative traits ($M = 1.25$, $SD = 0.46$), $t(118) = 40.89$, $P < .001$, Cohen's $d = 7.47$. The word length of the selected positive traits ($M = 7.78$, $SD = 1.09$) was not significantly different from the selected negative traits ($M = 7.68$, $SD = 1.07$), $t(118) = 0.51$, $P = .612$. We programmed and administered the experiment using PsychoPy (Version 2021.2.3; Peirce 2007).

Participants completed a modified emotional Stroop task in a quiet laboratory environment via computer and in the context of a larger study. (Participants also completed a Flanker task and a Monetary Incentive Delay Task. At the end, they completed a battery of personality questionnaires unrelated to the current investigation.) They were presented with a series of sentences and instructed to name the colour of each sentence, while ignoring its meaning, by pressing corresponding keys (V for red, B for blue, N for green) as quickly and accurately as possible. One third of these sentences described the authentic self (e.g. 'I am genuinely ingenious'), one third described the presented self (e.g. 'I am outwardly unkind'), and one third constituted the control condition (e.g. 'It is clearly honest') encompassing both positive and negative traits. We administered 360 trials across four blocks. Each block of 90 trials included an equal number of authentic self, presented self, and control trials. Each of these three trial sets consisted for an equal number of positive and negative traits. In all, there were 60 trials in each of the following bins: authentic self/positive, authentic self/negative, presented self/positive, presented self/negative, control/positive, control/negative. A trial started with the presentation of a central fixation cross for 800–1200 ms. Then, the colour sentence appeared on the screen until a response (key-pressing) occurred, followed by an 800 ms inter-stimulus interval (Fig. 1 for the trial event diagram). Prior to the formal task, participants underwent 12 practice trials to familiarize themselves with the colour-key mapping.

Data recording and data analysis

We collected the EEG data from 32 scalp sites using Ag/AgCl electrodes embedded in a flexible cap (Brain Products, Germany), with an online reference to Cz. We mounted a ground electrode positioned at Fpz. We recorded the vertical electrooculogram (VEOG) below the right eye, based the electrode cap on the 10–20 system, and kept electrode impedances below 10 k Ω . We amplified and sampled the signals at 500 Hz with an online bandpass filter from DC to 140 Hz (–6 dB point, half-amplitude cutoff).

In offline processing, we initially pre-processed the EEG data by using EEGLAB, an open-source toolbox running in the MATLAB environment (Delorme and Makeig 2004). We digitally filtered the EEG data with a band-pass filter (high pass: 0.10 Hz, low pass: 40 Hz, 50 Hz notch; –6 dB point, half-amplitude cutoff) with a zero-phase FIR (finite impulse response) filter via Hamming-windowed sinc filtering. The filter order was automatically determined based on the transition bandwidth by the EEGLAB defaults. We segmented the EEG data from 200 ms prior to 800 ms following the onset of each word, and baseline corrected them to the 200 ms pre-stimulus baseline along with re-referencing them to the mastoids average (TP9, TP10). We detected bad channels by visual inspection of the waveforms and replaced them with a spherical spline identified interpolation (SSI; Perrin et al. 1989). We corrected segments contaminated by blinks, eye movements, and other artefacts using an independent component analysis (ICA) algorithm (e.g. fixed-point ICA, joint approximate diagonalization of eigen-matrices; Delorme and Makeig 2004) and applied ICLabel—a proposed statistical model—to automatically label ICs (Pion-Tonachini et al. 2019). Also, we excluded bad segments where a voltage deviation on any channel of $\pm 75 \mu\text{V}$ from the average baseline voltage.

Following best practices (Luck and Gaspelin 2017) and similar lines of research, we quantified: (a) P1 as the average peak amplitude from 80 to 130 ms after stimulus onset over lateral occipital electrode cluster (O1, OZ, O2; Wieser and Moscovitch 2015, Jetha et al. 2021); (b) N170 as the average peak amplitude from 140 to 200 ms after stimulus onset over lateral posterior electrode cluster

(P3, P4, P7, P8; Keyes et al. 2010, Cai et al. 2016); (c) P2 as the average peak amplitude from 150 to 250 ms after stimulus onset over mid-line region (Fz, Cz, Pz; Fan et al. 2013, 2016); and (d) P3 as the average peak amplitude from 300 to 400 ms after stimulus onset over mid-line region (Fz, Cz, Pz; Gray et al. 2004, Wada et al. 2019, Riggins and Scott 2020).

The main dependent variables were reaction times (RT) and ERPs (P1, N170, P2, P3). We only used correct responses (Montalan et al. 2008). We created a RT data processing pipeline (Morís Fernández and Vadillo 2020). Specifically, we removed: (a) the 1% slowest and 1% fastest trials; (b) participants with more than 50% missing data; (c) participants with a mean RT exceeding ± 3 SDs. Further, we averaged the ERPs for each of the six experimental conditions. We analyzed the data in SPSS (Version 24), addressing multiple comparisons with Bonferroni corrections.

Results

We conducted 3 (self) \times 2 (valence) repeated measures Analyses of Variance on RT and ERPs (P1, N170, P2, P3).

Reaction times

The main effect of self was significant, $F(2, 145) = 4.40$, $P = .014$, $\eta_p^2 = .06$. Participants responded faster on authentic-self ($M = 625.05$, $SD = 108.79$) than presented-self ($M = 631.28$, $SD = 109.59$) traits, $P = .028$, 95% CI = [–11.97, –0.49]. They did not differ in their speed of responding to control ($M = 626.72$, $SD = 106.26$) and presented-self ($M = 631.28$, $SD = 109.59$) traits, $P = .071$, 95% CI = [–0.27, 9.38], or control and authentic-self traits, $P = .999$, 95% CI = [–7.60, 4.25]. Further, the main effect of valence was significant. As per the emotional Stroop effect (Williams et al. 1996), participants responded slower to negative ($M = 629.97$, $SD = 109.54$) than positive ($M = 625.39$, $SD = 106.22$) traits, $F(1, 146) = 4.19$, $P = .042$, $\eta_p^2 = .03$.

Crucially, the Self \times Valence interaction was significant, $F(2, 145) = 30.88$, $P < .001$, $\eta_p^2 = .30$ (Fig. 2). The prototypical emotional Stroop effect emerged on control trials: Participants were slower to respond to negative ($M = 630.96$, $SD = 109.53$) than positive ($M = 622.48$, $SD = 107.13$) traits, $P = .016$, 95% CI = [1.60, 15.37]. The emotional Stroop effect was amplified on presented-self trials: Participants were even slower to respond to negative ($M = 642.40$, $SD = 117.98$) than positive ($M = 620.16$, $SD = 105.17$) traits, $P < .001$, 95% CI = [15.10, 29.38]. Finally, the emotional Stroop effect was attenuated on authentic-self trials: Participants responded faster to negative ($M = 616.55$, $SD = 108.29$) than positive ($M = 633.54$, $SD = 112.77$) traits, $P < .001$, 95% CI = [–23.41, –10.58].

