University of Southampton Research Repository

University of Southampton

Faculty of Environmental and Life Sciences

School of Psychology

Authenticity and Self-Control: A Social Neuroscience Approach

by

Chengli Huang

ORCID ID 0000-0001-5215-6734

Thesis for the degree of Doctor of Philosophy

June 2025

University of Southampton

Abstract

Faculty of Environmental and Life Sciences
School of Psychology
Doctor of Philosophy
Authenticity and Self-Control: A Social Neuroscience Approach
by
Chengli Huang

This dissertation explores the essence of authenticity—the true, genuine self of an individual—from three distinct perspectives: self-accuracy, self-consistency, and self-enhancement. Additionally, it investigates the role of authenticity in the relationship between self-control and reward processing. In Chapter 1, I review the existing literature on the self and authenticity, systematically tracing their conceptual evolution across different historical periods and theoretical frameworks. I examine the three perspectives of authenticity along with their respective empirical evidence from psychological research, and contrast authenticity with another prominent form of self-representation—the presented self. I also discuss behavioural and neuroscientific methods used to study the self and authenticity. Furthermore, I investigate the role of trait authenticity in shaping self-control exertion over reward processing.

The empirical chapters present a series of studies designed to contribute to the current literature. In Chapter 2, I test the self-enhancement view of authenticity rigorously by comparing the authentic self with the presented self using the SR-valence task. Both behavioural and neuroscientific findings reveal that authenticity, albeit predominantly positive, allows room for negativity, providing support for the self-accuracy and self-consistency perspectives. In Chapter 3, I further examine the self-enhancement, self-accuracy, and self-consistency views by investigating the interference of negative information on self-evaluation through the self-referent emotional Stroop task. Both behavioural and neuroscience findings demonstrate that the presented self is more inclined toward positivity, whereas the authentic self exhibits greater tolerance for negativity. In Chapter 4, I first examine the neural basis of the reward responsivity hypothesis of self-control by a modified monetary incentive delay task (Part A), and then re-examine whether the reward responsivity following self-control exertion is potentially influenced by trait authenticity (Part B). The findings provide neurophysiological evidence supporting the reward responsivity hypothesis of self-control, although the enhancement of reward responsivity appears to be independent of trait authenticity.

Finally, in Chapter 5, I synthesize the key findings of the studies, discuss their theoretical and empirical implications, highlight the strengths and limitations of the research, and propose directions for future investigations.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Chengli Huang

Title of thesis: Authenticity and Self-Control: A Social Neuroscience Approach

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
   Chapter 2- Huang, C., Sedikides, C., Angus, D. J., Davis, W. E., Butterworth, J. W., Jeffers, A., ... & Kelley, N. J. (2025). Demystifying Authenticity: Behavioral and Neurophysiological Signatures of Self-Positivity for Authentic and Presented Selves. *NeuroImage*, 307, Article 121046. https://doi.org/10.1016/j.neuroimage.2025.121046
   Chapter 4-Part A- Huang, C., Zhou, Z., Angus, D. J., Sedikides, C., & Kelley, N. J. (2025). Exercising Self-Control Increases Responsivity to Hedonic and Eudaimonic Rewards. *Social Cognitive and Affective Neuroscience, 20*(1), Article nsaf016. https://doi.org/10.1093/scan/nsaf016

Signature: .... Date:........................

Acknowledgements

First and foremost, I would like to extend my sincerest gratitude to my exceptional supervisors, Nicholas Kelley and Constantine Sedikides, for their unwavering guidance, encouragement, and invaluable feedback throughout this entire process. Nick has been more than a supervisor to me; he is a mentor, a friend, and a constant source of inspiration. His approachable leadership style created an environment where I felt empowered to voice my ideas and explore my potential. His childlike enthusiasm and optimism breathed life into my research life, making every challenge feel like an exciting opportunity. Nick’s consistent support and belief in my abilities gave me the confidence to navigate obstacles with resilience and grace.

Constantine, on the other hand, is undoubtedly the most experienced and insightful professor I have ever had the privilege of working with. With remarkable patience, he guided me step-by-step in presenting scientifically rigorous and clear research findings when we were faced with the towering mountain of "Authenticity." His profound expertise and sharp intellect were instrumental in helping me overcome moments of confusion and doubt. Beyond his academic brilliance, Constantine’s youthful spirit and infectious laughter brought joy and lightness to the often-rigorous world of scientific research. I consider myself incredibly fortunate to have crossed paths with both Nick and Constantine at this pivotal stage of my life. The past three and a half years have been immensely fulfilling and enjoyable under their mentorship, and I will forever cherish the lessons they imparted and the memories we created together. As time flies, I know I will carry their teachings and the joy they brought me into my future career, striving to pass on their wisdom and guidance to my own students one day.

I am deeply indebted to my parents and my beloved sister, whose unconditional love and unwavering support have been my greatest strength throughout this journey. Their encouragement has been a constant source of motivation, especially during moments of doubt and homesickness. I am also profoundly grateful to my dear domestic friends—Chunrong, Jing, Cui, and Zirui—who have stood by me through thick and thin. Their companionship and encouragement have been a beacon of light, giving me the strength to thrive in a foreign land.

My heartfelt thanks also go to my wonderful friends—Wenjin, Qidi, Zhiwei, Hongyi, and Jin—as well as my amazing roommates—Yan, Chun, You, and Meng. Studying far from home can often feel lonely, but their presence turned my life in an unfamiliar place into a vibrant and fulfilling experience. From cooking together and hiking to playing badminton, every shared moment added color and richness to my life, making my time abroad truly unforgettable.

I would also like to express my gratitude to Marie Levorsen, Irem Ozdemir, Monica Sood, Paul Conway, Pascal Burgmer, Tim Wildschut, and every member of CSRI, as well as technical advisors, Jin Zhang and Antony Wood, for their invaluable support and contributions to my PhD journey. Their expertise, feedback, and encouragement have played a significant role in shaping my research and academic growth.

Lastly, I extend my deepest appreciation to the Economic and Social Research Council (ESRC) for providing the financial support that made my higher education in the UK possible. Their generosity enabled me to embark on this unforgettable journey at the University of Southampton, an experience that has profoundly shaped both my academic and personal growth.

Thank you all!

Definitions and Abbreviations

ANOVA	Analysis of Variance
BBC	British Broadcasting Corporation
BCE	Before the Common Era
EEG	electroencephalography
ERP	event-related potentials
ESEM	exploratory structural equation modelling
fERN	feedback error-related negativity
fMRI	functional magnetic resonance imaging
FN	feedback negativity
FRN	feedback-related negativity
HEOG	horizontal electrooculogram
HLM	hierarchical linear model
Hz	Hertz, the unit of frequency in the International System of Units.
ICA	independent component analysis
kΩ	kiloohm, a unit of electrical resistance in the International System of Units.
LPP	Late Positive Potential, an ERP characterized by a sustained positive deflection in the EEG signal. It typically emerges approximately 400–500 ms after stimulus onset and persists for several hundred milliseconds, with maximal amplitude observed over the midline centroparietal region of the scalp.
M	mean
MFN	medial frontal negativity
MID	monetary incentive delay task
MLM	multilevel model
ms	milliseconds
N170	An ERP that represents a negative deflection in the EEG signal, primarily observed over the lateral occipito-temporal regions of the

	scalp. It is a face-sensitive component that typically peaks at approximately 170 ms after stimulus onset.
P1	An ERP that represents the first positive deflection in the EEG signal, typically observed in the parieto-occipital region of the scalp. It is a visually evoked potential that emerges approximately 60–90 ms after the presentation of a visual stimulus, reaching its peak amplitude between 100 and 130 ms post-stimulus.
P2	An ERP characterized by a positive deflection in the EEG signal, typically observed over the frontier-central-parietal region of the scalp. It occurs within a time window of 150 to 250 ms after stimulus onset.
P3	An ERP characterized by a prominent positive wave in the EEG signal. It typically reaches its maximal amplitude around 300 ms after stimulus onset and is most prominently observed over the midline frontier-central-parietal region of the scalp.
RewP	Reward Positivity, an ERP that serves as a neural marker of reward responsivity. It typically peaks approximately 200 to 300 ms following the onset of performance-related feedback, with the most pronounced activity observed over fronto-central sites on the scalp.
RT	reaction time
SAS	Southampton Authenticity Scale
SCT	self-categorization theory
SD	standard deviation
SDT	self-determination theory
SEM	standard error of the mean
SIT	social identity theory
SR-valence	self-reference valence task
SSI	spherical spline identified interpolation
TSAG	true-self-as-guide framework
TV	television
μV	microvolt, a unit of electrical potential difference or electromotive force in the International System of Units.

VEOG	vertical electrooculogram
WEIRD	Western, Educated, Industrialized, Rich, and Democratic samples

Chapter 1 Literature Review

1.1 The Self

1.1.1 An Introduction to the Psychological Self

As an enigmatic and complex construct, the self has long been a focal point of inquiry across a wide range of disciplines. Indeed, the self has captivated philosophers, scientists, religious leaders, political figures, writers, and poets alike (Sedikides et al., 2006). The history of exploration of the self can be at least traced back to the ancient Greek times when the oracle on the Temple of Apollo directed people to “Know thyself.” This call to self-understanding has echoed through millennia, influencing not only philosophical and existential thought, but also scientific inquiry.

In psychology, the self has come under increasing scrutiny over the past century (Gallagher, 2000; Northoff et al., 2006), with various schools of thought offering their own views from psychoanalysts (e.g., Sigmund Freud) and humanists (e.g., Abraham Maslow, Carl Rogers) to developmentalists (e.g., Erik Erikson) and, more recently, cognitive and social psychologists, as well as cognitive neuroscientists.

Freud (1923) provided one of the earliest psychological models of the self. He developed a model of psychic structure, comprising three distinct yet interacting components: *id* (instinctual desires), *ego* (the rational, decision-making part), and *superego* (moral conscience). Freud’s theory posits that the self is shaped by the interactions among these parts and acts to maintain a homeostasis between biological drives and societal expectations. This model frames the self as a battleground between unconscious desires and conscious regulation, emphasizing the role of internal conflict within the tripartite structure in shaping personality and behaviour.

Rogers (1961), a leading figure in humanistic psychology, also pondered the idea of the self-concept, which he regarded as the collection of beliefs and perceptions one has about their personality. He distinguished between the *real self* (who one truly is) and the *ideal self* (who one aspires to be), arguing that psychological well-being is achieved when there is congruence between these two self-aspects. However, when a sizeable gap exists between the real and ideal selves, individuals may experience anxiety, inadequacy, and distress. Although Freud (1923) viewed the self as largely shaped by unconscious processes, Rogers (1961) offered a more positive and growth-oriented perspective, underscoring the importance of self-acceptance and personal improvement in attaining psychological well-being.

Erikson (1950) introduced the concept of *ego identity*, referring to an individual's dynamic and conscious sense of self that emerges through social interaction and experiences. Erikson regarded ego identity as central to psychological well-being, particularly during critical developmental periods such as adolescence and emerging adulthood. During these periods, individuals face the developmental crisis of achieving a stable and coherent sense of identity, a process that involves reconciling personal values with societal roles and expectations. He argued that successfully navigating this crisis leads to a stronger, more integrated sense of self, whereas failure can result in identity confusion or role diffusion. Erikson's theory positions the self not only as a personal construct, but also as one embedded in social contexts and relationships.

Cognitive psychologists have extensively examined how the self shapes attention and memory. An example is the cocktail party effect (Cherry, 1953; Moray, 1959), defined as the ability to focus one's auditory attention on a specific sound source and filter out other sounds amidst a cacophony of conversations and background noise, especially if that sound source is related to self, such as one's name (Newman, 2005; Wood & Cowan, 1995). Self-relevant information is of high priority and automatically elicits attention, indicating the privileged position of the self in the allocation of attention. This attentional prioritization extends beyond auditory domains. In visual attention, self-relevant stimuli such as one's own face (Bola et al., 2021; Keyes & Brady, 2010; Y. Ma & Han, 2010; Tacikowski & Nowicka, 2010) or even geometric shapes that match the self (Macrae et al., 2017; Sui et al., 2012, 2013) also garner enhanced processing.

Importantly, this self-prioritization is not confined to low-level perceptual mechanisms; it is also evident in higher-order cognitive functions. For instance, a great deal of evidence indicates better memory performance (e.g., recall, recognition) for stimuli in relation to the self versus another person (Sui & Humphreys, 2015; Symons et al., 1997). This self-referential memory advantage emerges early in life (Cunningham et al., 2014; Sui & Zhu, 2005) and remains stable across the lifespan, persisting in both younger and older adults (Glisky & Marquine, 2009; Gutchess et al., 2007, 2010; Leshikar et al., 2015). Beyond perceptual and mnemonic domains, self-referential stimuli can also facilitate self-control processes. For example, self-relevant stimuli are theorized to enhance self-control in identity relevant domains by amplifying the value of behaviours aligned with personal goals (Berkman et al., 2017). Consistent with this view, studies using the classical Stroop task have found that self-referential information diminishes the congruency effect relative to control conditions, providing evidence the influence of self on cognitive control (Dignath et al., 2022; Z. Li et al., 2024). Collectively, the culmination of various research streams suggests that cognitive processing—whether at higher or lower levels—is enhanced when the self is salient.

Social psychologists have also highlighted the dynamic interplay between the self and the external environment. According to social identity theory (SIT; Tajfel, 1978; Tajfel & Turner, 1979) and self-categorization theory (SCT; Turner, 1999; Turner & Reynolds, 2011), the self comprises both personal identity and social identity. Specifically, each individual is defined by a dual set of attributes: social characteristics derived from their membership in social groups (e.g., gender, ethnicity, occupation) and personal characteristics that are unique and idiosyncratic to them. Moreover, individuals evaluate their social identity by comparing the relative status of their ingroup (i.e., the group they identify with) to outgroups (i.e., groups they do not identify with), maintaining or enhancing positive self-evaluation through intergroup competition (Abrams & Hogg, 1988; Deschamps & Devos, 1998). These conceptual frameworks can be extended to broader societal and cultural contexts, such as the independent and interdependent self-constructions that emerge within individualistic and collectivistic cultures (Markus & Kitayama, 1991, 2010). Together, these theories emphasize the self as a construct defined by the continuous interaction between internal characteristics and external social environments.

More recently, cognitive neuroscientists have expanded understanding of the self by identifying key brain structures and mental processes that contribute to self-related behaviours, such as self-awareness and self-referential processing. Neuroimaging research has consistently shown activation in the brain’s cortical midline structures during self-related stimuli across multiple functional domains, including perception, emotion, memory, and motor activities (Northoff et al., 2006; Northoff & Bermpohl, 2004). These cortical midline structures encompass the medial prefrontal cortex that plays a critical role in evaluating the unique value or significance of self-related information (D’Argembeau, 2013; Kim & Johnson, 2015) and self-referential mental activities and emotional processing (Gusnard et al., 2001; Lieberman et al., 2019), posterior cingulate cortex that is central to self-referential cognition and autobiographical memory (Foster et al., 2023), and anterior cingulate cortex that is essential for self-control (Allman et al., 2001; Heilbronner & Hayden, 2016). These regions are part of the default mode network that is active when individuals engage in self-generated thought or consider important aspects of their self-concept (Andrews-Hanna et al., 2014). The findings offer a neural foundation for understanding self-related processes.

Building upon the relevant literature, Sui and Humphreys (2015) proposed a framework that conceptualizes the self as an integrative hub, facilitating the binding of diverse types of information and various stages of cognitive psychological processing. Within this framework, the self serves multiple functions: it facilitates perceptual integration, as evidenced by faster classification of self-faces compared to others’ faces; it enhances memory processes, with individuals demonstrating superior recollection for items judged in relation to themselves versus others; it integrates distinct stages of cognitive processing, such as heightened self-focused attention leading to increased certainty in decision-making; and it promotes neural interaction across different brain regions—for instance, self-related processing strengthens neural coupling between core self-representation areas (e.g., the ventromedial prefrontal cortex and domain-specific regions, such as the left posterior superior temporal sulcus that is involved in self-related attention. This interdisciplinary synthesis underscores the representation of the self as a dynamic construct shaped by cognitive, emotional, and neurobiological factors.

1.1.2 The Multifaceted Self

Like many key constructs in psychology, such as “consciousness,” “attention,” “perception,” and “memory,” the constructs of “self” is difficult to define. Perhaps it is this conceptual ambiguity that has made scholars’ explorations of self so diverse. Indeed, empirical progress resulted from the realization that the self cannot be approached as a unitary, monolithic entity (Sedikides & Spencer, 2007). Instead, it should be approached as a complex, multi-faceted construct.

William James (1890) distinguished among the material self (i.e., things that belong to a person, such as body, clothes, family), the social self (i.e., recognition that an individual received from others), and the spiritual self (i.e., one’s inner or subjective being, such as dispositions, core values, and conscience). Other theorists or researchers followed suit. As mentioned above, Freud (1923) distinguished among *id*, *ego*, and *superego*, and Roger (1961) between *ideal self* and *real self*. Higgins (1987) differentiated among *ideal*, *ought*, and *actual* selves: the *ideal* self contains attributes that someone (self or other) would like the person to possess, the *ought* self contains attributes that someone (self or other) believes the person should or ought to possess, and the *actual* self contains attributes that the person possesses. Other distinctions include those between the physical self and psychological self (Gillihan & Farah, 2005; Uddin, 2011), public and private self (Baumeister, 2012; Fenigstein et al., 1975), individual, relational, and collective self (Sedikides et al., 2011; Sedikides & Brewer, 2015), as well as past, present and future self (Peetz & Wilson, 2008; Sedikides et al., 2023). See Table 1.1.

Although the self can be partitioned and conceptualized in a number of ways, researchers generally agree that people have both an external self and an internal self that are distinct from one another (Sedikides & Gregg, 2003; Markus & Nurius, 1986). Even folk psychological theories highlight the potential discrepancies between external and internal selves (Johnson et al., 2004; Ratcliffe, 2006). Relying on this intellectual tradition, I split the self into an internal (authentic self) and external (presented self) component.

Table 1.1 *Mapping of Self-Concepts: Internal self (Authentic Self) vs. External self (Presented Self)*

	Internal self (Authentic self)	External self (Presented self)
James (1890)	spiritual self	material self, social self
Freud (1923)	id	ego, superego
Higgins (1987)	ideal self	ought self, actual self
Gillihan & Farah (2005); Uddin (2011)	psychological self	physical self
Baumeister (2012); Fenigstein et al. (1975)	private self	public
Sedikides et al. (2011); Sedikides & Brewer (2015)	individual self	relational self, collective self

1.2 The Nature of the Authentic Self

1.2.1 Historical and Modern Interpretations of the Authentic Self

The concept of authenticity is a blend of philosophical, psychological, and spiritual ideas, reflecting its rich history (Steiner & Reisinger, 2006).

The Ethics of Authenticity

Exploration of authenticity has long captivated human thought. In ancient Greece, Socrates famously declared, “the unexamined life is not worth living” (Plato, 1961, Apology, 38a), positing that continuous reflection and self-examination are essential for discovering one’s authentic self, ultimately leading to the highest good (Plato, 1961). In “The Republic,” Plato (2007) presents the concept of the tripartite soul, central to his understanding of the authentic self, in which the authentic self is identifies with the rational part of the soul—the logical, reasoning aspect that seeks truth and knowledge. Aristotle (2004) also asserted that the authentic self is intertwined with reason, and emerges through engagement in rational activities and the practice of virtue. Living in accordance with one’s authentic self, according to Aristotle (2004), facilitates the attainment of the ultimate life goal, “flourishing” or “well-being” (eudaimonia). In short, the authentic self transcends physical form and immediate desires, being actualized through reason, philosophical contemplation, and moral pursuit. By living in accordance with their true self, individuals can achieve internal harmony.

The Existentialism of Authenticity

Another high point in the exploration for the authentic self stems from the existential movement of the 19th century. As a precursor to this movement, Kierkegaard (1983) stated, ‘‘truth is subjectivity,’’ which asserted that truth and authenticity are inherently personal and cannot be derived from external authorities or objective facts. Heidegger’s (1962) view of authenticity is linked to his analysis of *Dasein*, a German word meaning “existence.” Inauthenticity arises when Dasein succumbs to the norms, values, and expectations imposed by an anonymous social collective, leading individuals to lose themselves in societal roles and identities; in contrast, authenticity is achieved when *Dasein* assumes ownership of its existence, recognizing its freedom and responsibility to shape its own life (Dreyfus, 1990). For Sartre (1956), an authentic existence is tethered to a person’s *choices*. Authenticity involves acknowledging one’s freedom and using this freedom to make choices. Moreover, the authentic individual should fully accept responsibility for the consequences of those choices. Taken together, although ancient Greek philosophers pondered authenticity in the context of achieving the highest good, existentialists emphasized self-discovery and the relationship between the authentic self and the external world.

The Psychology of Authenticity

Authenticity gained traction in the 20th century, largely due to the contributions of humanistic psychologists. Maslow placed self-actualization—the drive “to become everything that one is capable of becoming” and to “do what one is fitted for”—at the pinnacle of human motivation. Self-actualizing individuals share some good personal qualities (Maslow, 1954). For example, they are able to accept themselves, others, and the world as they are without chagrin or complaint, including the flaws and imperfections inherent in human nature. Moreover, these individuals are characterized by independence from external sources of satisfaction (e.g., other people and society); instead, they live in accordance with their true selves, and rely on their own potential and latent resources for growth and development.

Rogers (1961) introduced the construct of “real self” versus “ideal self.” The real self is the person one is, representing the core of one’s true identity, discovered through the genuine experience of one’s feelings rather than through societal impositions. However, the ideal self represents the person one wishes to be, shaped by external expectations and societal conditions of worth that are out of step with one’s own valuing. In his therapeutic approach, Rogers introduced “client-centred” therapy (1951) and the principle of “unconditional positive regard” (1957), which involves accepting clients without judgment, thus creating a safe space for them to explore and express their feelings and experiences. Unconditional acceptance allows individuals to express their authentic selves without fear of rejection, becoming a fully functioning person, thus attaining authenticity (Rogers, 1963). In summary, humanistic psychologists positioned authenticity as a fundamental aspect of personal growth, self-actualization, and the fulfilment of one's potential, underscoring its crucial role in living a fully realized life.

Despite humanistic psychologists' enduring fascination with the authentic self, empirical investigation into authenticity began only two decades ago. Drawing on self-determination theory (SDT; Ryan & Deci, 2000, 2002), which posits that authenticity is reflected in actions that align with one's true or core self—characterized by autonomy and self-determination—as well as humanistic psychologists' concepts of the self-actualizing and the fully functioning individual (Maslow, 1954; Rogers, 1963), Kernis and Goldman (Goldman, 2002; Kernis, 2003; Kernis & Goldman, 2005, 2006) provided a systematic definition and a multicomponent framework of authenticity. This framework comprises four distinct but interconnected components: awareness (i.e., being aware of one's strengths and weaknesses, trait characteristics, and emotions), unbiased processing (i.e., objectivity and acceptance with respect to one's strengths and weaknesses), behaviour (i.e., acting in accord with one's values, preferences, and needs), and relational orientation (i.e., being genuine and not "fake" in one's relationships with close others). These components encapsulate the trajectory of the authentic self from cognitive acknowledgment to behavioural expression. Since the introduction of this framework, the empirical study of authenticity has gained momentum in psychological science. Currently, researchers have interpreted the authentic self in various ways, with a predominant focus on its role as an expression of one's true or core self (Harter, 2002; Kernis & Goldman, 2005, 2006; Lenton, Bruder, et al., 2013; Sedikides et al., 2017, 2019). In other words, the authentic self is an underlying and potentially obscured aspect of the self (Strohminger et al., 2017).

1.2.2 Views of Authenticity

Next, I will review three major views of the subjective experience of authenticity (i.e., self-accuracy, self-consistency, self-enhancement) and discuss relevant empirical evidence (Sedikides & Schlegel, 2024)

Self-Accuracy View

A traditional view conceptualizes authenticity as *self-accuracy*, the motivation to form an accurate sense of the self or process self-relevant information in an unbiased manner (Kernis, 2003; Kernis & Goldman, 2006). Empirical evidence supports this view. For instance, individuals reporting higher levels of authenticity actively engage in exploring identity-relevant information while showing a decreased propensity to avoid acknowledging, deciding upon, or reconciling their identities (Kernis & Goldman, 2006). Defensiveness, a self-protective cognitive strategy employed to cope with self-threatening information, helps individuals maintain or enhance their self-esteem (Barrett et al., 2002). In this vein, some researchers observed that individuals with higher self-reported authenticity exhibit less defensive behaviour when confronted with evidence of actions that contradict their desired self-view (e.g., “Tell me about a time when you’ve done something unethical on an assignment”, Lakey et al., 2008). Indeed, everyday experiences of authenticity are associated with self-introspection and self-examination (Kernis & Goldman, 2006; Wood et al., 2008).

However, it is challenging to empirically verify this view, as accurate self-knowledge is difficult to achieve (Vazire & Carlson, 2010; Vazire & Wilson, 2012). This difficulty is pronounced in the current context, given the complexities involved in determining which aspects of the self are authentic. Also, individuals may defensively distort their responses or lack introspective access required to assess their authentic selves accurately (Koole, 2003). Indeed, individuals who self-report high levels of authenticity also report possessing more favourable attributes, raising questions about the objectivity of their self-assessments (Gillath et al., 2010; Newman et al., 2014).

Self-Consistency View

Another view conceptualizes authenticity as *self-consistency*, the motivation to align one’s behaviour in sync with internal standards (e.g., personality, attitudes, values, goals, or desires) and resist to external influences (Jongman-Sereno & Leary, 2019; Kernis & Goldman, 2006; Wood et al., 2008). This view can be tracked to existentialism (e.g., Kierkegaard, Heidegger) and humanistic psychology (e.g., Maslow, Rogers).

Supporting this view, evidence indicates that across-role self-concept consistency (i.e., self-concept consistency across various roles or contexts) is positively associated with self-reported authenticity (Boucher, 2011; Cross et al., 2003; Sheldon et al., 2012). Although this relationship is predominantly observed in Western cultures, another form of self-concept consistency within-role self-concept consistency (i.e., consistency within a single role or context over time), is also positively associated with subjective authenticity in samples from an Eastern culture (English & Chen, 2011). Additionally, elevated power enhances authenticity, mediated by increased self-concept consistency (Kraus et al., 2011). Other research distinguished between compartmentalized individuals (i.e., those who organise positive and negative self-concepts in a separate manner) and integrative individuals (i.e., those who organise positive and negative self-concepts in an integrated manner; Showers et al., 2015). The former feel that their outcomes are controlled by external events and their self-evaluations are contingent on the approval of others, whereas the latter focus on internal effort and choices. Integrative individuals experience greater authenticity than compartmentalized ones. Also, empirically, experimental manipulations that reduce identity integration increase inauthenticity

(Ebrahimi et al., 2020). Similarly, incongruence between one's gender identity (i.e., female) and experimentally assigned self-presentation (i.e., to present oneself in a masculine way) decreases authenticity, particularly in women with strong feminine identification and weak masculine identification (Dormanen et al., 2020).

However, people are inclined to perceive socially desirable behaviours as authentic even when those behaviours do not align with their underlying self-concept. For instance, participants rated themselves on the Big Five traits and authenticity as tethered to different roles (i.e., friend, student, employee, child, and partner); the more authentic participants felt within a particular role, the more positively they rated themselves, regardless of their dispositional Big-Five traits (Sheldon et al., 1997). A similar pattern was also observed in an ecological momentary assessment study, where state authenticity and state versions of traditional Big Five traits were measured as the behaviour and feelings were occurring; the positive enactment of certain states (e.g., acting agreeably) predicted authenticity: the more individuals acted agreeably, the more authentic they felt (Fleeson & Wilt, 2010). Likewise, moral behaviour, regarded as socially desirable, is consistently experienced and perceived as authentic, independently of participants' underlying traits (Christy et al., 2017; Newman et al., 2014). Finally, positive feelings, rather than trait-state consistency (i.e., behaving congruently with one's traits) predicts authenticity in daily life (Cooper et al., 2018). Taken together, authenticity is not necessarily aligned with the self-concept; rather, it is confounded by positivity.

Self-Enhancement View

An emerging view regards authenticity as *self-enhancement*, the motivation to pursue a tendentiously favourable view of oneself (Alicke & Sedikides, 2009; Sedikides, 2021; Sedikides & Gregg, 2008). Individuals generally perceive their authentic selves as inherently positive and morally good (Hicks et al., 2019; Strohminger et al., 2017). For instance, authenticity is positively associated with favourable self-views (Fleeson & Wilt, 2010; Sheldon et al., 1997; Wood et al., 2008). Moreover, individuals perceive behaviours reflecting positive versus negative personality traits as more authentic, even when both kinds of behaviours are congruent with their personal characteristics; additionally, both self-congruence and behavioural positivity influence perceptions of authenticity, indicating that self-perceived authenticity is confounded with positivity of one's actions (Jongman-Sereno & Leary, 2016). Also, daily self-enhancement strivings predict increased authenticity (Guenther et al., 2024). In addition, manipulating self-enhancement (e.g., receiving favourable feedback, describing an instance of exhibiting a positive trait—whether in the context of a past event or an imagined future scenario) elevates authenticity (Bailey & Iyengar, 2023; Guenther et al., 2024) and vice versa (Guenther et al., 2024). When individuals evaluate personal change in their lives more favourably, they are more likely to perceive the change as being guided by authenticity (Bench et al., 2015). The association between positivity and authenticity is further underscored by evidence that experimental manipulations of positive affect raise authenticity (Chen et al., 2023; Lenton, Slabu, et al., 2013).

The desire for a self-enhancing authentic self is consistent across both independent (e.g., United States,) and interdependent (e.g., China, Singapore) cultural contexts (Slabu et al., 2014), and even extends to misanthropes—those who generally hold negative views of humanity (De Freitas et al., 2018). Reinforcing this trend, both children and adults conceptualize goodness more than negativity as the essence of the self, suggesting a stable, cross-age perception of the fundamentally good authentic self (Heiphetz, 2019). Collectively, these findings suggest a strong connection between authenticity and self-enhancement.

In summary, the *self-accuracy/self-consistency* views propose that individuals strive to minimize distortion when processing self-relevant information, maintaining openness to both their strengths and weaknesses. However, the implied impartial self-perception is not met with strong empirical verification. In contrast, the *self-enhancement* view posits that the authentic self encompasses predominantly positive characteristics. Evidence for this view is compelling.

1.2.3 A Rigorous Test of the Self-enhancement View: Authentic Self versus Presented Self

As stated above, evidence is stronger for the *self-enhancement* view compared to the *self-accuracy* and *self-consistency* views of authenticity. But can the self-enhancement view account for the full conceptual range of authenticity? Is authenticity merely a manifestation of positivity or self-enhancement? The current thesis addressed these questions by comparing the authentic self with the self that individuals present to others, termed the “presented self.”

As the name implies, the presented self refers to the version of oneself that individuals present or project to others. Self-presentation is a fundamental aspect of human behaviour: even prehistoric peoples were conscious of the importance of presenting oneself to others (Tedeschi, 1981). In contemporary discussions, the concept of self-presentation is often associated with Goffman’s (1956) notion of “impression management.” He introduced a dramaturgical perspective of the self and other, suggesting that how people present themselves to others is a conscious process to reveal certain aspects of the self and to conceal others, in order to create a particular impression in social interactions. Similarly, in daily life, people knowingly and unwittingly manipulate the impression they give off, which is analogous to the relationship between a stage actor and their audience.

This deliberate and selective self-presentation positions the presented self as a benchmark of social favourability. Stakes are high for the presented self as it can facilitate or undermine cooperation, reputation, respect, status, and access to social groups, professional resources (e.g., jobs, promotions, housing), or personal resources (e.g., friends, partners; Dores Cruz et al., 2021; Vonasch et al., 2018). Consequently, self-presentation typically promotes a sanitized portrait of the individual, overemphasizing, if not extolling, one's strengths and underemphasizing, if not concealing, one's weaknesses (Baumeister, 1982; Roth et al., 1986). Humanistic theories and person-centred therapies suggest that people distort their social performances to align with perceived "conditions of worth" in their environment (Rogers, 1964). 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, may all be motivated by the desire to avoid being negatively judged by others. Research supports this notion, with findings indicating that self-presenting individuals often deny negative traits and endorse positive ones (Lee et al., 1999; Roth et al., 1986, 1988). Furthermore, the descriptors individuals select for their presented self (e.g., "who you are during most of your activities") are more socially desirable than those selected for their authentic self (e.g., "who you really are"; Schlegel et al., 2009).

Importantly, the presented self is not an ephemerality. Instead, it is internalized as part of the private self. Theoretical frameworks and empirical evidence bolster this assertion. People come to know themselves through interactions with others (Sluss & Ashforth, 2007). According to symbolic interactionism and role theory (Stryker & Statham, 1985), individuals construct their sense of self through social interactions, and in particular the behaviours they enact or roles they assume as well as others' reactions to these behaviours or roles. Similarly, the theory of reflected self-appraisal (Lundgren, 2004; Tice & Wallace, 2003) or the "looking-glass self" (Shrauger & Schoeneman, 1979), posits that self-perceptions influence judgments of others' perceptions, and, in reverse, judgments of others' perceptions, like looking glass, shape one's self-perception (Wallace & Tice, 2012). In other words, people project identities to one another and form identities out of the reactions of others to them. Research findings concur. Strategic self-presentations influence subsequent private self-views; that is, people shift both their overall evaluations of themselves and their evaluations of specific characteristics of themselves in the direction of their preceding self-presentations (Leary, 1995). Moreover, changes in self-evaluations that occurred in one context because of self-presentations can carry over to a new context in the absence of self-presentational pressures (Schlenker, 2003). In all, the presented self constitutes a mental representation that is as integral to one's identity as the authentic self.

1.2.4 Summary

The concept of authenticity has been increased empirical attention. In this section, I reviewed historical and contemporary interpretations of the authentic self and discussed three major views on what constitutes the experience of authenticity along with relevant empirical evidence. Also, I considered a highly positive of self-representation, the presented self. The first question I ask in this thesis is the following: When testing the authentic self against the presented self, will I observe unvarnished self-enhancement or traces of self-accuracy and self-consistency? In Chapters 2 and 3, I placed the concept of authenticity under empirical scrutiny behaviourally and neurophysiologically, comparing it with the presented self.

1.3 Behavioural Approaches to the Self

1.3.1 Self-reference Valence Task

The self-reference valence (SR-valence) task is a variant of the self-reference task, which indicates improved memory and faster reaction times for trait adjectives that are accompanied by self-referential instructions (“does the word describes you?”) relative to control, including other-referential, instructions (Northoff & Bermpohl, 2004). In the SR-valence task, participants judge whether positive versus negative traits are self-descriptive or non-self-descriptive (D’Argembeau et al., 2005). The task allows assessing the endorsement of positive versus negative traits (trait endorsement), and the speed of this endorsement (reaction time). Reaction time is used as a proxy for cognitive processing speed (Jensen, 2006). In the context of SR-valence tasks, it can reflect the cognitive accessibility of self-concept information (Schlegel et al., 2009); indeed, faster reaction times are indicative of stronger, more accessible associations with the self-concept (Cai et al., 2016; Rameson et al., 2010). In the SR-valence task, higher endorsement (i.e., judging more traits as self-descriptive), or faster reaction time thereof, of positive than negative traits (i.e., Valence × Endorsement interaction) is a signature of self-enhancement motivation (Cai et al., 2016). In Chapter 2 (Experimental Paper I), I use the SR-valence task to test the self-enhancement view of authenticity.

1.3.2 Emotional Stroop Task

The emotional Stroop task is a modified version of the traditional Stroop task (Mathews & Macleod, 1985; Watts et al., 1986). In this paradigm, participants attempt to identify the ink colour of words, but the words themselves are not colour-related but instead emotionally charged (e.g., related to the pathology of clinical patients) or neutral. Accumulated evidence across cognitive, social, and clinical psychology indicates a pronounced slowdown in responding to the ink colour of negative words compared to positive or neutral words, termed the emotional Stroop effect (Bar-Haim et al., 2007; Phaf & Kan, 2007; Williams et al., 1996). Although the Stroop effect quantifies the disparity in colour-naming performance between congruent (e.g., “red” printed in red) and incongruent (e.g., “red” printed in green) stimuli, the emotional Stroop effect delves into the discrepancy between emotional (e.g., “death” printed in red) and neutral (e.g., “door” printed in red) stimuli. Given that the dimensions in the emotional Stroop task lack the semantic conflict or agreement central to the classic Stroop effect, where word meaning interferes with colour naming, the emotional Stroop effect emerges as an independent phenomenon distinct from the traditional Stroop effect. In essence, it represents a generic slowdown driven by threat perception rather than a selective attention mechanism associated with the classic Stroop effect (Algom et al., 2004). Specifically, this threat-driven slowdown arises from the tendency to allocate attention preferentially toward threatening stimuli at the expense of concurrent tasks (Öhman, 1993; Öhman et al., 2001).

Self-enhancement motivation is potent (Baumeister, 1982; Leary & Kowalski, 1990; Sedikides & Gregg, 2008). As such, it is feasible to test the self-enhancement view via the interference of negative (vs. positive) self-evaluations on different self-representation (e.g., the authentic self, the presented self). In Chapter 3 (Experimental Paper II), I use a self-referent emotional Stroop task to test the self-enhancement view.

1.4 A Cognitive Neuroscientific Approach to the Self

Researchers commonly explore the self via both self-report questionnaires and experiments (e.g., SR-valence tasks where participants judge whether a list of traits are self-descriptive or not; Cai et al., 2016). However, individuals might defensively distort their answers or lack introspective access to their self, especially to the authentic self; thus, questions about the authentic self demand answers that cannot be given by these traditional methodologies (Koole, 2003). In the current thesis, I adopted a cognitive neuroscientific approach to examine the electrophysiological underpinnings underlying self by recording participants’ electroencephalography signal while they carried out the behavioural task. I considered event-related potentials (ERP) as covert measures independent of behavioural responses. Specifically, I focused on ERPs relevant to early-stage selective attention (P1) and attention allocation (N170) in response to emotionally evocative stimuli in testing the self-enhancement view. I also examined earlier (P2) and later (P3) stages of self-relevant information processing, as well as later elaborate processing and stimulus significance (Late Positive Potential or LPP).