Event-related potentials

We observed a pronounced emotional Stroop effect for the presented self and a diminished emotional Stroop effect for the

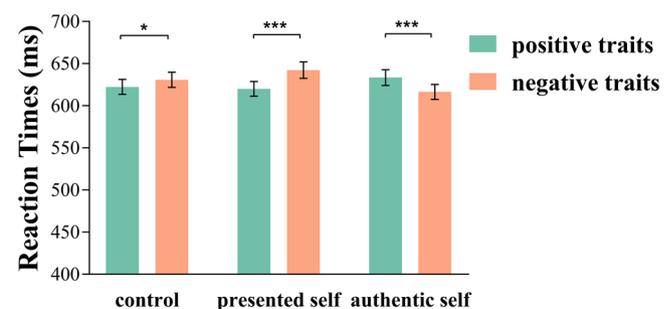


Figure 2. Reaction times to positive traits and negative traits for control, presented self, and authentic self.

authentic self. We next turned to ERPs relevant to early-stage selective attention (P1), attention allocation to emotionally evocative stimuli (N170), and early (P1) and late (P3) stages of self-relevant information processing in search of an explanation for this behavioural effect (the Self \times Valence interaction on RTs).

P1. The Self \times Valence interaction was significant, $F(2, 145) = 3.75$, $P = .026$, $\eta_p^2 = .05$ (Fig. 3a). Negative traits ($M = 4.19$, $SD = 3.77$) elicited a larger P1 than positive traits ($M = 3.91$, $SD = 3.71$) for the presented self, $P = .019$, 95% CI = [0.05, 0.53]. The P1 did not differ between negative ($M = 3.97$, $SD = 3.92$) and positive ($M = 3.98$, $SD = 3.74$) traits for the authentic self, $P = .919$, 95% CI = [-0.27, 0.30], nor did it differ between negative ($M = 3.67$, $SD = 3.67$) and positive ($M = 3.98$, $SD = 3.68$) control traits, $P = .052$, 95% CI = [0.00, 0.62]. Neither the main effect of self, $F(2, 145) = 1.50$, $P = .227$, $\eta_p^2 = .02$, nor that of valence, $F(2, 145) = 0.03$, $P = .855$, $\eta_p^2 < .001$, was significant.

N170. The main effect of self was significant, $F(2, 145) = 3.71$, $P = .027$, $\eta_p^2 = .05$ (Fig. 3b). The N170 was larger for the presented self ($M = -2.31$, $SD = 2.57$) than control ($M = -2.11$, $SD = 2.73$) traits, $P = .027$, 95% CI = [-0.39, -0.02]. Presented-self and authentic-self ($M = -2.28$, $SD = 2.56$) traits did not differ, $P = .999$, 95% CI = [-0.19, 0.14] and neither did authentic-self and control traits, $P = .088$, 95% CI = [-0.37, 0.02]. The main effect of valence, $F(1, 146) = 1.80$, $P = .182$, $\eta_p^2 = .01$, and the Self \times Valence interaction, $F(2, 145) = 0.37$, $P = .695$, $\eta_p^2 = .01$, were not significant.

P2. The main effect of self was significant, $F(2, 145) = 27.27$, $P < .001$, $\eta_p^2 = .27$ (Fig. 3c). The P2 was larger on control traits ($M = 6.58$, $SD = 3.53$) compared to both authentic-self ($M = 6.23$, $SD = 3.42$) and presented-self ($M = 5.93$, $SD = 3.47$) traits, $P = .004$, 95% CI = [0.09, 0.61] and $P < .001$, 95% CI = [0.43, 0.86]. These results are consistent with findings showing substantially reduced P2 amplitudes for self-relevant stimuli (Keyes et al. 2010, Liu et al. 2019). The P2 was also larger on authentic-self compared to presented-self traits, $P = .008$, 95% CI = [0.06, 0.53]. Thus, we observed a linear pattern where the P1 was largest for control traits, intermediate for authentic-self traits, and smallest for presented-self traits. Neither the main effect of valence, $F(1, 146) = 0.05$, $P = .993$, $\eta_p^2 < .001$, nor the Self \times Valence interaction, $F(2, 145) = 0.06$, $P = .940$, $\eta_p^2 = .001$, was significant.

P3. The main effect of self was significant, $F(2, 145) = 16.07$, $P < .001$, $\eta_p^2 = .18$. (For P3, the results of three electrodes separately differed somewhat from the analysis that averaged all three electrode sites together. Specifically, we obtained the Self \times Valence interaction at FZ and CZ, but not at PZ. For detailed results of each electrode, see the [Supplementary Materials](#) (pp. 2–3).) The P3 was larger on authentic-self ($M = 7.54$, $SD = 4.12$) than presented-self ($M = 7.20$, $SD = 4.28$) traits, $P = .026$, 95% CI = [0.03, 0.66] and control ($M = 6.87$, $SD = 4.20$), $P < .001$, 95% CI = [0.39, 0.97] traits. The P3 was larger on presented-self compared to control traits, $P = .021$, 95% CI = [0.04, 0.063]. The finding that P3 was larger for self-relevant stimuli is compatible with the literature (Gray et al. 2004, Miyakoshi et al. 2007, Tacikowski and Nowicka 2010). The main effect of valence was not significant, $F(1, 146) = 1.62$, $P = .205$, $\eta_p^2 = .01$. However, the important Self \times Valence interaction was significant, $F(2, 145) = 4.07$, $P = .019$, $\eta_p^2 = .05$ (Fig. 3d). The P3 did not differ between negative and positive traits in the control ($P = .346$, 95% CI = [-0.19, 0.54]) and authentic-self ($P = .187$, 95% CI = [-0.56, 0.11]) condition. Yet, the P3 was larger for positive ($M = 7.41$, $SD = 4.37$) than negative ($M = 6.99$, $SD = 4.18$) traits in the presented-self condition, $P = .007$, 95% CI = [0.12, 0.73]. (Although the Self \times Valence hypothesis posits a stronger P3 to negative words due to their lower expectedness or congruence (Pritchard 1981, Polich 2007), the evidence more robustly aligns with the self-enhancement view. First, the main effect of self revealed a larger P3 for both self conditions compared to control. This effect likely reflects enhanced attention and

processing of self-relevant information relative to control rather than increased surprise or unexpectedness for self-related stimuli. Second, the absence of a significant main effect of valence contradicts the unexpectedness account, which would predict larger P3 responses to negative words due to their presumed incongruence. Finally, simple effects analyses of the Self \times Valence interaction showed that for positive words, the P3 was larger for both the authentic self and the presented self compared to the control condition, whereas for negative words, the P3 was larger for the authentic self than for both the presented self and the control condition. Interpreted through the lens of unexpectedness, this would imply that positive and negative words in self-relevant conditions (authentic self and presented self) were more surprising than in the non-self-relevant condition (control condition), a counterintuitive proposition.) Alternatively, when processing positive words, the P3 was larger for both the authentic ($M = 7.43$, $SD = 4.02$; $P = .004$, 95% CI = [0.12, 0.83]) and presented ($M = 7.41$, $SD = 4.37$; $P = .047$, 95% CI = [0.004, 0.91]) self relative to the control condition ($M = 6.95$, $SD = 4.13$), with the authentic- and presented-self conditions not differing significantly, $P = .999$, 95% CI = [-0.41, 0.45]. In contrast, when processing negative words, the P3 was larger for the authentic-self ($M = 7.66$, $SD = 4.23$) than for both the presented-self ($M = 6.99$, $SD = 4.18$; $P < .001$, 95% CI = [0.26, 1.08]) and control ($M = 6.78$, $SD = 4.27$; $P < .001$, 95% CI = [0.42, 1.33]) condition, with the authentic- and presented-self conditions not differing significantly, $P = .650$, 95% CI = [-0.20, 0.61].

Discussion

We decomposed the self into two mental representations, presented and authentic. Accordingly, we examined competing views of authenticity (self-enhancement vs. self-consistency/self-accuracy) using behavioural and neurophysiological measures in a modified Emotional Stroop Task. Results largely favoured the self-consistency/self-accuracy views. On control trials, we demonstrated a prototypical emotional Stroop effect (slowdown for negative compared to positive information). This effect was amplified on presented-self trials, attributable to the potent self-enhancement motivation driving self-presentation (Schlenker and Pontari 2000, Alicke and Sedikides 2009, Sedikides et al. 2015). On authentic-self trials, however, the emotional Stroop effect was attenuated, as per the self-consistency/self-accuracy views.

The ERP results help to explain the behavioural effects. The earliest stages of selective attention (P1) largely echoed our behavioural findings. Negative (vs. positive) traits elicited a larger P1 for the presented self. However, there was no difference in P1 amplitudes between negative and positive traits for the authentic self, a pattern compatible with the self-consistency/self-accuracy views. The P1 is sensitive to emotional stimuli (Mueller et al. 2013, Schindler and Bublatzky 2020), especially threat-related information, representing an evolutionary adaptation for threat vigilance (Luo et al. 2010, Zhang et al. 2014). Therefore, these findings indicate that the presented self is strongly motivated by self-enhancement and is consequently more susceptible to threatening information. Accordingly, attentional resources are involuntarily allocated towards negative self-descriptive stimuli during the early stages of attentional allocation. In contrast, the authentic self acknowledges both strengths and weaknesses (Kernis and Goldman 2006), rendering it possible to distribute attentional resources more evenly between positive and negative traits. Moreover, the literature indicates that emotional stimuli are rapidly processed in the early attention stage, with self-referent processing arising later (Herbert et al. 2011, Zhu et al. 2016, Schäfer et al. 2020). Hence, our P1 findings may represent

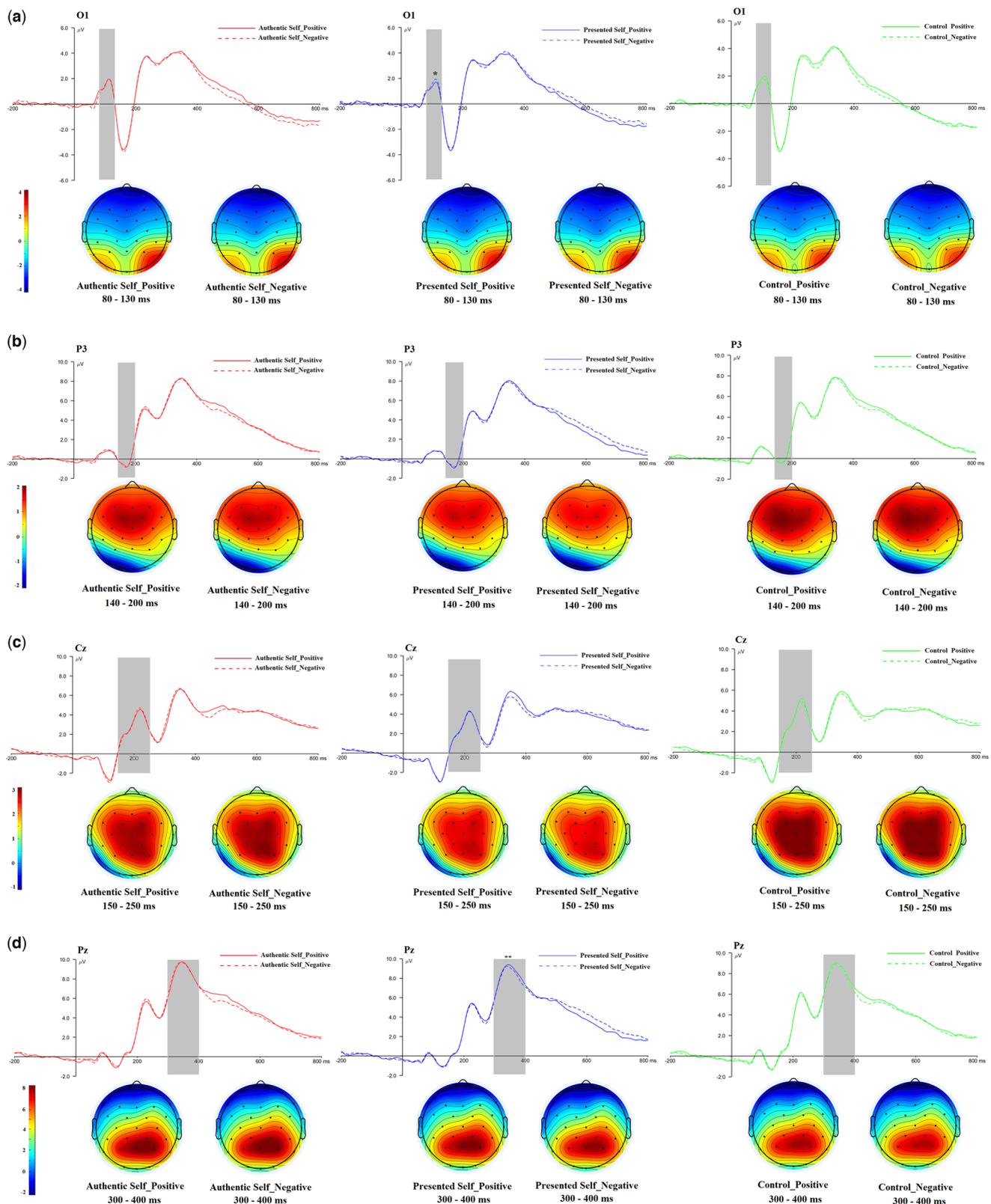


Figure 3. Grand averages waveforms and topographies of ERPs (P1: a; N170: b; P2: c; P3: d) in the authentic self, presented self, and control conditions.