1.4.1 P1

The P1 is the initial positive deflection in the parieto-occipital region, and typically emerges 60–90 ms post-stimulus with a peak between 100 and 130 ms (Luck, 2014). Originating in the visual cortex, the P1 has conventionally been conceptualized as an early sensory-evoked component in relation to sensory amplification and selective attention (Hillyard et al., 1998). Previous studies indicated that P1 is mainly sensitive to physical stimulus characteristics (visual contrast, spatial frequency, luminance, size, etc.) and reflects activity of striate and extrastriate visual areas (Dhond et al., 2001; Hauk and Pulvermuller, 2004). Recently, researchers also found that the P1 is sensitive to emotional stimuli (Mueller et al., 2013; Schindler & Bublatzky, 2020), suggesting its role in the rapid detection of affectively salient information. Moreover, some studies observed larger P1 amplitudes evoked by negative stimuli (e.g., faces, words) compared to neutral counterparts (Luo et al., 2010; Zhang et al., 2014), indicating that the P1 can differentiate between non-threatening and potentially threatening information (Zhang et al., 2014). This early emotional discrimination could be mediated by rapid, coarse visual processing via magnocellular pathways projecting to the amygdala cortex (Pourtois et al., 2013). Alternatively, some accounts propose that the P1 emotion effect arises from feedback signals from higher-order regions that modulate early visual processing (Pessoa & Adolphs, 2010). Critically, such early emotional response may signify rapid extraction of emotion-related information and may function—at least partly— independent of subsequent, more detailed emotional processes such as N170 (Vuilleumier & Pourtois, 2007), which supports the notion of a "quick and dirty" emotional processing system (LeDoux, 1996). This mechanism may facilitate survival by enabling swift behavioural responses (e.g., vigilance or avoidance) while conserving cognitive resources for subsequent elaborate evaluations.

1.4.2 N170

Following the P1 component, the N1 emerges as a negative deflection maximal over occipito-temporal regions, typically peaking at about 100 ms post-stimulus onset with a duration of approximately 100 ms (Näätänen & Picton 1987). This component can reflect at least six distinct cerebral processes originating from different neural generators, each supporting unique psychophysiological functions (Näätänen & Picton 1987). Notably, the N1 has gained particular prominence for its sensitivity to visual discrimination processes, manifesting being modulated by visual features and attention under different stimulus types (Luck, 2014). Crucially, it exhibits robust differentiation between orthographically structured stimuli (words, letters, and pseudofont strings) and visually matched but linguistically meaningless controls (symbols or non-character patterns), with significantly enhanced amplitudes for the former (Brem et al., 2005, 2006; Maurer et al., 2005). This selective enhancement suggests early neural specialization for word processing within the first 100 ms of visual analysis. While both P1 and N1 serve as reliable early neurophysiological markers of visual processing, they exhibit distinct functional profiles: the P1 primarily reflects early sensory encoding of physical stimulus characteristics, whereas the N1 is more strongly associated with relative higher-order feature discrimination and analysis (Brem et al., 2006).

The N170 is a special type of the posterior N1, a negative deflection that typically peaks at approximately 170 ms after stimulus onset over the lateral occipito-temporal regions (Luck, 2014). The N170 is well-known for being face-sensitive, evincing a larger peaking in response to face-elicited stimuli than non-face-elicited stimuli to reflect early rapid attention to visual stimuli (Rossion & Jacques, 2012). The face-related N170 can be modulated by the valence of facial expression, with a significantly enhanced negativity for emotional relative to neutral facial expression (Luo et al., 2010; Williams et al., 2006). Such modulation is also found in emotional word processing (Zhang et al., 2014), with larger N170 amplitudes for negative adjectives compared to positive adjectives (Montalan et al., 2008). In all, the N170 can reflect early attention to stimuli with emotional valence, especially for negative emotional stimuli, with a larger N170 amplitude representing the allocation of more attentional resources (Cai et al., 2016).

1.4.3 P2

The P2 is a positive deflection spanning from 150 to 250 ms over anterior-central region (Luck, 2014). Typically, it exhibits greater amplitude in response to stimuli containing target features, indicating early selective attention towards task-relevant stimuli (Potts, 2004; Potts et al., 2006). This effect is enhanced when the targets are relatively infrequent (Glazer & Nusslock, 2022; Luck & Hillyard, 1994). Furthermore, the P2 has been associated with emotional processing, suggesting its role in modulating selective attention influenced by emotional content (Hajcak et al., 2012; Kotz & Paulmann, 2011). 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 & Bublatzky, 2020). Unlike the earlier emotion-detection stages represented by the P1 and N170, the P2 is generally associated with higher-order, deeper, and conscious emotional processing (Nie et al., 2020; Prete et al., 2015, 2018). An issue is whether the P2 is implicated in processes related to self-referential processing. Results have been inconsistent, with some studies reporting a significant reduced P2 amplitude for self-related stimuli (Keyes et al., 2010; Liu et al., 2019), some producing the opposite pattern (Fan et al., 2016; Fields & Kuperberg, 2012; San Martín et al., 2016), and others reporting null findings (Yang et al., 2014). For these mixed P2 findings, I consider several theoretically meaningful interpretations: a) familiarity account: The P2 reduction may reflect efficient processing of self-relevant information (Caharel et al., 2002; Keyes et al., 2010), where familiar self-descriptives require fewer attentional resources; b) motivational account: P2 enhancement could indicate heightened motivational significance of self-relevant stimuli (Fields & Kuperberg, 2012), particularly for emotionally charged self-information; c) task account: The null findings suggest P2 effects may be sensitive to specific task demands (Yang et al., 2014). In the current project, it can clarify whether this component reflects: (a) familiarity of self-information, (b) motivational significance of self-information, or (c) task-specific account (e.g., P2 is modulated by different self-representation, i.e., authentic vs. presented self).

1.4.4 P3

The P3 is a maximal positive wave that typically peaks around 300 ms post-stimulus at the parietal midline region (Luck, 2014). As one of the most extensively studied ERP components, it serves as a multifaceted neural marker of higher-order cognitive and affective psychological processing. Most prominently, the P3 is a hot topic in the electroencephalography field for its iconic increased positivity following the presentation of infrequent and surprising (have low subjective probability) stimuli, with larger amplitudes elicited by rare or unexpected target stimuli (Polich, 2012; Pritchard, 1981). Although both P2 and P3 are larger for infrequent stimuli, modulation of the P2 occurs only when the target is defined by fairly simple stimulus features, whereas modulation of the P3 can occur for complex target categories (Barkaszi et al., 2013; Luck, 2014; Song et al., 2005). This temporal and functional dissociation highlights the P3's role in higher-order cognitive operations rather than low-level sensory processing.

Of particular relevance to this project, numerous ERP studies of self-referential processing have showed that P3 is frequently associated with the discrimination of self from others, that is, a larger P3 wave has followed the presentation of self-related objects, words, names, and faces relative to the same stimuli of others (Knyazev, 2013; Gray et al., 2004; Miyakoshi et al., 2007; Tacikowski & Nowicka, 2010). These findings indicate that the amplitude of the P300 reflects increased attention or deeper processing of self-relevant stimuli (Porter et al., 2021). Moreover, the P3 also serves as a sensitive index of emotional processes, showing heightened amplitudes in response to emotional stimuli (both pleasant and unpleasant stimuli) compared to neutral ones (for a review, see Hajcak et al., 2010). These affective modulations can reflect the engagement of motivated attention systems (Bradley et al., 2003), as emotionally salient stimuli inherently capture attentional resources due to their motivational significance.

Furthermore, the P3 has been closely linked to decision-making and confidence evaluation. Its amplitude is positively modulated by decision confidence, exhibiting larger deflections as confidence increases (Boldt & Yeung, 2015; Selimbeyoglu et al., 2012; Ye et al., 2019), an effect that persists independent of objective accuracy (Eimer & Mazza, 2005). Besides, P3 amplitude is also related with task difficulty and effort, however, this relationship remains equivocal: while some studies report amplitude reduction with increased effort, others demonstrate enhancement or no significant modulation (for a review, see Ghani et al., 2020).

The P3 is also sensitive to reward-related processes. Its amplitude scales with both actual and anticipated reward magnitude (Bellebaum et al., 2010; Glazer et al., 2018; Goldstein et al., 2006; Yeung & Sanfey, 2004), and reward-specific P3 modulations correlate with ventral striatal BOLD activity (Pfabigan et al., 2014)—a key hub in the brain's reward circuitry (Schultz, 2000).

Collectively, the P3's dual sensitivity to cognitive and affective dimensions renders it a uniquely valuable tool for probing their interplay, particularly in contexts where emotional salience and self-relevance shape information processes.

1.4.5 Late Positive Potential

The LPP manifests as a sustained positive deflection typically observed approximately 400–500 ms post-stimulus presentation, persisting for several hundred milliseconds at the midline centroparietal region (Hajcak et al., 2012). It should be noted that “LPP” is not a standardized ERP component and may reflect different neural processes across studies. Nevertheless, it is consistently a late-emerging component (i.e., ~300 ms post-stimulus and extend for many hundreds of milliseconds) and primarily sensitive to emotional stimuli (Hajcak et al., 2012; Luck, 2014). Notably, the LPP exhibits an augmented amplitude in response to emotionally arousing stimuli when compared to neutral stimuli, spanning various modalities such as pictures, faces, hand gestures, and words (Hajcak & Foti, 2020). Moreover, the LPP is notably sensitive to self-referential information, exhibiting greater amplitudes for self-reference relative to non-self-reference content (Hudson et al., 2020; Jordan et al., 2022; Żochowska et al., 2021). Furthermore, the LPP can capture both the emotional and evaluative processing with respect to the self. For instance, some studies have found augmented LPP responses to negative (vs. positive) stimuli when participants refer to themselves (Cai et al., 2016; Herbert et al., 2011), whereas other studies report the reverse, with greater LPP responses following positive versus negative words during the self-reference task (Auerbach et al., 2015; Shestyuk & Deldin, 2010). Although findings remain mixed, the amplified LPP in self-referential tasks may reflect deeper processing of self-relevant information, in line with the LPP’s broader role in sustained attention and elaborative processing (Auerbach et al., 2015; Hajcak et al., 2012), as well as in marking stimulus significance and motivational relevance (i.e., activation of appetitive or aversive motivational systems; Hajcak & Foti, 2020), with a larger LPP value indicating a greater significance of the stimulus. Specifically, as self-relevant stimuli inherently carry a higher motivational salience due to their connection with personal identity and goals (Leary, 2007), this amplified LPP thus indicate the prioritization of self-related stimuli in cognitive and emotional processing.

1.4.6 Summary

In this section, I reviewed ERPs associated with emotional arousal, specifically P1 and N170, to evaluate the self-enhancement view of authenticity. Additionally, I reviewed ERPs linked directly to self-relevant information processing, including P2, P3, and LPP. In Chapters 2 and 3, I use these ERPs to investigate the neurophysiological underpinnings of self-enhancement in relation to authenticity.

1.5 Authenticity, Self-Control, and Reward Responsivity

Building upon the three fundamental perspectives of authenticity and its measurement approaches (behavioural and neural) discussed in the preceding sections, I will review its psychological functions - particularly its role in modulating self-control and reward processing. Emerging evidence suggests authenticity serves as a self-regulatory resource that may influence reward valuation (e.g., Ge & Hou, 2021; Kokkoris et al., 2019; Li et al., 2023), and there is a close relationship between authenticity and eudaimonia reward (e.g., Disabato et al., 2016; Huta & Waterman, 2014; Ryan & Deci, 2001). By elaborating these relationships further in this section, it allows me to systematically evaluate how trait authenticity moderates the impact of self-control exertion on reward responsivity.

1.5.1 A Simple Summary of Self-Control

The ability to override or alter motivated responses (i.e., self-control) is crucial for goal-directed behaviour and contributes to many consequential outcomes including physical health, psychological well-being, ethical decision making, and successful interpersonal relationships (Vohs & Baumeister, 2016). Conversely, failures in self-control have negative consequences in these and other domains. Self-control has thus been of keen interest to psychologists, neuroscientists, philosophers, and the public.

The Resource Model of Self-Control

Self-control has been extensively investigated through the lens of the resource model (Baumeister et al., 1998). For 30 years, this model has enjoyed widespread influence in social/personality psychology and psychological science in general. According to it, the capacity to override or alter one's responses depends on limited inner resource or strength (Baumeister et al., 1998; Baumeister, Vohs, et al., 2007). Acts of self-control are theorized to consume (i.e., deplete) this strength, resulting in temporary decline in the capacity for self-control (i.e., ego depletion). In support, numerous studies have found that engaging in a taxing (or depleting) self-control task undermines performance on subsequent demanding tasks (Baumeister et al., 2007, 2018, 2023). Mechanistically, these effects were thought to be driven by glucose, which posits that limited self-control resources can be replenished through metabolic means (Gailliot et al., 2007), though meta-analyses have cast doubt on this interpretation (Dang, 2016).

Nevertheless, empirical challenges, controversies, and debates related to the validity of the resource model have arisen. An initial meta-analysis of the relevant literature reported evidence for consistent and large effects (Hagger et al., 2010), but more recent meta-analyses have concluded that the effect is negligible after adjusting for publication bias (Carter et al., 2015; Carter & McCullough, 2014). However, the bias-correction statistical techniques used in these latter meta-analyses, which aimed to address publication bias, were untested and demonstrated variable efficacy across different contexts, thereby raising methodological concerns (Inzlicht et al., 2015). Multi-laboratory experiments obtained non-significant aftereffects of self-control exertion (Hagger et al., 2016; Vohs et al., 2021), whereas other preregistered, large-sample experiments obtained statistically significant, albeit smaller than expected, effects (Dang et al., 2017; Garrison et al., 2019). Collectively, the mechanisms and aftereffects of self-control exertion remain poorly understood.

Reward Responsivity Hypothesis of Self-Control

The reward responsivity hypothesis of self-control (Kelley et al., 2019) was a response to controversies and challenges to the resource model. According to this hypothesis, irrespective of self-control success, exercising self-control is aversive and engenders negative affect (David et al., 2024; Kurzban, 2016). To countermand this discomfort, reward seeking behaviour may be augmented after bouts of self-control, bringing individuals back to a mildly positive baseline state. In contrast, the resource model does not explicitly predict that exercising self-control increases subsequent reward-related impulse strength. Yet, several studies inspired by the resource model have reported evidence that exercising self-control increases subsequent reward-seeking behaviour, including eating, spending, and sexual behaviour (Baumeister, Schmeichel, et al., 2007). These behavioural outcomes could be due to a reduction in the capacity for control (as the resource model of self-control initially assumed) or increases in reward responsivity (as the reward responsivity hypothesis proposed). Several studies in line with the reward responsivity hypothesis of self-control have circumvented this interpretational ambiguity by instructing participants to complete reward-related tasks requiring little to no self-control. These studies find that self-control exertion enhances self-reported approach motivation (Schmeichel et al., 2010) and positive emotional reactivity (Finley & Schmeichel, 2019).

Several theoretical models of self-control are also consistent with the reward responsivity hypothesis. The process model proposes that shifts in motivation and attention steer individuals away from the further engagement of control and toward rewards (Inzlicht et al., 2014; Inzlicht & Schmeichel, 2012). Integrative self-control theory posits that, because control efforts are taxing and dependent upon finite resources, exercising self-control eventually tips the scales toward greater desire-driven reward-seeking behaviour (Kotabe & Hofmann, 2015). Value-based choice models of self-control suggest that exercising self-control shifts valued-based calculations in favour of more immediate, hedonic options over more effortful options in accord with one's long-term goals (Berkman et al., 2017). Similarly, both cognitive dissonance (Aronson & Mills, 1959; Harmon-Jones & Mills, 2019) and psychological contrast (Zentall, 2010) accounts of effort suggest that aversive states elicited by the exertion of effort make the end-result or reward appear more valuable. Furthermore, research on counter-regulation indicates that individuals more easily process information that is incongruent with their prevailing emotional state (Rothermund et al., 2008). Insofar as exercising self-control is aversive, counter-regulation would predict greater ease in processing reward-relevant information, leading to a compensatory attentional shift towards rewards after exerting self-control. Finally, opponent-process theories of motivation assume that organisms have a fundamental motivation to maintain homeostasis and shifts in affective states over time are presumably consequences of this motivation (Solomon, 1980; Solomon & Corbit, 1974). From this perspective, strong emotional states can have strong opposing emotional aftereffects. Self-control exertion is aversive, so prolonged exertion may eventually trigger an internal threshold that temporarily shifts attention toward sources of reward or good feelings until homeostasis is attained.

1.5.2 Self-Control and Reward Responsivity

Given the potential involvement of reward mechanisms in self-control processes, research has begun to address reward responsivity following acts of self-control. In this thesis, I examine Reward Positivity as a key neural marker of reward responsivity.

Reward Positivity

The Reward Positivity (RewP), a neural marker of reward responsivity, typically peaks approximately 200 to 300 ms following the onset of performance-related feedback, with the most pronounced activity over fronto-central sites (Glazer et al., 2018; Holroyd et al., 2008, 2011; Miltner et al., 1997; Walsh & Anderson, 2012). This response also has been referred to as feedback error-related negativity (fERN), feedback-related negativity (FRN), medial frontal negativity (MFN), or feedback negativity (FN) (Glazer et al., 2018; Proudfit, 2015). The RewP is sensitive to feedback reflecting the outcome of an action (Carlson et al., 2011; Foti et al., 2011; Walsh & Anderson, 2012), and is modulated by the delivery of advantageous versus disadvantageous or neutral outcomes (Harmon-Jones, Clarke, et al., 2020; Harmon-Jones, Willoughby, et al., 2020; Luo et al., 2022; Ma et al., 2014; San Martín et al., 2016). Consequently, the RewP is often examined by computing the difference between the ERPs elicited by advantageous and disadvantageous or neutral outcomes (Harmon-Jones, Clarke, et al., 2020).

Recently, converging evidence indicates that this ERP difference wave is primarily driven by reward-related rather than by loss- or error-related cues (Proudfit, 2015). For example, the RewP amplitude correlates with both the likelihood and magnitude of reward (Sambrook & Goslin, 2015) and is linked to self-reported trait reward responsiveness and behavioural indices of reward sensitivity (Bress & Hajcak, 2013). Additionally, the RewP is thought to reflect activity in reward-related subcortical and cortical regions, including the ventral striatum/nucleus accumbens, medial prefrontal cortex (e.g., anterior cingulate cortex, orbital frontal cortex), and amygdala (Becker et al., 2014; Carlson et al., 2011, 2015; Crane et al., 2022; Foti et al., 2011, 2014).

Self-Control and Reward Positivity

Evidence suggests that increased effort amplifies the RewP. For instance, Ma et al. (2014) demonstrated that high-effort tasks (e.g., multiplication) elicited larger RewP responses to the reward and non-reward discrepancy, whereas low-effort tasks (e.g., addition) did not produce such differentiation. Similarly, Pan et al. (2023) demonstrated that tasks requiring greater effort (e.g., pressing the target key as many times as possible) generated a more pronounced RewP amplitude compared to lower-effort tasks (e.g., pressing the target key once), a pattern observed in both reward and non-reward conditions. Consistent with these findings, Harmon-Jones et al. (2024) found that in an effortful task-switching paradigm, high-effort (versus low-effort) condition yielded greater RewP amplitudes when participants believed their effort can result in a reward. Additionally, Bogdanov et al. (2022) reported significantly elevated RewP responses in trials requiring higher cognitive effort compared to those requiring less, and effort levels significantly predicted participants' subjective rating of the effort demanding in each trial. These neural findings were also corroborated by self-reports data, where subjective effort exertion correlates with larger RewP difference waves (Harmon-Jones, Clarke, et al., 2020; Harmon-Jones, Willoughby, et al., 2020). In summary, the literature indicates that effort exertion modulates the RewP. Given that effort constitutes an integral component of self-control (Kotabe & Hofmann, 2015), it is plausible that exerting self-control may enhance the RewP.

Type of Reward

Reward is essential to the sense of well-being for everyday human behaviour (Berridge & Kringelbach, 2008). One of the earliest and most enduring frameworks in pursuit of well-being distinguishes between hedonic well-being (hedonia) and eudaimonic well-being (eudaimonia), which was originally grounded in ancient Greek philosophers such as Aristotle and Aristippus (4th Century BCE). Specifically, hedonic well-being is characterized by the pursuit of pleasure and comfort (the "pleasant life"), whereas eudaimonic well-being is centred on meaning and self-actualization (the "meaningful life"; Huta & Waterman, 2014; Ryan & Deci, 2001; Telzer et al., 2014).

Although both hedonism and eudaimonia are fundamental to well-being (Huta & Waterman, 2014), and are positively associated with life satisfaction, meaning, and flourishing (Henderson et al., 2013), as well as with each other (Disabato et al., 2016; Goodman et al., 2018), substantial research highlights their relative independence. For example, Huta and Waterman (2014) reviewed previous studies on hedonia and eudaimonia, identifying key differences in their core elements—the most common core elements in definition of eudaimonia are growth, authenticity, meaning, and excellence, whereas pleasure, enjoyment, comfort, absence of distress are central to the definition of hedonia. Also, hedonia is typically framed in terms of experiences, whereas eudaimonia is discussed in terms of orientations or functioning. Joshanloo (2016), employing exploratory structural equation modelling (ESEM), similarly found that hedonic and eudaimonic factors are correlated yet largely independent from each other. Furthermore, eudaimonic behaviours (e.g., give money to a person in need) have stronger and more lasting associations with well-being (e.g., meaning in life, life satisfaction, positive affect) than hedonic behaviours (e.g., attending a party), with the effects of eudaimonic behaviours persisting the following day (Steger et al., 2008).

Along these lines, reward can also be classified into two distinct yet complementary categories: hedonic and eudaimonic rewards. Specifically, hedonic rewards are extrinsically pleasure-driven and self-focused, such as enjoying material goods or playing video games, while eudaimonic rewards are intrinsically meaningful and purposeful, such as helping strangers or donating to charity (Shizgal, 1999; Steger et al., 2008; Telzer et al., 2014). These reward types differentially influence well-being. For instance, in a longitudinal study (Telzer et al., 2014), neural activation associated with eudaimonic rewards (e.g., donating money to family) predicted decreases in depressive symptoms over time, whereas activation linked to hedonic rewards (e.g., keeping money for themselves) related to longitudinal increases in depressive symptoms. Similarly, neural indices (e.g., P3, RewP) of eudaimonic reward processing, but not hedonic reward processing, positively predicted longitudinal improvement in well-being (e.g., increased positive emotions; Luo et al., 2019, 2022).

Crucially, some recent studies indicate that hedonic rewards and eudaimonic rewards also influence reward responsivity, although findings remain inconsistent. For instance, one study reported that hedonic rewards (i.e., winning rewards for oneself) elicited a larger RewP difference wave compared to eudaimonic rewards (i.e., winning rewards for charity; Luo et al., 2019). However, other studies found comparable RewP between hedonic rewards and eudaimonic rewards (Luo et al., 2022; Zhang et al., 2023). These discrepancies highlight the need for further research to clarify the distinct neural processes underlying these different forms of reward.

1.5.3 Authenticity and Self-control

Authenticity has been established as a critical factor in psychological functioning (Sedikides & Schlegel, 2024). Emerging research highlights its pivotal role in improving human psychological health and well-being. For example, authenticity relates to, predicts, and increases meaning in life (Hong et al., 2024; Lutz et al., 2023; Schlegel et al., 2011, 2012) as well as life satisfaction (Boyraz et al., 2014; Lutz et al., 2023; Rathi & Lee, 2021). Furthermore, authenticity serves as a protective factor against adverse mental health symptoms, such as anxiety and depression (Asher & Aderka, 2021; Bryan et al., 2017), stress (Maffly-Kipp et al., 2020; Zou et al., 2023), and burnout (Ockerman, Mardourian, Han, Petrauskis, et al., 2024; Ockerman, Mardourian, Han, Sorice-Virk, et al., 2024). In the current thesis, I focus specifically on the relationship between authenticity and self-control.

This relationship is intricate. Does self-control suppress who people truly are, or does it facilitate becoming one's authentic self? According to self-determination theory, pursuing goals aligned with one's genuine interests and values could promote authenticity (Ryan & Deci, 2000, 2002). In this context, self-control, defined as the ability to override momentary impulses in favour of goal-directed and long-term outcomes (De Ridder et al., 2012), is assumed to enhance authenticity. Empirical evidence supports this notion. For instance, self-reported self-control is positively associated with authenticity, and experimental manipulations of self-control (e.g., resisting the temptation to buy a favourite cake) increase authenticity (Ge & Hou, 2021). Resisting (vs. indulging) temptation leads to heightened authenticity for individuals with a rationalist orientation—those who prioritize reason over feelings (Kokkoris et al., 2019). Longitudinal data from a large sample of Chinese adolescents (*N* = 2,982) reveal a reciprocal relationship: self-control predicted increases in authenticity over time, and vice versa (Li et al., 2023). Interestingly, an actor-observer asymmetry has been observed: individuals perceived self-control as more authentic for themselves, whereas they perceive impulsive actions as more authentic for others (Garrison et al., 2023). This divergence suggests that self-control is an intrinsic component of the authentic self-construction, whereas it is distinct from perceptions of others' authenticity. Self-control also plays a crucial role in the pursuit of authentic goals. Specifically, individuals with higher (vs. lower) self-control are more likely to prioritize goals that reflect their true selves, and in turn increased goal authenticity predicts greater goal attainment (Stavrova et al., 2019). Taken together, this stream of research indicates that, when people exert self-control, they feel more authentic.

However, the alternative perspective suggests that exerting self-control may be perceived as suppressing desires and conforming to external pressures (Kokkoris, 2024), thus decreasing authenticity. In daily life, perceptions of authenticity often include qualities such as "uninhibited" and "unaffected by others" (Garrison et al., 2023), implying that impulsive actions are more authentic. Supporting this notion, an internal meta-analysis controlling for the positivity of self-control revealed that exerting self-control was seen as less authentic than acting on impulse, and this effect was stronger when evaluating others (Garrison et al., 2023). Also, the habitual use of emotional suppression (i.e., chronic use of self-control to inhibit or override emotional responses) often results in negative consequences (e.g., weaker relationship satisfaction, lower social support), driven by inauthenticity (English & John, 2013). This suggests that self-control, by creating a disconnect between inner experience and outward behaviour, undermines authenticity, whereas more spontaneous or uncontrolled responses strengthen authenticity. Finally, when self-control is made salient, individuals are less likely to perceive their choices as reflecting their preferences, which could undermine authenticity (Sela et al., 2017). Collectively, this body of research suggests that, although self-control may serve long-term goals, it can sometimes diminish authenticity.

Another question is: does authenticity enhance self-control? A lay theory, the true-self-as-guide (TSAG) framework, posits that the true self guides behaviour (Rivera et al., 2019). Acting in alignment with one's authentic self, then, serves as an internal norm (e.g., follow who you real are), promoting congruence between individual choices and the true self. Adhering to one's true self is seen as an effective strategy for navigating conflict situations that require self-control, thereby contributing to well-being. The self-concordance model similarly posits that pursuing self-concordant goals (i.e., goals consistent with one's intrinsic interests and core values) enables individuals to invest sustained effort in achieving them (Sheldon & Elliot, 1999). Empirical evidence provides preliminary and indirect support for the notion that a clear and consistent self-concept enables individuals to identify and prioritize self-initiated and personally valued goals, thereby enhancing effective self-control in goal pursuit. Self-concept clarity reflects the certainty and coherence of one’s self-concept, and people with low self-concept clarity rely on external factors to maintain a coherent identity, making them more dependent on, and susceptible to, external influences (Campbell, 1990). In in contrast, those with higher self-concept clarity—possessing a clear and coherent self-view—are more likely to exhibit effective self-control (Jiang et al., 2023). Similarly, self-concept clarity and grit—defined as perseverance and passion for long-term goals (Duckworth et al., 2007), a construct closely related to self-control—have been shown to positively reinforce on each other over time (Wong & Vallacher, 2018). Indeed, authenticity predicts increases in self-control over time, and vice versa (Li et al., 2023). In summary, although research on this topic is limited, initial findings suggest that authenticity may foster self-control. In summary, the relationship between self-control and authenticity is more sophisticated than suggested by earlier research (Kokkoris, 2024).

1.5.4 Authenticity and Eudaimonia

Reward responsivity following self-control exertion may also be modulated by authenticity. This possibility is underscored by the close connection between authenticity and eudaimonia, which extends as far back as ancient Hellenic philosophy. In his *Nicomachean Ethics*, Aristotle (1985) introduces the concept of “*daimon*,” which refers to the inherent potentialities within each individual. The realization of these potentialities represents the highest form of human fulfilment, and striving to live in accordance with one’s *daimon*—essentially, to achieve self-realization—leads to a state known as *eudaimonia*. Eudaimonia, therefore, calls for individuals to live authentically, aligning their lives with their true selves to achieve personal growth (Disabato et al., 2016; Ryan & Deci, 2001; Waterman, 1993). Additionally, authenticity is not only integral to eudaimonia, but it is also considered one of its core elements (Huta & Waterman, 2014; Smallenbroek et al., 2017). Research has established a positive relationship between authenticity and meaning in life (Lutz et al., 2023; Schlegel et al., 2009; Schlegel & Hicks, 2011), and this meaning-making function of the authentic self further reinforces its unique association with eudaimonia, distinguishing it from hedonia (Schlegel et al., 2013; Smallenbroek et al., 2017).

Taken together, both authenticity and eudaimonia emphasize living in harmony with one’s true self and actualizing one’s potential. Both concepts prioritize aligning actions with deeply held values and purpose, fostering meaning in life and personal development. Consequently, individuals who live authentically are more likely to experience eudaimonia rather than hedonia, which is focused primarily on immediate pleasure and satisfaction. In the context of reward processing, such a relationship between authenticity and eudaimonia may also shape reward responsivity following self-control exertion based on reward types (e.g., hedonic vs. eudaimonic rewards).

1.5.5 Summary

In this section, I reviewed the literature on the resource model and the reward responsivity hypothesis of self-control, focusing on the neural correlates of reward responsivity, specifically the RewP following self-control exertion. Additionally, I examined the relationship between authenticity and self-control, as well as the connection between authenticity and eudaimonia. Although much of the self-control literature centres on hedonic rewards, its influence on eudaimonic rewards remains unclear. Moreover, given the established links between authenticity and both self-control and eudaimonic well-being, it is essential to investigate whether trait authenticity influences the effect of self-control on different reward types. I will do so in Chapter 4 (Experimental Paper III, Parts A and B).

1.6 Summary

This dissertation aims to explore the essence of authenticity—the true, genuine self of an individual—from three distinct perspectives: self-accuracy, self-consistency, and self-enhancement. Additionally, it further investigates the role of authenticity in the relationship between self-control and reward processing.

1.6.1 Empirical Paper I

Authenticity has been predominantly conceptualized as self-accuracy, self-consistency, and self-enhancement, with the last view gaining evidentiary ground. Empirical Paper I put the self-enhancement view of authenticity to a rigorous test by comparing the authentic self against another a highly positive self-representation, the presented self. I ask whether the authentic self, compared to the presented self, is a fierce denouncer of undesirable information and an unabashed consumer of desirable information (self-enhancement view), or, alternatively, whether the authentic self is prone to exploring or accepting the possibility of having some undesirable attributes (self-accuracy and self-consistency views). I address this issue by using the SR-valence task in which participants are shown a list of positive and negative traits and need to judge whether each trait describe themselves or not while undergoing EEG recording. It yields behavioural measures (reaction times, trait endorsements) and neural markers (N170, P3, LPP) of self-referential processing.

1.6.2 Empirical Paper II

Building on the theoretical framework and findings established in Empirical Paper I, I test the self-enhancement, self-accuracy, and self-consistency views further in Empirical Paper II through the implementation of the self-referent emotional Stroop task, and in particular the interference of negative information on the self-evaluation. Specifically, participants view in coloured text positive or negative traits exemplifying themselves and need to identify the colour of these sentences while undergoing EEG recording. It yields behavioural measure (reaction times) and neural markers (P1, N170, P2, P3) of self-referential processing.

1.6.3 Empirical Paper III

According to the reward responsivity hypothesis of self-control, exercising self-control is aversive and engenders negative affect. To countermand this discomfort, reward seeking behaviours may be amplified after bouts of self-control, bringing individuals back to the baseline state. Research and theory indicate that exercising self-control enhances the neural basis of the reward responsivity (i.e., RewP). However, it is unclear whether this effect occurs for hedonic rewards, eudaimonic rewards, or both. Moreover, individuals who live authentically are likely to have a high level of self-control and experience eudaimonia. Taken together, the purpose of Empirical Paper III Part A is to (1) examine the neural basis of the reward responsivity hypothesis of self-control by assessing how self-control exertion impacts the Reward Positivity, and (2) expand this hypothesis by testing the extent to which exercising self-control influences the reward system differently for hedonic versus eudaimonic rewards. To investigate these questions, I employ a modified monetary incentive delay task in which participants need to complete a speeded reaction time task where they exercised self-control or not, and then have the opportunity to win money for themselves (hedonic rewards) or a charity (eudaimonic rewards) while EEG is recorded. It yields behavioural measures (hit rate, reaction times) and neural marker of reward processing, RewP. The purpose of Empirical Paper III Part B is to examine whether the reward responsivity following self-control exertion is potentially influenced by authenticity at the trait level. To investigate this question, participants' trait authenticity is assessed alongside behavioural and neural measures of reward processing.

References

Abrams, D., & Hogg, M. A. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. Routledge.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: a generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General*, *133*(3), 323–338. https://doi.org/10.1037/0096-3445.133.3.323

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*(1), 1–48. https://doi.org/10.1080/10463280802613866

Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The anterior cingulate cortex: The evolution of an interface between emotion and cognition. *Annals of the New York Academy of Sciences*, *935*(1), 107–117. https://doi.org/10.1111/j.1749-6632.2001.tb03476.x

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*(1), 29–52. https://doi.org/10.1111/nyas.12360

Aristotle. (1985). *Nicomachean ethics*. (T. Irwin, Trans.). Indianapolis: Hackett. (Original work published ca. 350 BCE)

Aristotle, Tredinnick, H., & Barnes, J. (2004). *Nicomachean ethics* (J. A. K. Thomson, Trans.; Revised by H. Trennidick; Introduction by J. Barnes). London, England: Penguin. (Original work published ca. 340 BCE)

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, *59*(2), 177–181. https://doi.org/10.1037/h0047195

Asher, M., & Aderka, I. M. (2021). How real do you feel? Self-and partner-authenticity in social anxiety disorder. *Journal of Abnormal Psychology*, *130*(2), 166.

Auerbach, R. P., Stanton, C. H., Proudfit, G. H., & Pizzagalli, D. A. (2015). Self-referential processing in depressed adolescents: A high-density event-related potential study. *Journal of Abnormal Psychology*, *124*(2), 233–245. https://doi.org/10.1037/abn0000023

Bailey, E. R., & Iyengar, S. S. (2023). Positive—More than unbiased—Self-perceptions increase subjective authenticity. *Journal of Personality and Social Psychology*, *125*(6), 1351–1372. https://doi.org/10.1037/pspa0000359

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25–34. https://doi.org/10.1016/j.biopsycho.2011.01.010

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Barkaszi, I., Czigler, I., & Balázs, L. (2013). Stimulus complexity effects on the event-related potentials to task-irrelevant stimuli. *Biological Psychology*, *94*(1), 82–89. https://doi.org/10.1016/j.biopsycho.2013.05.007

Barrett, L. F., Williams, N. L., & Fong, G. T. (2002). Defensive verbal behaviour assessment. *Personality and Social Psychology Bulletin*, *28*(6), 776–788. https://doi.org/10.1177/0146167202289007

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3–26. https://doi.org/10.1037/0033-2909.91.1.3

Baumeister, R. F. (2012). *Public self and private self*. Springer Science & Business Media.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Baumeister, R. F., Schmeichel, B. J., & Vohs, K. D. (2007). Self-regulation and the executive function: The self as controlling agent. In *Social psychology: Handbook of basic principles* (Vol. 2, pp. 516–539). Guilford Press.

Baumeister, R. F., Tice, D. M., & Bushman, B. J. (2023). A review of multisite replication projects in social psychology: is it viable to sustain any confidence in social psychology’s knowledge base?. *Perspectives on Psychological Science, 18*(4), 912–935. https://doi.org/10.1177/17456916221121815

Baumeister, R. F., Tice, D. M., & Vohs, K. D. (2018). The strength model of self-regulation: Conclusions from the second decade of willpower research. *Perspectives on Psychological Science*, *13*(2), 141–145. https://doi.org/10.1177/1745691617716946

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355. https://doi.org/10.1111/j.1467-8721.2007.00534.x

Becker, M. P., Nitsch, A. M., Miltner, W. H., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: the feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia, 48*(11), 3343–3350. https://doi.org/10.1016/j.neuropsychologia.2010.07.023

Bench, S. W., Schlegel, R. J., Davis, W. E., & Vess, M. (2015). Thinking about change in the self and others: The role of self-discovery metaphors and the true self. *Social Cognition*, *33*(3), 169–185. https://doi.org/10.1521/soco.2015.33.3.2

Berkman, E. T., Hutcherson, C. A., Livingston, J. L., Kahn, L. E., & Inzlicht, M. (2017). Self-control as value-based choice. *Current Directions in Psychological Science*, *26*(5), 422–428. https://doi.org/10.1177/0963721417704394

Berridge, K. C., & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, *199*(3), 457–480. https://doi.org/10.1007/s00213-008-1099-6

Bogdanov, M., Renault, H., LoParco, S., Weinberg, A., & Otto, A. R. (2022). Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cerebral Cortex*, *32*, 4255–4270. https://doi.org/10.1093/cercor/bhab480

Bola, M., Paź, M., Doradzińska, Ł., & Nowicka, A. (2021). The self-face captures attention without consciousness: Evidence from the N2pc ERP component analysis. *Psychophysiology*, *58*(4), e13759. https://doi.org/10.1111/psyp.13759

Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience, 35*(8), 3478–3484. https://doi.org/10.1523/JNEUROSCI.0797-14.2015

Boucher, H. C. (2011). The dialectical self-concept II: Cross-role and within-role consistency, well-being, self-certainty, and authenticity. *Journal of Cross-Cultural Psychology*, *42*(7), 1251–1271. https://doi.org/10.1177/0022022110383316

Boyraz, G., Waits, J. B., & Felix, V. A. (2014). Authenticity, life satisfaction, and distress: A longitudinal analysis. *Journal of Counseling Psychology*, *61*(3), 498–505. https://doi.org/10.1037/cou0000031

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience, 117*(2), 369.

Brem, S., Lang-Dullenkopf, A., Maurer, U., Halder, P., Bucher, K., & Brandeis, D. (2005). Neurophysiological signs of rapidly emerging visual expertise for symbol strings. *Neuroreport, 16*(1), 45–48.

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., & Brandeis, D. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage, 29*(3), 822–837.

Bress, J. N., & Hajcak, G. (2013). Self-report and behavioural measures of reward sensitivity predict the feedback negativity. *Psychophysiology*, *50*(7), 610–616. https://doi.org/10.1111/psyp.12053

Bryan, J. L., Baker, Z. G., & Tou, R. Y. (2017). Prevent the blue, be true to you: Authenticity buffers the negative impact of loneliness on alcohol-related problems, physical symptoms, and depressive and anxiety symptoms. *Journal of Health Psychology*, *22*(5), 605–616. https://doi.org/10.1177/1359105315609090

Caharel, S., Poiroux, S., Bernard, C., Thibaut, F., Lalonde, R., & Rebai, M. (2002). ERPs associated with familiarity and degree of familiarity during face recognition. *International Journal of Neuroscience, 112*(12), 1499–1512.

Cai, H., Wu, L., Shi, Y., Gu, R., & Sedikides, C. (2016). Self-enhancement among Westerners and Easterners: A cultural neuroscience approach. *Social Cognitive and Affective Neuroscience*, *11*(10), 1569–1578. https://doi.org/10.1093/scan/nsw072

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, *59*(3), 538–549. https://doi.org/10.1037/0022-3514.59.3.538

Carlson, J. M., Foti, D., Harmon-Jones, E., & Proudfit, G. H. (2015). Midbrain volume predicts fMRI and ERP measures of reward reactivity. *Brain Structure and Function*, *220*(3), 1861–1866. https://doi.org/10.1007/s00429-014-0725-9

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *Neuroimage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796–815. https://doi.org/10.1037/xge0000083

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00823

Chen, K., Zhang, H., & Schlegel, R. J. (2023). Does being angry feel authentic? A test of how affective valence and motivational direction differentially influence state authenticity. *Motivation and Emotion*, *47*(5), 828–841. https://doi.org/10.1007/s11031-023-10026-1

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979.

Christy, A. G., Kim, J., Vess, M., Schlegel, R. J., & Hicks, J. A. (2017). The reciprocal relationship between perceptions of moral goodness and knowledge of others' true selves. *Social Psychological and Personality Science*, *8*(8), 910–917. https://doi.org/10.1177/1948550617693061

Cooper, A. B., Sherman, R. A., Rauthmann, J. F., Serfass, D. G., & Brown, N. A. (2018). Feeling good and authentic: Experienced authenticity in daily life is predicted by positive feelings and situation characteristics, not trait-state consistency. *Journal of Research in Personality*, *77*, 57–69. https://doi.org/10.1016/j.jrp.2018.09.005

Crane, N. A., Burkhouse, K. L., Gorka, S. M., Klumpp, H., & Phan, K. L. (2022). Electrocortical measures of win and loss processing are associated with mesocorticolimbic functional connectivity: A combined ERP and rs-fMRI study. *Psychophysiology*, *59*(12), e14118. https://doi.org/10.1111/psyp.14118

Cross, S. E., Gore, J. S., & Morris, M. L. (2003). The relational-interdependent self-construal, self-concept consistency, and well-being. *Journal of Personality and Social Psychology*, *85*(5), 933.

Cunningham, S. J., Brebner, J. L., Quinn, F., & Turk, D. J. (2014). The self-reference effect on memory in early childhood. *Child Development*, *85*(2), 808–823. https://doi.org/10.1111/cdev.12144

Dang, J. (2016). Testing the role of glucose in self-control: A meta-analysis. *Appetite*, 107, 222–230. https://doi.org/10.1016/j.appet.2016.07.021

Dang, J., Liu, Y., Liu, X., & Mao, L. (2017). The ego could be depleted, providing initial exertion is depleting. *Social Psychology*, *48*(4), 242–245. https://doi.org/10.1027/1864-9335/a000308

D’Argembeau, A. (2013). On the role of the ventromedial prefrontal cortex in self-processing: The valuation hypothesis. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00372

D’Argembeau, A., Comblain, C., & Van der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology*, *96*(4), 457–466. https://doi.org/10.1348/000712605X53218

David, L., Vassena, E., & Bijleveld, E. (2024). The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect. *Psychological Bulletin*. Advance online publication. https://doi.org/10.1037/bul0000443

De Freitas, J., Sarkissian, H., Newman, G. E., Grossmann, I., De Brigard, F., Luco, A., & Knobe, J. (2018). Consistent belief in a good true self in misanthropes and three interdependent cultures. *Cognitive Science*, *42*(S1), 134–160. https://doi.org/10.1111/cogs.12505

De Ridder, D. T. D., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviours. *Personality and Social Psychology Review, 16*(1), 76–99. https://doi.org/10.1177/1088868311418749

Deschamps, J.-C., & Devos, T. (1998). *Regarding the relationship between social identity and personal identity.* Sage Publications, Inc. https://psycnet.apa.org/record/1998-07662-001

Dhond, R. P., Buckner, R. L., Dale, A. M., Marinkovic, K., & Halgren, E. (2001). Spatiotemporal maps of brain activity underlying word generation and their modification during repetition priming. *Journal of Neuroscience, 21*(10), 3564–3571.

Dignath, D., Eder, A. B., Herbert, C., & Kiesel, A. (2022). Self-related primes reduce congruency effects in the Stroop task. *Journal of Experimental Psychology: General*, *151*(11), 2879–2892. https://doi.org/10.1037/xge0001210

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471–482. https://doi.org/10.1037/pas0000209

Dores Cruz, T. D., Thielmann, I., Columbus, S., Molho, C., Wu, J., Righetti, F., de Vries, R. E., Koutsoumpis, A., Van Lange, P. A. M., Beersma, B., & Balliet. D. (2021). Gossip and reputation in everyday life. *Philosophical Transactions of the Royal Society B, 376*(1838), Article 20200301. https://doi.org/10.1098/rstb.2020.0301

Dormanen, R., Sanders, C. S., Maffly-Kipp, J., Smith, J. L., & Vess, M. (2020). Assimilation undercuts authenticity: A consequence of women's masculine self-presentation in masculine contexts. *Psychology of Women Quarterly*, *44*(4), 488–502. https://doi.org/10.1177/0361684320947648

Dreyfus, H. L. (1990). *Being-in-the-world: A commentary on Heidegger's being in time, division I*. Mit Press.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087. https://doi.org/10.1037/0022-3514.92.6.1087

Ebrahimi, M., Kouchaki, M., & Patrick, V. M. (2020). Juggling work and home selves: Low identity integration feels less authentic and increases unethicality. *Organizational Behaviour and Human Decision Processes*, *158*, 101–111.

Eimer, M., & Mazza, V. (2005). Electrophysiological correlates of change detection. *Psychophysiology, 42*(3), 328-342. https://doi.org/10.1111/j.1469-8986.2005.00285.x

English, T., & Chen, S. (2011). Self-concept consistency and culture: The differential impact of two forms of consistency. *Personality and Social Psychology Bulletin*, *37*(6), 838–849. https://doi.org/10.1177/0146167211400621

English, T., & John, O. P. (2013). Understanding the social effects of emotion regulation: The mediating role of authenticity for individual differences in suppression. *Emotion*, *13*(2), 314–329. https://doi.org/10.1037/a0029847

Erikson, E. H. (1950). *Childhood and society*. W. W. Norton & Company.

Fan, W., Chen, J., Wang, X.-Y., Cai, R., Tan, Q., Chen, Y., Yang, Q., Zhang, S., Wu, Y., Yang, Z., Wang, X.-A., & Zhong, Y. (2013). Electrophysiological correlation of the degree of self-reference effect. *PLOS ONE*, *8*(12), e80289. https://doi.org/10.1371/journal.pone.0080289

Fan, W., Zhong, Y., Yang, Z., Li, J., Ouyang, Y., Cai, R., Li, H., & Fu, X. (2016). The degree of self-reference effect of extroversion individuals. *Acta Psychologica Sinica*, *48*(8), 1002–1012. https://doi.org/10.3724/SP.J.1041.2016.01002

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*(4), 522.

Fields, E. C., & Kuperberg, G. R. (2012). It’s all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage*, *62*(1), 562–574. https://doi.org/10.1016/j.neuroimage.2012.05.003

Finley, A. J., & Schmeichel, B. J. (2019). Aftereffects of self-control on positive emotional reactivity. *Personality and Social Psychology Bulletin*, *45*(7), 1011–1027. https://doi.org/10.1177/0146167218802836

Fleeson, W., & Wilt, J. (2010). The relevance of Big Five trait content in behaviour to subjective authenticity: Do high levels of within-person behavioural variability undermine or enable authenticity achievement? *Journal of Personality*, *78*(4), 1353–1382. https://doi.org/10.1111/j.1467-6494.2010.00653.x

Foster, B. L., Koslov, S. R., Aponik-Gremillion, L., Monko, M. E., Hayden, B. Y., & Heilbronner, S. R. (2023). A tripartite view of the posterior cingulate cortex. *Nature Reviews Neuroscience*, *24*(3), 173–189. https://doi.org/10.1038/s41583-022-00661-x

Foti, D., Carlson, J. M., Sauder, C. L., & Proudfit, G. H. (2014). Reward dysfunction in major depression: Multimodal neuroimaging evidence for refining the melancholic phenotype. *NeuroImage*, *101*, 50–58. https://doi.org/10.1016/j.neuroimage.2014.06.058

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping*, *32*(12), 2207–2216. https://doi.org/10.1002/hbm.21182

Freud, S. (1923). *The Ego and the Id*. Standard edition of the complete psychological works of Sigmund Freud, Volume XIX.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*(2), 325–336. https://doi.org/10.1037/0022-3514.92.2.325

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21. https://doi.org/10.1016/S1364-6613(99)01417-5

Garrison, K. E., Finley, A. J., & Schmeichel, B. J. (2019). Ego depletion reduces attention control: Evidence from two high-powered preregistered experiments. *Personality and Social Psychology Bulletin*, *45*(5), 728–739. https://doi.org/10.1177/0146167218796473

Garrison, K. E., Rivera, G. N., Schlegel, R. J., Hicks, J. A., & Schmeichel, B. J. (2023). Authentic for thee but not for me: Perceived authenticity in self-control conflicts. *Personality and Social Psychology Bulletin*, *49*(12), 1646–1662. https://doi.org/10.1177/01461672221118187

Ge, X., & Hou, Y. (2021). Confucian ideal personality traits (Junzi personality) and mental health: The serial mediating roles of self-control and authenticity. *Acta Psychologica Sinica*, *53*(4), 374–386. https://doi.org/10.3724/SP.J.1041.2021.00374

Ghani, U., Signal, N., Niazi, I. K., & Taylor, D. (2020). ERP based measures of cognitive workload: A review. *Neuroscience & Biobehavioral Reviews, 118*, 18–26. https://doi.org/10.1016/j.neubiorev.2020.07.020

Gillath, O., Sesko, A. K., Shaver, P. R., & Chun, D. S. (2010). Attachment, authenticity, and honesty: Dispositional and experimentally induced security can reduce self- and other-deception. *Journal of Personality and Social Psychology*, *98*(5), 841–855. https://doi.org/10.1037/a0019206

Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, *131*(1), 76–97. https://doi.org/10.1037/0033-2909.131.1.76

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Glazer, J., & Nusslock, R. (2022). Outcome valence and stimulus frequency affect neural responses to rewards and punishments. *Psychophysiology*, *59*(3), e13981. https://doi.org/10.1111/psyp.13981

Glisky, E. L., & Marquine, M. J. (2009). Semantic and self-referential processing of positive and negative trait adjectives in older adults. *Memory*, *17*(2), 144–157. https://doi.org/10.1080/09658210802077405

Goffman, E. (1956). The presentation of self in everyday life. *American Sociological Review*, *21*(5), 631–640. https://doi.org/10.2307/2089106

Goldman, B. M. (2002). The role of authenticity in healthy psychological functioning and subjective well-being. *Annals of the American Psychotherapy Association*, *5*(6), 18–20.

Goldstein, R. Z., Cottone, L. A., Jia, Z., Maloney, T., Volkow, N. D., & Squires, N. K. (2006). The effect of graded monetary reward on cognitive event-related potentials and behavior in young healthy adults. *International Journal of Psychophysiology, 62*(2), 272–279. https://doi.org/10.1016/j.ijpsycho.2006.05.006

Goodman, F. R., Disabato, D. J., Kashdan, T. B., & Kauffman, S. B. (2018). Measuring well-being: A comparison of subjective well-being and PERMA. *The Journal of Positive Psychology*, *13*(4), 321–332. https://doi.org/10.1080/17439760.2017.1388434

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, *40*(2), 216–224. https://doi.org/10.1016/S0022-1031(03)00092-1

Guenther, C. L., Zhang, Y., & Sedikides, C. (2024). The authentic self is the self-enhancing self: A self-enhancement framework of authenticity. *Personality and Social Psychology Bulletin*, *50*(8), 1182–1196. https://doi.org/10.1177/01461672231160653

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(7), 4259–4264. https://doi.org/10.1073/pnas.071043098

Gutchess, A. H., Kensinger, E. A., & Schacter, D. L. (2010). Functional neuroimaging of self-referential encoding with age. *Neuropsychologia*, *48*(1), 211–219. https://doi.org/10.1016/j.neuropsychologia.2009.09.006

Gutchess, A. H., Kensinger, E. A., Yoon, C., & Schacter, D. L. (2007). Ageing and the self-reference effect in memory. *Memory*, *15*(8), 822–837. https://doi.org/10.1080/09658210701701394

Hagger, Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G., Bruyneel, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T., Crowell, A., De Ridder, D. T. D., Dewitte, S., ... Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*(4), 495–525. https://doi.org/10.1037/a0019486

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7), e13570. https://doi.org/10.1111/psyp.13570

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology, 35*(2), 129–155. https://doi.org/10.1080/87565640903526504

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In *The Oxford handbook of event-related potential components* (Vol. 441, pp. 441–472). Oxford University Press.

Harmon-Jones, E., Clarke, D., Paul, K., & Harmon-Jones, C. (2020). The effect of perceived effort on reward valuation: Taking the reward positivity (RewP) to dissonance theory. *Frontiers in Human Neuroscience*, *14*, 157. https://doi.org/10.3389/fnhum.2020.00157

Harmon-Jones, E., Matis, S., Angus, D. J., & Harmon-Jones, C. (2024). Does effort increase or decrease reward valuation? Considerations from cognitive dissonance theory. *Psychophysiology*, *61*(6), e14536. https://doi.org/10.1111/psyp.14536

Harmon-Jones, E., & Mills, J. (2019). *Cognitive dissonance: Reexamining a pivotal theory in psychology* (Second Editon). American Psychological Association.

Harmon-Jones, E., Willoughby, C., Paul, K., & Harmon-Jones, C. (2020). The effect of perceived effort and perceived control on reward valuation: Using the reward positivity to test a dissonance theory prediction. *Biological Psychology*, *154*, 107910. https://doi.org/10.1016/j.biopsycho.2020.107910

Harter, S. (2002). Authenticity. In *Handbook of positive psychology*. Oxford University Press.

Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping, 21*(3), 191–201.

Heidegger, M. (1962). *Being and Time* (J. Macquarrie & E. Robinson, Trans.). Harper & Row. (Original work published 1927).

Heilbronner, S. R., & Hayden, B. Y. (2016). Dorsal anterior cingulate cortex: A bottom-up view. *Annual Review of Neuroscience*, *39*(Volume 39, 2016), 149–170. https://doi.org/10.1146/annurev-neuro-070815-013952

Heiphetz, L. (2019). Moral essentialism and generosity among children and adults. *Journal of Experimental Psychology: General*, *148*(12), 2077–2090. https://doi.org/10.1037/xge0000587

Henderson, L. W., Knight, T., & Richardson, B. (2013). An exploration of the well-being benefits of hedonic and eudaimonic behaviour. *The Journal of Positive Psychology*, *8*(4), 322–336. https://doi.org/10.1080/17439760.2013.803596

Herbert, C., Pauli, P., & Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Social Cognitive and Affective Neuroscience*, *6*(5), 653–661. https://doi.org/10.1093/scan/nsq082

Hicks, J. A., Schlegel, R. J., & Newman, G. E. (2019). Introduction to the special issue: Authenticity: Novel insights into a valued, yet elusive, concept. *Review of General Psychology*, *23*(1), 3–7. https://doi.org/10.1177/1089268019829474

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *353*(1373), 1257–1270. https://doi.org/10.1098/rstb.1998.0281

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *NeuroReport*, *22*(5), 249. https://doi.org/10.1097/WNR.0b013e328345441d

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Hong, E. K., Zhang, Y., & Sedikides, C. (2024). Future self-continuity promotes meaning in life through authenticity. *Journal of Research in Personality*, *109*, 104463. https://doi.org/10.1016/j.jrp.2024.104463

Hudson, A., Wilson, M. J. G., Green, E. S., Itier, R. J., & Henderson, H. A. (2020). Are you as important as me? Self-other discrimination within trait-adjective processing. *Brain and Cognition*, *142*, 105569. https://doi.org/10.1016/j.bandc.2020.105569

Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies*, *15*(6), 1425–1456. https://doi.org/10.1007/s10902-013-9485-0

Inzlicht, M., Gervais, W., & Berkman, E. (2015). Bias-correction techniques alone cannot determine whether ego depletion is different from zero: Commentary on Carter, Kofler, Forster, & McCullough, 2015. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2659409

Inzlicht, M., & Schmeichel, B. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, *7*, 450–463. https://doi.org/10.1177/1745691612454134

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*(3), 127–133. https://doi.org/10.1016/j.tics.2013.12.009

James, W. (1890). *The principles of psychology*. Henry Holt.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

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. https://doi.org/10.1177/01461672221109664

Johnson, J. T., Robinson, M. D., & Mitchell, E. B. (2004). Inferences about the authentic self: When do actions say more than mental states? *Journal of Personality and Social Psychology*, *87*(5), 615–630. https://doi.org/10.1037/0022-3514.87.5.615

Jongman-Sereno, K. P., & Leary, M. R. (2016). Self-perceived authenticity is contaminated by the valence of one's behaviour. *Self and Identity*, *15*(3), 283–301. https://doi.org/10.1080/15298868.2015.1128964

Jongman-Sereno, K. P., & Leary, M. R. (2019). The enigma of being yourself: A critical examination of the concept of authenticity. *Review of General Psychology*, *23*(1), 133–142. https://doi.org/10.1037/gpr0000157

Jordan, C. D., Stewart, R. A., Brush, C. J., Cougle, J. R., & Hajcak, G. (2022). Appearance concerns are uniquely associated with LPP amplitude to pictures of oneself. *Social Cognitive and Affective Neuroscience*, *17*(4), 430–436. https://doi.org/10.1093/scan/nsab115

Joshanloo, M. (2016). Revisiting the empirical distinction between hedonic and eudaimonic aspects of well-being using exploratory structural equation modeling. *Journal of Happiness Studies*, *17*(5), 2023–2036. https://doi.org/10.1007/s10902-015-9683-z

Kelley, N. J., Finley, A. J., & Schmeichel, B. J. (2019). After-effects of self-control: The reward responsivity hypothesis. *Cognitive, Affective, & Behavioural Neuroscience*, *19*(3), 600–618. https://doi.org/10.3758/s13415-019-00694-3

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, *14*(1), 1–26. https://doi.org/10.1207/S15327965PLI1401_01

Kernis, M. H., & Goldman, B. M. (2005). From thought and experience to behaviour and interpersonal relationships: A multicomponent conceptualization of authenticity. In *Building,* *defending and regulating the self: A psychological perspective* (pp. 31–52). Psychology Press. https://doi.org/10.4324/9780203998052

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 283–357). Academic Press. https://doi.org/10.1016/S0065-2601(06)38006-9

Keyes, H., & Brady, N. (2010). Self-face recognition is characterized by “bilateral gain” and by faster, more accurate performance which persists when faces are inverted. *Quarterly Journal of Experimental Psychology*, *63*(5), 840–847. https://doi.org/10.1080/17470211003611264

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and Cognition*, *72*(2), 244–254. https://doi.org/10.1016/j.bandc.2009.09.006

Kierkegaard, S. (1983). *Fear and Trembling* (H. V. Hong & E. H. Hong, Trans.). Princeton University Press. (Original work published 1843).

Kim, K., & Johnson, M. K. (2015). Activity in ventromedial prefrontal cortex during self-related processing: Positive subjective value or personal significance? *Social Cognitive and Affective Neuroscience*, *10*(4), 494–500. https://doi.org/10.1093/scan/nsu078

Knyazev, G. G. (2013). EEG correlates of self-referential processing. *Frontiers in Human Neuroscience*, *7*, 264. https://doi.org/10.3389/fnhum.2013.00264

Kokkoris, M. D. (2024). Self-control and self-expression. *Current Opinion in Psychology*, 101846. https://doi.org/10.1016/j.copsyc.2024.101846

Kokkoris, M. D., Hoelzl, E., & Alós-Ferrer, C. (2019). True to which self? Lay rationalism and decision satisfaction in self-control conflicts. *Journal of Personality and Social Psychology*, *117*(2), 417–447. https://doi.org/10.1037/pspp0000242

Koole, S. L. (2003). In search of the real self: A functional perspective on optimal self-esteem and authenticity. *Psychological Inquiry*, *14*(1), 43–48.

Kotabe, H. P., & Hofmann, W. (2015). On integrating the components of self-control. *Perspectives on Psychological Science*, *10*(5), 618–638. https://doi.org/10.1177/1745691615593382

Kraus, M. W., Chen, S., & Keltner, D. (2011). The power to be me: Power elevates self-concept consistency and authenticity. *Journal of Experimental Social Psychology*, *47*(5), 974–980.

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, *7*, 67–70. https://doi.org/10.1016/j.copsyc.2015.08.003

Lakey, C. E., Kernis, M. H., Heppner, W. L., & Lance, C. E. (2008). Individual differences in authenticity and mindfulness as predictors of verbal defensiveness. *Journal of Research in Personality*, *42*(1), 230–238. https://doi.org/10.1016/j.jrp.2007.05.002

Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behaviour*. Brown & Benchmark Publishers.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, *58*(1), 317–344. https://doi.org/10.1146/annurev.psych.58.110405.085658

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*(1), 34.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* Simon & Schuster.

Lee, S.-J., Quigley, B. M., Nesler, M. S., Corbett, A. B., & Tedeschi, J. T. (1999). Development of a self-presentation tactics scale. *Personality and Individual Differences*, *26*(4), 701–722. https://doi.org/10.1016/S0191-8869(98)00178-0

Lenton, A. P., Bruder, M., Slabu, L., & Sedikides, C. (2013). How does “being real” feel? The experience of state authenticity: state authenticity. *Journal of Personality*, *81*(3), 276–289. https://doi.org/10.1111/j.1467-6494.2012.00805.x

Lenton, A. P., Slabu, L., Sedikides, C., & Power, K. (2013). I feel good, therefore I am real: Testing the causal influence of mood on state authenticity. *Cognition and Emotion*, *27*(7), 1202–1224. https://doi.org/10.1080/02699931.2013.778818

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Aging, Neuropsychology, and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Li, Q., Ren, X., Zhou, Z., & Wang, J. (2023). Reciprocal relationships between self-control and self-authenticity: A two-wave study. *Frontiers in Psychology*, *14*, 1207230. https://doi.org/10.3389/fpsyg.2023.1207230

Li, Z., Chen, Y., Yin, S., & Chen, A. (2024). Self-referential information optimizes conflict adaptation. *Memory & Cognition*, *52*, 648–662. https://doi.org/10.3758/s13421-023-01490-8

Lieberman, M. D., Straccia, M. A., Meyer, M. L., Du, M., & Tan, K. M. (2019). Social, self,(situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neuroscience & Biobehavioural Reviews*, *99*, 311–328.

Liu, L., Li, W., Li, J., Lou, L., & Chen, J. (2019). Temporal features of psychological and physical self-representation: An ERP study. *Frontiers in Psychology*, *10*, 785. https://doi.org/10.3389/fpsyg.2019.00785

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed). MIT Press.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x

Lundgren, D. C. (2004). Social feedback and self-appraisals: Current status of the Mead-Cooley hypothesis. *Symbolic Interaction*, *27*(2), 267–286. https://doi.org/10.1525/si.2004.27.2.267

Luo, W., Feng, W., He, W., Wang, N., & Luo, Y. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, *49*(2), 1857–1867. https://doi.org/10.1016/j.neuroimage.2009.09.018

Luo, Y., Jiang, H., Chen, X., Zhang, Y., & You, X. (2019). Temporal dynamics of hedonic and eudaimonic reward processing: An event-related potentials (ERPs) study. *International Journal of Psychophysiology*, *137*, 63–71. https://doi.org/10.1016/j.ijpsycho.2018.12.009

Luo, Y., Zhang, X., Jiang, H., & Chen, X. (2022). The neural habituation to hedonic and eudaimonic rewards: Evidence from reward positivity. *Psychophysiology*, *59*, Airtle e13977. https://doi.org/10.1111/psyp.13977

Lutz, P. K., Newman, D. B., Schlegel, R. J., & Wirtz, D. (2023). Authenticity, meaning in life, and life satisfaction: A multicomponent investigation of relationships at the trait and state levels. *Journal of Personality*, *91*(3), 541–555. https://doi.org/10.1111/jopy.12753

Ma, Q., Meng, L., Wang, L., & Shen, Q. (2014). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research*, *261*, 1–7. https://doi.org/10.1016/j.bbr.2013.11.045

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 619.

Macrae, C. N., Visokomogilski, A., Golubickis, M., Cunningham, W. A., & Sahraie, A. (2017). Self-relevance prioritizes access to visual awareness. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(3), 438.

Maffly-Kipp, J., Flanagan, P., Kim, J., Schlegel, R. J., Vess, M., & Hicks, J. A. (2020). The role of perceived authenticity in psychological recovery from collective trauma. *Journal of Social and Clinical Psychology*, *39*(5), 419–448. https://doi.org/10.1521/jscp.2020.39.5.419

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, *5*(4), 420–430. https://doi.org/10.1177/1745691610375557

Maslow, A. H. (1954). *Motivation and personality* (pp. xiv, 411). Harpers.

Mathews, A., & Macleod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, *23*(5), 563–569. https://doi.org/10.1016/0005-7967(85)90104-4

Maurer, U., Brem, S., Bucher, K., & Brandeis, D. (2005). Emerging neurophysiological specialization for letter strings. *Journal of Cognitive Neuroscience, 17*(10), 1532–1552.

Miyakoshi, M., Nomura, M., & Ohira, H. (2007). An ERP study on self-relevant object recognition. *Brain and Cognition*, *63*(2), 182–189. https://doi.org/10.1016/j.bandc.2006.12.001

Montalan, B., Caharel, S., Personnaz, B., Le Dantec, C., Germain, R., Bernard, C., Lalonde, R., & Rebaï, M. (2008). Sensitivity of N170 and late positive components to social categorization and emotional valence. *Brain Research*, *1233*, 120–128. https://doi.org/10.1016/j.brainres.2008.07.055

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*(1), 56–60. https://doi.org/10.1080/17470215908416289

Mueller, A. K., Falter, C. M., & Tucha, O. (2013). Distinct attentional mechanisms for social information processing as revealed by event-related potential studies. *International Journal of Brain and Cognitive Sciences*, *2*(2), 23–37. https://doi.org/10.5923/j.ijbcs.20130202.03

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*(4), 375–425.

Newman, G. E., Bloom, P., & Knobe, J. (2014). Value judgments and the true self. *Personality and Social Psychology Bulletin*, *40*(2), 203–216. https://doi.org/10.1177/0146167213508791

Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, *41*(2), 352–362. https://doi.org/10.1037/0012-1649.41.2.352

Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, *8*, 102–107. https://doi.org/10.1016/j.tics.2004.01.004

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. https://doi.org/10.1016/j.neuroimage.2005.12.002

Ockerman, K. M., Mardourian, M., Han, S. H., Petrauskis, M., Sorice-Virk, S., & Ching, J. (2024). Authenticity: Protection against burnout and depression in emergency medicine physicians. *The American Journal of Emergency Medicine*, *83*, 138–141. https://doi.org/10.1016/j.ajem.2024.07.021

Ockerman, K. M., Mardourian, M., Han, S. H., Sorice-Virk, S., & Ching, J. (2024). Protective effects of authenticity against depression, suicide, and burnout among surgeons. *Journal of the American College of Surgeons*, *239*(5), 485. https://doi.org/10.1097/XCS.0000000000001128

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In *Handbook of emotions* (pp. 511–536). Guilford Press.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466

Pan, W., Lu, J., Wu, L., Kou, J., & Lei, Y. (2023). Expending effort may share neural responses with reward and evokes high subjective satisfaction. *Biological Psychology*, *177*, 108480. https://doi.org/10.1016/j.biopsycho.2022.108480

Peetz, J., & Wilson, A. E. (2008). The temporally extended self: The relation of past and future selves to current identity, motivation, and goal pursuit. *Social and Personality Psychology Compass*, *2*(6), 2090–2106. https://doi.org/10.1111/j.1751-9004.2008.00150.x

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience, 11*(11), 773–783. https://doi.org/10.1038/nrn2920

Pfabigan, D. M., Seidel, E. M., Sladky, R., Hahn, A., Paul, K., Grahl, A., ... & Lamm, C. (2014). P300 amplitude variation is related to ventral striatum BOLD response during gain and loss anticipation: an EEG and fMRI experiment. *NeuroImage, 96*, 12–21. https://doi.org/10.1016/j.neuroimage.2014.03.077

Phaf, R. H., & Kan, K.-J. (2007). The automaticity of emotional Stroop: A meta-analysis. *Journal of Behaviour Therapy and Experimental Psychiatry*, *38*(2), 184–199. https://doi.org/10.1016/j.jbtep.2006.10.008

Plato. (1961). *Apology*. In E. Hamilton & H. Cairns (Eds.), *The collected dialogues of Plato* (pp. 3-26). Princeton University Press.

Plato. (2007). *The Republic.* (D. Lee, Trans.; Introduction by M. Lane, 3rd ed.). London: Penguin Classics. (Original work published ca. 375 BCE)

Polich, J. (2012). Neuropsychology of P300. In *The Oxford handbook of event-related potential components* (pp. 159–188). Oxford University Press.

Porter, N. A., Fields, E. C., Moore, I. L., & Gutchess, A. (2021). Late frontal positivity effects in self-referential memory: Unique to the self? *Social Neuroscience*, *16*(4), 406–422. https://doi.org/10.1080/17470919.2021.1929460

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology, 92*(3), 492–512. https://doi.org/10.1016/j.biopsycho.2012.02.007

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain and Cognition*, *56*(1), 5–13. https://doi.org/10.1016/j.bandc.2004.03.006

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: The medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, *18*(7), 1112–1119.

Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, *89*(3), 506–540. https://doi.org/10.1037/0033-2909.89.3.506

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449–459. https://doi.org/10.1111/psyp.12370

Rameson, L. T., Satpute, A. B., & Lieberman, M. D. (2010). The neural correlates of implicit and explicit self-relevant processing. *NeuroImage*, *50*(2), 701–708. https://doi.org/10.1016/j.neuroimage.2009.12.098

Ratcliffe, M. (2006). “Folk psychology” is not folk psychology. *Phenomenology and the Cognitive Sciences*, *5*(1), 31–52. https://doi.org/10.1007/s11097-005-9010-y

Rathi, N., & Lee, K. (2021). Does it pay to be authentic? Implications of authenticity for life satisfaction and psychological well-being in a collectivist culture. *Journal of Happiness Studies*, *22*(1), 147–161. https://doi.org/10.1007/s10902-020-00223-x

Rivera, G. N., Christy, A. G., Kim, J., Vess, M., Hicks, J. A., & Schlegel, R. J. (2019). Understanding the relationship between perceived authenticity and well-being. *Review of General Psychology*, *23*(1), 113–126. https://doi.org/10.1037/gpr0000161

Rogers, C. (1951). *Client centreed therapy: Its current practice, implications and theory*. Houghton Mifflin Company.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, *21*(2), 95–103.

Rogers, C. R. (1961). *On becoming a person*. Houghton Mifflin Harcourt.

Rogers, C. R. (1963). The concept of the fully functioning person. *Psychotherapy: Theory, Research & Practice*, *1*(1), 17–26. https://doi.org/10.1037/h0088567

Rogers, C. R. (1964). Toward a modern approach to values: The valuing process in the mature person. *The Journal of Abnormal and Social Psychology, 68*(2), 160–167. https://doi.org/10.1037/h0046419

Rossion, B., & Jacques, C. (2012). The N170: Understanding the time course of face perception in the human brain. In *The Oxford handbook of event-related potential components* (pp. 115–141). Oxford University Press.

Roth, D. L., Harris, R. N., & Snyder, C. R. (1988). An individual differences measure of attributive and repudiative tactics of favourable self-presentation. *Journal of Social and Clinical Psychology*, *6*(2), 159–170. https://doi.org/10.1521/jscp.1988.6.2.159

Roth, D. L., Snyder, C. R., & Pace, L. M. (1986). Dimensions of favourable self-presentation. *Journal of Personality and Social Psychology*, *51*(4), 867–874.

Rothermund, K., Voss, A., & Wentura, D. (2008). Counter-regulation in affective attentional biases: A basic mechanism that warrants flexibility in emotion and motivation. *Emotion*, *8*(1), 34–46. https://doi.org/10.1037/1528-3542.8.1.34

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166. https://doi.org/10.1146/annurev.psych.52.1.141

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In *Handbook of self-determination research* (Vol. 2, pp. 3–33). University of Rochester Press.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. https://doi.org/10.1037/bul0000006

San Martín, R., Kwak, Y., Pearson, J. M., Woldorff, M. G., & Huettel, S. A. (2016). Altruistic traits are predicted by neural responses to monetary outcomes for self vs charity. *Social Cognitive and Affective Neuroscience*, *11*(6), 863–876. https://doi.org/10.1093/scan/nsw026

Sartre, J.P. (1956). *Being and Nothingness: An Essay on Phenomenological Ontology* (Hazel E. Barnes, Trans.). New York: Washington Square Press.

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, *130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Schlegel, R. J., & Hicks, J. A. (2011). The true self and psychological health: Emerging evidence and future directions. *Social and Personality Psychology Compass*, *5*(12), 989–1003. https://doi.org/10.1111/j.1751-9004.2011.00401.x

Schlegel, R. J., Hicks, J. A., Arndt, J., & King, L. A. (2009). Thine own self: True self-concept accessibility and meaning in life. *Journal of Personality and Social Psychology*, *96*(2), 473–490. https://doi.org/10.1037/a0014060

Schlegel, R. J., Hicks, J. A., King, L. A., & Arndt, J. (2011). Feeling like you know who you are: Perceived true self-knowledge and meaning in life. *Personality and Social Psychology Bulletin*, *37*(6), 745–756. https://doi.org/10.1177/0146167211400424

Schlegel, R. J., Hirsch, K. A., & Smith, C. M. (2013). The importance of who you really are: The role of the true self in eudaimonia. In *The best within us: Positive psychology perspectives on eudaimonic functioning* (pp. 207–226). American Psychological Association. https://psycnet.apa.org/record/2012-24003-011

Schlegel, R. J., Vess, M., & Arndt, J. (2012). To discover or to create: Metaphors and the true self. *Journal of Personality, 80*(4), 969–993. https://doi.org/10.1111/j.1467-6494.2011.00753.x

Schlenker, B. R. (2003). Self-presentation. *Handbook of Self and Identity*, *2*, 542–570.

Schmeichel, B. J., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology*, *99*(1), 162–173. https://doi.org/10.1037/a0019797

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience, 1*(3), 199–207. https://doi.org/10.1038/35044563

Sedikides, C. (2021). Self-construction, self-protection, and self-enhancement: A homeostatic model of identity protection. *Psychological Inquiry*, *32*(4), 197–221. https://doi.org/10.1080/1047840X.2021.2004812

Sedikides, C., & Brewer, M. B. (Eds.). (2015). *Individual self, relational self, collective self*. Psychology Press. https://doi.org/10.4324/9781315783024

Sedikides, C., Gaertner, L., & O’Mara, E. M. (2011). Individual self, relational self, collective self: Hierarchical ordering of the tripartite self. *Psychological Studies*, *56*(1), 98–107. https://doi.org/10.1007/s12646-011-0059-0

Sedikides, C., & Gregg, A. P. (2003). Portraits of the self. In *Sage handbook of social psychology* (pp. 110–138). Sage Publications. https://doi.org/10.4135/9781848608221.n5

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3*(2), 102–116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

Sedikides, C., Hong, E. K., & Wildschut, T. (2023). Self-continuity. *Annual Review of Psychology*, *74*, 333–361. https://doi.org/10.1146/annurev-psych-032420-032236

Sedikides, C., Lenton, A. P., Slabu, L., & Thomaes, S. (2019). Sketching the contours of state authenticity. *Review of General Psychology*, *23*(1), 73–88. https://doi.org/10.1037/gpr0000156

Sedikides, C., & Schlegel, R. J. (2024). Distilling the concept of authenticity. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-024-00323-y

Sedikides, C., Skowronski, J. J., & Dunbar, R. I. (2006). When and why did the human self evolve. In *Evolution and social psychology: Frontiers in social psychology* (pp. 55–80). Psychology Press.