an initial demonstration of self-relevant modulation of emotional processing in this early attention stage. This modulation, if replicated by future work, would mark a novel addition to understanding of the interplay between the self and emotional processing in this early attention stage. Although some psycholinguistic research (Serenio et al. 2003) suggests minimal influence of lexical-semantic

properties on early ERP components, we observed significant P1 modulation, which may be attributed to the Emotional Stroop task we employed. Specifically, by restricting task demands to low-level perceptual processing (colour identification) while presenting emotionally charged words, we ensured that emotional processing remained implicit—an automatic, unconscious processing that

may enhance the ability to detect early-stage automatic attentional biases towards emotional stimuli. Researchers could conduct independent replications of the P1 effect to verify its robustness.

Although the N170 component did not differentiate between positive and negative self-descriptiveness, N170 amplitudes were larger for presented self than control trials, a pattern that warrants further exploration. Overall, these findings suggest that the presented self selectively heightens attention to negativity during early processing, whereas the authentic self lowers selective attention to negativity. We carefully matched stimuli across emotion conditions for basic lexica properties known to influence early ERPs (Hauk and Pulvermüller 2004). For example, the word length of the selected positive and negative traits was not significantly different. Moreover, the implementation of within-subjects design addresses potential issues arising from orthographic neighbourhood and bigram frequency.

We observed an interaction between self and valence at the later processing stage, the P3 (but not P2). Whereas there was no difference between the P3 to negative versus positive traits for the authentic self (as per the self-consistency/self-accuracy views), the P3 was larger for positive than negative traits for the presented self (as per the self-enhancement view). Also, although both P2 and P3 are larger for infrequent and salient stimuli, modulation of the P2 occurs only when the target is defined by simple stimuli, but modulation of the P3 can occur for complex stimuli (Song et al. 2005, Barkaszi et al. 2013, Luck 2014), as stated earlier. The presence of the Self × Valence interaction for the P3 (and not the P2) bolsters the representational richness of the self (Kihlstrom et al. 1988, Sedikides and Gregg 2003, McConnell 2011). Moreover, the lack of difference in the P3 response to negative and positive traits for the authentic self suggests that, at this later processing stage, negativity and positivity are comparably relevant to the authentic self. Similarly, the larger P3 for positive (vs. negative) traits for the presented self indicates that, at this stage, positivity is novel and salient to the presented self. This shift contrasts with the earlier stage (P1), where the presented self exhibited heightened sensitivity to negativity. This pattern can be accounted for by the mobilization-minimization hypothesis (Taylor 1991; see also Sedikides et al. 2016), according to which negative or threatening information triggers swift physiological, cognitive, emotional, and social responses (i.e. mobilization), followed by counteractions to minimize, undo, or even reverse these initial responses (i.e. minimization). In our research, negative self-descriptive information received preferential processing initially (mobilization; P1), followed by preferential processing of positive, self-descriptive information (minimization; P3).

We observed the reverse P3 pattern in the frontal-central rather than parietal region, in line with neuroimaging evidence that self-referential processing predominantly activates cortical midline structures (CMS; Northoff and Bermpohl 2004, Uddin et al. 2007). Specifically, self-referential emotional processing involves anterior CMS regions such as the ventromedial prefrontal cortex, anterior cingulate cortex, and dorsomedial prefrontal cortex (Northoff et al. 2006). Thus, the frontal-central distribution of our reverse P3 effects in the Emotional Stroop Task with self-referential stimuli provides electrophysiological support for this CMS framework.

The main effects of self enrich understanding of the P3. ERP studies of self-referential processing show that the P3 is frequently larger following presentation of self-relevant objects, words, names, and faces relative to identical stimuli describing another person (Gray et al. 2004, Miyakoshi et al. 2007, Tacikowski and Nowicka 2010). Consequently, the P3's amplitude might reflect increased attention or deeper processing of self-relevant stimuli (Porter et al. 2021). We

replicated this finding by demonstrating that P3 was larger for the authentic and presented self than in the control trials. We then extended this finding by illustrating that the effects of self-reference on the P3 are stronger for the authentic self. Hence, deeper processing of self-relevant stimuli (Porter et al. 2021, Levorsen et al. 2023) may be driven more by the authentic than presented self. Future investigations might address this possibility.

Limitations

We did not provide participants with operational definitions of the terms 'genuinely' and 'outwardly', instead relying on their lexical comprehension. Although these terms lexically map onto the constructs of 'authentic self' and 'presented self' ('Genuine' denotes authenticity: 'Having the character or origin represented; real, true, not counterfeit, unfeigned, unadulterated' (Oxford University Press, 2025a). 'Outward' denotes the presented self: 'Of or relating to thoughts, attitudes, actions, etc, which manifest externally, as opposed to those which are experienced internally; merely exterior; public. Also: of or relating to the outer or visible form of something' (Oxford University Press, 2025b).), respectively, this approach may have introduced variability in how participants interpreted them. To increase methodological precision, future research could clarify such manipulations by defining the terms and ensuring their comprehension via pretesting.

Also, we used whole-sentence rather than word-by-word presentation. This approach may increase ocular artefacts compared to word-by-word presentation and introduce variability in the precise timing of word processing. Given EEG's temporal precision (Cohen 2017), even a 100ms difference in when participants read the critical stimuli could impact both the selection of analysis time windows and the interpretation of observed components as reflecting specific neural processes. However, alternative sequential presentation of self-referential and emotional words would create its own confounds: congruent colours between words might produce facilitation effects, whereas incongruent colours could lead to interference in colour judgement. Future studies might implement more sophisticated designs, such as rapid serial visual presentation or masked priming approaches, to simultaneously address both temporal precision and artefact minimization concerns.

Conclusion

Distinct self-representations, the authentic and presented selves, are differential susceptibility to negative self-relevant information, behaviourally and neurophysiologically. From the self-enhancement view, the presented self is particularly vulnerable to negative self-descriptors. In contrast, from the self-consistency/self-accuracy views, the authentic self integrates both positive and negative self-aspects with comparable weight. The findings aligned with the latter views, suggesting that authenticity extends beyond mere self-enhancement.