Sedikides, C., Slabu, L., Lenton, A., & Thomaes, S. (2017). State authenticity. *Current Directions in Psychological Science*, *26*(6), 521–525. https://doi.org/10.1177/0963721417713296

Sedikides, C., & Spencer, S. (2007). *The self: Frontiers in social psychology*. Psychology Press.

Selimbeyoglu, A., Keskin-Ergen, Y., & Demiralp, T. (2012). What if you are not sure? Electroencephalographic correlates of subjective confidence level about a decision. *Clinical Neurophysiology, 123*(6), 1158–1167. https://doi.org/10.1016/j.clinph.2011.10.037

Sela, A., Berger, J., & Kim, J. (2017). How self-control shapes the meaning of choice. *Journal of Consumer Research*, *44*(4), 724–737. https://doi.org/10.1093/jcr/ucx069

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*(3), 482–497.

Sheldon, K. M., Gunz, A., & Schachtman, T. R. (2012). What does it mean to be in touch with oneself? Testing a social character model of self-congruence. *Self and Identity*, *11*(1), 51–70. https://doi.org/10.1080/15298868.2010.503130

Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of Personality and Social Psychology*, *73*(6), 1380–1393. https://doi.org/10.1037/0022-3514.73.6.1380

Shestyuk, A. Y., & Deldin, P. J. (2010). Automatic and strategic representation of the self in major depression: Trait and state abnormalities. *American Journal of Psychiatry*, *167*(5), 536–544. https://doi.org/10.1176/appi.ajp.2009.06091444

Shizgal, P. (1999). On the neural computation of utility: Implications from studies of brain stimulation reward. In *Well-being: The foundations of hedonic psychology* (pp. 502–526). Russell Sage Foundation.

Showers, C. J., Ditzfeld, C. P., & Zeigler-Hill, V. (2015). Self-concept structure and the quality of self-knowledge. *Journal of Personality*, *83*(5), 535–551. https://doi.org/10.1111/jopy.12130

Shrauger, J. S., & Schoeneman, T. J. (1979). Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, *86*(3), 549–573. https://doi.org/10.1037/0033-2909.86.3.549

Slabu, L., Lenton, A. P., Sedikides, C., & Bruder, M. (2014). Trait and state authenticity across cultures. *Journal of Cross-Cultural Psychology*, *45*(9), 1347–1373. https://doi.org/10.1177/0022022114543520

Sluss, D. M., & Ashforth, B. E. (2007). Relational identity and identification: Defining ourselves through work relationships. *Academy of Management Review*, *32*(1), 9–32. https://doi.org/10.5465/amr.2007.23463672

Smallenbroek, O., Zelenski, J. M., & Whelan, D. C. (2017). Authenticity as a eudaimonic construct: The relationships among authenticity, values, and valence. *The Journal of Positive Psychology*, *12*(2), 197–209. https://doi.org/10.1080/17439760.2016.1187198

Solomon, R. L. (1980). The opponent-process theory of acquired motivation: The costs of pleasure and the benefits of pain. *American Psychologist*, *35*(8), 691–712.

Solomon, R. L., & Corbit, J. D. (1974). An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychological Review*, *81*(2), 119–145.

Song, S., Zilverstand, A., Song, H., d'Oleire Uquillas, F., Wang, Y., Xie, C., Cheng, L., & Zou, Z. (2017). The influence of emotional interference on cognitive control: A meta-analysis of neuroimaging studies using the emotional Stroop task. *Scientific Reports*, *7*(1), 2088. https://doi.org/10.1038/s41598-017-02266-2

Song, Y., Ding, Y., Fan, S., Qu, Z., Xu, L., Lu, C., & Peng, D. (2005). Neural substrates of visual perceptual learning of simple and complex stimuli. *Clinical Neurophysiology*, *116*(3), 632–639. https://doi.org/10.1016/j.clinph.2004.09.019

Stavrova, O., Pronk, T., & Kokkoris, M. D. (2019). Choosing goals that express the true self: A novel mechanism of the effect of self-control on goal attainment. *European Journal of Social Psychology*, *49*(6), 1329–1336. https://doi.org/10.1002/ejsp.2559

Steger, M. F., Kashdan, T. B., & Oishi, S. (2008). Being good by doing good: Daily eudaimonic activity and well-being. *Journal of Research in Personality*, *42*(1), 22–42. https://doi.org/10.1016/j.jrp.2007.03.004

Steiner, C. J., & Reisinger, Y. (2006). Understanding existential authenticity. *Annals of Tourism Research*, *33*(2), 299–318. https://doi.org/10.1016/j.annals.2005.08.002

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, *12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Stryker, S., & Statham, A. (1985). *Symbolic interaction and role theory*. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 311–378). Random House.

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1105–1117. https://doi.org/10.1037/a0029792

Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences*, *19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015

Sui, J., Rotshtein, P., & Humphreys, G. W. (2013). Coupling social attention to the self forms a network for personal significance. *Proceedings of the National Academy of Sciences*, *110*(19), 7607–7612. https://doi.org/10.1073/pnas.1221862110

Sui, J., & Zhu, Y. (2005). Five-year-olds can show the self-reference advantage. *International Journal of Behavioural Development*, *29*(5), 382–387. https://doi.org/10.1080/01650250500172673

Symons, C. S., CCoolllege, H., & JJoohhnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394.

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318–324. https://doi.org/10.1016/j.biopsycho.2010.03.009

Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press.

Tajfel, H., & Turner, J. (1979). *An integrative theory of intergroup conflict*. In J. A. Williams, & S. Worchel (Eds.), The social psychology of intergroup relations (pp. 33–47). Belmont, CA: Wadsworth.

Tedeschi, J. T. (Ed.). (1981). *Impression management theory and social psychological research*. Academic Press.

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galván, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proceedings of the National Academy of Sciences*, *111*(18), 6600–6605. https://doi.org/10.1073/pnas.1323014111

Tice, D. M., & Wallace, H. M. (2003). The reflected self: Creating yourself as (you think) others see you. In *Handbook of self and identity* (pp. 91–105). Guilford Press.

Tunnel, G. (1984). The discrepancy between private and public selves: Public self-consciousness and its correlates. *Journal of Personality Assessment*, *48*(5), 549–555. https://doi.org/10.1207/s15327752jpa4805_15

Turner, J. C. (1999). Some current issues in research on social identity and self-categorization theories. *Social Identity: Context, Commitment, Content/Blackwell*.

Turner, J. C., & Reynolds, K. J. (2011). Self-categorization theory. *Handbook of Theories in Social Psychology*, *2*(1), 399–417.

Uddin, L. Q. (2011). The self in autism: An emerging view from neuroimaging. *Neurocase*, *17*(3), 201–208. https://doi.org/10.1080/13554794.2010.509320

Vazire, S., & Carlson, E. N. (2010). Self-knowledge of personality: Do people know themselves? *Social and Personality Psychology Compass*, *4*(8), 605–620. https://doi.org/10.1111/j.1751-9004.2010.00280.x

Vazire, S., & Wilson, T. D. (2012). *Handbook of self-knowledge*. Guilford Press. https://books.google.com/books?hl=zh-CN&lr=&id=nPtJN2Dr8QsC&oi=fnd&pg=PP2&dq=Vazire+%26+Wilson,+2012&ots=Ui5OkBsqdk&sig=Px8x4TubxSaW0mk52McPaGILUHo

Vohs, K. D., & Baumeister, R. F. (2016). *Handbook of self-regulation: Research, theory, and applications* (Third Edition). The Guilford Press.

Vohs, K. D., Schmeichel, B. J., Lohmann, S., Gronau, Q. F., Finley, A. J., Ainsworth, S. E., Alquist, J. L., Baker, M. D., Brizi, A., Bunyi, A., Butschek, G. J., Campbell, C., Capaldi, J., Cau, C., Chambers, H., Chatzisarantis, N. L. D., Christensen, W. J., Clay, S. L., Curtis, J., … Albarracín, D. (2021). A multisite preregistered paradigmatic test of the ego-depletion effect. *Psychological Science*, *32*(10), 1566–1581. https://doi.org/10.1177/0956797621989733

Vonasch, A. J., Reynolds, T., Winegard, B. M., & Baumeister, R. F. (2018). Death before dishonor: Incurring costs to protect moral reputation. *Social Psychological and Personality Science, 9*(5), 604–613. https://doi.org/10.1177/1948550617720271

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194. https://doi.org/10.1016/j.neuropsychologia.2006.06.003

Wallace, H. M., & Tice, D. M. (2012). Reflected appraisal through a 21st-century looking glass. In *Handbook of self and identity* (Second edition, pp. 124–140). Guilford Press.

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioural choice. *Neuroscience & Biobehavioural Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, *64*(4), 678–691.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology*, *77*(1), 97–108. https://doi.org/10.1111/j.2044-8295.1986.tb01985.x

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24. https://doi.org/10.1037/0033-2909.120.1.3

Williams, L. M., Palmer, D., Liddell, B. J., Song, L., & Gordon, E. (2006). The 'when' and 'where' of perceiving signals of threat versus non-threat. *NeuroImage*, *31*(1), 458–467. https://doi.org/10.1016/j.neuroimage.2005.12.009

Wong, A. E., & Vallacher, R. R. (2018). Reciprocal feedback between self-concept and goal pursuit in daily life. *Journal of Personality*, *86*(3), 543–554. https://doi.org/10.1111/jopy.12334

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the Authenticity Scale. *Journal of Counseling Psychology*, *55*(3), 385–399. https://doi.org/10.1037/0022-0167.55.3.385

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 255.

Yang, J., Qi, M., & Guan, L. (2014). Self-esteem modulates the latency of P2 component in implicit self-relevant processing. *Biological Psychology*, *97*, 22–26. https://doi.org/10.1016/j.biopsycho.2014.01.004

Ye, M., Lyu, Y., Sclodnick, B., & Sun, H. J. (2019). The P3 reflects awareness and can be modulated by confidence. *Frontiers in Neuroscience, 13*, 510. https://doi.org/10.3389/fnins.2019.00510

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience, 24*(28), 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Zentall, T. R. (2010). Justification of effort by humans and pigeons: Cognitive dissonance or contrast? *Current Directions in Psychological Science*, *19*(5), 196–300. https://doi.org/10.1177/0963721410383381

Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., Li, H., & Luo, Y. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, *9*(12), 1897–1903. https://doi.org/10.1093/scan/nst188

Zhang, Y., Rong, Y., & Wei, P. (2023). Mothers exhibit higher neural activity in gaining rewards for their children than for themselves. *Social Cognitive and Affective Neuroscience*, *18*(1), nsad048. https://doi.org/10.1093/scan/nsad048

Żochowska, A., Nowicka, M. M., Wójcik, M. J., & Nowicka, A. (2021). Self-face and emotional faces—are they alike?. *Social Cognitive and Affective Neuroscience, 16*(6), 593–607. https://doi.org/10.1093/scan/nsab020

Zou, X., Sedikides, C., & Wildschut, T. (2023). What good is organizational nostalgia in the time of pandemic? Unpacking a pathway from COVID-related stress to authenticity at work. *Self and Identity*, *22*(4), 620–638. https://doi.org/10.1080/15298868.2022.2154258

Chapter 2 Demystifying Authenticity: Behavioural and Neurophysiological Signatures of Self-Positivity for Authentic and Presented Selves

Chengli Huang, Nicholas J. Kelley, and Constantine Sedikides

Department of Psychology, University of Southampton

This is a jointly authored work, and the following is a statement of authors contributions:

Conceptualization: Constantine Sedikides, Nicholas J. Kelley, Chengli Huang

Methodology: Chengli Huang, Douglas J Angus

Investigation: Chengli Huang, James W Butterworth, William E Davis, Alexiss Jeffers

Visualization: Chengli Huang, Douglas J Angus

Funding acquisition: Chengli Huang

Project administration: Nicholas J. Kelley, Constantine Sedikides

Supervision: Nicholas J. Kelley, Constantine Sedikides

Writing – original draft: Chengli Huang

Writing – review & editing: Chengli Huang, Constantine Sedikides, Nicholas J. Kelley, Rebecca Schlegel, Douglas J Angus, William E Davis.

Publication details: Huang, C., Sedikides, C., Angus, D. J., Davis, W. E., Butterworth, J. W., Jeffers, A., Schlegel R., & Kelley, N. J. (2025). Demystifying Authenticity: Behavioural and Neurophysiological Signatures of Self-Positivity for Authentic and Presented Selves. *NeuroImage, 307,* Article 121046. https://doi.org/10.1016/j.neuroimage.2025.121046

Corresponding author: Nicholas J. Kelley, Centre for Research on Self and Identity, School of Psychology, University of Southampton; Southampton, SO17 1BJ, United Kingdom; email: n.j.kelley@soton.ac.uk.

Abstract

Authenticity has captivated scholars. But what is it? An emerging view considers it exaggerated favourability (self-enhancement), whereas traditional views regard it as self-accuracy and self-consistency. We tested these theoretical views by contrasting the authentic self with the presented self, a highly desirable representation. Behaviourally, participants ascribed less positivity to the authentic self: They endorsed more negative traits and were faster to admit having them; also, they endorsed fewer positive traits and were slower to admit having them. Neurally, participants manifested preferential processing of threatening information (P1), followed by preferential processing of favourable information (N170), about the presented self (than authentic self), indicating its brittleness. At a later stage (LPP), participants engaged in more elaborate processing of threatening and favourable information about the authentic self, indicating its subjective importance. Authenticity, albeit mostly positive, allows room for negativity.

Keywords: authenticity, authentic self, presented self, self-reference valence task, neuroscience of authenticity

2.1 Introduction

The concept of authenticity has been gathering traction. Commentators have hailed the rise of the age of authenticity (Wilkinson, 2018), and the term was declared word of the year in 2023 by Merriam-Webster (BBC, 2023). Institutions (e.g., educational centres, mental health and wellness organizations, workplaces) encourage authenticity, as do art (e.g., expressionist art, folks art and cultural crafts, street art and graffiti), fashion, literature, TV shows, movies, sports coaches, song, magazine articles, blogs, and self-help books. Individuals, across ages, walks of life, and cultures, are normatively prescribed to pursue it (Bauer, 2017; Ferrara, 1993; Guignon, 2004).

Despite its seemingly recent appeal, the concept has a long history. It was articulated by Aristotle (384/322 BCE; Tredennick & Thomson, 1976) and pondered by existential philosophers (Golomb, 1995) and sociologists (Erickson, 1995). Intrigued, psychologists have joined in, prioritizing it in their research agendas (Sedikides & Schlegel, 2024; Sutton, 2020). Yet, the nature of authenticity remains elusive (Baumeister, 2019; Hicks et al., 2019).

In this article, we placed the concept under empirical scrutiny. Following other scholars, we define authenticity as the perception of being one’s true self (Kernis & Goldman, 2006). But what is the nature of this perception? Authenticity has been predominantly conceptualized as self-accuracy, self-consistency, and self-enhancement, with the last view gaining evidentiary ground (Sedikides & Schlegel, 2024). In two experiments, we put the self-enhancement view of authenticity to a rigorous test (Platt, 1964). We did so by comparing the authentic self-concept against another highly positive self-concept, the self that is presented to others (i.e., the presented self). If the experience of authenticity is only associated with self-enhancement, we would expect to see this pattern reflected in the content of true self concepts, such that they are just as positive as the presented self-concept. However, if the experience of authenticity is also associated with self-accuracy or self-consistency, we would expect to observe a more mixed valence in true self-concepts compared to the presented self-concept. We implemented both behavioural and neuroscientific techniques. We asked if the authentic self, compared to the presented self, is a fierce denouncer of undesirable information and an unabashed consumer of desirable information (self-enhancement view), or, alternatively, if the authentic self is prone to exploring or accepting the possibility of having some undesirable attributes (self-accuracy and self-consistency views).

2.1.1 Views of Authenticity

One view of authenticity focuses on self-accuracy, the veracious representation or unbiased processing of characteristics and beliefs that comprise one's identity (Kernis & Goldman, 2006). Indeed, people high in authenticity report eagerness, rather than avoidance, to explore identity-relevant information (Kernis & Goldman, 2006), and are less defensive when encounter evidence that their prior behaviour does not faithfully reflect their ideals (Lakey et al., 2008). However, self-accuracy is difficult to attain or empirically verify (Vazire & Wilson, 2012), particularly in the case of the authentic self. It is not clear how the authentic self could be measured directly, and both self and observer reports risk being erroneous. Additionally, individuals who believe they are unbiased in the processing of self-relevant information report that they possess more favourable than unfavourable attributes, thus calling into question how unbiased they are (Gillath et al., 2010).

Authenticity has also been viewed as self-consistency, the alignment of one's behaviour with internal standards, goals, or values (Kernis & Goldman, 2006; Wood et al., 2008). In accord with this view, authenticity is related to self-rated overlap across aspects or roles of one's life (Boucher, 2011), experimentally-induced identity integration across roles increases authenticity (Ebrahimi et al., 2020), and incongruence between one's gender identity (female) and experimentally assigned self-presentation (masculine) decreases authenticity (Dormenan et al., 2020). However, people consider their socially desirable behaviours authentic regardless of whether these behaviours are congruent or incongruent with their self-concept (Sheldon et al., 1997), appraise themselves as authentic when their behaviours align with positive (than negative) behaviours regardless of whether they have traits that underlie these behaviours (Fleeson & Wilt, 2010), and deem enacted desirable (than undesirable) behaviours as more authentic (Jongman-Sereno & Leary, 2016).

Additionally, authenticity has been viewed as self-enhancement, the subjectively exaggerated favourability of one's self-attributes. People regard their true self as positive and moral (Strohminger et al., 2017), endorsing highly positive traits is associated with endorsing authenticity (Bailey & Iyengar, 2023), and daily self-aggrandizement predicts rises in daily authenticity (Guenther et al., 2024). Further, the more favourably people judge a personal change in their lives, the more likely they are to believe the change was guided by authenticity (Bench et al., 2015), and people consider more authentic the times in which they expressed behaviourally a positive (than negative) trait (Bailey & Iyengar, 2023). Lastly, favourable (vs. unfavourable) feedback, and future behavioural positivity (expressing much higher caring, understanding, and kindness than currently held) versus future behavioural negativity (expressing much lower caring, understanding, and kindness than currently held), heighten authenticity, while induced authenticity (thinking of a time in which one felt true to themselves) versus inauthenticity (thinking of a time in which one felt untrue to themselves) heightens self-aggrandizement (Guenther et al., 2024). The link between valence and authenticity is so strong that experimental manipulations of positive affect increase authenticity (Chen et al., 2023; Lenton et al., 2013). In addition, individuals who self-report as being high on authenticity are more prone to *appear* to be authentic. For example, self-proclaimed authentic individuals try to strategically convey authenticity to others, even when such behaviours were inconsistent with their objective experiences (Hart et al., 2020). Taken together, there is enough evidence to suggest that authenticity judgments are a form of self-enhancement, leading some researchers to question whether authenticity has any meaning at all beyond valence (Jongman-Sereno & Leary, 2019).

2.1.2 The Authentic Self and the Presented Self

As stated above, evidence is stronger for the self-enhancement view compared to the self-accuracy and self-consistency views of authenticity. But can the self-enhancement view account for the full conceptual range of authenticity? Is authenticity just positivity or self-enhancement?

We addressed these questions by comparing the authentic self with the self that individuals present to others. The presented self is the benchmark of positive self-presentation[1]. Stakes are high for the presented self as it can facilitate or undermine cooperation, reputation, respect, status, and access to social groups, professional resources (e.g., jobs, promotions, housing), or personal resources (e.g., friends, partners; Dores Cruz et al., 2021; Vonasch et al., 2018). Consequently, self-presentations typically promote a sanitized portrait of the individual, overemphasizing, if not extolling, one’s strengths and underemphasizing, if not concealing, one’s weaknesses (Baumeister, 1982; Roth et al., 1986; see Study S1, Appendix A). Indeed, the

---

[1] Self-presentation can serve various goals beyond favourability, including the projection of both positive and negative attributes (Schlenker, 1980). However, self-enhancement remains a potent motive, particularly in contexts where individuals seek to maintain or enhance their social image. Research has established that individuals are more likely to engage in self-enhancing presentations to be perceived favourably, boost their self-esteem, achieve social approval, and make the best possible impression (Leary, 2007; Paulhus et al., 2003; Schlenker & Leary, 1982; Sedikides & Gregg, 2008). Furthermore, individuals who self-derogate are enhanced by others (e.g., increased numbers of “likes” and comments from their network friends; Bareket-Bojmel et al., 2016; Lee et al., 2014). Therefore, the current study regards the presented self as the “benchmark of positive self-presentation.”

words people select to describe their true self are less socially desirable than the words they select to describe their presented self (Schlegel et al., 2009).

The presented self is not an ephemerality. Instead, it is internalized as part of the private self. Theory and empirical findings bolster this assertion. According to symbolic interactionism and role theory (Stryker & Statham, 1985), people construct their sense of self though social interactions, and in particular the behaviours they enact or roles they play as well as others' reactions to these behaviours or roles. Research findings concur. Strategic self-presentations influence subsequent private self-views; that is, people shift both their overall evaluations of themselves and their evaluations of specific characteristics of themselves in the direction of their preceding self-presentations (Leary, 1995). Also, changes in self-evaluations that occurred in one context because of self-presentations carry over to a new context in the absence of self-presentational pressures (Schlenker, 2003). Taken together, the presented self constitutes a mental representation, just like the authentic self. To clarify, we do not argue that the presented self is inauthentic, and we do not contrast the authentic with the presented self. Indeed, a given trait can be endorsed as part of both selves. Rather, we examine whether the authentic self is inherently positive by comparing it to the benchmark of favourability, the presented self.

2.1.3 A Combination of Behavioural with Event Related Potential Assessment to Examine Authenticity

We collected behavioural data (Experiments 1–2) by means of the self-reference valence (SR-valence) task. This is a variant of the self-reference task, which indicates improved memory and faster reaction times for trait adjectives that are accompanied by self-referential instructions ("does the word describes you?") relative to control, including other-referential, instructions (Northoff & Bermpohl, 2004). In the SR-valence task, participants judge whether positive versus negative traits are self-descriptive or non-self-descriptive (D'Argembeau et al., 2005). The task allows assessing the endorsement of positive versus negative traits (trait endorsement), and the speed of this endorsement (reaction time). Reaction time is used as a proxy for cognitive processing speed (Jensen, 2006). In the context of self-reference tasks, it can reflect the cognitive accessibility of self-concept information (Schlegel et al., 2009); indeed, faster reaction times are indicative of stronger, more accessible associations with the self-concept (Cai et al., 2016; Rameson et al., 2010). In the SR-valence task, higher endorsement (i.e., judging more traits as self-descriptive), or faster reaction time thereof, of positive than negative traits (i.e., Valence × Endorsement interaction) is a signature of self-positivity (Cai et al., 2016).

We also collected neuropsychological data (Experiment 2) to examine the extent to which neurocognitive processes tracked behavioural performance on the SR-valence task. Although research directly exploring the neural underpinnings of authenticity is scant (Sedikides & Schlegel, 2024), there is a growing body of literature examining related constructs, such as self-referential processing. This research often focuses on how the brain processes emotionally salient stimuli, including emotional word tasks, offering insights into mechanisms that may overlap with the experience of authenticity.

Prior event related potential (ERP) studies have identified distinct stages of emotional word processing: the P1, which differentiates between non-threatening and threatening information; the N170 and early posterior negativity, which reflect emotional and non-emotional discrimination; and the late positive potential (LPP), which distinguishes between positive and negative words (Zhang et al., 2014). Similar stages of emotional processing have also been observed in facial recognition studies (Luo et al., 2010). These three stages of emotional processing provide a useful framework for understanding self-reference responses to stimuli of varying emotional valence.

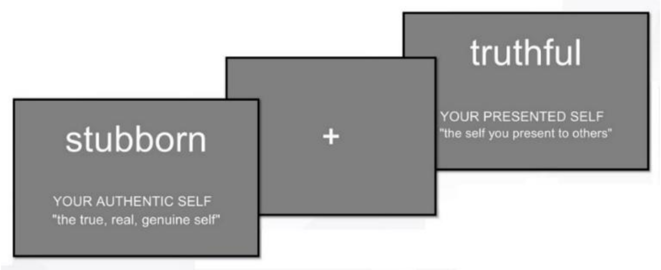
Recently, a stream of EEG literature has addressed self-reference processing in emotional word contexts via the SR-valence task. In one study, negative traits elicited larger N170 responses in East-Asian (but not Western) participants, and self-descriptive traits, particularly negative ones, produced larger LPP responses compared to non-self-descriptive traits (Cai et al., 2016). In another study, P1 and LPP effectively captured biased self-reference processing in female adolescents with depression (Auerbach et al., 2015). Specifically, depressed participants (vs. non-depressed controls) exhibited greater P1 amplitudes following negative words. Non-depressed controls showed greater LPP activity following positive (vs. negative) words, whereas depressed participants demonstrated the opposite pattern. Further, in yet another study, emotional content rapidly captured attention (reflected in augmented early posterior negativity for unpleasant and pleasant nouns vs. neutral ones), followed by higher-order self-referential processing (manifested as augmented LPPs for unpleasant nouns only when preceded by personal pronouns; Herbert et al., 2011). However, self-referential processing may occur earlier than emotional processing, with self-other discrimination emerging as early as the P1, and interactions between self-reference and emotional valence appearing later, manifested in the LPP (Zhou et al., 2017). Despite variations in prioritizing self-referential versus emotional processing, this literature indicates that self-referential processing in emotional contexts operates through multiple stages, and it is possible to identify distinct markers of it at different stages.

Informed by these findings, we considered three ERP components as covert measures of attention allocated independently of behavioural responses: P1, N170, and LPP. We offer a detailed description of them in the introduction to Experiment 2.

2.1.4 Pitting the Authentic Self Against the Presented Self via Self-Positivity

We subjected the favourability of the authentic self to a litmus test, comparing it to the presented self. Specifically, we examined the relative strength of self-positivity for the authentic and presented selves. We offered two competing hypotheses (Platt, 1964). To test them, participants responded to a series of positive and negative traits, indicating whether each trait described their authentic and presented self while reaction time was being recorded (Figure 2.1). First, in line with the self-enhancement view, we hypothesized that the strength of self-positivity would be comparable for the authentic and presented selves. Self-enhancement is thought to operate broadly, manifesting across self-representations (Sedikides, 2020, 2021; Sedikides & Gregg, 2008). This view anticipates an interaction between valence (positive vs. negative traits) and endorsement (self-descriptiveness vs. non-self-descriptiveness) that remains independent of self (authentic vs. presented). Second, in line with the self-accuracy and self-consistency views, we hypothesized that the strength of self-positivity would be weaker for the authentic compared to the presented self. These views highlight the importance of recognizing both the genuinely positive and genuinely negative aspects of oneself, as doing so contributes to greater accuracy or self-consistency (Kernis & Goldman, 2006; Lakey et al., 2008; Wood et al., 2008). However, this recognition may not extend to the presented self, where accuracy and consistency are not directly relevant. Consequently, these views anticipate an interaction involving valence, endorsement, and self.

Figure 2.1 *The Trial Event Diagram*



Note. For each trial, participants made a binary judgement (yes vs. no) as to whether a trait described their authentic self and presented self. We randomized, separately for each participant, the order of traits and blocks of traits referring to the authentic self or presented self. Each trait was displayed on the computer screen until a response (key-pressing) occurred but no longer than 6 seconds, or the screen would automatically switch to the next trial. We randomized interstimulus intervals between 800 and 1200 ms, during which we presented a central fixation.

2.1.5 Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and we follow Journal Article Reporting Standards (Appelbaum et al., 2018). All data, analysis code, and research materials will be made available on request. We analysed the data using Jamovi, version 2.3.21 (Şahin & Aybek, 2020), R, version 4.3.1 (R Core Team, 2023) and the package *ggplot*, version 3.4.3 (Wickham & Wickham, 2016). We addressed the issue of multiple comparisons using Bonferroni corrections[2]. Neither experiment was preregistered.

2.2 Experiment 1

In Experiment 1, we tested behaviourally the strength of self-positivity for the authentic self versus the presented self. We implemented a 2 (self: authentic self vs. presented self) × 2 (valence: positive traits vs. negative traits) × 2 (endorsement: self-descriptiveness vs. non-self-descriptiveness) within-subjects design.

2.2.1 Method

Participants and Design

We focused our power analysis on the hypothesis derived from the self-accuracy and self-consistency views because they require a significant three-way interaction, whereas the self-enhancement view only requires a significant Valence × Endorsement interaction. We used *Superpower* (Lakens & Caldwell, 2021) to conduct a simulation-based power analysis. We carried out 2,000 Monte Carlo simulations, assuming a correlation among within-subject factors of .5 and a common standard deviation of 1.00. We sought to have sufficient power to detect small-to-moderate (d = .20) reductions in self-positivity for the authentic (vs. presented) self. Based on these parameters, 50 participants were needed to detect a significant three-way interaction with 80% power. We considered this our minimum sample size and proceeded to recruit 339 Wittenberg University introductory psychology students (from the corresponding

---

[2] The corrected p-value is calculated by multiplying the original p-value by the number of comparisons. For instance, in post-hoc tests of a 2×2 interaction effect, the adjusted p-value becomes 6 times the original value.

participant pool) throughout the semester in exchange for course credit. We excluded six participants for the following reasons (see data processing pipeline under "Data Recording and Data Analysis"): (a) two did not complete the whole task, (b) two evinced more than 50% missing data after we removed "no response" trials (i.e., longer than 6 seconds), "impossibly fast" trials (i.e., less than 200ms)[3], as well as the 1% slowest and 1% fastest trials, and (c) two manifested mean reaction time that exceeded ± 3 standard deviations (Morís Fernández & Vadillo, 2020). All procedures used in the current experiment were approved by the Wittenberg University ethics committee (No. IRB2021-1268M).

We used a multilevel model (MLM)/hierarchical linear model (HLM) to analyse reaction times. This model requires a minimal threshold of five observations when testing fixed effects (McNeish & Stapleton, 2016); here, these effects pertained to self, valence, and endorsement. Thus, we excluded an additional 111 participants, because they engaged in fewer than five trials in at least one condition; for example, we excluded participants who only endorsed two negative traits as their presented self (participants excluded per condition: negative descriptive traits for the presented self, *n* = 83; negative descriptive traits for the authentic self, *n* = 29; positive non-descriptive traits for the presented self, *n* = 22; positive non-descriptive traits for the authentic self, *n* = 20)[4]. The final sample consisted of 222 participants (131 women, 89 men, 2 unknow) ranging in age from 18 to 22 years (*M* = 18.68, *SD* = 0.92). Of them, 124 identified as White, 54 as Latinx, 19 as Asian, 14 as mixed race, and 9 as Black (two did not indicate their ethnicity).

Stimuli and Procedure

Anderson (1968) introduced a list of 555 personality traits rated for likableness and meaningfulness. Chandler (2018) tested the replicability of Anderson's list by asking participants to rate each trait's likableness (0 = *least favourable or desirable*, 6 = *most favourable or desirable*). The resulting ratings were highly correlated with Anderson's ratings (*r* = 0.96, *p* < 0.001). We selected 85 positive traits and 85 negative traits from Chandler's list[5]. The

---

[3] Two participants evinced more than 50% missing data after we removed the 1% slowest and 1% fastest trials, rather than the "no response" trials (i.e., longer than 6 seconds) and "impossibly fast" trials (i.e., less than 200ms). To ensure consistency between the "Participants and Design" (pp. 9-10) and "Data Recording and Data Analysis" (p. 11) section, we included references to "no response" trials and "impossibly fast" trials under "Participants and Design."

[4] Some participants had fewer than 5 trials in more than one condition; for example, one participant could judge fewer than 5 negative traits as self-descriptive of the presented self, and the same participant could also judge fewer than 5 negative traits as self-descriptive of the authentic self; hence the total number of participants is greater than 111.

[5] These traits are listed in both Anderson's and Chandler's lists, and represent extremes in terms of likability ratings—either occupying the lower end (i.e., negative traits) or the upper end (i.e., positive traits) of the spectrum.

likableness of the selected positive traits ($M = 4.71$, $SD = 0.53$) was much higher than the likableness of the selected negative traits ($M = 1.22$, $SD = 0.42$), $t(168) = 47.99$, $p < 0.001$, Cohen's $d = 7.36$.

Participants completed the SR-valence task in a quiet laboratory room via computer. They were shown a list of positive and negative traits, and judged whether each trait was self-descriptive or non-self-descriptive (D'Argembeau et al., 2005). Participants evaluated each trait twice: once for the authentic self and once for the presented self. We defined the authentic self as "the true, real, genuine self," and the presented self as "the self you present to others" (see Figure 2.1 for the trial event diagram). We programmed and administered the experiment using jsPsych (Version 6.3; de Leeuw, 2015).

Data Recording and Data Analysis

The main dependent variables were trait endorsement (judgments of positive vs. negative traits as self-descriptive or non-self-descriptive) and reaction time (RT; speed of trait endorsement). We created a reaction time data processing pipeline based on Morís Fernández and Vadillo's (2020) suggestions. First, we excluded "no response" trial (i.e., longer than 6 seconds) or "impossibly fast" trials (i.e., less than 200ms). Second, we removed the 1% slowest and 1% fastest trials. Third, we removed participants with more than 50% missing data. Finally, we computed the mean reaction time. We did not log transform the RT data, because they were normally distributed (Skew and Kurtosis < ± 2 for each trial type in each study; Byrne, 2013).

2.2.2 Results

Trait Endorsement

We entered the number of trait endorsements (i.e., self-descriptive vs. non-self-descriptive) into a three-way Analysis of Variance (ANOVA) with self (authentic self vs. presented self), valence (positive trait vs. negative trait), and endorsement (self-descriptive vs. non-self-descriptive) as within-subjects factors.

The Valence × Endorsement interaction was significant, $F(1, 221) = 3478.21$, $p < 0.001$, $\eta_p^2 = 0.94$. Participants endorsed more positive traits (66.09 ± 8.31) than negative traits (15.34 ± 7.57) as self-descriptive, $t(221) = 58.45$, $p < 0.001$, 95% $CI = [49.05, 52.47]$, *Cohen's d* = 3.92, but judged more negative traits (66.79 ± 8.65) than positive traits (15.78 ± 6.95) as non-self-descriptive, $t(221) = 59.22$, $p < 0.001$, 95% $CI = [49.31, 52.70]$, *Cohen's d* = 3.97. This pattern replicates self-positivity (Cai et al., 2015; Shi et al., 2017). Moreover, self-positivity was evident for both the authentic self and the presented self (Appendix A).

Crucially, the Self × Valence × Endorsement interaction was significant, $F(1, 221) = 47.85$, $p < 0.001$, $\eta_p^2 = 0.18$. We examined the Self × Endorsement interaction separately for positive traits and negative traits, testing whether self-positivity was stronger for one kind of self versus another (Figure 2.2a). In the case of positive traits, participants showed no significant difference when endorsing positive traits as descriptive of the presented self (66.71 ± 8.76) versus the authentic self (65.48 ± 9.74), $t(221) = 2.23$, $p = .746$, 95% *CI* = [-0.14, 2.31], *Cohen's d* = 0.15, and judged positive traits as equally non-descriptive of the authentic self (16.16 ± 8.04) and the presented self (15.40 ± 7.81), $t(221) = 1.49$, $p = 1.000$, 95% *CI* = [-0.24, 1.77], *Cohen's d* = 0.10. However, in the case of negative traits, participants endorsed more negative traits as descriptive of the authentic self (18.09 ± 9.63) than the presented self (12.58 ± 8.06), $t(221) = 8.83$, $p < .001$, 95% *CI* = [4.28, 6.74], *Cohen's d* = 0.59, and judged more negative traits as non-descriptive of the presented self (69.79 ± 8.86) than the authentic self (63.78 ± 10.83), $t(221) = 9.33$, $p < .001$, 95% *CI* = [4.74, 7.28], *Cohen's d* = 0.63. The authentic self evinced weaker self-positivity than the presented self.

Finally, the Self × Endorsement interaction was significant, $F(1, 221) = 65.18$, $p < 0.001$, $\eta_p^2 = 0.23$. Participants endorsed more traits as descriptive of the authentic self (41.79 ± 5.61) than the presented self (39.64 ± 4.74), $t(221) = 6.72$, $p < 0.001$, 95% *CI* = [1.52, 2.77], *Cohen's d* = 0.45, but judged more traits as non-descriptive of the presented self (42.60 ± 4.62) than the authentic self (39.97 ± 5.51), $t(221) = 8.41$, $p < 0.001$, 95% *CI* = [2.01, 3.24], *Cohen's d* = 0.56. The authentic self appeared to be more inclusive than the presented self.