Author contributions

Chengli Huang (Data Curation-Lead, Formal Analysis-Lead, Funding Acquisition-Lead, Investigation-Lead, Methodology-Lead, Project Administration-Equal, Visualization-Lead, Writing Original Draft-Lead), Emily K. Penney (Data Curation-Equal, Investigation-Equal, Writing Original Draft-Supporting, Writing Review Editing-Supporting), Constantine Sedikides (Conceptualization-Equal, Project Administration-Equal, Supervision-Supporting, Writing Review Editing-Lead), and Nicholas J. Kelley (Conceptualization-Equal,

Project Administration-Equal, Supervision-Lead, Writing Original Draft-Equal, Writing Review Editing-Lead).

Supplementary Material

See the [online supplementary material](#) for detailed stimulus validation procedures, additional analyses, and figures.

Conflict of interest

None.

Funding

This research was supported by Economic and Social Research Council (ESRC) South Coast Doctoral Training Partnership (SCDTP) Grant ES/P000673/1.

Data availability

All data, code, and materials used in the analysis are available on Open Science Framework (https://osf.io/64sq8/?view_only=99c43f1d13774d30ace7879a5cecbfcc).

References

- Algom D, Chajut E, Lev S. A rational look at the emotional Stroop phenomenon: a generic slowdown, not a Stroop effect. *J Exp Psychol Gen* 2004;**133**:323–38. <https://doi.org/10.1037/0096-3445.133.3.323>
- Alicke MD, Sedikides C. Self-enhancement and self-protection: What they are and what they do. *Eur Rev Soc Psychol* 2009;**20**:1–48. <https://doi.org/10.1080/10463280802613866>
- Anderson NH. Likableness ratings of 555 personality-trait words. *J Pers Soc Psychol* 1968;**9**:272–9. <https://doi.org/10.1037/h0025907>
- Anderson AK, Christoff K, Panitz D et al. Neural correlates of the automatic processing of threat facial signals. *J Neurosci* 2003;**23**:5627–33. <https://doi.org/10.1523/JNEUROSCI.23-13-05627.2003>
- Bailey ER, Iyengar SS. Positive—more than unbiased—self-perceptions increase subjective authenticity. *J Pers Soc Psychol* 2023;**125**:1351–72. <https://doi.org/10.1037/pspa0000359>
- Bar-Haim Y, Lamy D, Pergamin L et al. Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol Bull* 2007;**133**:1–24. <https://doi.org/10.1037/0033-2909.133.1.1>
- Barkaszi I, Czizler I, Balázs L. Stimulus complexity effects on the event-related potentials to task-irrelevant stimuli. *Biol Psychol* 2013;**94**:82–9. <https://doi.org/10.1016/j.biopsycho.2013.05.007>
- Batty M, Taylor MJ. Early processing of the six basic facial emotional expressions. *Brain Res Cogn Brain Res* 2003;**17**:613–20. [https://doi.org/10.1016/S0926-6410\(03\)00174-5](https://doi.org/10.1016/S0926-6410(03)00174-5)
- Baumeister RF. A self-presentational view of social phenomena. *Psychol Bull* 1982;**91**:3–26. <https://doi.org/10.1037/0033-2909.91.1.3>
- Baumeister RF. Public Self and Private Self. New York: Springer, 1986.
- Baumeister RF. Stalking the true self through the jungles of authenticity: problems, contradictions, inconsistencies, disturbing findings—and a possible way forward. *Rev Gen Psychol* 2019;**23**:143–54. <https://doi.org/10.1177/1089268019829472>
- Baumeister RF, Bratslavsky E, Finkenauer C et al. Bad is stronger than good. *Rev Gen Psychol*. 2001;**5**:323–70. <https://doi.org/10.1037/1089-2680.5.4.323>
- Bench SW, Schlegel RJ, Davis WE et al. Thinking about change in the self and others: the role of self-discovery metaphors and the true self. *Soc Cogn* 2015;**33**:169–85. <https://doi.org/10.1521/soco.2015.33.3.2>
- Cai H, Wu L, Shi Y et al. Self-enhancement among westerners and easterners: a cultural neuroscience approach. *Soc Cogn Affect Neurosci* 2016;**11**:1569–78. <https://doi.org/10.1093/scan/nsw072>
- Cohen MX. Where does EEG come from and what does it mean? *Trends Neurosci* 2017;**40**:208–18. <https://doi.org/10.1016/j.tins.2017.02.004>
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 2004;**134**:9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dores Cruz TD, Thielmann I, Columbus S et al. Gossip and reputation in everyday life. *Philos Trans R Soc Lond B Biol Sci* 2021;**376**:20200301. <https://doi.org/10.1098/rstb.2020.0301>
- Dormanen R, Sanders CS, Maffly-Kipp J et al. Assimilation undercuts authenticity: a consequence of women's masculine self-presentation in masculine contexts. *Psychol Women Q* 2020;**44**:488–502. <https://doi.org/10.1177/0361684320947648>
- Fan W, Chen J, Wang X-Y et al. Electrophysiological correlation of the degree of self-reference effect. *PLoS One* 2013;**8**:e80289. <https://doi.org/10.1371/journal.pone.0080289>
- Fan W, Zhong Y, Yang Z et al. The degree of self-reference effect of extroversion individuals. *Acta Psychol Sin* 2016;**48**:1002–12. <https://doi.org/10.3724/SPJ.1041.2016.01002>
- Faul F, Erdfelder E, Buchner A et al. Statistical power analyses using GPower 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60. <https://doi.org/10.3758/BRM.41.4.1149>
- Fenigstein A. Private and public self-consciousness. In: Leary MR, Hoyle RH (eds.), *Handbook of Individual Differences in Social Behavior*. New York: The Guilford Press, 2009, 495–511.
- Fields EC, Kuperberg GR. It's all about you: an ERP study of emotion and self-relevance in discourse. *Neuroimage* 2012;**62**:562–74. <https://doi.org/10.1016/j.neuroimage.2012.05.003>
- Fields EC, Kuperberg GR. Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Soc Cogn Affect Neurosci* 2015;**10**:1202–9. <https://doi.org/10.1093/scan/nsv004>
- Gillath O, Sesko AK, Shaver PR et al. Attachment, authenticity, and honesty: dispositional and experimentally induced security can reduce self- and other-deception. *J Pers Soc Psychol* 2010;**98**:841–55. <https://doi.org/10.1037/a0019206>
- Glazer J, Nusslock R. Outcome valence and stimulus frequency affect neural responses to rewards and punishments. *Psychophysiology* 2022;**59**:e13981. <https://doi.org/10.1111/psyp.13981>
- Gray HM, Ambady N, Lowenthal WT et al. P300 as an index of attention to self-relevant stimuli. *J Exp Soc Psychol* 2004;**40**:216–24. [https://doi.org/10.1016/S0022-1031\(03\)00092-1](https://doi.org/10.1016/S0022-1031(03)00092-1)
- Guenther CL, Sedikides C. Authenticity as self-enhancement. In: Vess M, Schlegel B, Hicks J (eds.), *The Oxford Handbook of Authenticity*. Oxford University Press, in press.
- Guenther CL, Zhang Y, Sedikides C. The authentic self is the self-enhancing self: a self-enhancement framework of authenticity. *Pers Soc Psychol Bull* 2024;**50**:1182–96. <https://doi.org/10.1177/01461672231160653>
- Hajcak G, Weinberg A, MacNamara A, Foti D. ERPs and the study of emotion. In: *The Oxford Handbook of Event-Related Potential Components*, Vol. 441. Oxford: Oxford University Press, 2012, 441–472.
- Hancock JT, Toma CL. Putting your best face forward: the accuracy of online dating photographs. *J Commun* 2009;**59**:367–86. <https://doi.org/10.1111/j.1460-2466.2009.01420.x>
- Hauk O, Pulvermüller F. Effects of word length and frequency on the human event-related potential. *Clin Neurophysiol* 2004;**115**:1090–103. <https://doi.org/10.1016/j.clinph.2003.12.020>
- Herbert C, Herbert BM, Ethofer T et al. His or mine? The time course of self-other discrimination in emotion processing. *Soc Neurosci* 2011;**6**:277–88. <https://doi.org/10.1080/17470919.2010.523543>