Reaction Times

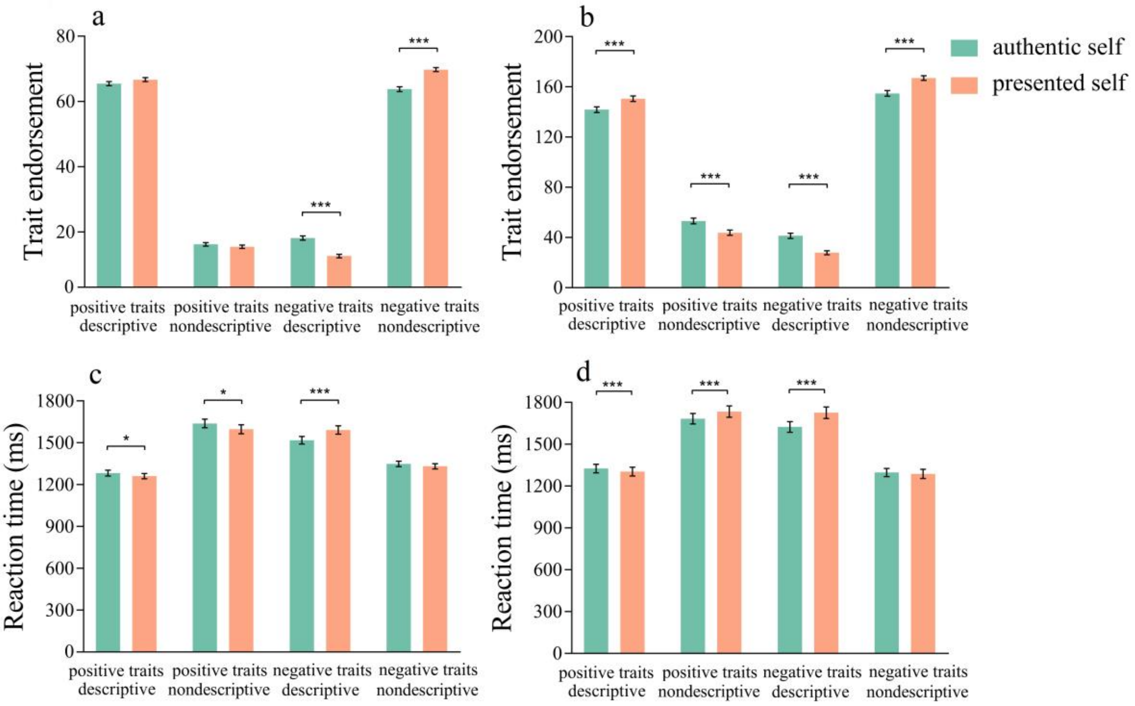
We used MLM to analyse reaction time (RT) and employed the R package 'lme4' (Bates et al., 2015) to fit it (for the model settings see Appendix A).

The Valence × Endorsement interaction was significant, $B = -126.53$, $t_{72690} = -35.40$, $p < 0.001$, $\beta = -0.16$, 95% CI [-0.17, -0.15]. We followed up with simple slope analyses in MLM (Curran et al., 2015). Participants were faster to endorse positive traits than negative traits as self-descriptive ($\gamma = -158.28$, $z = -31.88$, $p < 0.001$), but were faster to reject (i.e., non-endorse) negative traits than positive traits as self-descriptive ($\gamma = 94.78$, $z = 18.60$, $p < 0.001$). This pattern of results replicates self-positivity (Cai et al., 2016; Shi et al., 2017). Moreover, self-positivity was evident for both the authentic self and the presented self (Appendix A).

Crucially, the Self × Valence × Endorsement interaction was significant on RT, $B = 8.76$, $t_{72627} = 2.48$, $p = 0.013$, $\beta = 0.01$, 95% CI [0.002, 0.02]. We examined the Self × Endorsement interaction separately for positive traits and negative traits, testing whether self-positivity was stronger for one kind of self versus another (Figure 2.2c). In the case of positive traits, participants endorsed positive traits faster for the presented self than the authentic self (γ =

9.85, $z = 2.29$, $p = 0.022$), and rejected positive traits faster for the presented self than the authentic self ($\gamma = -20.73$, $z = -2.35$, $p = 0.019$). In the case of negative traits, participants endorsed negative traits faster for the authentic self than the presented self ($\gamma = -38.07$, $z = -4.17$, $p < 0.001$), and did not differ in their rejection of negative traits for the two selves ($\gamma = 7.85$, $z = 1.83$, $p = 0.067$). We display in Table S2.1 detailed results of the fixed effects of the MLM.

Figure 2.2 *Behavioural Manifestations of Self-Positivity for the Authentic Self and Presented Self in Experiments 1 and 2*



Note. (**a**) Endorsement of self-descriptive traits and non-self-descriptive traits in Experiment 1. (**b**) Endorsement of self-descriptive traits and non-self-descriptive traits in Experiment 2 (see Table S2.3 and Table S2.4 for fixed effects of self, valence, endorse, and their interactions on endorsement). (**c**) Reaction time of self-descriptive traits and non-self-descriptive traits in Experiment 1 (see Table S2.1 for fixed effects of self, valence, endorse, and their interactions on reaction time). (**d**) Reaction time of self-descriptive traits and non-self-descriptive traits in Experiment 2 (see Table S2.5 and Table S2.6 for fixed effects of self, valence, endorse, and their interactions on reaction time). Error bars represent SEM; $^{***}p < .001$. $^{**}p < .01$. $^{*}p < .05$.

2.2.3 Discussion

Participants endorsed an equivalent number of positive traits as descriptive of the authentic and presented self, while judging an equivalent number of such traits as nondescriptive of the two selves. However, participants judged more negative traits as descriptive of the authentic than presented self and judged more such traits as nondescriptive of the presented than authentic self. Further, they were faster to endorse positive traits for the presented than authentic self and were faster to endorse negative traits for the authentic than presented self. Overall, self-positivity was weaker for the authentic than presented self, in line with the self-accuracy and self-consistency views. Next, we sought to re-test these behavioural findings and explore pertinent neuropsychological underpinnings.

2.3 Experiment 2

In Experiment 2, we tested the strength of self-positivity for the authentic versus presented self not only behaviourally (with an identical procedure to Experiment 1's), but also neuropsychologically. We examined neural mechanisms via EEG in a 2 (self: authentic self vs. presented self) × 2 (valence: positive traits vs. negative traits) × 2 (endorsement: self-descriptiveness vs. non-self-descriptiveness) within-subjects design. Specifically, we considered P1, N170, and LPP as covert measures of attention allocated independently of behavioural responses.

The P1 is the initial positive deflection in the parieto-occipital region, and typically emerges 60–90 ms post-stimulus with a peak between 100 and 130 ms (Luck, 2014). Originating in the visual cortex, the P1 has conventionally been conceptualized as an early sensory-evoked component in relation to sensory amplification and selective attention (Hillyard et al., 1998). Multiple studies have indicated that the P1 is sensitive to emotional stimuli (for reviews, see: Mueller et al., 2013; Schindler & Bublatzky, 2020). Moreover, some studies observed larger P1 amplitudes evoked by negative stimuli (e.g., faces, words) compared to neutral counterparts (Luo et al., 2010; Zhang et al., 2014), indicating that the P1 can differentiate between non-threatening and potentially threatening information (Zhang et al., 2014). Such early emotional response may signify rapid extraction of emotion-related information and may function—at least partly— independent of subsequent, more detailed emotional processes such as N170 (Vuilleumier & Pourtois, 2007). Relevant to Experiment 2, in a SR-valence task, depressed female adolescents (compared to healthy counterparts), who generally maintain a negative self-view (Auerbach et al., 2014), displayed heightened P1 amplitudes in response to negative words, but not to positive words (Auerbach et al., 2015). Building on this finding and initial sensory encoding function of P1 (Brem et al., 2006), we proposed two competing hypotheses regarding early perceptual processing of self-relevant information. Aligning with the self-enhancement view—positivity is reassuring, whereas negativity is threatening, to the self—we hypothesized that P1 would show enhanced sensory gain for negative versus positive self-descriptive traits and would be comparable across both selves. This pattern would be reflected in a Valence × Endorsement interaction, indicating stimulus-driven amplification and heightened sensitivity to negative self-relevant information at the initial sensory stage for both the authentic and presented selves. Alternatively, aligning with the self-accuracy and self-consistency views—where positivity is reassuring to self, and negativity is not threatening to the authentic self—we hypothesized that P1 responses would still be larger for negative versus positive self-descriptive traits, but that this effect would be attenuated for the authentic self compared to the presented self. This pattern would be reflected in a Self × Valence × Endorsement interaction indicating reduced sensitivity to negative self-relevant information specifically for the authentic self, suggesting early sensory modulation by self representation.

The N170 is a negative deflection that typically peaks at approximately 170 ms after stimulus onset over the lateral occipito-temporal regions, especially over the right hemisphere (Luck, 2014). The N170 is known for being face-sensitive: it manipulates a larger peaking in response to face-elicited stimuli than non-face-elicited stimuli (Rossion & Jacques, 2012). In addition, the N170 can be modulated by the valence of the facial expression, with a significantly augmented negativity for emotional relative to neutral facial expression (Luo et al., 2010). Such modulation has also been found in emotional word processing (Zhang et al., 2014). Moreover, negative adjectives elicit larger N170 amplitudes than positive adjectives (Montalan et al., 2008). The above findings indicate that the N170 can reflect early attention to stimuli with emotional valence, especially for negative emotional stimuli, with a larger N170 amplitude representing the allocation of more attentional resources (Cai et al., 2016). Building on these findings and relative higher-order feature discrimination function of N170 (Brem et al., 2006), we offered two competing hypotheses regarding subsequent processing of self-relevant information, similar to those for P1. First, aligning with the self-enhancement view—positivity is reassuring to the self, but negativity is threatening to the self—we hypothesized that N170 responses would be larger for negative versus positive self-descriptive traits, and comparable across both selves. This pattern would be reflected in a Valence × Endorsement interaction, evincing earlier attentional discrimination to negative self-relevant information for both selves. Second, aligning with the self-accuracy and self-consistency views—positivity is reassuring to the self, and negativity is not threatening to the authentic self—we hypothesized that N170 responses would be larger for negative versus positive self-descriptive traits, but that this effect would be attenuated for the authentic compared to the presented self. This pattern would be reflected in a Self × Valence × Endorsement interactions, evincing reduced earlier attention to negative self-relevant information, in particular for the authentic self, suggesting higher-order feature modulation by self representation.

The LPP manifests as a sustained positive deflection typically observed approximately 400–500 ms post-stimulus presentation, persisting for several hundred milliseconds at the midline centroparietal region (Hajcak et al., 2012). Notably, the LPP exhibits an augmented amplitude in response to emotionally arousing stimuli when compared to neutral stimuli, spanning various modalities such as pictures, faces, hand gestures, and words (Hajcak & Foti, 2020). Moreover, the LPP is sensitive to self-referent information, exhibiting greater amplitudes for self-referent relative to non-self-referent content (Hudson et al., 2020; Jordan, et al., 2022; Żochowska et al., 2021). In addition, the LPP can capture both emotional and evaluative processing with respect to the self. For instance, some studies have found augmented LPP responses to negative (vs. positive) stimuli when participants refer to themselves (Cai et al., 2016; Herbert et al., 2011), whereas other studies report the reverse, with greater LPP responses following positive versus negative words during the self-reference task (Auerbach et al., 2015; Shestyuk & Deldin, 2010). Although findings remain mixed, the amplified LPP in self-referential tasks may reflect deeper processing of self-relevant information, in line with the LPP's broader role in sustained attention and elaborative processing (Auerbach et al., 2015; Hajcak et al., 2012), as well as in signalling stimulus significance and motivational relevance (i.e., activation of appetitive or aversive motivational systems; Hajcak & Foti, 2020). These variations in LPP response may imply underlying factors, such as differences in self-representation (e.g., presented vs. authentic self), that influence how self-relevant information is processed.

Building on these findings, we offered two competing hypotheses. First, congruent with the self-enhancement view—positivity is reassuring, but negativity is threatening, to self—we hypothesized that LPP responses would be larger for positive versus negative self-descriptive traits and would be comparable across the two selves. This pattern would be reflected in a Valence × Endorsement interaction, manifesting more elaborative processing and stimulus significance of positivity for both the authentic and presented selves. Alternatively, congruent with the self-accuracy and self-consistency views—positivity is reassuring to self, while negativity is not threatening to the authentic self—we hypothesized that LPP responses would be larger for positive versus negative self-descriptive traits, but that this effect would be weaker for the authentic versus presented self. This pattern would be reflected in a Self × Valence × Endorsement interaction, manifesting elaborative processing and stimulus significance of positivity, specifically weaker for the authentic self.

2.3.1 Method

Participants and Design

Based on the power analysis from Experiment 1, we sought to test at least 50 participants. We recruited, until the end of the academic year, 157 University of Southampton introductory psychology students (from the participant pool) for course credit. We excluded seven participants for the following reasons: one did not complete the whole task, five encountered equipment failures (e.g., keyboard, EEG acquisition equipment; Tacikowski & Nowicka, 2010), one manifested mean reaction time exceeded ± 3 *SD*s (Cai et al., 2016). Additionally, we excluded 29 participants, because they failed to meet the requirement for ERP analysis due to insufficient (< 5) EEG trials[6] (participants excluded per condition: negative descriptive traits for the presented self, *n* = 24; negative descriptive traits for the authentic self, *n* = 9; positive non-descriptive traits for the presented self, *n* = 6). The final sample comprised 121 participants (97 women, 24 men) aged between 18 and 46 years (*M* = 19.83, *SD* = 3.45). We did not collect ethnicity information, but we note that over 90% of the sponsoring University's undergraduates are White. All procedures used in the current experiment were approved by the University of Southampton ethics committee (No. 67233).

Stimuli and Procedure

The stimulus materials were 200 positive traits and 200 negative traits from Anderson's personality list. We increased the number of traits due to requirements of EEG experiments. Based on Chandler's (2018) ratings, the likableness of the selected positive traits (*M* = 4.74, *SD* = 0.50) was higher than that of the selected negative traits (*M* = 1.33, *SD* = 0.48), *t*(398) = 67.88, *p* < 0.001, Cohen's *d* = 6.79. We programmed the experiment using PsychoPy (Version 2021.2.3; Peirce, 2007).

Data Recording and Data Analysis

We collected the EEG data continuously from 64 scalp sites using Ag/AgCl electrodes mounted in an elastic cap (Neuroscan, NC), with an online reference to the left mastoid and off-line algebraic re-reference to the average of left and right mastoids. We mounted a ground electrode, i.e., AFz. We recorded the vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) from two pairs of electrodes, with one placed above and below the left eye, and another placed 10 mm from the outer canthi of each eye. We based the electrode cap on the 10–20 system. We kept electrode impedances below 5 kΩ. We amplified and sampled the signals at 1000 Hz with an online bandpass filter from 0.10–100 Hz.

---

[6] Researchers have provided guidelines for the number of trials for ERP experiments, with 20 trials suggested for the P300 (Cahn & Polich, 2006) and 8-12 trials for LPP (Moran et al., 2013). However, apart from a specific number of trials, researchers must also consider factors that may influence the ability to obtain a "stable" ERP waveform, such as sample size, anticipated effect magnitude, and noise level (Boudewyn et al., 2018). Moreover, the representativeness of the sample can vary substantially based on number of trials. This was a crucial consideration in the current study. Specifically, with 5 trials per condition, 121 participants remain, out of the original 150 (80.67%). With 8 trials per condition, 107 participants remain (71.33%), and, with 20 trials per condition, only 61 participants remain (40.67%). Here, a greater number of trials excluded corresponds to reduced sample representativeness. As such, participants likely to evince strong self-positivity may be excluded due to insufficient trials, especially the ones in the presented self, negative traits, self-descriptiveness condition. Consequently, we opted for 5 trials per condition to maximize participant inclusion. Importantly, the results were comparable across 121, 107, and 61 participants.

In offline processing, we initially pre-processed the EEG data by using EEGLAB, an open-source toolbox running in the MATLAB environment (Delorme & 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), segmented them from 200 ms prior to 1000 ms following the onset of each word, and baseline corrected them to the -200–0 ms. We identified bad channels by visual inspection of the waveforms and replaced them by using a spherical spline identified interpolation (SSI; Perrin et al., 1989). We corrected segments contaminated by blinks, eye movements, and other artifacts using an independent component analysis (ICA) algorithm (Delorme & Makeig, 2004) and ICLabel, a proposed statistical model, to automatically label ICA components (Pion-Tonachini et al., 2019). We also excluded bad segments where a voltage deviation on any channel of ± 75 μV.

Then, we averaged the ERPs for each of the eight conditions (2 [self: authentic self vs. presented self] × 2 [valence: positive traits vs. negative traits] × 2 [endorsement: self-descriptiveness vs. non-self-descriptiveness]). We excluded data from trials where a participant had not responded (reaction time > 6 seconds) or provided an improper response (in less than 200 ms). There was an average of 739.51 trials per participant. We display information on the number of retained EEG trials per condition in Appendix A (Table S2.2).

Our ERPs of interest were quantified following best practices (Luck & Gaspelin, 2017). For each ERP, we employed a collapsed localizer approach, in which a grand average of all conditions is created and used to identify where each component is spatially and temporally maximal. For the P1, we measured the mean amplitude between 90 ms and 130 ms over 9 parieto-occipital sites: P3, P4, Pz, PO3, PO4, POZ, O1, O2, and OZ. For the N170, we measured the mean amplitude between 120 ms and 200 ms over 16 temporal-parieto-occipital sites: TP7, TP8, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, O1, and O2. Finally, for the LPP we measured the mean amplitude between 350 ms and 800 ms over 15 frontal-central-parietal sites: F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4, CP3 CPZ, CP4, P3, P4, and Pz. These measurement locations and time windows are consistent with previous literature on P1 (e.g., Luo et al., 2010), N170 (e.g., Hinojosa et al., 2015), and LPP (e.g., Webber et al., 2022). The main dependent variables were trait endorsement, reaction time (RT), and ERPs (P1, N170, LPP). For the RT, we adopted the same preprocessing steps as in Experiment 1 to reduce the false-positive rate.

2.3.2 Results

Trait Endorsement

We entered the number of trait endorsements into a three-way ANOVA. The Valence × Endorsement interaction was significant, $F(1, 120) = 1172.22$, $p < .001$, $\eta_p^2 = .91$. Participants endorsed more positive traits (146.22 ± 22.55) than negative traits (34.62 ± 18.61) as self-descriptive, $t(120) = 33.77$, $p < 0.001$, 95% $CI = [105.03, 118.15]$, *Cohen's d* = 3.07, but judged more negative traits (160.95 ± 20.07) than positive traits (48.48 ± 21.48) as non-self-descriptive, $t(120) = 34.53$, $p < 0.001$, 95% $CI = [106.01, 118.91]$, *Cohen's d* = 3.14. This pattern replicates self-positivity (Cai et al., 2015; Shi et al., 2017). Moreover, self-positivity was evident for both the authentic self and the presented self (Appendix A). We provide in Table 2.1 the 30 most commonly endorsed positive and negative traits, and we display in Figure 2.3 positive and negative self-portraits based on trait frequency.

Crucially, the Self × Valence × Endorsement interaction was significant, $F(1, 120) = 57.50$, $p < .001$, $\eta_p^2 = 0.32$. We examined the Self × Endorsement interaction separately for positive traits and negative traits, testing whether self-positivity was stronger for one kind of self versus another (Figure 2.2b). In the case of positive traits, participants endorsed more such traits as descriptive of the presented self (150.61 ± 24.03) than the authentic self (141.83 ± 25.01), $t(120) = 5.02$, $p < 0.001$, 95% $CI = [5.31, 12.24]$, *Cohen's d* = 0.46, but judged more such traits as non-descriptive of the authentic self (53.14 ± 24.72) than the presented self (43.83 ± 22.31), $t(120) = 5.31$, $p < 0.001$, 95% $CI = [5.84, 12.79]$, *Cohen's d* = 0.48. In the case of negative traits, participants endorsed more such traits as descriptive of the authentic self (41.35 ± 23.02) than the presented self (27.88 ± 17.66), $t(120) = 8.58$, $p < 0.001$, 95% $CI = [10.35, 16.57]$, *Cohen's d* = 0.78, but judged more such traits as non-descriptive of the presented self (167.08 ± 19.27) than the authentic self (154.81 ± 24.90), $t(120) = 7.01$, $p < 0.001$, 95% $CI = [8.81, 15.74]$, *Cohen's d* = 0.64. Again, as in Experiment 1, the authentic self manifested weaker self-positivity than the presented self: Participants endorsed fewer positive traits for the authentic self than the presented self and endorsed more negative traits for the authentic self than the presented self.

Lastly, the Self × Endorsement interaction was significant, $F(1, 120) = 5.24$, $p = 0.024$, $\eta_p^2 = 0.04$. Participants endorsed more traits as descriptive of the authentic self (91.59 ± 10.76) than the presented self (89.25 ± 11.08), $t(120) = 2.73$, $p = 0.043$, 95% $CI = [0.65, 4.04]$, *Cohen's d* = 0.25, but judged an equivalent number of traits as non-descriptive of the presented self (105.45 ± 11.68) and the authentic self (103.98 ± 11.73), $t(120) = 1.60$, $p = 0.668$, 95% $CI = [-0.35, 3.31]$, *Cohen's d* = 0.15. As in Experiment 1, the authentic self was more inclusive than the presented self. The above results were comparable to those we obtained for *N*s of 107 and 61 (Appendix A, Tables S2.3 and S2.4).

Table 2.1 *The Most Common Endorsed Positive and Negative Traits in Experiment 2*

Positive traits	Count	Negative traits	Count
loyal	233	nervous	159
respectful	232	clumsy	152
good-natured	232	insecure	148
kind	232	headstrong	144
friendly	231	stubborn	142
polite	230	nosey	142
nice	230	oversensitive	141
decent	229	overcritical	140
helpful	229	gossipy	127
well-mannered	229	lazy	119
considerate	228	childish	119
reliable	228	jumpy	118
moral	227	messy	118
likable	227	moody	111
grateful	227	noisy	109
understanding	227	complaining	103
pleasant	227	fault-finding	101
trustworthy	227	frustrated	99
kind-hearted	227	touchy	98
kindly	226	irritable	98
reasonable	226	superstitious	97
trustful	226	jealous	95
thoughtful	226	bossy	94
appreciative	226	untidy	93
sympathetic	226	ultra-critical	89

Positive traits	Count	Negative traits	Count
warm-hearted	225	mediocre	86
open-minded	225	unhealthy	83
educated	224	petty	82
able	223	unproductive	81
good	223	loud-mouthed	80

Figure 2.3 *Positive and Negative Self-Portraits Based on Trait Endorsement in Experiment 2*

Positive Self-Portrait

Negative Self-Portrait

Reaction Times

We analysed the RT data via MLM applying the same model as in Experiment 1 (Appendix A). The Valence × Endorsement interaction was significant, $B = -166.93$, $t_{94375} = -58.22$, $p < 0.001$, $\beta = -0.22$, 95% CI [-0.23, -0.21]. We followed up with simple slope analyses. In replication of self-positivity, participants exhibited quicker endorsement of positive traits than negative traits as self-descriptive ($\gamma = -177.56$, $z = -47.46$, $p < 0.001$), but exhibited faster rejection of negative traits than positive traits as self-descriptive ($\gamma = 156.31$, $z = 36.52$, $p < 0.001$). Moreover, self-positivity was evident for both the authentic self and the presented self (Appendix A).

Crucially, the Self × Valence × Endorsement interaction was significant, $B = 21.75$, $t_{94329} = 7.74$, $p < 0.001$, $\beta = 0.03$, 95% CI [0.02, 0.04]. We examined the Self × Endorsement interaction separately for positive traits and negative traits, testing whether self-positivity was stronger for one kind of self versus another (Figure 2.2d). In the case of positive traits, participants endorsed positive traits faster for the presented self than the authentic self ($\gamma = 14.65$, $z = 4.00$, $p < 0.001$), and rejected positive traits faster for the authentic self than the presented self ($\gamma = -33.23$, $z = -5.19$, $p < 0.001$). In the case of negative traits, participants endorsed negative traits faster for the authentic self than the presented self ($\gamma = -37.43$, $z = -4.87$, $p < 0.001$), and did not differ in their rejection of negative traits for the authentic self and the presented self ($\gamma = 1.69$, $z = 0.48$, $p = 0.629$). As in Experiment 1, self-positivity was weaker for the authentic self than the presented self. The results were comparable for *N*s of 107 and 61 participants (Table S2.5 and Table S2.6).

ERP

We depict the amplitudes of P1, N170, and LPP while participants underwent the SR-valence task (see Figure 2.7 and Table 2.2).

We used MLM to analyse ERP data (for the model settings see Appendix A).

P1. The model revealed a significant Valence × Endorsement interaction, $B = -0.04$, $t_{7616} = -3.12$, $p = 0.002$, $\beta = -0.01$, 95% CI [-0.02, -0.004], supporting the self-enhancement hypothesis. We followed up with simple slope tests (Curran et al., 2015). The elicited P1 was larger when endorsing negative traits as self-descriptive (vs. non-self-descriptive) ($\gamma = 0.07$, $z = 3.42$, $p < 0.001$), whereas the elicited P1 was equivalent when endorsing positive traits as self-descriptive and non-self-descriptive ($\gamma = -0.02$, $z = -0.99$, $p = 0.320$). P1 can reflect the processing of threatening information (Zhang et al., 2014). As such, the threat potential of having negative, self-descriptive traits emerged very early during processing of self-relevant information.

This threat potential was linked distinctly to the authentic and presented self, as evinced by the critical Self × Valence × Endorsement interaction, $B = 0.06$, $t_{7616} = 3.93$, $p < 0.001$, $\beta = 0.01$, 95% CI [0.01, 0.02] (Figure 2.4a, Figure 2.5. a1-a2, and Figure 2.6a). We then examined the presented-authentic contrasts by examining the Self × Endorsement interaction separately for positive and negative traits, testing whether self-positivity was stronger for one kind of self versus another (i.e., the self-accuracy and self-consistency views). The P1 was larger when participants endorsed negative traits as descriptive of the presented than authentic self ($\gamma = -0.26$, $z = -9.20$, $p < 0.001$), but it was equivalent when they endorsed positive traits as descriptive of the presented than authentic self ($\gamma = -0.004$, $z = -0.14$, $p = 0.882$). (For the results of P1 in judging the non-self-descriptiveness of positive and negative traits, see Appendix A.) The modulation of the P1 suggested preferential processing of negative information referring to the presented (vs. authentic) self. The authentic self exhibited weaker sensitivity to potentially threatening information at the very initial stage of processing, in line with the self-accuracy and self-consistency views.

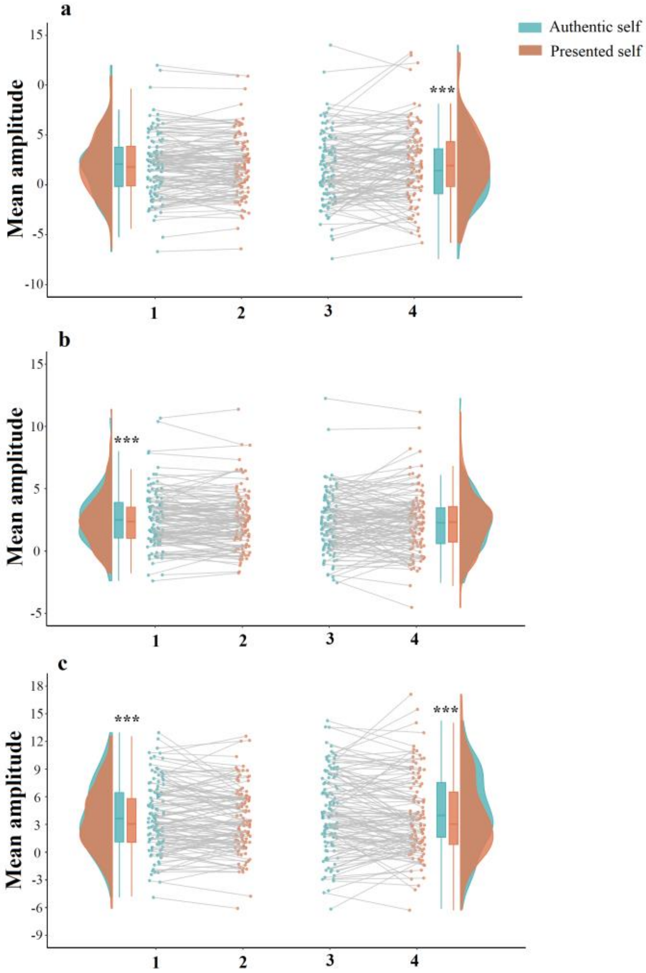
N170. Although the Valence × Endorsement interaction was not significant, $B = 0.004$, $t_{13544} = -0.40$, $p = 0.693$, $\beta = -0.001$, 95% CI [-0.007, 0.004], the crucial Self × Valence × Endorsement interaction was significant, $B = 0.03$, $t_{13544} = 2.94$, $p = 0.003$, $\beta = 0.01$, 95% CI [0.003, 0.01] (Figure 2.4b, Figure 2.5. b1-b2, and Figure 2.6b). We then examined the presented-authentic contrasts by examining the Self × Endorsement interaction separately for positive and negative traits, testing whether self-positivity was stronger for one kind of self versus another (i.e., the self-accuracy and self-consistency views). The N170 was larger when participants endorsed positive traits as descriptive of the presented than authentic self ($\gamma = 0.06$, $z = 3.35$, $p < 0.001$), but the N170 was not larger when participants endorsed negative traits as descriptive of the authentic versus presented self ($\gamma = -0.03$, $z = -1.72$, $p = 0.085$). (For the results of N170 in judging the non-self-descriptiveness of positive and negative traits, see Appendix A.) Given that the N170 reflects early attentional resource allocation to emotional stimuli (Cai et al., 2016; Montalan et al., 2008; Zhang et al., 2014), we inferred preferential processing of positive information referring to the presented (vs. authentic) self. That is, the presented self showed greater sensitivity to positive information in this subsequent stage of processing, a pattern opposite to the hypotheses derived from all three theoretical views. We provided an explanation for the conflicting result patterns of P1 and N170 in the General Discussion.

LPP. The model revealed a significant Valence × Endorsement interaction, $B = -0.31$, $t_{12698} = -20.57$, $p < 0.001$, $\beta = -0.07$, 95% CI [-0.08, -0.06], supporting the self-enhancement hypothesis. The LPP was larger when participants endorsed negative (vs. positive) traits as self-descriptive ($\gamma = -0.17$, $z = -7.78$, $p < 0.001$), and was larger when they rejected positive (vs. negative) traits as self-descriptive ($\gamma = 0.45$, $z = 21.30$, $p < 0.001$). This result is in line with prior findings (Cai et al., 2016; Herbert et al., 2011).

More importantly, the Self × Valence × Endorsement interaction was also significant, $B = 0.04$, $t_{12698} = 2.54$, $p = 0.011$, $\beta = 0.01$, 95% CI [0.002, 0.02] (Figure 2.4c, Figure 2.5. c1-c2, and Figure 2.6c). We then examined the presented-authentic contrasts by examining the Self × Endorsement interaction separately for positive and negative traits, testing whether self-positivity was stronger for one kind of self versus another (i.e., the self-accuracy and self-consistency views). The LPP was larger when participants endorsed positive traits as descriptive of the authentic (than presented) self ($\gamma = 0.14$, $z = 4.72$, $p < 0.001$), and was also larger when participants endorsed negative traits as descriptive of the authentic (than presented) self ($\gamma = 0.19$, $z = 6.35$, $p < 0.001$). (For the results of LPP in judging the non-self-descriptiveness of positive and negative traits, see Appendix A.)

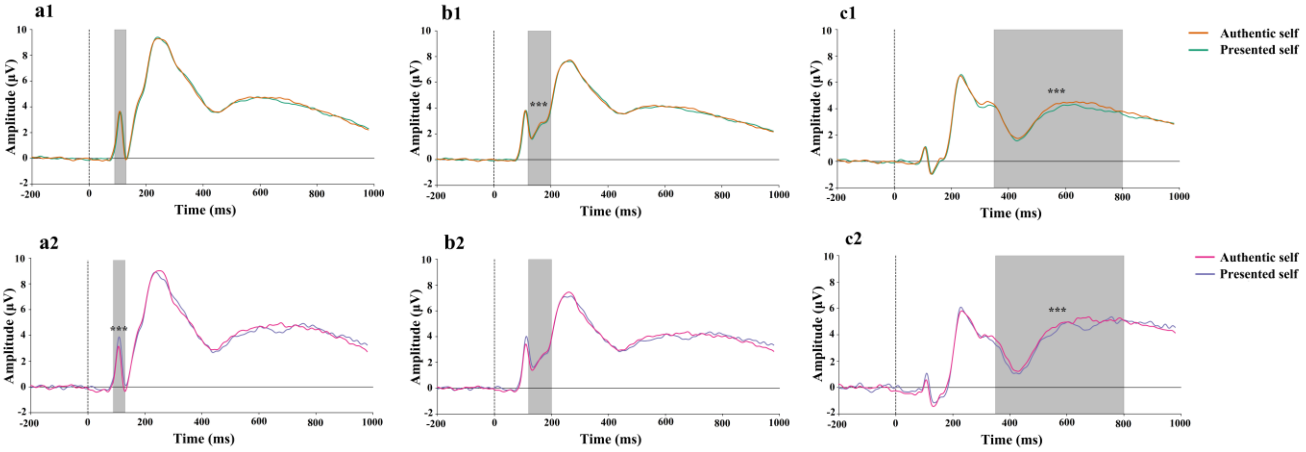
These findings did not fully align with our initial hypotheses derived from the three theoretical views. According to a burgeoning perspective, the LPP reflects sustained attention and elaborative processing (Auerbach et al., 2015; Hajcak et al., 2012), along with stimulus significance, with larger LPP responses observed in conjunction with significant stimuli that demand more elaborative processing (Hajcak & Foti, 2020). Therefore, our results indicated that individuals allocate more sustained attention and engage in deeper processing for both favoured authentic self (i.e., positive and self-descriptive traits) and disfavoured authentic self (i.e., negative and self-descriptive traits), a pattern somewhat compatible with the self-accuracy and self-consistency views. We revisit the issue in General Discussion.

Figure 2.4 *Neural Manifestations of Self-positivity for the Authentic Self and Presented Self in Experiment 2*



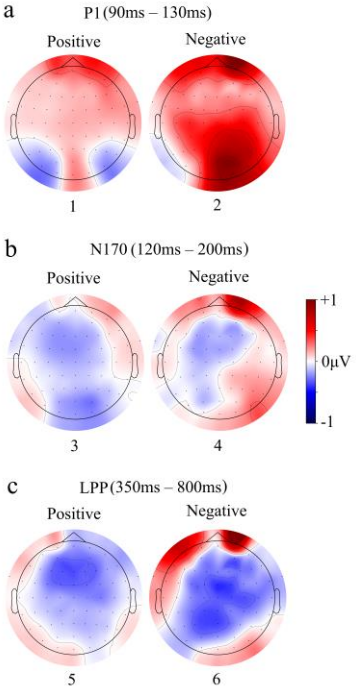
Note. (**a**) P1 mean amplitude for the authentic self and presented self in judging the self-descriptiveness of positive and negative traits. (**b**) N170 mean amplitude for the authentic self and presented self in judging the self-descriptiveness of positive and negative traits. (**c**) LPP mean amplitude for the authentic self and presented self in judging the self-descriptiveness of positive and negative traits (see Table S2.7 to S2.12 for fixed effects of self, valence, endorse, and their interactions on P1, N170, and LPP; see Figure S2.1 for the mean amplitude of P1, N170, and LPP for the authentic self and presented self in judging the non-self-descriptiveness of positive and negative traits). 1 = authentic self, positive traits, self-descriptiveness; 2 = presented self, positive traits, self-descriptiveness; 3 = authentic self, negative traits, self-descriptiveness; 4 = presented self, negative traits, self-descriptiveness. $^{***}p < .001$.

Figure 2.5 *Grand Averages for the ERPs of Self-positivity for the Authentic Self and Presented Self in Experiment 2*



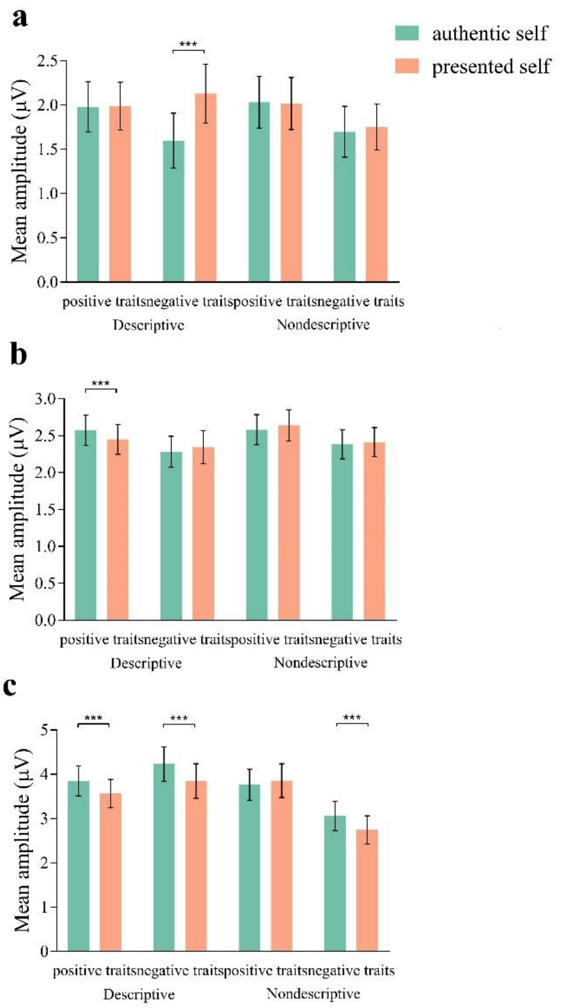
Note. (**a1**) Grand averages of P1 for the authentic self and presented self in judging the self-descriptiveness of positive traits. (**a2**) Grand averages of P1 for the authentic self and presented self in judging the self-descriptiveness of negative traits. (**b1**) Grand averages of N170 for the authentic self and presented self in judging the self-descriptiveness of positive traits. (**b2**) Grand averages of N170 for the authentic self and presented self in judging the self-descriptiveness of negative traits. (**c1**) Grand averages of LPP for the authentic self and presented self in judging the self-descriptiveness of positive traits. (**c2**) Grand averages of LPP for the authentic self and presented self in judging the self-descriptiveness of negative traits (see Figure S2.3 for grand averages of P1, N170, and LPP for the authentic self and presented self in judging the non-self-descriptiveness of positive and negative traits). The grand-averaged waveforms represent the grand average across component-specific electrode clusters: for the P1 component, signals were averaged across nine posterior sites (P3, P4, Pz, PO3, PO4, POz, O1, O2, and Oz); for the N170 component, across 16 temporo-parieto-occipital sites (TP7, TP8, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, O1, and O2); and for the LPP component, across 15 fronto-centro-parietal sites (F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4, CP3 CPZ, CP4, P3, P4, and Pz). $^{***}p < .001$.

Figure 2.6 *Topological Maps of Self-positivity for the Authentic Self and Presented Self in Experiment 2*



Note. (**a**) P1 amplitude difference between the presented self and authentic self in judging the self-descriptiveness of positive and negative traits. The time window of P1 is 90 ms – 130 ms. (**b**) N170 amplitude difference between the presented self and authentic self in judging the self-descriptiveness of positive and negative traits. The time window of N170 is 120 ms – 200 ms. (**c**) LPP amplitude difference between the presented self and authentic self in judging the self-descriptiveness of positive and negative traits. The time window of LPP is 350 ms – 800 ms (see Figure S2.2 for amplitude difference of P1, N170, and LPP between the presented self and authentic self in judging the non-self-descriptiveness of positive and negative traits). 1 = P1 amplitude of descriptiveness judgments on positive traits for the presented self minus P1 amplitude of descriptiveness judgments on positive traits for the authentic self; 2 = P1 amplitude of descriptiveness judgments on negative traits for the presented self minus P1 amplitude of descriptiveness judgments on negative traits for the authentic self; 3 = N170 amplitude of descriptiveness judgments on positive traits for the presented self minus N170 amplitude of descriptiveness judgments on positive traits for the authentic self; 4 = N170 amplitude of descriptiveness judgments on negative traits for the presented self minus N170 amplitude of descriptiveness judgments on negative traits for the authentic self; 5 = LPP amplitude of descriptiveness judgments on positive traits for the presented self minus LPP amplitude of descriptiveness judgments on positive traits for the authentic self; 6 = LPP amplitude of descriptiveness judgments on negative traits for the presented self minus LPP amplitude of descriptiveness judgments on negative traits for the authentic self.

Figure 2.7 *Mean Amplitude of ERPs in terms of Self-positivity for the Authentic Self and Presented Self in Experiment 2*



Note. (**a**) P1 mean amplitude for the authentic self and presented self in judging of positive and negative traits. (**b**) N170 mean amplitude for the authentic self and presented self in judging of positive and negative traits. (**c**) LPP mean amplitude for the authentic self and presented self in judging of positive and negative traits. $^{***}p < .001$.

Table 2.2 *Means and Standard Deviations for the Mean Amplitudes (μV) of P1, N170, and LPP*

			P1			
	Self-descriptive			Non-self-descriptive		
	Positive Traits	Negative Traits	Average	Positive Traits	Negative Traits	Average
Authentic Self	1.98 (3.11)	1.60 (3.40)	1.80 (3.26)	2.03 (3.22)	1.70 (3.16)	1.87 (3.19)
Presented Self	1.99 (2.98)	2.13 (3.65)	2.06 (3.32)	2.02 (3.23)	1.75 (2.85)	1.89 (3.04)
Average	1.98 (3.04)	1.86 (3.53)	1.92 (3.29)	2.03 (3.22)	1.73 (3.00)	1.88 (3.11)
			N170			
	Self-descriptive			Non-self-descriptive		
	Positive Traits	Negative Traits	Average	Positive Traits	Negative Traits	Average
Authentic Self	2.58 (2.24)	2.28 (2.31)	2.43 (2.28)	2.58 (2.24)	2.38 (2.19)	2.48 (2.21)
Presented Self	2.45 (2.22)	2.35 (2.48)	2.40 (2.35)	2.64 (2.34)	2.41 (2.17)	2.53 (2.25)
Average	2.51 (2.23)	2.32 (3.39)	2.41 (2.31)	2.61 (2.29)	2.40 (2.18)	2.51 (2.23)
			LPP			
	Self-descriptive			Non-self-descriptive		
	Positive Traits	Negative Traits	Average	Positive Traits	Negative Traits	Average
Authentic Self	3.85 (3.71)	4.23 (4.30)	4.04 (4.01)	3.77 (3.87)	3.06 (3.64)	3.41 (3.76)
Presented Self	3.57 (3.52)	3.85 (4.28)	3.71 (3.90)	3.86 (4.19)	2.75 (3.52)	3.30 (3.85)
Average	3.71 (3.61)	4.04 (4.29)	3.88 (3.95)	4.81 (4.03)	2.90 (3.58)	3.36 (3.80)

2.3.3 Discussion

Experiment 2 replicated behavioural evidence that self-positivity is weaker for the authentic than presented self. We found neurophysiological differences for the P1, N170, and LPP. We also found asymmetrical neural patterns during the earlier processing stages (P1, N170). Participants showed a larger P1 for negative traits descriptive of the presented (vs. authentic) self. This patter was reversed at the subsequent processing stage (N170). Finally, we observed a larger LPP for the authentic (vs. presented) self.

2.4 General Discussion

What does it mean to be authentic? We tested the emerging view of authenticity as self-enhancement against more traditional views of it as self-accuracy and self-consistency. To do so, we placed authenticity under the behavioural and neuropsychological microscope, comparing it with a highly positive mental representation, the presented self. We tested two competing hypotheses (Platt, 1964). First, in line with the self-enhancement view, we hypothesized that the strength of self-positivity would be comparable for the authentic and presented selves. Alternatively, in line with the self-accuracy and self-consistency views, we hypothesized that the strength of self-positivity would be weaker for the authentic compared to presented self.

2.4.1 Summary of Findings

Behavioural Evidence

Across two experiments, we replicated self-positivity (the Valence × Endorsement interaction). In terms of trait endorsement, participants overall endorsed more positive than negative traits as self-descriptive but judged more negative than positive traits as non-self-descriptive. Further, in both experiments, participants evinced self-positivity for both the authentic and presented self (Appendix A). In regard to reaction times, in both experiments, participants showed faster endorsement of positive than negative traits as self-descriptive and showed faster rejection of negative than positive traits as self-descriptive. Likewise, in both experiments, participants manifested self-positivity for both the authentic and presented self (Appendix A).

Our main interest was in the relative strength of self-positivity tethered to the authentic versus the presented self (the three-way interaction among endorsement, valence, and self). The results were similar across experiments. In terms of trait endorsement, in both experiments, participants judged more negative traits as descriptive of the authentic than presented self, but judged more such traits as nondescriptive of the presented than authentic self. In Experiment 2, participants further endorsed more positive traits as descriptive of the presented than authentic self and endorsed more such traits as nondescriptive of the authentic than presented self. In regard to reaction times, in both experiments, participants were faster to endorse positive traits for the presented than authentic self but were faster to endorse negative traits for the authentic than presented self.

In summary, participants endorsed a higher number of negative traits, and a lower number of positive traits, as part of the authentic than presented self. Also, participants were speedier in endorsing negative traits, but slower in endorsing (or speedier in denouncing) positive traits, for their authentic than presented self. Taken together, the behavioural results across experiments suggest that the self-positivity was weaker for the authentic than presented self. This is a challenge to the self-enhancement view of authenticity, which would anticipate equal degree of favourability for the authentic and presented self-concepts. In contract, the behavioural results are consistent with the self-consistency and self-accuracy views. Therefore, the authentic self, albeit positive on its own, is less positive than the presented self.

Neuropsychological Evidence

P1, N170, and LPP amplitudes constituted the neuropsychological evidence. Initially, participants exhibited augmented P1 responses when endorsing negative traits as self-descriptive versus non-self-descriptive) (Valence × Endorsement interaction), and this effect was attenuated for the authentic self compared to the presented self (Self × Valence × Endorsement interaction). Subsequently, participants exhibited augmented N170 responses when endorsing positive traits as descriptive of the presented than authentic self (a three-way interaction among endorsement, valence, and self). Finally, participants exhibited augmented LPP responses when endorsing both positive and negative traits as self-descriptive (vs. non-self-descriptive) (a two-way interaction between endorsement and valence), and these effects were more pronounced for the authentic than presented self (Self × Valence × Endorsement interaction). Collectively, these neuropsychological results were compatible with the self-consistency and self-accuracy views. We provided a detailed interpretation in later sections.

2.4.2 Empirical Implications

Here, the neuropsychological findings were nuanced, manifesting intricate processing sequences. At the very early processing stage, the P1 component was heightened when participants endorsed negative (but not positive) traits as self-descriptive versus non-self-descriptive. Moreover, this pattern was attenuated for the authentic than presented self, which aligns with the self-consistency and self-accuracy views. Previous research indicates that P1 reflects early attentional allocation (Hillyard et al., 1998) and is sensitive to negative information

(Luo et al., 2010; Zhang et al., 2014), with stronger P1 responses evoked by negative (compared to neutral) stimuli. Therefore, our results indicate that participants allocated more attentional resources to negative traits about the presented than authentic self. That is, they evinced preferential processing of information that posed a threat to their presented (vs. authentic) self. Alternatively, negative information is less threatening for the authentic self, indicating the authentic self was more unperturbable than the presented self. Although prior studies typically report null interactions between self-relevance and valence in the P1 response within SR-valence tasks (Ding et al., 2020; Fields & Kuperberg, 2012; Hudson et al., 2020; Wieser et al., 2014; Zhou et al., 2017), our findings may introduce a novel direction for exploring the P1's role in SR-valence processing.

However, at the subsequent processing stage, this pattern reversed: the N170 amplitude was heightened when participants endorsed positive traits as descriptive of the presented than authentic self. Previous research indicates that the N170 component reflects early attentional resource allocation to emotional stimuli (Zhang et al., 2014), especially negative ones (Cai et al., 2016; Montalan et al., 2008). Therefore, our results suggest that participants allocated more attentional resources to positive traits about the presented than authentic self. Stated otherwise, participants manifested preferential processing of positive (but not negative) information referring to their presented than authentic self, which diverged from hypotheses offered by all three theoretical views. However, both the authentic and presented selves demonstrated a more negative N170 deflection in response to negative traits than to positive ones (Table S2.9, Appendix A; main effect of Valence: $\beta = 0.10$, $t_{13544} = 11.11$, $p < 0.001$, negative valence = 2.36 ± 3.39 μV, positive valence = 2.56 ± 3.31μV), which is still consistent with N170's broad sensitivity to negativity.

Interestingly, the differentiation between the authentic self and the presented self emerged with positive rather than negative traits—a pattern opposite to that observed in the P1 stage. Two explanations may account for this pattern. First, this differentiation reflects the distinct stages of emotional processing (Luo et al., 2010; Pourtois et al., 2013; Zhang et al., 2014). Specifically, early modulation of the P1 by emotion rapidly distinguishes between non-threatening and threatening information, which can facilitate swift detection of threatening stimuli. Then, at a later stage (N170), more differentiated emotional processing occurs, enabling refined feature recognition and emotion assessment. In accord with this reasoning, the earliest component (P1) showed initial selective attention to and rapid detection of negative traits, whereas the subsequent N170 reflected more nuanced processing for the authentic and presented self on positive traits. Although the N170 is typically linked to negative information processing (Rossion & Jacques, 2012), a meta-analysis revealed that N170's sensitivity to emotional stimuli is heterogeneous, with both negative (e.g., angry, fearful) and positive (e.g., happy) faces eliciting heightened N170 amplitudes compared to neutral faces (Hinojosa et al., 2015). The meta-analytic finding suggests that the N170 stage also involves attentional resources for positive stimuli. Considering that this stage likely entails more complex self-reference processing, it is possible that the broadly positive content of the self (Alicke & Sedikides, 2009; Sedikides & Gregg, 2008) contributed to greater differentiation between the authentic and presented selves, specifically for positive traits. Moreover, this positivity is more pronounced for the presented self, which is compatible, to some extent, with the self-consistency and self-accuracy views.

Second, the asymmetrical neural patterns during the earlier processing stages (P1, N170) are partially accounted for by the mobilization-minimization hypothesis (Taylor, 1991; see also Sedikides et al., 2016). According to it, 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 the context of our research, negative self-descriptive information received preferential processing initially (mobilization; P1), followed by preferential processing of positive, self-descriptive information (minimization; N170). Moving beyond this hypothesis, the mobilization-minimization dynamic was more strongly associated with the presented than authentic self. The presented self is more brittle (e.g., changeable, malleable, fluctuating, and shifting) and so needed to be defended more strongly; alternatively, the authentic self is more robust or stable and so in less need of defence (Study S2, Appendix A). From this vantage point, the findings of N170 were also in line with the self-consistency and self-accuracy views.

Past research has indicated that self-reference processing can elicit an augmented N170 (Caharel et al., 2007; Keyes et al., 2010; Shi, 2016). However, few studies have examined how the self-reference sensitivity of the N170 interacts with its emotion-sensitive properties, with most of them reporting a null Self-Reference × Emotional Valence interaction (McCrackin & Itier, 2018; Qun et al., 2018; Wieser et al., 2014). Our findings may thus open a promising new direction for exploring the N170's role in processing self-referential valence.

Lastly, at the ensuing processing stage, participants exhibited augmented LPP responses when endorsing both positive and negative traits as self-descriptive (vs. non-self-descriptive) of the authentic than presented self. Although these findings were incompatible with our original hypotheses derived from the three theoretical views, they were largely congruent with the self-accuracy and self-consistency views. According to an emerging literature, the amplified LPP amplitudes reflect sustained attention and elaborative processing (Auerbach et al., 2015; Hajcak et al., 2012) as well as stimulus significance (Hajcak & Foti, 2020). Within this framework, our findings suggest that participants regarded the authentic self as more significant and engaged in more elaborative processing, as demonstrated by their stronger responses (LPPs) to both threatening (negative, descriptive) and non-threatening (positive, descriptive) information about the authentic self. Participants may have considered both positive and negative aspects as integral to their authentic self, largely in line with the self-accuracy and self-consistency views. Further, most LPP experiments select stimuli based on normative valence or arousal rather than stimulus significance. We asked participants to judge whether an identical set of traits (thus holding valence and arousal constant) represents the authentic self and presented self. Insofar as participants regard their authentic self as more important, valuable, and significant than their presented self (Study S3, Appendix A), our findings provide a rigorous test of and strong support for the stimulus significance perspective of LPP.

2.4.3 Theoretical Implications

Our findings help to clarify the three theoretical views on authenticity. Some researchers conceptualized authenticity through the lens of self-accuracy, the candid and unbiased processing of identity relevant information (Kernis & Goldman, 2006; Lakey et al., 2008). In part because self-accuracy is difficult to empirically verify (Vazire & Wilson, 2012), others conceptualized authenticity as self-consistency, the alignment of one’s behaviour with internal standards, goals, or values (Kernis & Goldman, 2006; Wood et al., 2008). Still, other researchers considered authenticity as self-enhancement (Bailey & Iyengar, 2023; Bench et al., 2015; Guenther et al., 2024; Strohminger et al., 2017). Although evidence is stronger for the self-enhancement view (Sedikides & Schlegel, 2024), our results pose a challenge to it. While still positive, the authentic self allows for some acknowledgment of negativity, a results pattern more compatible with the self-accuracy and self-consistency views. It appears as if people know they have some negative traits that they are unwilling to share with others (top panel of Figure 2.2; see also: Cheung et al., 2014; Preuss & Alicke, 2017). Future research would benefit from examining how the two pathways to authenticity—reflected in the theoretical views—function both independently and jointly.

The findings contribute a novel perspective to the literature on authenticity as self-enhancement by incorporating the processing of negative traits. Although prior research has largely emphasized the connection between authenticity and the endorsement of positive traits (Guenther & Sedikides, in press), our findings underscore the crucial role of distancing oneself from negative traits in shaping authenticity.

2.4.4 Limitations and Future Directions

Our goal was to establish the internal validity of our findings, and hence our use of convenience samples was justified (Mook, 1983; Sherman, 2024). Yet, for generalizability, future studies would do well to test non-WEIRD samples.

Our method primarily compared the authentic self with the presented self in terms of self-positivity. We did not directly measure each motive (i.e., self-enhancement, self-accuracy, self-consistency). Nevertheless, our findings provide indirect evidence of how these motives may shape the expression of authenticity. Let us take the case of self-enhancement. The discrepancy in self-positivity between the authentic and the presented self indicates that the presented self may be influenced by self-enhancing concerns, whereas the authentic self reflects a more balanced or realistic appraisal. Let us now consider self-accuracy. If individuals acknowledge both strengths and weaknesses in their authentic self, this practice could point to a more unerring self-perception. In contrast, preference to endorse strengths but reject weaknesses in the presented self could suggest that this self is influenced by self-enhancement concerns, thereby lacking a degree of realistic appraisal. Finally, let us focus on self-consistency. Showing smaller self-positivity in one's authentic self might be driven by strong internal alignment, acknowledging one's weakness, whereas larger self-positivity in the presented self might indicate inconsistencies, potentially driven by external pressures to conform to social expectations. Future research would benefit from more direct measurement in testing the relation between authenticity and these motives.

We defined the presented self for participants as "the version of the self you present to others." It is possible that participants found the presented self more difficult to process since they define various presented selves across varying contexts. In line with this possibility, it has been reported that the true self is slightly easier to describe than the actual self, suggesting that the true self might also be easier to process than the presented self (Schlegel et al., 2011). However, the effect size in the relevant study was very small (Cohen's d = .11), indicating that the size of processing differences is negligible at best. Our findings are also compatible with a lack of significant processing differences. Although we found evidence that the authentic self is more robust, more significant, and less sanitized (Study S1-S3, Appendix A), we obtained no evidence to suggest that the authentic self is more difficult to process; that is, we observed no difference in reaction times when participants made decisions on the authentic versus presented self. Specifically, participants endorsed positive traits faster for the presented than authentic self. Moreover, participants endorsed negative traits faster for the authentic than presented self. Finally, they did not differ in their rejection of negative traits for the authentic and presented selves. These result patterns emerged in both experiments. Nonetheless, future research should delve into the intricacies of how individuals define their presented self across varying contexts to increase understanding of differences between the two selves.

Further, our neural evidence relied on EEG, which has excellent temporal resolution but poor spatial resolution (Cohen, 2017). Functional magnetic resonance imaging (fMRI) studies of reward-relevant brain regions may complement our findings. Reward-related brain regions like the striatum are critical to self-processing (Berridge & Kringelbach, 2013; Delgado, 2007). Thinking about the self feels good and activates parts of the striatum (Enzi et al., 2009). Evidence of decreased striatal activation when making judgments about the authentic (vs.

presented) self would bolster our findings. However, increased striatal activation when making judgments about the authentic (vs. presented) self would support the self-enhancement view of authenticity. Other neuroscientific studies link self-enhancement to both structural (Chavez & Heatherton, 2015; Chester et al., 2016) and functional (Chavez & Heatherton, 2015) connectivity between self-relevant (medial prefrontal cortex) and reward relevant (striatum) brain regions. If self-enhancement underlies the authentic self, we would expect particularly strong connectivity between medial prefrontal cortex and the striatum when participants make judgments about the authentic self. However, based on our findings, we would obtain weaker connectivity between medial prefrontal cortex and the striatum when participants make judgements about the authentic self. No fMRI studies have so far distinguished between the authentic and presented self. Such studies would complement our findings and enrich the emerging neuroscience of authenticity.

2.5 Conclusion

Authenticity has held an enduring fascination with intellectuals, researchers, and the public. We aimed to capture the essence of it. Although authenticity is largely self-enhancing, it also entails a willingness to explore the possibility of unfavourable pockets of selfhood or even admit one’s weakness. Authenticity entails the notion that the self is highly positive, but this notion appears to be secure enough to tolerate partial negativity or inconsistency.

References

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology, 20*(1), 1–48. https://doi.org/10.1080/10463280802613866

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology, 9*(3), 272–279. https://doi.org/10.1037/h0025907

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. http://dx.doi.org/10.1037/amp0000191

Auerbach, R. P., Ho, M. H. R., & Kim, J. C. (2014). Identifying cognitive and interpersonal predictors of adolescent depression. *Journal of Abnormal Child Psychology, 42*(6), 913-924. https://doi.org/10.1007/s10802-013-9845-6

Auerbach, R. P., Stanton, C. H., Proudfit, G. H., & Pizzagalli, D. A. (2015). Self-referential processing in depressed adolescents: A high-density event-related potential study. *Journal of Abnormal Psychology, 124*(2), 233–245. http://dx.doi.org/10.1037/abn0000023

Bailey, E. R., & Iyengar, S. S. (2023). Positive—more than unbiased—self-perceptions increase subjective authenticity. *Journal of Personality and Social Psychology, 125*(6), 1351–1372. https://doi.org/10.1037/pspa0000359

Bareket-Bojmel, L., Moran, S., & Shahar, G. (2016). Strategic self-presentation on Facebook: Personal motives and audience response to online behaviour. *Computers in Human Behaviour, 55*, 788-795. https://doi.org/10.1016/j.chb.2015.10.033

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, K. (2017). To be or not to be authentic: In defense of authenticity as an ethical ideal. *Ethical Theory and Moral Practice, 20*(3), 567–580. https://doi.org/10.1007/s10677-017-9803-4

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*(1), 3–26. https://doi.org/10.1037/0033-2909.91.1.3

Baumeister, R. F. (2019). Stalking the true self through the jungles of authenticity: Problems, contradictions, inconsistencies, disturbing findings—and a possible way forward. *Review of General Psychology, 23*(1), 143–154. https://doi.org/10.1177/1089268019829472

BBC (2023, November, 27). *Authentic: Merriam-Webster's word of the year*. https://www.bbc.co.uk/news/world-us-canada-67543895

Berridge, K. C., & Kringelbach, M. L. (2013). Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology*, *23*(3), 294–303. https://doi.org/10.1016/j.conb.2013.01.017

Bench, S. W., Schlegel, R. J., Davis, W. E., & Vess, M. (2015). Thinking about change in the self and others: The role of self-discovery metaphors and the true self. *Social Cognition, 33*(3), 169–185. https://doi.org/10.1521/soco.2015.33.3.2

Boucher, H. C. (2011). The dialectical self-concept II: Cross-role and within-role consistency, well-being, self-certainty, and authenticity. *Journal of Cross-Cultural Psychology, 42*(7), 1251–1271. https://doi.org/10.1177/0022022110383316

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology, 55*(6), Article e13049. https://doi.org/10.1111/psyp.13049

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.

Caharel, S., Bernard, C., Thibaut, F., Haouzir, S., Di Maggio-Clozel, C., Allio, G., Fouldrin, G., Petit, M., Lalonde, R., & Rebaï, M. (2007). The effects of familiarity and emotional expression on face processing examined by ERPs in patients with schizophrenia. *Schizophrenia Research, 95*(1-3), 186–196. https://doi.org/10.1016/j.schres.2007.06.015

Cahn, B. R., & Polich, J. (2006). Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin, 132*(2), 180–211. https://doi.org/10.1037/0033-2909.132.2.180

Cai, H., Wu, L., Shi, Y., Gu, R., & Sedikides, C. (2016). Self-enhancement among Westerners and Easterners: a cultural neuroscience approach. *Social Cognitive and Affective Neuroscience, 11*(10), 1569–1578. https://doi.org/10.1093/scan/nsw072

Chandler, J. (2018). Likeableness and meaningfulness ratings of 555 (+487) person-descriptive words. *Journal of Research in Personality, 72*, 50–57. https://doi.org/10.1016/j.jrp.2016.07.005

Chavez, R. S., & Heatherton, T. F. (2015). Multimodal frontostriatal connectivity underlies individual differences in self-esteem. *Social Cognitive and Affective Neuroscience*, *10*(3), 364–370. https://doi.org/10.1093/scan/nsu063

Chen, K., Zhang, H., & Schlegel, R. J. (2023). Does being angry feel authentic? a test of how affective valence and motivational direction differentially influence state authenticity. *Motivation and Emotion*, *47*(5), 828–841. https://doi.org/10.1007/s11031-023-10026-1

Chester, D. S., Lynam, D. R., Powell, D. K., & DeWall, C. N. (2016). Narcissism is associated with weakened frontostriatal connectivity: A DTI study. *Social Cognitive and Affective Neuroscience*, *11*(7), 1036–1040. https://doi.org/10.1093/scan/nsv069

Cheung, W. Y., Wildschut, T., Sedikides, C., & Pinter, B. (2014). Uncovering the multifaceted-self in the domain of negative traits: On the muted expression of negative self-knowledge. *Personality and Social Psychology Bulletin, 40*(4), 513–525. https://doi.org/10.1177/0146167213518224

Cohen, M. X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences, 40*(4), 208–218. https://doi.org/10.1016/j.tins.2017.02.004

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2015). Testing and probing interactions in hierarchical linear growth models. In C. S. Bergeman & S. M. Boker (Eds.), *Methodological issues in aging research* (1st ed., pp. 99–129). Psychology Press.

D'Argembeau, A., Comblain, C., Van der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology, 96*(4), 457–466. https://doi.org/10.1348/000712605X53218.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioural experiments in a Web browser. *Behaviour Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*(1), 70–88. https://doi.org/10.1196/annals.1390.002

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. https://doi.org/10.1016/j.jneumeth.2003.10.009

Ding, X. B., Wang, R., Kang, T. J., Liu, J. Y., & Zhou, J. N. (2020). Consistency between self-reference and mother-reference in emotional perception of others' faces: Evidence from ERP. *Acta Psychologica Sinica*, *52*(6), 694–705. https://doi.org/10.3724/SP.J.1041.2020.00694

Dores Cruz, T. D., Thielmann, I., Columbus, S., Molho, C., Wu, J., Righetti, F., de Vries, R. E., Koutsoumpis, A., Van Lange, P. A. M., Beersma, B., & Balliet. D. (2021). Gossip and reputation in everyday life. *Philosophical Transactions of the Royal Society B,* 376(1838), Article 20200301. https://doi.org/10.1098/rstb.2020.0301

Dormanen, R., Sanders, C. S., Maffly-Kipp, J., Smith, J. L., & Vess, M. (2020). Assimilation undercuts authenticity: A consequence of women’s masculine self-presentation in masculine contexts. *Psychology of Women Quarterly, 44*(4), 488–502. https://doi.org/10.1177/0361684320947648

Ebrahimi, M., Kouchaki, M., & Patrick, V. M. (2020). Juggling work and home selves: Low identity integration feels less authentic and increases unethicality. *Organizational Behaviour and Human Decision Processes, 158*, 101–111. https://doi.org/10.1016/j.obhdp.2019.02.005

Enzi, B., de Greck, M., Proesch, U., Tempelmann, C., & Northoff, G. (2009). Is our self nothing but reward? Neuronal overlap and distinction between reward and personal relevance and its relation to human personality. *PLoS One*, *4*(12), Article e8429. https://doi.org/10.1371/journal.pone.0008429

Erickson, R. J. (1995). The importance of authenticity for self and society. *Symbolic Interactionism, 18*(2), 121–144. https://doi.org/10.1525/si.1995.18.2.121

Feingold, A. (2015). Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *Journal of Consulting and Clinical Psychology, 83*(1), 157–168. https://doi.org/10.1037/a0037721

Ferrara, A. (1993). *Modernity and authenticity: A study of the social and ethical thought of Jean-Jacques Rousseau*. Sunny Press.

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage, 62*(1), 562–574. https://doi.org/10.1016/j.neuroimage.2012.05.003

Fleeson, W., & Wilt, J. (2010). The relevance of Big Five trait content in behaviour to subjective authenticity: Do high levels of within-person behavioural variability undermine or enable authenticity achievement? *Journal of Personality, 78*(4), 1353–1382. https://doi.org/10.1111/j.1467-6494.2010.00653.x

Gillath, O., Sesko, A. K., Shaver, P. R., & Chun, D. S. (2010). Attachment, authenticity, and honesty: dispositional and experimentally induced security can reduce self-and other-deception. *Journal of Personality and Social Psychology, 98*(5), 841–855. https://doi.org/10.1037/a0019206

Golomb, J. (1995). *In search of authenticity: From Kierkergaard to Camus*. Routledge.

Guenther, C. L., & Sedikides, C. (in press). Authenticity as self-enhancement. In M. Vess, B. Schlegel, & J. Hicks (Eds.), *The Oxford handbook of authenticity*. Oxford University Press.

Guenther, C. L., Zhang, Y., & Sedikides, C. (2024). The authentic self is the self-enhancing self: A self-enhancement framework of authenticity. *Personality and Social Psychology Bulletin, 50*(8), 1182–1196. https://doi.org/10.1177/101461672231160653

Guignon, C. (2004). *On being authentic*. Routledge.

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology, 57*(7), Article e13570. https://doi.org/10.1111/psyp.13570

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In *The Oxford handbook of event-related potential components* (Vol. 441, pp. 441–472). Oxford University Press.

Hart, W., Richardson, K., Breeden, C. J., & Kinrade, C. (2020). To be or to appear to be: Evidence that authentic people seek to appear authentic rather than be authentic. *Personality and Individual Differences*, *166*, Article 110165. https://doi.org/10.1016/j.paid.2020.110165

Herbert, C., Pauli, P., & Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Social Cognitive and Affective Neuroscience*, 6(5), 653–661. https://doi.org/10.1093/scan/nsq082

Hudson, A., Wilson, M. J., Green, E. S., Itier, R. J., & Henderson, H. A. (2020). Are you as important as me? Self-other discrimination within trait-adjective processing. *Brain and Cognition, 142*, Article 105569. https://doi.org/10.1016/j.bandc.2020.105569

Hicks, J. A., Schlegel, R. J., & Newman, G. E. (2019). Introduction to the special issue: Authenticity: Novel insights into a valued, yet elusive, concept. *Review of General Psychology*, *23*(1), 3–7. https://doi.org/10.1177/108926801982947

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence.

Philosophical Transactions of the Royal Society of London B: Biological Sciences, 353(1373), 1257–1270. https://doi.org/10.1098/rstb.1998.0281

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience & Biobehavioural Reviews, 55*, 498–509. https://doi.org/10.1016/j.neubiorev.2015.06.002

Hudson, A., Wilson, M. J., Green, E. S., Itier, R. J., & Henderson, H. A. (2020). Are you as important as me? Self-other discrimination within trait-adjective processing. *Brain and Cognition, 142*, Article 105569. https://doi.org/10.1016/j.bandc.2020.105569

Jordan, C. D., Stewart, R. A., Brush, C. J., Cougle, J. R., & Hajcak, G. (2022). Appearance concerns are uniquely associated with LPP amplitude to pictures of oneself. *Social Cognitive and Affective Neuroscience, 17*(4), 430–436. https://doi.org/10.1093/scan/nsab115

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

Jongman-Sereno, K. P., & Leary, M. R. (2016). Self-perceived authenticity is contaminated by the valence of one's behaviour. *Self and Identity, 15*(3), 283–301. https://doi.org/10.1080/15298868.2015.1128964

Jongman-Sereno, K. P., & Leary, M. R. (2019). The enigma of being yourself: A critical examination of the concept of authenticity. *Review of General Psychology*, *23*(1), 133–142. https://doi.org/10.1037/gpr0000157

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. *Advances in Experimental Social Psychology, 38*, 283–357. https://doi.org/10.1016/S0065-2601(06)38006-9

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and Cognition, 72*(2), 244–254. https://doi.org/10.1016/j.bandc.2009.09.006

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920951503. https://doi.org/10.1177/2515245920951503

Lakey, C. E., Kernis, M. H., Heppner, W. L., & Lance, C. E. (2008). Individual differences in authenticity and mindfulness as predictors of verbal defensiveness. *Journal of Research in Personality, 42*(1), 230–238. https://doi.org/10.1016/j.jrp.2007.05.002

Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behaviour.* Brown & Benchmark Publishers.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology, 58*(1), 317–344. https://doi.org/10.1146/annurev.psych.58.110405.085658

Lee, E., Kim, Y. J., & Ahn, J. (2014). How do people use Facebook features to manage social capital? *Computers in Human Behaviour, 36*, 440–445. https://doi.org/10.1016/j.chb.2014.04.007

Lenton, A. P., Slabu, L., Sedikides, C., & Power, K. (2013). I feel good, therefore I am real: Testing the causal influence of mood on state authenticity. *Cognition & Emotion*, *27*(7), 1202–1224. https://doi.org/10.1080/02699931.2013.778818

Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-scale Assessments in Education,* 6(1), 1–11. https://doi.org/10.1186/s40536-018-0061-2

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed). MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn’t). *Psychophysiology, 54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luo, W., Feng, W., He, W., Wang, N., & Luo, Y. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage, 49*(2), 1857–1867. https://doi.org/10.1016/j.neuroimage.2009.09.018

McCrackin, S. D., & Itier, R. J. (2018). Is it about me? Time-course of self-relevance and valence effects on the perception of neutral faces with direct and averted gaze. *Biological Psychology, 135*, 47–64. https://doi.org/10.1016/j.biopsycho.2018.03.003

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review, 28*, 295–314. https://doi.org/10.1007/s10648-014-9287-x

Montalan, B., Caharel, S., Personnaz, B., Le Dantec, C., Germain, R., Bernard, C., Lalonde, R., & Rebaï, M. (2008). Sensitivity of N170 and late positive components to social categorization and emotional valence. *Brain Research, 1233*, 120–128. https://doi.org/10.1016/j.brainres.2008.07.055

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*(4), 379–387. https://doi.org/10.1037/0003-066X.38.4.379

Moran, T. P., Jendrusina, A. A., & Moser, J. S. (2013). The psychometric properties of the late positive potential during emotion processing and regulation. *Brain Research, 1516*, 66–75. https://doi.org/10.1016/j.brainres.2013.04.018

Morís Fernández, L., & Vadillo, M. A. (2020). Flexibility in reaction time analysis: many roads to a false positive?. *Royal Society Open Science*, *7*(2), Article 190831. https://doi.org/10.1098/rsos.190831

Mueller, A. K., Falter, C. M., & Tucha, O. (2013). Distinct attentional mechanisms for social information processing as revealed by event-related potential studies. *International Journal of Brain and Cognitive Sciences, 2*(2), 23–37. https://doi.org/10.5923/j.ijbcs.20130202.03

Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences, 8*(3), 102–107. https://doi.org/10.1016/j.tics.2004.01.00

Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology, 84*(4), 890–904. https://doi.org/10.1037/0022-3514.84.4.890

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72*(2), 184–187. https://doi.org/10.1016/0013-4694(89)90180-6

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage, 198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Platt, J. R. (1964). Strong inference. *Science, 146*(3642), 347–353.

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology, 92*(3), 492–512. https://doi.org/10.1016/j.biopsycho.2012.02.007

Press, G. S., & Alicke, M. D. (2017). My worst faults and misdeeds: Self-criticism and self-enhancement can co-exist. *Self and Identity, 16*(6), 645–663. https://doi.org/10.1080/15298868.2017.1296019

Qun, T., Yueyang, Y., Shen, L., Shangfeng, H., Qiang, X., & Lin, Z. H. A. N. G. (2018). The processing advantage of self-positive expression: evidence from an ERPs study. *Acta Psychologica Sinica, 50*(10), 1120–1130. https://doi.org/10.3724/SP.J.1041.2018.01120

Rameson, L. T., Satpute, A. B., & Lieberman, M. D. (2010). The neural correlates of implicit and explicit self-relevant processing. *NeuroImage, 50*(2), 701–708. https://doi.org/10.1016/j.neuroimage.2009.12.098

R Core Team (2023). *R (4.3.1): A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Rossion, B., & Jacques, C. (2012). The N170: Understanding the time course of face perception in the human brain. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 115–141). Oxford University Press.

Roth, M. T. Y. M. E., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S., & Goddard, R. (1986). CAMDEX: a standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *The British Journal of Psychiatry, 149*(6), 698–709. https://doi.org/10.1192/bjp.149.6.69

Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Brooks/Cole.

Schlenker, B. R. (2003). Self-presentation. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 492–518). Guilford Press.

Schlenker, B. R., & Leary, M. R. (1982). Audiences' reactions to self-enhancing, self-denigrating, and accurate self-presentations. *Journal of Experimental Social Psychology, 18*(1), 89-104. https://doi.org/10.1016/0022-1031(82)90083-X

Sedikides, C. (2020). On the doggedness of self-enhancement and self-protection: How constraining are reality constraints? *Self and Identity, 19*(3), 251–271. https://doi.org/10.1080/15298868.2018.1562961

Sedikides, C. (2021). Self-construction, self-protection, and self-enhancement: A homeostatic model of identity protection. *Psychological Inquiry, 32*(4), 197–221. https://doi.org/10.1080/1047840X.2021.2004812

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one’s faults. *European Review of Social Psychology, 27*(1), 1–62. https://doi.org/10.1080/10463283.2016.1183913

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3*(2), 102-116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

Sedikides, C., & Schlegel, R. (2024). Distilling the concept of authenticity. *Nature Reviews Psychology, 3*, 509–523. https://doi.org/10.1038/s44159-024-00323-y

Şahin, M., & Aybek, E. (2020). Jamovi: An easy to use statistical software for the social scientists. *International Journal of Assessment Tools in Education, 6*(4), Article 4. https://doi.org/10.21449/ijate.661803

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex, 130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Schlegel, R. J., Hicks, J. A., Arndt, J., & King, L. A. (2009). Thine own self: True self-concept accessibility and meaning in life. *Journal of Personality and Social Psychology, 96*(2), 473–490. https://doi.org/10.1037/a0014060

Sedikides, C. (2020). On the doggedness of self-enhancement and self-protection: How constraining are reality constraints? *Self and Identity, 19*(3), 251–271. https://doi.org/10.1080/15298868.2018.1562961

Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of Personality and Social Psychology, 73*(6), 1380–1393. https://doi.org/10.1037/0022-3514.73.6.1380

Sherman, J. W. (2024). There is nothing WEIRD about basic research: The critical role of convenience samples in psychological science. *American Psychologist*. Advance online publication. https://dx.doi.org/10.1037/amp0001443

Shestyuk, A. Y., & Deldin, P. J. (2010). Automatic and strategic representation of the self in major depression: trait and state abnormalities. *American Journal of Psychiatry, 167*(5), 536–544. https://doi.org/10.1176/appi.ajp.2009.06091444

Shi, Z. (2016). My name or yours? Event-related potential correlates of self-name processing. *NeuroReport, 27*(7), 542–547. https://doi.org/10.1097/WNR.0000000000000582

Shi, Y., Sedikides, C., Cai, H., Liu, Y., & Yang, Z. (2017). Disowning the self: The cultural value of modesty can attenuate self-positivity. *The Quarterly Journal of Experimental Psychology, 70*(6), 1023–1032. https://doi.org/10.1080/17470218.2015.1099711

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science, 12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Stryker, S., & Statham, A. (1985). *Symbolic interaction and role theory*. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 311–378). Random House.