- Hicks JA, Schlegel RJ, Newman GE. Introduction to the special issue: authenticity: novel insights into a valued, yet elusive, concept. *Rev Gen Psychol* 2019;**23**:3–7. <https://doi.org/10.1177/1089268019829474>
- Hillyard SA, Vogel EK, Luck SJ. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos Trans R Soc Lond B Biol Sci* 1998;**353**:1257–70. <https://doi.org/10.1098/rstb.1998.0281>
- Huang C, Sedikides C, Angus DJ et al. Demystifying authenticity: behavioral and neurophysiological signatures of self-positivity for authentic and presented selves. *Neuroimage* 2025;**307**:121046. <https://doi.org/10.1016/j.neuroimage.2025.121046>
- Imbir KK, Duda-Golawska J, Pastwa M et al. Event-related potential correlates of valence, arousal, and subjective significance in processing of an emotional Stroop task. *Front Hum Neurosci* 2021;**15**:617861. <https://doi.org/10.3389/fnhum.2021.617861>
- James W. *The Principles of Psychology*. New York: Henry Holt, 1890.
- Jetha MK, Segalowitz SJ, Gatzke-Kopp LM. The reliability of visual ERP components in children across the first year of school. *Dev Psychobiol* 2021;**63**:e22150. <https://doi.org/10.1002/dev.22150>
- Jongman-Sereno KP, Leary MR. Self-perceived authenticity is contaminated by the valence of one's behavior. *Self Identity* 2016;**15**:283–301. <https://doi.org/10.1080/15298868.2015.1128964>
- Kernis MH, Goldman BM. A multicomponent conceptualization of authenticity: theory and research. *Adv Exp Soc Psychol* 2006;**38**:283–357. [https://doi.org/10.1016/S0065-2601\(06\)38006-9](https://doi.org/10.1016/S0065-2601(06)38006-9)
- Keyes H, Brady N, Reilly RB et al. My face or yours? Event-related potential correlates of self-face processing. *Brain Cogn* 2010;**72**:244–54. <https://doi.org/10.1016/j.bandc.2009.09.006>
- Kihlstrom JF, Cantor N, Albright JS, et al. (1988). Information processing and the study of the self. *Adv Exp Soc Psychol* **21**, 145–178. [https://doi.org/10.1016/S0065-2601\(08\)60226-9](https://doi.org/10.1016/S0065-2601(08)60226-9)
- Kotz SA, Paulmann S. Emotion, language, and the brain. *Lang Linguist Compass* 2011;**5**:108–25. <https://doi.org/10.1111/j.1749-818X.2010.00267.x>
- Lakey CE, Kernis MH, Heppner WL et al. Individual differences in authenticity and mindfulness as predictors of verbal defensiveness. *J Res Pers* 2008;**42**:230–8. <https://doi.org/10.1016/j.jrp.2007.05.002>
- Leary MR. *Self-Presentation: Impression Management and Interpersonal Behavior*. Boulder: Westview Press, 1995.
- Leary MR, Kowalski RM. Impression management: a literature review and two-component model. *Psychol Bull* 1990;**107**:34–47. <https://doi.org/10.1037/0033-2909.107.1.34>
- Lee SJ, Quigley BM, Nesler MS et al. Development of a self-presentation tactics scale. *Pers Individ Dif* 1999;**26**:701–22. [https://doi.org/10.1016/S0191-8869\(98\)00178-0](https://doi.org/10.1016/S0191-8869(98)00178-0)
- Levorsen M, Aoki R, Matsumoto K et al. The self-concept is represented in the media prefrontal cortex in terms of self-importance. *J Neurosci* 2023;**43**:3675–86. <https://doi.org/10.1523/JNEUROSCI.2178-22.2023>
- Liu L, Li W, Li J et al. Temporal features of psychological and physical self-representation: an ERP study. *Front Psychol* 2019;**10**:785. <https://doi.org/10.3389/fpsyg.2019.00785>
- Luck SJ. *An Introduction to the Event-Related Potential Technique*, 2nd edn. Cambridge: MIT Press, 2014.
- Luck SJ, Gaspelin N. How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology* 2017;**54**:146–57. <https://doi.org/10.1111/psyp.12639>
- Luck SJ, Hillyard SA. Electrophysiological correlates of feature analysis during visual search. *Psychophysiology* 1994;**31**:291–308. <https://doi.org/10.1111/j.1469-8986.1994.tb02218.x>
- Luo W, Feng W, He W et al. Three stages of facial expression processing: ERP study with rapid serial visual presentation. *Neuroimage* 2010;**49**:1857–67. <https://doi.org/10.1016/j.neuroimage.2009.09.018>
- Marsh HW. The structure of academic self-concept: the Marsh/Shavelson model. *J Educ Psychol* 1990;**82**:623–36. <https://doi.org/10.1037/0022-0663.82.4.623>
- Mathews A, MacLeod C. Selective processing of threat cues in anxiety states. *Behav Res Ther* 1985;**23**:563–9. [https://doi.org/10.1016/0005-7967\(85\)90104-4](https://doi.org/10.1016/0005-7967(85)90104-4)
- McConnell AR. The multiple self-aspects framework: self-concept representation and its implications. *Pers Soc Psychol Rev* 2011;**15**:3–27. <https://doi.org/10.1177/1088868310371101>
- Miyakoshi M, Nomura M, Ohira H. An ERP study on self-relevant object recognition. *Brain Cogn* 2007;**63**:182–9. <https://doi.org/10.1016/j.bandc.2006.12.001>
- Montalan B, Caharel S, Personnaz B et al. Sensitivity of N170 and late positive components to social categorization and emotional valence. *Brain Res* 2008;**1233**:120–8. <https://doi.org/10.1016/j.brainres.2008.07.055>
- Morís Fernández L, Vadillo MA. Flexibility in reaction time analysis: many roads to a false positive? *R Soc Open Sci* 2020;**7**:190831. <https://doi.org/10.1098/rsos.190831>
- Mueller AK, Falter CM, Tucha O. Distinct attentional mechanisms for social information processing as revealed by event-related potential studies. *Int J Brain Cogn Sci* 2013;**2**:23–37. <https://doi.org/10.5923/j.ijbcs.20130202.03>
- Nie L, Zhou S, Wang Z et al. Impaired perceptual processing of facial expression categorization in social anhedonia. *Psychophysiology* 2020;**57**:e13682. <https://doi.org/10.1111/psyp.13682>
- Northoff G, Bermpohl F. Cortical midline structures and the self. *Trends Cogn Sci* 2004;**8**:102–7. <https://doi.org/10.1016/j.tics.2004.01.004>
- Northoff G, Heinzel A, De Greck M et al. Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage* 2006;**31**:440–57. <https://doi.org/10.1016/j.neuroimage.2005.12.002>
- Öhman A. Fear and anxiety as emotional phenomena: clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In: Barrett LF, Lewis M, Haviland-Jones JM (eds.), *Handbook of Emotions*. New York: Guilford Press, 1993, 511–36.
- Öhman A, Flykt A, Esteves F. Emotion drives attention: detecting the snake in the grass. *J Exp Psychol Gen* 2001;**130**:466–78. <https://doi.org/10.1037/AXJ96-3445.130.3.466>
- Peirce JW. PsychoPy—psychophysics software in python. *J Neurosci Methods* 2007;**162**:8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Perrin F, Pernier J, Bertrand O et al. Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* 1989;**72**:184–7. [https://doi.org/10.1016/0013-4694\(89\)90180-6](https://doi.org/10.1016/0013-4694(89)90180-6)
- Phaf RH, Kan K-J. The automaticity of emotional Stroop: a meta-analysis. *J Behav Ther Exp Psychiatry* 2007;**38**:184–99. <https://doi.org/10.1016/j.jbtep.2006.10.008>
- Pion-Tonachini L, Kreutz-Delgado K, Makeig S. ICLLabel: an automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* 2019;**198**:181–97. <https://doi.org/10.1016/j.neuroimage.2019.05.026>
- Polich J. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 2007;**118**:2128–48. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Porter NA, Fields EC, Moore IL et al. Late frontal positivity effects in self-referential memory: unique to the self? *Soc Neurosci* 2021;**16**:406–22. <https://doi.org/10.1080/17470919.2021.1929460>
- Potts GF. An ERP index of task relevance evaluation of visual stimuli. *Brain Cogn* 2004;**56**:5–13. <https://doi.org/10.1016/j.bandc.2004.03.006>
- Potts GF, Martin LE, Burton P et al. When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *J Cogn Neurosci* 2006;**18**:1112–9. <https://doi.org/10.1162/jocn.2006.18.7.1112>