Sutton, A. (2020). Living the good life: A meta-analysis of authenticity, well-being and engagement. *Personality and Individual Differences, 153*, Article 109645. https://doi.org/10.1016/j.paid.2019.109645

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology, 84*(2), 318–324. https://doi.org/10.1016/j.biopsycho.2010.03.009

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin, 110*(1), 67–85. https://doi.org/10.1037/0033-2909.110.1.67

Tredennick, H. & Thomson, J. A. K. (1976). *Aristotle ethics*. Penguin Classics.

Vazire, S., & Wilson, T. D. (Eds.) (2012). *Handbook of self-knowledge*. Guilford Press.

Vonasch, A. J., Reynolds, T., Winegard, B. M., & Baumeister, R. F. (2018). Death before dishonor: Incurring costs to protect moral reputation. *Social Psychological and Personality Science, 9*(5), 604–613. https://doi.org/10.1177/1948550617720271

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia, 45*(1), 174–194. https://doi.org/10.1016/j.neuropsychologia.2006.06.003

Webber, H. E., de Dios, C., Kessler, D. A., Schmitz, J. M., Lane, S. D., & Suchting, R. (2022). Late positive potential as a candidate biomarker of motivational relevance in substance use: Evidence from a meta-analysis. *Neuroscience & Biobehavioural Reviews, 141*, 104835. https://doi.org/10.1016/j.neubiorev.2022.104835

Wickham, H., & Wickham, H. (2016). *Data analysis* (pp. 189-201). Springer International Publishing.

Wieser, M. J., Gerdes, A. B., Büngel, I., Schwarz, K. A., Mühlberger, A., & Pauli, P. (2014). Not so harmless anymore: How context impacts the perception and electrocortical processing of neutral faces. *NeuroImage, 92*, 74–82. https://doi.org/10.1016/j.neuroimage.2014.01.022

Wilkinson, M. (2018). The transformation of religion and the self in the age of authenticity. *Pneuma, 40*(1–2), 91–108. https://doi.org/10.1163/15700747-04001032

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the Authenticity Scale. *Journal of Counseling Psychology, 55*(3), 385–399. https://doi.org/10.1037/0022-0167.55.3.385

Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., & Luo, Y. J. (2014). Three stages of emotional word processing: an ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience, 9*(12), 1897–1903. https://doi.org/10.1093/scan/nst188

Zhou, H., Guo, J., Ma, X., Zhang, M., Liu, L., Feng, L., Yang, J., Wang, Z., Wang, G., & Zhong, N. (2017). Self-reference emerges earlier than emotion during an implicit self-referential emotion processing task: Event-related potential evidence. *Frontiers in Human Neuroscience, 11*, Article 451. https://doi.org/10.3389/fnhum.2017.00451

Żochowska, A., Nowicka, M. M., Wójcik, M. J., & Nowicka, A. (2021). Self-face and emotional faces—are they alike?. *Social Cognitive and Affective Neuroscience, 16*(6), 593–607. https://doi.org/10.1093/scan/nsab020

Chapter 3 Authenticity is More Than Self-Enhancement: Behavioural and Neurophysiological Evidence

Chengli Huang, Nicholas J. Kelley, and Constantine Sedikides

Department of Psychology, University of Southampton

This is a jointly authored work, and the following is a statement of authors contributions:

Conceptualization: Constantine Sedikides, Nicholas J. Kelley, Chengli Huang

Methodology: Chengli Huang

Investigation: Chengli Huang, Emily Penney

Visualization: Chengli Huang

Funding acquisition: Chengli Huang

Project administration: Nicholas J. Kelley, Constantine Sedikides

Supervision: Nicholas J. Kelley, Constantine Sedikides

Writing – original draft: Chengli Huang

Writing – review & editing: Chengli Huang, Constantine Sedikides, Nicholas J. Kelley, Emily Penney

This work is now revising for resubmission at *Social Cognitive and Affective Neuroscience*.

Corresponding author: Nicholas J. Kelley, Centre for Research on Self and Identity, School of Psychology, University of Southampton; Southampton, SO17 1BJ, United Kingdom; email: n.j.kelley@soton.ac.uk.

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 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 both 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

3.1 Introduction

The self-concept is multifaceted (Baumeister, 1998; 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 construed the distinction as delineating the private and public self, respectively (Baumeister, 1986; Fenigstein, 2009; Tetlock & Manstead, 1985). Individuals are motivated to perceive the external expressions of the self (henceforth "the presented self") in a positive light, extolling strengths and underemphasizing weaknesses (Baumeister, 1982; Leary, 1995; Roth et al., 1986). 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 & Schlegel, 2024). We test these competing views by examining how negative (vs. positive) self-relevant information captures attention—both behaviourally and neurophysiologically—when the presented and authentic self are salient.

3.1.1 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 & Thomson, 1976), its meaning has been a matter of controversy (Baumeister, 2019; Beer & Brandler, 2021; 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 process unbiasedly self-relevant information (Kernis & Goldman, 2006). Aligned with this view, individuals high on authenticity are less defensive when faced with evidence that their behaviour is short of their ideals (Lakey et al., 2008). However, self-accuracy is difficult to attain or measure (Vazire & Wilson, 2012). Also, the more unbiased people believe they are, the more likely they are to report that they have more positive than negative characteristics, calling into question 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, attitudes, and behaviours (Kernis & 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 (Dormenan et al., 2020). However, individuals appraise their socially desirable behaviours as authentic regardless of whether these are consistent or inconsistent with their self-concept (Sheldon et al., 1997), and individuals consider their positive behaviours as more authentic than their negative behaviours (Jongman-Sereno & Leary, 2016). These latter findings pose a challenge to the self-consistency view, but are easily accounted for by the self-enhancement view.

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

3.1.2 The Presented Self

The outwardly articulated self, or 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, and access to social networks, professional opportunities (e.g., jobs, promotions, housing), or personal resources (e.g., friends, partners; Dores Cruz et al., 2021; Vonasch et al., 2018). Hence, self-presentation promotes a sanitized portrait of the individual, exaggerating, if not glorifying strengths, while minimizing, if not concealing, weaknesses (Baumeister, 1982; Hancock & 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 engaging in strategic self-presentation, denying negative traits and drawing attention to positive traits (Lee et al., 1999; Roth et al., 1986, 1988).

3.1.3 The Current Investigation

Emotional Stroop Task

To test self-enhancement versus self-consistency/self-accuracy views of authenticity, we recorded behavioural and neurophysiological (i.e., event-related potential or ERP) responses to emotionally charged self-evaluations in a modified Stroop task, the Emotional Stroop Task (Mathews & MacLeod, 1985; Watts et al., 1986; Figure 3.1), in which participants identify the ink colour of emotionally evocative (or neutral) words. The emotional Stroop effect refers to the slowdown in responding to the ink colour of negative (vs. positive or neutral) words (Bar-Haim et al., 2007; Phaf & Kan, 2007; Williams et al., 1996). This effect is a threat-driven slowdown (Algom et al., 2004) caused by automatic attention allocation to threatening stimuli at the expense of concurrent task demands (Öhman, 1993; Öhman et al., 2001).

The presented self is strongly influenced by self-enhancement motivation (Baumeister, 1982; Leary & Kowalski, 1990; Sedikides & Gregg, 2008). Consequently, we hypothesized that interference of negative (vs. positive) self-evaluations would be more pronounced for the presented self (i.e., an amplified emotional Stroop effect). Although the presented self is robustly and consistently driven by self-enhancement motivation, the evidence for the authentic self is more equivocal. On the one hand, self-consistency/self-accuracy motivations might drive the authentic self. If so, then the interference of negative (vs. positive) self-evaluations would be less pronounced for the authentic self (i.e., an attenuated emotional Stroop effect). On the other hand, self-enhancement motivation might drive the authentic self, much like the presented self. If so, then the interference of negative (vs. positive) self-evaluations would be on par with the effect observed for the presented self (i.e., an amplified emotional Stroop effect).

Event-Related Potential Assessment

We also recorded participants' electroencephalography (EEG) activity during the emotional Stroop task and focused on the ERPs at early-stage selective attention (P1; Batty & Taylor, 2003; Pourtois et al., 2005; Zhang et al., 2014), attention allocation to emotionally evocative stimuli (N170; Cai et al., 2016; Williams et al., 2006; Zhang et al., 2014), and early (P2; Fan et al., 2016; Fields & Kuperberg, 2012; San Martín et al., 2016) as well as late (P3; Gray et al., 2004; Tacikowski & 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 60ms 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 & Bublatzky, 2020) with larger P1 amplitudes evoked by negative compared to neutral stimuli (e.g., faces, words; Batty & Taylor, 2003; Luo et al., 2010; van Hooff et al., 2008), which indicates that the P1 can differentiate between threatening and non-threatening information (Zhang et al., 2014). Such early emotional response might signify rapid extraction of emotion-related information and might function—at least in part—independently of subsequent, emotional processes (N170; Vuilleumier & 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 (e.g., face; Rossion & Jacques, 2012), with larger N170 amplitudes representing the allocation of more attentional resources (Cai et al., 2016). Moreover, the N170 is modulated by the emotional stimuli (e.g., faces, words), with a substantially enhanced amplitude for emotional relative to neutral stimuli (Luo et al., 2010; Schindler et al., 2023; Williams et al., 2006; Zhang et al., 2014), especially for negative (vs. positive) ones (Cai et al., 2016; Montalan et al., 2008).

P2. The P2 is a positive deflection spanning from 150 to 250 ms over anterior-central region (Luck, 2014). Typically, it exhibits greater amplitudes in response to stimuli containing target features, indicating early selective attention toward task-relevant stimuli (Potts, 2004; Potts et al., 2006). This amplification is pronounced when the targets are relatively infrequent (Glazer & Nusslock, 2022; Luck & Hillyard, 1994). Furthermore, the P2 has been associated with emotional processing, suggesting its role in modulating selective attention influenced by emotional content (Hajcak et al., 2012; Kotz & Paulmann, 2011). 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 & Bublatzky, 2020). Unlike the earlier emotion-detection stages represented by the P1 and N170, the P2 is generally associated with higher-order, deeper, and conscious emotional processing (Nie et al., 2020; Prete et al., 2015, 2018). Moreover, research has implicated the P2 self-referential processing, despite the findings remaining inconsistent. Some studies reported reduced P2 amplitudes for self-related stimuli (Keyes et al., 2010; Liu et al., 2019), whereas others obtained increased P2 responses (Fan et al., 2016; Fields & Kuperberg, 2012; 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 in the EEG field for its iconic increased positivity following the presentation of infrequent and surprising (i.e., low probability) stimuli (Polich, 2007; Pritchard, 1981). Although both 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 (Barkaszi et al., 2013; Luck, 2014; Song et al., 2005). Furthermore, ERP studies of self-referential processing show that P3 is often associated with the discrimination of self from others; that is, 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, 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) as well as lack of electrophysiological studies on authenticity (P3).

3.2 Method

3.2.1 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 Southampton undergraduate psychology student volunteers. Based on a-priori criteria, we excluded 15 participants: five encountered EEG acquisition device failures (Tacikowski & 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 *SD*s (Cai et al., 2016), and six had more than 50% of their trials rejected due to artifacts in the EEG data (Imbir et al., 2021). The final sample comprised 147 participants (114 women, 31 men, 2 non-binary) aged between 18 and 46 years ($M = 19.56$, $SD = 2.87$). We did not collect ethnicity information, but over 90% of the university's undergraduates are White. Sensitivity analyses (G*Power; Faul et al., 2009) indicated 80% power to detect effects as small as Cohen's $f = 0.086$ ($\eta^2 = .007$).

3.2.2 Stimuli and Procedure

For the generation of 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, $p < 0.001$, Cohen's $d = 7.47$. We programmed and administered the experiment using PsychoPy (Version 2021.2.3; Peirce, 2007).

Participants completed a modified emotional Stroop task (Figure 3.1) in a quiet laboratory environment via computer and in the context of a larger study[7]. Specifically, 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 with their dominant hand as quickly and accurately as possible. For example, right-handed participants used: index finger (V/red), middle finger (B/blue), and ring finger (N/green). 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 the 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. Prior to the formal task, participants underwent 12 practice trials to familiarize themselves with the colour-key mapping.

Figure 3.1 *Trial Event Diagram*

Note: 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.

---

[7] Participants also completed a Flanker task and a Monetary Incentive Delay Task. At the end of the session, they filled out a battery of personality questionnaires unrelated to the current investigation.

3.2.3 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. Also, 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 kΩ. Further, we amplified and sampled the signals at 500 Hz with an online bandpass filter from 0.10–100 Hz.

In offline processing, we initially pre-processed the EEG data by using EEGLAB, an open-source toolbox running in the MATLAB environment (Delorme & 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), segmented them 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 (i.e., 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 artifacts using an independent component analysis algorithm (Delorme & Makeig, 2004). Also, we excluded bad segments where a voltage deviation on any channel of ± 75 μV.

Following best practices (Luck & Gaspelin, 2017) and similar lines of research, we quantified: (a) P1 as the average peak amplitude from 80–130 ms after stimulus onset over lateral occipital electrode cluster (i.e., O1, OZ, O2; Jetha et al., 2021; Wieser & Moscovitch, 2015); (b) N170 as the average peak amplitude from 140–200 ms after stimulus onset over lateral posterior electrode cluster (i.e., P3, P4, P7, P8; Cai et al., 2016; Keyes et al., 2010); (c) P2 as the average peak amplitude from 150–250 ms after stimulus onset over frontier-central-parietal electrode cluster (i.e., F3, Fz, F4, C3, Cz, C4, P3, Pz, P4; Fan et al., 2013, 2016); and (d) P3 as the average peak amplitude from 300–400 ms after stimulus onset over frontier-central-parietal electrode cluster (i.e., F3, Fz, F4, C3, Cz, C4, P3, Pz, P4; Gray et al., 2004; Riggins & Scott, 2020; Wada et al., 2019).

The main dependent variables were reaction times (RT) and ERPs (i.e., P1, N170, P2, P3). We took only correct responses into account (Montalan et al., 2008). We created a RT data processing pipeline (Morís Fernández & 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 *SD*s. Further, we averaged the ERPs for each of the six experimental conditions. We analysed the data in SPSS (Version 24), addressing multiple comparisons with Bonferroni corrections[8].

3.3 Results

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

3.3.1 Reaction Times

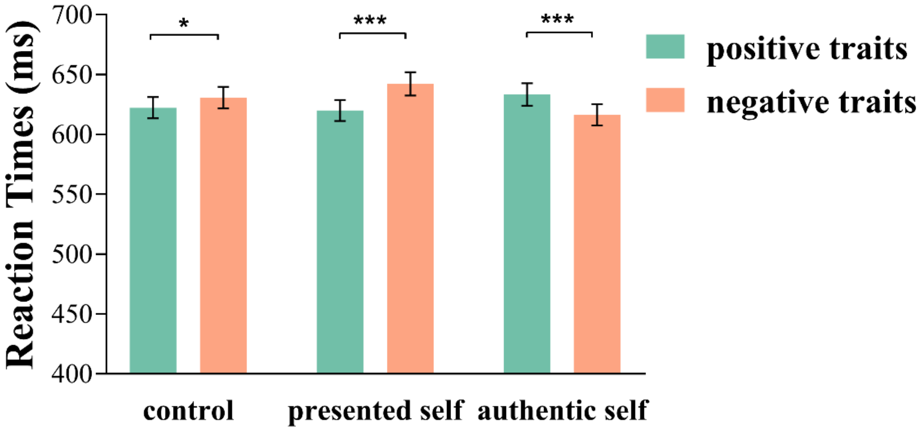
The main effect of self was significant, $F(2, 145) = 4.40$, $p = .014$, $\eta_p^2 = .06$. Participants responded faster on authentic-self ($M = 625.05$, $SD = 108.79$) than presented-self ($M = 631.28$, $SD = 109.59$) traits, $p = .028$, 95% $CI = [-11.97, -0.49]$. They did not differ in their speed 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 × Valence interaction was significant, $F(2, 145) = 30.88$, $p < .001$, $\eta_p^2 = .30$ (Figure 3.2). 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 reversed 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% $CI = [-23.41, -10.58]$.

---

[8] The corrected *p*-value is calculated by multiplying the original *p*-value by the number of comparisons. For instance, in post-hoc tests of a 3×2 interaction effect, the adjusted *p*-value becomes 15 times the original value.

Figure 3.2 *Reaction Times to Positive Traits and Negative Traits for Control, Presented Self, and Authentic Self*



Note. Error bars represent Standard Error of the Mean; $^{***}p < .001$, $^{*}p < .05$.

3.3.2 Event-Related Potentials

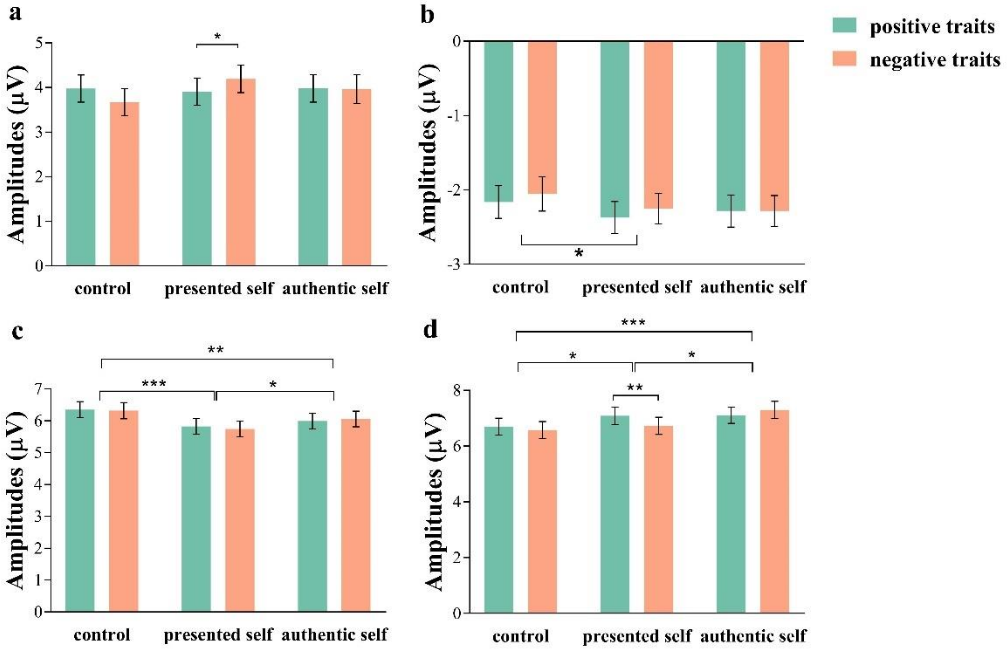
We observed a pronounced emotional Stroop effect for the presented self and a diminished emotional Stroop effect for the authentic self. We next turned to ERPs relevant to early-stage selective attention (P1), attention allocation to emotionally evocative stimuli (N170), and early (P1) and late (P3) stages of self-relevant information processing in search of an explanation for this behavioural effect (i.e., the Self × Valence interaction on RTs). Detailed descriptive statistics are presented in Figure 3.3 and Table 3.1.***P1.*** The Self × Valence interaction was significant, *F*(2, 145) = 3.75, *p* = .026, η_p^2 = .05 (Figure 3.4a). Negative traits (*M* = 4.19 μV, *SD* = 3.77 μV) elicited a larger P1 than positive traits (*M* = 3.91 μV, *SD* = 3.71 μV) for the presented self, *p* = .019, 95% *CI* = [0.05, 0.53]. The P1 did not differ between negative (*M* = 3.97 μV, *SD* = 3.92 μV) and positive (*M* = 3.98 μV, *SD* = 3.74 μV) traits for the authentic self, *p* = 0.919, 95% *CI* = [-0.27, 0.30], nor did it differ between negative (*M* = 3.67 μV, *SD* = 3.67 μV) and positive (*M* = 3.98 μV, *SD* = 3.68 μV) control traits, *p* = .052, 95% *CI* = [0.00, 0.62]. Neither the main effect of self, *F*(2, 145) = 1.50, *p* = .227, η_p^2 = .02, nor that of valence, *F*(2, 145) = 0.03, *p* = .855, η_p^2 < .001, was significant.

N170. The main effect of self was significant, *F*(2, 145) = 3.71, *p* = .027, η_p^2 = .05 (Figure 3.4b). The N170 was larger for the presented self (*M* = -2.31 μV, *SD* = 2.57 μV) than control (*M* = -2.11 μV, *SD* = 2.73 μV) traits, *p* = .027, 95% *CI* = [-0.39, -0.02]. Presented-self and authentic-self (*M* = -2.28 μV, *SD* = 2.56 μV) 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, η_p^2 = .01, and the Self × Valence interaction, *F*(2, 145) = 0.37, *p* = .695, η_p^2 = .01], were not significant.

P2. The main effect of self was significant, $F(2, 145) = 22.48$, $p < .001$, $\eta_p^2 = .24$ (Figure 3.4c). The P2 was larger on control traits ($M = 6.33$ μV, $SD = 3.08$ μV) compared to both authentic-self ($M = 6.02$ μV, $SD = 3.08$ μV) and presented-self ($M = 5.79$ μV, $SD = 3.01$ μV) traits, $p = .004$, 95% $CI = [0.08, 0.54]$ and $p < .001$, 95% $CI = [0.35, 0.75]$, respectively. 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, $p = .017$, 95% $CI = [0.03, 0.44]$. 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, $F(1, 146) = 0.05$, $p = .832$, $\eta_p^2 < .001$, nor the Self × Valence interaction, $F(2, 145) = 0.27$, $p = .763$, $\eta_p^2 = .01$, was significant.

P3. The main effect of self was significant, $F(2, 145) = 15.12$, $p < .001$, $\eta_p^2 = .17$. The P3 was larger on authentic-self ($M = 7.20$ μV, $SD = 3.64$ μV) than control ($M = 6.64$ μV, $SD = 3.67$ μV), $p < .001$, 95% $CI = [0.32, 0.81]$ and presented-self ($M = 6.91$ μV, $SD = 3.76$ μV), $p = .033$, 95% $CI = [0.02, 0.57]$ traits. The P3 was larger on presented-self compared to control traits, $p = .036$, 95% $CI = [0.01, 0.53]$. The finding that P3 was larger for self-relevant stimuli is compatible with the literature (Gray et al., 2004; Miyakoshi et al., 2007; Tacikowski & Nowicka, 2010). The main effect of valence was not significant, $F(1, 146) = 1.34$, $p = .249$, $\eta_p^2 = .01$. Crucially, the Self × Valence interaction was significant, $F(2, 145) = 3.91$, $p = .022$, $\eta_p^2 = .05$ (Figure 3.4d). The P3 did not differ between negative and positive traits in the control condition ($p = 0.422$, 95% $CI = [-0.44, 0.19]$) or in the case of the authentic self ($p = 0.188$, 95% $CI = [-0.48, 0.10]$). However, the P3 was larger for positive ($M = 7.09$ μV, $SD = 3.82$ μV) than negative ($M = 6.73$ μV, $SD = 3.70$ μV) traits in the case of the presented self, $p = 0.010$, 95% $CI = [0.09, 0.63]$.

Figure 3.3 *Peak Amplitude of ERPs to Positive and Negative Traits for Control, Presented Self, and Authentic Self*



Note. (**a**) P1 peak amplitude to positive and negative traits for control, authentic self and presented self. (**b**) N170 peak amplitude to positive and negative traits for control, authentic self and presented self. (**c**) P2 peak amplitude to positive and negative traits for control, authentic self and presented self. (**d**) P3 peak amplitude to positive and negative traits for control, authentic self and presented self. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Table 3.1 *Means and Standard Deviations for the Peak Amplitudes (μV) of P1, N170, P2, and P3*

P1

	Authentic Self	Presented Self	Control	Average
Positive Traits	3.98 (3.74)	3.91 (3.71)	3.98 (3.68)	3.95 (3.71)
Negative Traits	3.97 (3.92)	4.19 (3.77)	3.67 (3.67)	3.94 (3.78)
Average	3.97 (3.83)	4.05 (3.74)	3.82 (3.67)	

N170

	Authentic Self	Presented Self	Control	Average
Positive Traits	-2.28 (2.61)	-2.37 (2.62)	-2.16 (2.66)	-2.27 (2.63)
Negative Traits	-2.28 (2.52)	-2.25 (2.51)	-2.05 (2.79)	-2.19 (2.61)
Average	-2.28 (2.56)	-2.31 (2.57)	-2.11 (2.73)	

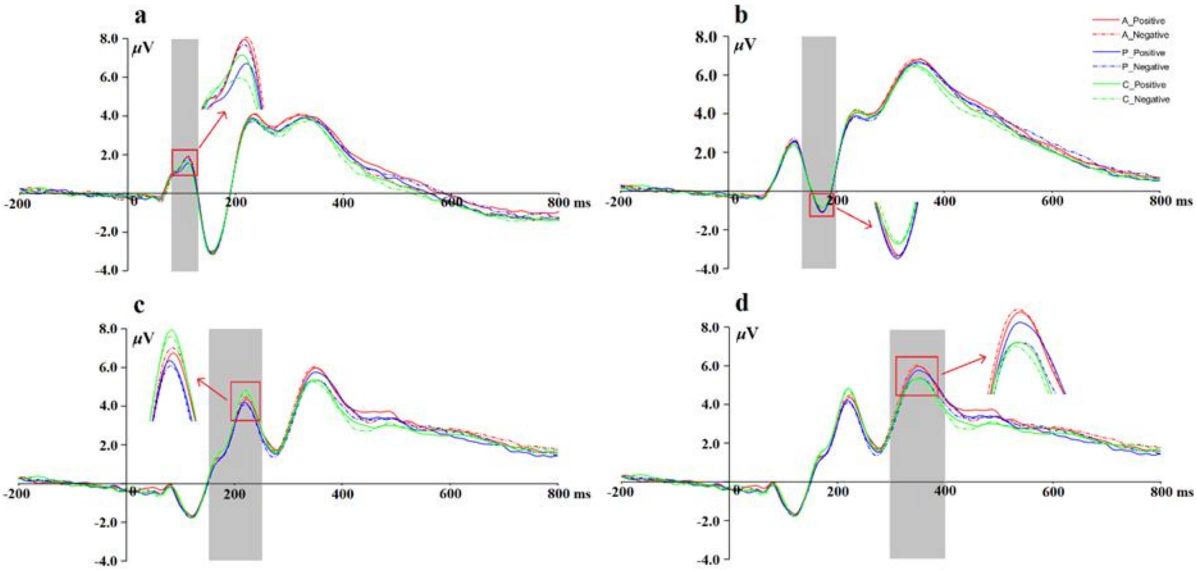
P2

	Authentic Self	Presented Self	Control	Average
Positive Traits	5.99 (2.92)	5.83 (3.01)	6.35 (3.07)	6.06 (3.00)
Negative Traits	6.06 (2.97)	5.75 (3.01)	6.32 (3.09)	6.04 (3.02)
Average	6.02 (2.95)	5.79 (3.01)	6.33 (3.08)	

P3

	Authentic Self	Presented Self	Control	Average
Positive Traits	7.11 (3.55)	7.09 (3.82)	6.70 (3.62)	6.97 (3.66)
Negative Traits	7.30 (3.72)	6.73 (3.70)	6.58 (3.73)	6.87 (3.72)
Average	7.20 (3.64)	6.91 (3.76)	6.64 (3.67)	

***Figure 3.4** Grand Averages for ERPs in the Authentic Self, Presented Self, and Control Conditions*



Note. (**a**) Grand averages for P1 in the authentic-self, presented-self, and control conditions in judging the colours of sentences. (**b**) Grand averages for N170 in the authentic-self, presented-self, and control conditions in judging the colour of sentences. (**c**) Grand averages for P2 in the authentic-self, presented-self, and control conditions in judging the colour of sentences. (**d**) Grand averages for P3 in the authentic-self, presented-self, and control conditions in judging the colour of sentences. A: authentic-self traits; P: presented-self traits; C: control traits. The grand-averaged waveforms represent the grand average across component-specific electrode clusters: for the P1 component, signals were averaged across 3 occipital sites (O1, O2, and Oz); for the N170 component, across 4 lateral posterior sites (P3, P4, P7, P8); for the P2 component, across 9 frontier-central-parietal sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4); and for the P3 component, across 9 frontier-central-parietal sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4).

3.4 Discussion

We decomposed the self into two mental representations, the presented and authentic selves. 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 (i.e., slowdown for negative compared to positive information). This effect was amplified on presented-self trials, which is attributable to the potent self-enhancement motivation driving self-presentation (Alicke & Sedikides, 2009; Schlenker & Pontari, 2000; 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 & Bublatzky, 2020), especially to threat-related information (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 toward negative self-descriptive stimuli during the early stages of attentional allocation. In contrast, the authentic self acknowledges both strengths and weaknesses (Kernis & 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 typically arising later (Herbert et al., 2011; Schäfer et al., 2020; Zhu et al., 2016). Hence, our P1 findings may represent 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 self-concept and emotional processing in this early attention stage. Although the N170 component did not differentiate between positive and negative self-descriptiveness at this stage, 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 also 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 (which was in line with the self-consistency/self-accuracy views), the P3 was larger for positive than negative traits for the presented self (which was in line with 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 (Barkaszi et al., 2013; Luck, 2014; Song et al., 2005), 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; McConnell, 2011; Sedikides & Gregg, 2003). 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 later stage, positivity is novel and salient to the presented self. Notably, 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).

The main effects of self that we observed 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 & Nowicka, 2010), as mentioned above. As such, 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 these findings by illustrating on the effects of self-reference on the P3 are stronger for the authentic self. Hence, deeper processing of self-relevant stimuli (Levorsen et al., 2023; Porter et al., 2021) may be driven more by the authentic self than the presented self. Future investigations will do well to address this possibility.

To conclude, 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.

References

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: a generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General*, *133*(3), 323–338. https://doi.org/10.1037/0096-3445.133.3.323

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology, 20*(1), 1–48. https://doi.org/10.1080/10463280802613866

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*(3), 272–279. https://doi.org/10.1037/h0025907

Bailey, E. R., & Iyengar, S. S. (2023). Positive—more than unbiased—self-perceptions increase subjective authenticity. *Journal of Personality and Social Psychology, 125*(6), 1351–1372. https://doi.org/10.1037/pspa0000359

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Barkaszi, I., Czigler, I., & Balázs, L. (2013). Stimulus complexity effects on the event-related potentials to task-irrelevant stimuli. *Biological Psychology*, *94*(1), 82–89. https://doi.org/10.1016/j.biopsycho.2013.05.007

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, *17*(3), 613–620. https://doi.org/10.1016/S0926-6410(03)00174-5

Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 680–740). McGraw-Hill.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*(1), 3–26. https://doi.org/10.1037/0033-2909.91.1.3

Baumeister, R. F., (1986). *Public self and private self*. Springer.

Baumeister, R. F. (2019). Stalking the true self through the jungles of authenticity: Problems, contradictions, inconsistencies, disturbing findings—And a possible way forward. *Review of General Psychology, 23*(1), 143–154. https://doi.org/10.1177/1089268019829472

Beer, J. S., & Brandler, S. (2021). Why don't we know more about the minds of authentic people? *Social and Personality Psychology Compass, 15*(10), Article e12637. https://doi.org/10.1111/spc3.12637

Bench, S. W., Schlegel, R. J., Davis, W. E., & Vess, M. (2015). Thinking about change in the self and others: The role of self-discovery metaphors and the true self. *Social Cognition, 33*(3), 169–185. https://doi.org/10.1521/soco.2015.33.3.2

Cai, H., Wu, L., Shi, Y., Gu, R., & Sedikides, C. (2016). Self-enhancement among Westerners and Easterners: A cultural neuroscience approach. *Social Cognitive and Affective Neuroscience*, *11*(10), 1569–1578. https://doi.org/10.1093/scan/nsw072

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. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dores Cruz, T. D., Thielmann, I., Columbus, S., Molho, C., Wu, J., Righetti, F., de Vries, R. E., Koutsoumpis, A., Van Lange, P. A. M., Beersma, B., & Balliet. D. (2021). Gossip and reputation in everyday life. *Philosophical Transactions of the Royal Society B, 376*(1838), Article 20200301. https://doi.org/10.1098/rstb.2020.0301

Dormanen, R., Sanders, C. S., Maffly-Kipp, J., Smith, J. L., & Vess, M. (2020). Assimilation undercuts authenticity: A consequence of women's masculine self-presentation in masculine contexts. *Psychology of Women Quarterly, 44*(4), 488–502. https://doi.org/10.1177/0361684320947648

Fan, W., Chen, J., Wang, X.-Y., Cai, R., Tan, Q., Chen, Y., Yang, Q., Zhang, S., Wu, Y., Yang, Z., Wang, X.-A., & Zhong, Y. (2013). Electrophysiological correlation of the degree of self-reference effect. *PLoS One*, *8*(12), Article e80289. https://doi.org/10.1371/journal.pone.0080289

Fan, W., Zhong, Y., Yang, Z., Li, J., Ouyang, Y., Cai, R., Li, H., & Fu, X. (2016). The degree of self-reference effect of extroversion individuals. *Acta Psychologica Sinica*, *48*(8), 1002–1012. https://doi.org/10.3724/SP.J.1041.2016.01002

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behaviour Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fenigstein, A. (2009). Private and public self-consciousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behaviour* (pp. 495–511). The Guilford Press.

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage*, *62*(1), 562–574. https://doi.org/10.1016/j.neuroimage.2012.05.003

Gillath, O., Sesko, A. K., Shaver, P. R., & Chun, D. S. (2010). Attachment, authenticity, and honesty: Dispositional and experimentally induced security can reduce self- and other-deception. *Journal of Personality and Social Psychology, 98*(5), 841–855. https://doi.org/10.1037/a0019206

Glazer, J., & Nusslock, R. (2022). Outcome valence and stimulus frequency affect neural responses to rewards and punishments. *Psychophysiology*, *59*(3), Article e13981. https://doi.org/10.1111/psyp.13981

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, *40*(2), 216–224. https://doi.org/10.1016/S0022-1031(03)00092-1

Guenther, C. L., Zhang, Y., & Sedikides, C. (2024). The authentic self is the self-enhancing self: A self-enhancement framework of authenticity. *Personality and Social Psychology Bulletin*, *50*(8), 1182–1196. https://doi.org/10.1177/01461672231160653

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In *The Oxford handbook of event-related potential components* (Vol. 441, pp. 441–472). Oxford University Press.

Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication, 59*(2), 367–386. https://doi.org/10.1111/j.1460-2466.2009.01420.x

Herbert, C., Herbert, B. M., Ethofer, T., & Pauli, P. (2011). His or mine? The time course of self–other discrimination in emotion processing. *Social Neuroscience*, *6*(3), 277–288. https://doi.org/10.1080/17470919.2010.523543

Hicks, J. A., Schlegel, R. J., & Newman, G. E. (2019). Introduction to the special issue: Authenticity: Novel insights into a valued, yet elusive, concept. *Review of General Psychology*, *23*(1), 3–7. https://doi.org/10.1177/1089268019829474

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *353*(1373), 1257–1270. https://doi.org/10.1098/rstb.1998.0281

Imbir, K. K., Duda-Goławska, J., Pastwa, M., Jankowska, M., & Żygierewicz, J. (2021). Event-related potential correlates of valence, arousal, and subjective significance in processing of an emotional Stroop task. *Frontiers in Human Neuroscience*, *15*, Article 617861. https://doi.org/10.3389/fnhum.2021.617861

James, W. (1890). The principles of psychology. *Henry Holt*.