- Pourtois G, Dan ES, Grandjean D et al. Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. *Hum Brain Mapp* 2005;**26**:65–79. <https://doi.org/10.1002/hbm.20130>
- Pratto F, John OP. Automatic vigilance: the attention-grabbing power of negative social information. *J Pers Soc Psychol* 1991;**61**:380–91.
- Prete G, Capotosto P, Zappasodi F et al. The cerebral correlates of subliminal emotions: an electroencephalographic study with emotional hybrid faces. *Eur J Neurosci* 2015;**42**:2952–62. <https://doi.org/10.1111/ejn.13078>
- Prete G, Capotosto P, Zappasodi F et al. Contrasting hemispheric asymmetries for emotional processing from event-related potentials and behavioral responses. *Neuropsychology* 2018;**32**:317–28. <http://dx.doi.org/10.1037/neu0000443>
- Pritchard WS. Psychophysiology of P300. *Psychol Bull* 1981;**89**:506–40. <https://doi.org/10.1037/0033-2909.89.3.506>
- Ratcliff R. Methods for dealing with reaction time outliers. *Psychol Bull* 1993;**114**:510–32. <https://doi.org/10.1037/0033-2909.114.3.510>
- Riggins T, Scott LS. P300 development from infancy to adolescence. *Psychophysiology* 2020;**57**:Article13346. <https://doi.org/10.1111/psyp.13346>
- Rogers C. Client-Centered Therapy: Its Current Practice, Implications, and Theory. Boston: Houghton Mifflin Company, 1951.
- Rossion B, Jacques C. The N170: understanding the time course of face perception in the human brain. In: Luck SJ, Kappenman ES (eds.), *The Oxford Handbook of Event-Related Potential Components*. Oxford: Oxford University Press, 2012, 115–41.
- Roth DL, Snyder CR, Pace LM. Dimensions of favorable self-presentation. *J Pers Soc Psychol* 1986;**51**:867–74. <https://doi.org/10.1037/0022-3514.51.4.867>
- San Martín R, Kwak Y, Pearson JM et al. Altruistic traits are predicted by neural responses to monetary outcomes for self vs charity. *Soc Cogn Affect Neurosci* 2016;**11**:863–76. <https://doi.org/10.1093/scan/nsw026>
- Sande GN. The multifaceted self. In: Olson JM, Zanna MP (eds.), *Self-Inference Processes: The Ontario Symposium*, Vol. 6. Mahwah: Lawrence Erlbaum, 1990, 1–16.
- Schäfer S, Wentura D, Frings C. Creating a network of importance: the particular effects of self-relevance on stimulus processing. *Atten Percept Psychophys* 2020;**82**:3750–66. <https://doi.org/10.3758/s13414-020-02070-7>
- Schindler S, Bruchmann M, Straube T. Beyond facial expressions: a systematic review on effects of emotional relevance of faces on the N170. *Neurosci Biobehav Rev* 2023;**153**:105399. <https://doi.org/10.1016/j.neubiorev.2023.105399>
- Schindler S, Bublatzky F. Attention and emotion: an integrative review of emotional face processing as a function of attention. *Cortex* 2020;**130**:362–86. <https://doi.org/10.1016/j.cortex.2020.06.010>
- Schlenker BR, Pontari BA. The strategic control of information: impression management and self-presentation in daily life. In: Tesser A, Felson RB, Suls JM (eds.), *Psychological Perspectives on Self and Identity*. Washington: American Psychological Association, 2000, 199–232. <https://doi.org/10.1037/10357-008>
- Sedikides C. Self-protection. In: Leary MR, Tangney JP (eds.), *Handbook of Self and Identity*, 2nd edn. New York: Guilford Press, 2012, 327–53.
- Sedikides C, Green JD, Saunders J et al. Mnemonic neglect: selective amnesia of one's faults. *Eur Res Soc Psychol* 2016;**27**:1–62. <https://doi.org/10.1080/10463283.2016.1183913>
- Sedikides C, Gregg AP. Portraits of the self. In: Hogg MA, Cooper J (eds.), *Sage Handbook of Social Psychology*. Thousand Oaks: Sage Publications, 2003, 110–138.
- Sedikides C, Gregg AP. Self-enhancement: food for thought. *Perspect Psychol Sci* 2008;**3**:102–16. <https://doi.org/10.1111/j.1745-6916.2008.00068.x>
- Sedikides C, Hoorens V, Dufner M. Self-enhancing self-presentation: interpersonal, relational, and organizational implications. In: Guay F, McInerney DM, Craven R, Marsh HW (eds.), *Self-Concept, Motivation and Identity: Underpinning Success with Research and Practice. International Advances in Self Research*, Vol. 5. Charlotte: Information Age Publishing, 2015, 29–55.
- Sedikides C, Lenton AP, Slabu L et al. Sketching the contours of state authenticity. *Rev. Gen. Psychol.* 2019;**23**:73–88. <https://doi.org/10.1037/gpr0000156>
- Sedikides C, Schlegel R. Distilling the concept of authenticity. *Nat Rev Psychol* 2024;**3**:509–523. <https://doi.org/10.1038/s44159-024-00323-y>
- Sereno SC, Brewer CC, O'Donnell PJ. Context effects in word recognition: evidence for early interactive processing. *Psychol Sci* 2003;**14**:328–33. <https://doi.org/10.1111/1467-9280.14471>
- Sheldon KM, Ryan RM, Rawsthorne LJ et al. Trait self and true self: cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *J Pers Soc Psychol* 1997;**73**:1380–93. <https://doi.org/10.1037/0022-3514.73.6.1380>
- Song Y, Ding Y, Fan S et al. Neural substrates of visual perceptual learning of simple and complex stimuli. *Clin Neurophysiol* 2005;**116**:632–9. <https://doi.org/10.1016/j.clinph.2004.09.019>
- Strohming N, Knobe J, Newman G. The true self: a psychological concept distinct from the self. *Perspect Psychol Sci* 2017;**12**:551–60. <https://doi.org/10.1177/1745691616689495>
- Tacikowski P, Nowicka A. Allocation of attention to self-name and self-face: an ERP study. *Biol Psychol* 2010;**84**:318–24. <https://doi.org/10.1016/j.biopsycho.2010.03.009>
- Taylor SE. Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychol Bull* 1991;**110**:67–85. <https://doi.org/10.1037/0033-2909.110.1.67>
- Tredennick H, Thomson JAK. *Aristotle Ethics*. London: Penguin Classics, 1976.
- Tunnel G. The discrepancy between private and public selves: public self-consciousness and its correlates. *J Pers Assess* 1984;**48**:549–55. https://doi.org/10.1207/s15327752jpa4805_15
- Uddin LQ, Iacoboni M, Lange C et al. The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn Sci* 2007;**11**:153–7. <https://doi.org/10.1016/j.tics.2007.01.001>
- Van Hooff JC, Dietz K, Sharma D et al. Neural correlates of intrusion of emotion words in a modified Stroop task. *Int J Psychophysiol* 2008;**67**:23–34. <https://doi.org/10.1016/j.ijpsycho.2007.09.002>
- Vazire S, Wilson TD (eds.). *Handbook of Self-Knowledge*. New York: Guilford Press, 2012.
- Vonasch AJ, Reynolds T, Winegard BM et al. Death before dishonor: incurring costs to protect moral reputation. *Soc Psychol Pers Sci* 2018;**9**:604–13. <https://doi.org/10.1177/1948550617720271>
- Vuilleumier P, Pourtois G. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia* 2007;**45**:174–94. <https://doi.org/10.1016/j.neuropsychologia.2006.06.003>
- Wada M, Kurose S, Miyazaki T et al. The P300 event-related potential in bipolar disorder: a systematic review and meta-analysis. *J Affect Disord* 2019;**256**:234–49. <https://doi.org/10.1016/j.jad.2019.06.010>
- Watts FN, McKenna FP, Sharrock R et al. Colour naming of phobia-related words. *Br J Psychol* 1986;**77** (Pt 1):97–108. <https://doi.org/10.1111/j.2044-8295.1986.tb01985.x>
- Wieser MJ, Moscovitch DA. The effect of affective context on visuo-cortical processing of neutral faces in social anxiety. *Front Psychol* 2015;**6**:1824. <https://doi.org/10.3389/fpsyg.2015.01824>
- Williams JMG, Mathews A, MacLeod C. The emotional Stroop task and psychopathology. *Psychol Bull* 1996;**120**:3–24. <https://doi.org/10.1037/0033-2909.120.1.3>

- Williams LM, Palmer D, Liddell BJ et al. The 'when' and 'where' of perceiving signals of threat versus non-threat. *Neuroimage* 2006;**31**:458–67. <https://doi.org/10.1016/j.neuroimage.2005.12.009>
- Wood AM, Linley PA, Maltby J et al. The authentic personality: a theoretical and empirical conceptualization and the development of the authenticity scale. *J Couns Psychol* 2008;**55**:385–99. <https://doi.org/10.1037/0022-0167.55.3.385>
- Yang J, Qi M, Guan L. Self-esteem modulates the latency of P2 component in implicit self-relevant processing. *Biol Psychol* 2014;**97**:22–6. <http://dx.doi.org/10.1016/j.biopsycho.2014.01.004>
- Zald DH. The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev* 2003;**41**:88–123. [https://doi.org/10.1016/S0165-0173\(02\)00248-5](https://doi.org/10.1016/S0165-0173(02)00248-5)
- Zhang D, He W, Wang T et al. Three stages of emotional word processing: an ERP study with rapid serial visual presentation. *Soc Cogn Affect Neurosci* 2014;**9**:1897–903. <https://doi.org/10.1093/scan/nst188>
- Zhu M, Luo J, Zhao N et al. The temporal primacy of self-related stimuli and negative stimuli: an ERP-based comparative study. *Soc Neurosci* 2016;**11**:507–14. <https://www.tandfonline.com/doi/full/10.1080/17470919.2015.1114021>