Jetha, M. K., Segalowitz, S. J., & Gatzke-Kopp, L. M. (2021). The reliability of visual ERP components in children across the first year of school. *Developmental Psychobiology*, *63*(6), Article e22150. https://doi.org/10.1002/dev.22150

Jongman-Sereno, K. P., & Leary, M. R. (2016). Self-perceived authenticity is contaminated by the valence of one's behaviour. *Self and Identity*, *15*(3), 283–301. https://doi.org/10.1080/15298868.2015.1128964

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. *Advances in Experimental Social Psychology, 38*, 283–357. https://doi.org/10.1016/S0065-2601(06)38006-9

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and Cognition*, *72*(2), 244–254. https://doi.org/10.1016/j.bandc.2009.09.006

Kihlstrom, J. F., Cantor, N., Albright, J. S., Chew, B. R., Klein, S. B., & Niedenthal, P. M. (1988). Information processing and the study of the self. *Advances in Experimental Social Psychology, 21,* 145–178. https://doi.org/10.1016/S0065-2601(08)60226-9

Kotz, S. A., & Paulmann, S. (2011). Emotion, language, and the brain. *Language and Linguistics Compass*, *5*(3), 108–125. https://doi.org/10.1111/j.1749-818X.2010.00267.x

Lakey, C. E., Kernis, M. H., Heppner, W. L., & Lance, C. E. (2008). Individual differences in authenticity and mindfulness as predictors of verbal defensiveness. *Journal of Research in Personality, 42*, 230–238. https://doi.org/10.1016/j.jrp.2007.05.002

Leary, M. R. (1995). *Self-presentation: Impression Management and Interpersonal Behaviour*. Westview.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin, 107*(1), 34–47. https://doi.org/10.1037/0033-2909.107.1.34

Lee, S. J., Quigley, B. M., Nesler, M. S., Corbett, A. B., & Tedeschi, J. T. (1999). Development of a self-presentation tactics scale. *Personality and Individual Differences*, *26*(4), 701–722. https://doi.org/10.1016/S0191-8869(98)00178-0

Levorsen, M., Aoki, R., Matsumoto, K., Sedikides, C., & Izuma, K. (2023). The self-concept is represented in the media prefrontal cortex in terms of self-importance. *Journal of Neuroscience, 43*(20), 3675–3686. https://doi.org/10.1523/JNEUROSCI.2178-22.2023

Liu, L., Li, W., Li, J., Lou, L., & Chen, J. (2019). Temporal features of psychological and physical self-representation: An ERP study. *Frontiers in Psychology, 10*, Article 785. https://doi.org/10.3389/fpsyg.2019.00785

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed). MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x

Luo, W., Feng, W., He, W., Wang, N., & Luo, Y. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, *49*(2), 1857–1867. https://doi.org/10.1016/j.neuroimage.2009.09.018

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*(4), 623–663. https://doi.org/10.1037/0022-0663.82.4.623

Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, *23*(5), 563–569. https://doi.org/10.1016/0005-7967(85)90104-4

McConnell, A. R. (2011). The multiple self-aspects framework: Self-concept representation and its implications. *Personality and Social Psychology Review, 15*(1), 3–27. https://doi.org/10.1177/1088868310371101

Miyakoshi, M., Nomura, M., & Ohira, H. (2007). An ERP study on self-relevant object recognition. *Brain and Cognition*, *63*(2), 182–189. https://doi.org/10.1016/j.bandc.2006.12.001

Montalan, B., Caharel, S., Personnaz, B., Le Dantec, C., Germain, R., Bernard, C., Lalonde, R., & Rebaï, M. (2008). Sensitivity of N170 and late positive components to social categorization and emotional valence. *Brain Research*, *1233*, 120–128. https://doi.org/10.1016/j.brainres.2008.07.055

Morís Fernández, L., & Vadillo, M. A. (2020). Flexibility in reaction time analysis: Many roads to a false positive? *Royal Society Open Science*, *7*(2), Article 190831. https://doi.org/10.1098/rsos.190831

Mueller, A. K., Falter, C. M., & Tucha, O. (2013). Distinct attentional mechanisms for social information processing as revealed by event-related potential studies. *International Journal of Brain and Cognitive Sciences*, *2*(2), 23–37. https://doi.org/10.5923/j.ijbcs.20130202.03

Nie, L., Zhou, S., Wang, Z., Wang, M., Wang, Y., & Zheng, Y. (2020). Impaired perceptual processing of facial expression categorization in social anhedonia. *Psychophysiology*, *57*(12), Article e13682. https://doi.org/10.1111/psyp.13682

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 511–536). Guilford Press.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184–187. https://doi.org/10.1016/0013-4694(89)90180-6

Phaf, R. H., & Kan, K.-J. (2007). The automaticity of emotional Stroop: A meta-analysis. *Journal of Behaviour Therapy and Experimental Psychiatry*, *38*(2), 184–199. https://doi.org/10.1016/j.jbtep.2006.10.008

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Porter, N. A., Fields, E. C., Moore, I. L., & Gutchess, A. (2021). Late frontal positivity effects in self-referential memory: Unique to the self? *Social Neuroscience*, *16*(4), 406–422. https://doi.org/10.1080/17470919.2021.1929460

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain and Cognition*, *56*(1), 5–13. https://doi.org/10.1016/j.bandc.2004.03.006

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, *18*(7), 1112–1119. https://doi.org/10.1162/jocn.2006.18.7.1112

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping*, *26*(1), 65–79. https://doi.org/10.1002/hbm.20130

Prete, G., Capotosto, P., Zappasodi, F., Laeng, B., & Tommasi, L. (2015). The cerebral correlates of subliminal emotions: an electroencephalographic study with emotional hybrid faces. *European Journal of Neuroscience*, *42*(11), 2952–2962. https://doi.org/10.1111/ejn.13078

Prete, G., Capotosto, P., Zappasodi, F., & Tommasi, L. (2018). Contrasting hemispheric asymmetries for emotional processing from event-related potentials and behavioural responses. *Neuropsychology*, *32*(3), 317–328. http://dx.doi.org/10.1037/neu0000443

Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, *89*(3), 506–540. https://doi.org/10.1037/0033-2909.89.3.506

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Riggins, T., & Scott, L. S. (2020). P300 development from infancy to adolescence. *Psychophysiology*, *57*(7), Article e13346. https://doi.org/10.1111/psyp.13346

Rogers, C. (1951). *Client-centreed therapy: its current practice, implications, and theory*. Houghton Mifflin Company.

Rossion, B., & Jacques, C. (2012). The N170: Understanding the time course of face perception in the human brain. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 115–141). Oxford University Press.

Roth, D. L., Harris, R. N., & Snyder, C. R. (1988). An individual differences measure of attributive and repudiative tactics of favourable self-presentation. *Journal of Social and Clinical Psychology*, *6*(2), 159–170. https://doi.org/10.1521/jscp.1988.6.2.159

Roth, D. L., Snyder, C. R., & Pace, L. M. (1986). Dimensions of favourable self-presentation. *Journal of Personality and Social Psychology, 51*(4), 867–874. https://doi.org/10.1037/0022-3514.51.4.867

San Martín, R., Kwak, Y., Pearson, J. M., Woldorff, M. G., & Huettel, S. A. (2016). Altruistic traits are predicted by neural responses to monetary outcomes for self vs charity. *Social Cognitive and Affective Neuroscience*, *11*(6), 863–876. https://doi.org/10.1093/scan/nsw026

Sande, G. N. (1990). The multifaceted self. In J. M. Olson & M. P. Zanna (Eds.), *Self-inference processes: The Ontario symposium* (Vol. 6, pp. 1-16). Lawrence Erlbaum.

Schäfer, S., Wentura, D., & Frings, C. (2020). Creating a network of importance: The particular effects of self-relevance on stimulus processing. *Attention, Perception, & Psychophysics*, *82*(7), 3750–3766. https://doi.org/10.3758/s13414-020-02070-7

Schindler, S., Bruchmann, M., & Straube, T. (2023). Beyond facial expressions: a systematic review on effects of emotional relevance of faces on the N170. *Neuroscience & Biobehavioural Reviews*, *153*, Article 105399. https://doi.org/10.1016/j.neubiorev.2023.105399

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, *130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Schlenker, B. R., & Pontari, B. A. (2000). The strategic control of information: Impression management and self-presentation in daily life. In A. Tesser, R. B. Felson, & J. M. Suls (Eds.), *Psychological perspectives on self and identity* (pp. 199–232). American Psychological Association. https://doi.org/10.1037/10357-008

Sedikides, C., Hoorens, V., & Dufner, M. (2015). Self-enhancing self-presentation: Interpersonal, relational, and organizational implications. In F. Guay, D. M. McInerney, R. Craven, & H. W. Marsh (Eds.), *Self-concept, motivation and identity: Underpinning success with research and practice*. International Advances in Self Research (Vol. 5, pp. 29–55). Information Age Publishing.

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one's faults. *European Review of Social Psychology, 27*(1), 1–62. https://doi.org/10.1080/10463283.2016.1183913

Sedikides, C., & Gregg, A. P. (2003). Portraits of the self. In M. A. Hogg & J. Cooper (Eds.), *Sage handbook of social psychology* (pp. 110–138). Sage Publications.

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3*(2), 102–116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

Sedikides, C., Lenton, A. P., Slabu, L., & Thomaes, S. (2019). Sketching the contours of state authenticity. *Review of General Psychology*, *23*(1), 73–88. https://doi.org/10.1037/gpr0000156

Sedikides, C., & Schlegel, R. (2024). Distilling the concept of authenticity. *Nature Reviews Psychology*. Advance online publication. https://doi.org/10.1038/s44159-024-00323-y

Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of Personality and Social Psychology*, *73*(6), 1380–1393. https://doi.org/10.1037/0022-3514.73.6.1380

Song, Y., Ding, Y., Fan, S., Qu, Z., Xu, L., Lu, C., & Peng, D. (2005). Neural substrates of visual perceptual learning of simple and complex stimuli. *Clinical Neurophysiology*, *116*(3), 632–639. https://doi.org/10.1016/j.clinph.2004.09.019

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, *12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318–324. https://doi.org/10.1016/j.biopsycho.2010.03.009

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*(1), 67–85. https://doi.org/10.1037/0033-2909.110.1.67

Tetlock, P. E., & Manstead, A. S. (1985). Impression management versus intrapsychic explanations in social psychology: A useful dichotomy? *Psychological Review, 92*(1), 59–77. https://doi.org/10.1037/0033-295X.92.1.59

Tredennick, H. & Thomson, J. A. K. (1976). *Aristotle ethics*. Penguin Classics.

Tunnel, G. (1984). The discrepancy between private and public selves: Public self-consciousness and its correlates. *Journal of Personality Assessment, 48*(5), 549–555. https://doi.org/10.1207/s15327752jpa4805_15

Van Hooff, J. C., Dietz, K., Sharma, D., & Bowman, H. (2008). Neural correlates of intrusion of emotion words in a modified Stroop task. *International Journal of Psychophysiology*, *67*(1), 23–34. https://doi.org/10.1016/j.ijpsycho.2007.09.002

Vazire, S., & Wilson, T. D. (Eds.). (2012). *Handbook of self-knowledge*. Guilford Press.

Vonasch, A. J., Reynolds, T., Winegard, B. M., & Baumeister, R. F. (2018). Death before dishonor: Incurring costs to protect moral reputation. *Social Psychological and Personality Science, 9*(5), 604–613. https://doi.org/10.1177/1948550617720271

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194. https://doi.org/10.1016/j.neuropsychologia.2006.06.003

Wada, M., Kurose, S., Miyazaki, T., Nakajima, S., Masuda, F., Mimura, Y., Nishida, H., Ogyu, K., Tsugawa, S., Mashima, Y., Plitman, E., Chakravarty, M. M., Mimura, M., & Noda, Y. (2019). The P300 event-related potential in bipolar disorder: A systematic review and meta-analysis. *Journal of Affective Disorders*, *256*, 234–249. https://doi.org/10.1016/j.jad.2019.06.010

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology*, *77*(1), 97–108. https://doi.org/10.1111/j.2044-8295.1986.tb01985.x

Wieser, M. J., & Moscovitch, D. A. (2015). The effect of affective context on visuocortical processing of neutral faces in social anxiety. *Frontiers in Psychology*, *6*, Article 167112. https://doi.org/10.3389/fpsyg.2015.01824

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24. https://doi.org/10.1037/0033-2909.120.1.3

Williams, L. M., Palmer, D., Liddell, B. J., Song, L., & Gordon, E. (2006). The ‘when’ and ‘where’ of perceiving signals of threat versus non-threat. *NeuroImage*, *31*(1), 458–467. https://doi.org/10.1016/j.neuroimage.2005.12.009

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the Authenticity Scale. *Journal of Counseling Psychology*, *55*(3), 385–399. https://doi.org/10.1037/0022-0167.55.3.385

Yang, J., Qi, M., & Guan, L. (2014). Self-esteem modulates the latency of P2 component in implicit self-relevant processing. *Biological Psychology*, *97*, 22–26. http://dx.doi.org/10.1016/j.biopsycho.2014.01.004

Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., Li, H., & Luo, Y. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, *9*(12), 1897–1903. https://doi.org/10.1093/scan/nst188

Zhu, M., Luo, J., Zhao, N., Hu, Y., Yan, L., & Gao, X. (2016). The temporal primacy of self-related stimuli and negative stimuli: An ERP-based comparative study. *Social Neuroscience, 11*(5), 507–514. https://www.tandfonline.com/doi/full/10.1080/17470919.2015.1114021

Chapter 4 Self-Control, Reward Responsivity, and Authenticity

Chengli Huang, Nicholas J. Kelley, and Constantine Sedikides

Department of Psychology, University of Southampton

Chapter 4 includes 2 parts. Part A explores whether exercising self-control increases responsivity to hedonic and eudaimonic rewards. Part B re-examined this process by examining whether trait authenticity moderates the effect of self-control exertion on reward responsivity.

This is a jointly authored work, and the following is a statement of authors contributions:

Conceptualization: Constantine Sedikides, Nicholas J. Kelley, Chengli Huang

Methodology: Chengli Huang

Investigation: Chengli Huang, Zhiwei Zhou

Visualization: Chengli Huang, Douglas J Angus

Funding acquisition: Chengli Huang

Project administration: Nicholas J. Kelley, Constantine Sedikides

Supervision: Nicholas J. Kelley, Constantine Sedikides

Writing – original draft: Chengli Huang

Writing – review & editing: Chengli Huang, Constantine Sedikides, Nicholas J. Kelley, Zhiwei Zhou.

Publication details: Huang, C., Zhou, Z., Angus, D. J., Sedikides, C., & Kelley, N. J. (2025). Exercising Self-Control Increases Responsivity to Hedonic and Eudaimonic Rewards. Social Cognitive and Affective Neuroscience. 20(1), Article nsaf016. https://doi.org/10.1093/scan/nsaf016

Corresponding author: Nicholas J. Kelley, Centre for Research on Self and Identity, School of Psychology, University of Southampton; Southampton, SO17 1BJ, United Kingdom; email: n.j.kelley@soton.ac.uk.

4.1 Exercising Self-Control Increases Responsivity to Hedonic and Eudaimonic Rewards

Abstract

The reward responsivity hypothesis of self-control proposes that, irrespective of self-control success, exercising self-control is aversive and engenders negative affect. To countermand this discomfort, reward-seeking behaviour may be amplified after bouts of self-control, bringing individuals back to a mildly positive baseline state. Previous studies indicated that effort—an integral component of self-control—can increase reward responsivity. We sought to test and extend the reward responsivity hypothesis by asking if exercising self-control increases a neural marker of reward responsivity (Reward Positivity) differentially for hedonic rewards or eudaimonic rewards. We instructed participants (*N* = 114) to complete a speeded reaction time task where they exercised self-control (incongruent Stroop trials) or not (congruent Stroop trials), and then had the opportunity to win money for themselves (hedonic rewards) or a charity (eudaimonic rewards) while EEG was recorded. Consistent with the reward responsivity hypothesis, participants evinced a larger RewP after exercising self-control (vs. not exercising self-control). Participants also showed a larger RewP for hedonic over eudaimonic rewards. Self-control and reward type did not interactively modulate RewP, suggesting that self-control increases the reward responsivity in a domain-general manner. The findings provide a neurophysiological mechanism for the reward responsivity hypothesis of self-control and promise to revitalize the relevant literature.

Keywords: self-control, effort, reward positivity, hedonic rewards, eudaimonic rewards

4.1.1 Introduction

The ability to override or alter motivated responses (i.e., self-control) is crucial for goal-directed behaviour and contributes to many consequential outcomes including physical health, psychological well-being, ethical decision making, and successful interpersonal relationships (Vohs & Baumeister, 2016). Conversely, failures in self-control have negative consequences in these and other domains. Self-control has thus been of keen interest to psychologists, neuroscientists, philosophers, and the public. The most influential model of self-control, the resource model (Baumeister et al., 1998), though generative, has come under intense scrutiny in recent years (Carter et al., 2015; Carter & McCullough, 2014; Hagger et al., 2016; Vohs et al., 2021). In response to empirical challenges to this model, the reward responsivity hypothesis of self-control proposes that exercising self-control does not influence behaviour generally, but influences the reward system specifically (Kelley et al., 2019). The purpose of this study is to (1) examine the neural basis of the reward responsivity hypothesis of self-control by assessing how self-control exertion impacts the Reward Positivity, and (2) expand this hypothesis by testing the extent to which exercising self-control influences the reward system differently for hedonic versus eudaimonic rewards.

The Resource Model of Self-Control

Self-control has been extensively investigated through the lens of the resource model (Baumeister et al., 1998). For approximately 25 years, this model has enjoyed widespread influence in social/personality psychology and psychological science in general. According to it, the capacity to override or alter one’s responses depends on limited inner resource or strength (Baumeister et al., 1998, 2007). Acts of self-control are theorized to consume (i.e., deplete) this strength, resulting in temporary decline in the capacity for self-control (i.e., ego depletion). In support, numerous studies have found that engaging in a taxing (or depleting) self-control task undermines performance on subsequent demanding tasks (Baumeister et al., 2007; 2018; 2023). Mechanistically, these effects were thought to be driven by glucose (Gailliot et al., 2007), though meta-analyses have cast doubt on this interpretation (Dang, 2016).

Nevertheless, empirical challenges, controversies, and debates related to the validity of the resource model have arisen. An initial meta-analysis of the relevant literature reported evidence for consistent and large effects (Hagger et al., 2010), but more recent meta-analyses have concluded that the effect is negligible after adjusting for publication bias (Carter et al., 2015; Carter & McCullough, 2014). Multi-laboratory experiments obtained non-significant aftereffects of self-control exertion (Hagger et al., 2016; Vohs et al., 2021), whereas other preregistered experiments obtained statistically significant, albeit smaller than expected effects (Dang et al., 2017; Garrison et al., 2019). Collectively, the mechanisms and aftereffects of self-control exertion remain poorly understood.

Reward Responsivity Hypothesis of Self-Control

The reward responsivity hypothesis of self-control (Kelley et al., 2019) was a response to controversies and challenges to the resource model. According to this hypothesis, irrespective of self-control success, exercising self-control is aversive and engenders negative affect (David et al., 2024; Kurzban, 2016). To countermand this discomfort, reward seeking behaviour may be augmented after bouts of self-control, bringing individuals back to a mildly positive baseline state. In other words, the reward responsivity hypothesis of self-control states that exercising self-control does not influence behaviour generally, but it influences specifically the reward system (Kelley et al., 2019). The latter aligns with the core tenet of the process model of self-control, which suggests that self-control shifts attention and motivation toward rewards (Inzlicht et al., 2014). In contrast, the resource model does not explicitly predict that exercising self-control increases subsequent reward-related impulse strength. Instead, it posits that engaging in taxing self-control tasks depletes limited resources, leading to impaired performance on subsequent demanding tasks in general. Yet, several studies inspired by the resource model have reported evidence that exercising self-control increases subsequent reward-seeking behaviour, including eating, spending, and sexual behaviour (Baumeister et al., 2007). These behavioural outcomes could be due to a reduction in the capacity for self-control (as the resource model initially assumed) or increases in reward responsivity (as the reward responsivity hypothesis proposed). Several studies in line with the reward responsivity hypothesis of self-control have circumvented this interpretational ambiguity by instructing participants to complete reward-related tasks requiring little to no self-control. For example, Finley and Schmeichel (2019) observed that self-control exertion enhances self-reported approach motivation and positive emotional reactivity. Our primary goal here was to examine whether exercising self-control would enhance a neural marker of reward responsivity: an ERP, known as the Reward Positivity (RewP).

Self-Control and Reward Positivity

The RewP (Carlson et al., 2011; Foti et al., 2011; Walsh & Anderson, 2012) is sensitive to feedback signalling the outcome of an action. The RewP peaks approximately 200 to 300 ms after feedback onset (Glazer et al., 2018), is most pronounced over fronto-central sites (Holroyd et al., 2008, 2011; Miltner et al., 1997) and is modulated by the delivery of advantageous versus neutral or disadvantageous outcomes (Harmon-Jones et al., 2020a, 2020b; Luo et al., 2022; Ma et al., 2014; San Martín et al., 2016). The RewP is partly driven by activity in reward-related subcortical regions such as the striatum (Becker et al., 2014; Carlson et al., 2011, 2015; Foti et al., 2011, 2014).

We conceptualize *effort* as the mobilization of general resources—both mental and physical—to execute behaviour (Gendolla & Wright, 2009). It involves the allocation of energy toward achieving any goal requiring energy, regardless of whether self-control is needed. Therefore, self-control is a specific form of effort that entails overriding impulses and resisting temptation[9]. In fact, training in effort enhances general self-control capacity (for a review, see: Smith et al., 2019). Moreover, effort constitutes an integral component of self-control and can determine self-control behaviour (Kotabe & Hofmann, 2015). Convergent evidence indicates that effort increases the RewP. For example, Pan et al. (2023) found that higher effort conditions evoke greater RewP neural amplitude response. Similarly, Bogdanov et al. (2022) reported that the RewP is significantly elevated in trials requiring more versus less cognitive effort. Furthermore, Ma et al. (2014) demonstrated that demanding mental arithmetic problems, but not simpler ones, are associated with larger RewP amplitudes. Similarly, Harmon-Jones et al. (2024), using an effortful task-switching paradigm, observed that high effort, compared to low effort, yields a larger RewP amplitude when participants believe that their effort led to the reward. These findings were corroborated by self-reports, where self-reported effort exertion was associated with larger RewP differences (Harmon-Jones et al., 2020a, 2020b). In summary, the literature suggests that effort exertion modules the RewP. Given that effort constitutes an integral component of self-control (Kotabe & Hofmann, 2015), we hypothesized that exerting self-control would enhance the RewP.

Rewards can take many forms. One of the earliest and most enduring conceptualizations of rewards distinguishes between hedonic and eudaimonic ones. Hedonic rewards are defined in terms of pleasure and comfort, whereas eudaimonic rewards are defined in terms of meaning and self-realization (Huta & Waterman, 2014; Ryan & Deci, 2001; Telzer et al., 2014). Thus, hedonic rewards are very pleasurable and self-focused, such as enjoying material goods and playing video games, whereas eudaimonic rewards are intrinsically meaningful and purposeful, such as helping strangers and donating to charity (Shizgal, 1999; Steger et al., 2008a; Telzer et al., 2014). Although hedonism and eudaimonia are positively associated (Disabato et al., 2016; Goodman et al., 2018; Kashdan et al., 2008), a good deal of studies highlight their relative independence and differentiation (Gallagher et al., 2009; Henderson et al., 2013; Huta & Waterman, 2014; Joshanloo, 2016). Neural activity associated with eudaimonic rewards (e.g., donating money to family) predicts increases in well-being, whereas neural activity associated with hedonic rewards (e.g., keeping money for oneself) predicts decreases in well-being (Luo et al., 2019, 2022; Telzer et al., 2014). Crucially, some recent studies indicate that hedonic and eudaimonic rewards also influence reward responsivity differently, although the findings are

---

[9] Several studies have operationalized self-control as effort (e.g., “How much effort did you exert on …?”; Muraven et al., 1998, 2006).

inconsistent. For instance, one study reported that hedonic rewards (i.e., winning rewards for oneself) elicited a larger RewP difference wave compared to eudaimonic rewards (i.e., winning rewards for charity; Luo et al., 2019). However, other studies found comparable RewP amplitudes between hedonic rewards and eudaimonic rewards (Luo et al., 2022; Zhang et al., 2023). The inconsistent findings highlight the need to clarify the distinct neural processes underlying these different forms of reward. Thus, our secondary goal was to examine whether the effects of self-control exertion on the RewP would differ for hedonic versus eudaimonic rewards.

Overview

Research and theory indicate that exercising self-control enhances the RewP. However, it is unclear whether this effect occurs for hedonic rewards, eudaimonic rewards, or both. On the one hand, exercising self-control may increase hedonic reward responsivity. Afterall, the majority of studies examining the effects of self-control exertion on reward responsivity have focused on hedonic rewards (Kelley et al., 2019), and hedonic rewards (vs. rewards for others) more strongly activate the ventral striatum (Morelli et al., 2015) which is a neural generator of the RewP (Carlson et al., 2011). On the other hand, self-control may increase eudaimonic reward responsivity. In support of this view, recent research suggests that effort exertion increases meaning in life (Campbell et al., 2024). Insofar as meaning is more strongly tied to eudaimonic than hedonic rewards, then exercising self-control may increase the RewP more so for eudaimonic rewards. Still another option is that exercising self-control increases the RewP similarly for hedonic and eudaimonic rewards. Such a perspective is consistent with the strong links between the two types of rewards (Disabato et al., 2016; Goodman et al., 2018; Kashdan et al., 2008) and common neural processes across them (Liu et al., 2011; Morelli et al., 2015; Sescousse et al., 2013). To test these competing viewpoints, participants exerted self-control (incongruent Stroop trials) or not (congruent Stroop trials) in a speeded reaction time task where they had the opportunity to win money for themselves (a hedonic reward) or a charity of their choosing (a eudaimonic reward) while EEG was recorded. We measured participants' reward responsivity via the RewP.

4.1.2 Method

Participants and Design

Following past research on the RewP to hedonic and eudaimonic rewards (Luo et al., 2019, 2022), we used G*Power (Faul et al., 2009) assuming a small effect size (Cohen's $f = .10$), six measures (deriving from a 2 × 3 within-subjects design), $\alpha = .05$, power $(1-\beta) = .80$, and a moderate relation among repeated measures ($r = .50$). Based on these parameters, 109 participants were required. We oversampled assuming data loss and recruited 121 participants from the University of Southampton psychology participant pool in exchange for course credit and task winnings. We tested them in private cubicles and via computer. We excluded seven participants from EEG analyses because more than 50% of their trials had been rejected due to artifacts or wrong response, leaving insufficient (< 30) trials, and thus failing to meet the requirement for ERP analysis (Cai et al., 2016). The final sample comprised 114 participants (93 women, 18 men, 3 non-binary), aged between 18 and 37 years (*M* = 19.63, *SD* = 2.99). We did not collect ethnicity information, but over 90% of the University of Southampton undergraduates are White. The experimental protocol was approved by the Ethics Committee of the University of Southampton (No. 79802). We used a 3 (reward: hedonic, eudaimonic, control) × 2 (congruency: congruent, incongruent) within-subjects design. We addressed the issue of multiple comparisons using Bonferroni corrections[10].

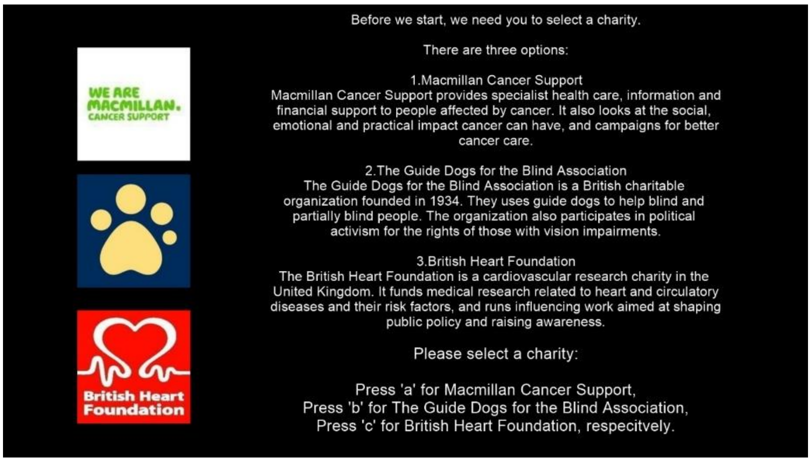
Procedure

All participants were familiarized with the electrophysiology laboratory and EEG recording procedure before providing informed consent. Participants were then fitted with recording electrodes and seated in a comfortable armchair approximately 80 cm away from a 60 cm × 33.5 cm monitor in a quiet laboratory room. They engaged in two core assessments: an 8-minute resting-state EEG session (as part of a different project) and a modified monetary incentive delay (MID) task (Knutson et al., 2001). Following Luo et al. (2022), participants first read a brief description of three representative charities: Macmillan Cancer Support, Guide Dogs for the Blind Association, and British Heart Foundation (Figure 4.1). Subsequently, they chose one of the three charities as the donation target. In the hedonic condition, the money they won belonged to them, whereas, in the eudaimonic condition, the money they won belonged to their chosen target.

---

[10] The corrected *p*-value is calculated by multiplying the original *p*-value by the number of comparisons. For instance, in post-hoc tests of a 3×2 interaction effect, the adjusted *p*-value becomes 15 times the original value.

Figure 4.1 *Charity Target Choice Prior to the Monetary Incentive Delay Task*



Before we start, we need you to select a charity.

There are three options:

1.Macmillan Cancer Support
Macmillan Cancer Support provides specialist health care, information and financial support to people affected by cancer. It also looks at the social, emotional and practical impact cancer can have, and campaigns for better cancer care.

2.The Guide Dogs for the Blind Association
The Guide Dogs for the Blind Association is a British charitable organization founded in 1934. They uses guide dogs to help blind and partially blind people. The organization also participates in political activism for the rights of those with vision impairments.

3.British Heart Foundation
The British Heart Foundation is a cardiovascular research charity in the United Kingdom. It funds medical research related to heart and circulatory diseases and their risk factors, and runs influencing work aimed at shaping public policy and raising awareness.

Please select a charity:

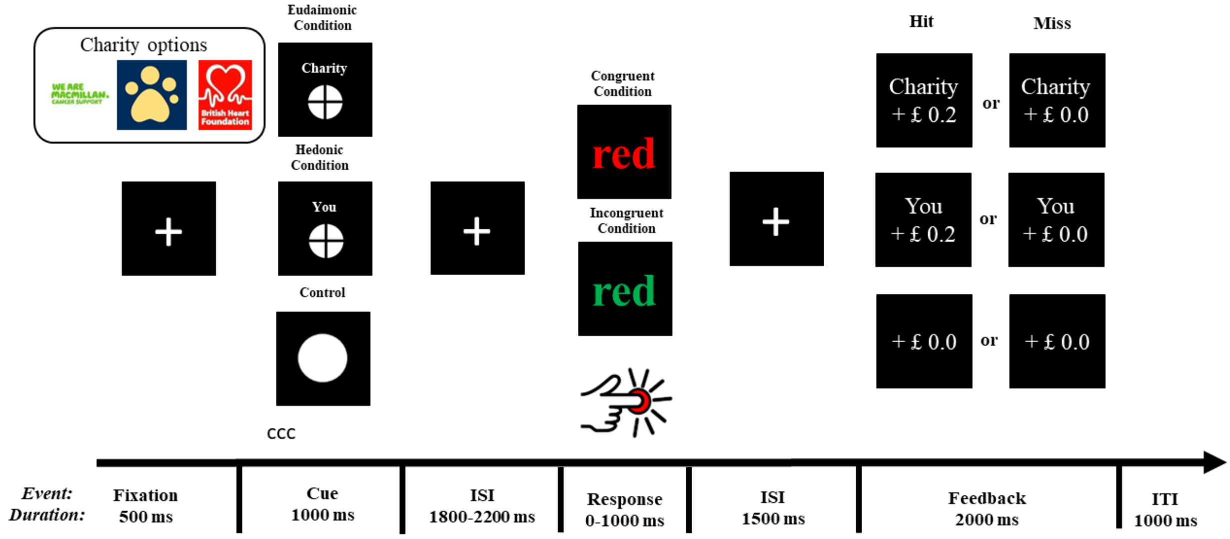
Press 'a' for Macmillan Cancer Support,
Press 'b' for The Guide Dogs for the Blind Association,
Press 'c' for British Heart Foundation, respecitvely.

Note. We described the function of each charity and how they use donations. We took descriptions from Wikipedia and edited them down to 35-40 words.

We report the trial structure in Figure 4.2. Each trial began with a 500 ms fixation-cross in the centre of the screen. Thereafter, we presented participants with an incentive cue for 1000 ms. There were three cue types in each session that prompted the object of the win money: self (i.e., hedonic condition), charity (i.e., eudaimonic condition), nobody (i.e., neutral control condition). In the hedonic condition (signalled by a circle with a cross inside labelled with “You” above), we informed participants of the potential monetary win for themselves. In the eudaimonic condition (signalled by a circle with a cross inside labelled with “Charity” above), we informed participants of the potential monetary win for the charity. In the neutral control condition (signalled by a circle), we informed participants that they would win money neither for themselves nor for the charity regardless of their efforts. We presented these cues with equal probability and in a random order. We followed the cue with a fixation cross appearing 1800 ms – 2200 ms. Then, we presented participants with the target stimulus, a colour word with either a congruent (i.e., congruent trials) or an incongruent (i.e., incongruent trials) ink colour. We instructed them to ignore the meaning of the word and identify the ink colour of the word as quickly and concretely as possible with their dominant hand by pressing the keyboard. We presented each word stimulus on the screen until a response (key-pressing) occurred, but no longer than 1000 ms. Lastly, after a 1500 ms fixation-cross, we signalled the outcome of each trial by feedback stimulus presented for 2000 ms. There were two types of feedback in each condition. In the hedonic condition, the feedback of “Self + £ 0.2” would be present if the response were correct and fast enough; otherwise, the feedback would be “Self + £ 0.0.” In the eudaimonic condition, the feedback of “Charity + £ 0.2” would be present if the response were correct and fast enough; otherwise, the feedback would be “Charity + £ 0.0.” In the neutral control condition, the feedback would always be “+ £ 0.0” regardless of the response.

Participants completed a practice block of 27 trials prior to the experimental blocks to allow them to learn the association between each cue and experimental condition. The experiment consisted of 324 trials and was divided into 6 blocks of 54 trials. Each block involved a randomized distribution of three conditions. Participants received a self-paced break after each block. We programmed and administered the experiment using PsychoPy (Version 2021.2.3; Peirce, 2007). At the end of the study, we compensated participants with £10[11] (in addition to course credits and irrespective of task performance) and gave them the donation website for each of three charities.

Figure 4.2 *Trial Structure of the Monetary Incentive Delay Task*



Note. ISI = inter-stimulus interval; ITI = inter-trial interval.

Data Recording and Data Analysis

We collected the EEG data continuously from 64 scalp sites using Ag/AgCl electrodes mounted in an elastic cap (Neuroscan, NC), with an online reference to the left mastoid and an off-line algebraic re-reference to the average of left and right mastoids. We mounted a ground electrode midway between FPz and Fz. We recorded the vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) from two pairs of electrodes, with one placed above and below the left eye, and another placed 10 mm from the outer canthi of each eye. We based the electrode cap on the 10–20 system. We kept electrode impedances below 5 kΩ. Also, we amplified and sampled the signals at 1000 Hz with an online bandpass filter from 0.10–100 Hz.

---

[11] $12.77 or €11.69 or ¥91.51

In offline processing, we initially pre-processed the EEG data by using EEGLAB, an open-source toolbox running in the MATLAB environment (Delorme & 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), segmented them from 200 ms prior to 800 ms following the onset of feedback, and baseline corrected them to the -200–0 ms. We identified bad channels by visual inspection of the waveforms and replaced them by using a spherical spline identified interpolation (SSI; Perrin et al., 1989). We corrected segments contaminated by blinks, eye movements, and other artifacts using an independent component analysis (ICA) algorithm (Delorme & Makeig, 2004) and ICLabel, a proposed statistical model, to automatically label ICA components (Pion-Tonachini et al., 2019). We also excluded bad segments where a voltage deviation on any channel of ± 100 μV. Finally, we used extracted average waveforms for each participant and condition to calculate grand average waveforms.

Following best practices (i.e., to employ multiple comparisons correction, to average across the electrode sites, to use difference scores, that is, RewP difference wave; Luck & Gaspelin, 2017), previous studies (Harmon-Jones et al., 2020b; Luo et al., 2022), and inspection of the grand average waveforms, we quantified the RewP as the mean amplitude on a 100 ms window (i.e., 280 ms - 380 ms) after feedback onset over frontal-central sites (i.e., Fz, FCz, Cz). Also, we calculated the RewP difference wave as the difference between the ERP response to gains (i.e., rewards) minus the ERP response to neutral[12] (Luo et al., 2019; Ma et al., 2014).

4.1.3 Results

Hit Rate and Reaction Time

We excluded data from trials where participants provided an improper response (< 200 ms). All participants' mean hit rate and reaction time were within three standard deviations from the mean. We conducted a 3 (reward: hedonic, eudaimonic, control) × 2 (congruency: congruent, incongruent) repeated Analysis of Variance (ANOVA) on hit rate and reaction time. The main effects of congruency were significant, as participants had a higher hit rate, $F(1, 113)$ =

---

[12] In previous studies, the RewP effect was calculated as the difference between the ERP in response to gains minus the ERP in response to loss (Harmon-Jones et al., 2020a, 2024; Luo et al., 2022; San Martín et al., 2016) or gains minus the ERP in response to neutral (Luo et al., 2019; Ma et al., 2014). However, we included no loss condition (i.e., a condition in which participants would lose money) in the current study. Considering that prior work has found that ERPs to neutral feedback and loss feedback are equivalent in this type of task (Holroyd et al., 2006, Experiment 5), we calculated the RewP effect as the difference between the ERP in response to gains minus the ERP in response to neutral.

166.79, $p < .001$, $\eta_p^2 = .60$, and were faster, $F(1, 113) = 372.61$, $p < .001$, $\eta_p^2 = .77$, on congruent than incongruent trials. The main effects of Reward were significant for both hit rate, $F(2, 112) = 9.08$, $p < .001$, $\eta_p^2 = .14$, and reaction time $F(2, 112) = 15.59$, $p < .001$, $\eta_p^2 = .22$. Compared to control trials, participants had a higher hit rate on hedonic trials ($p < .001$) and tended to have a higher hit rate on eudaimonic trials ($p = .092$); also, participants had a higher hit rate on hedonic than eudaimonic trials ($p = .007$). The pattern was similar for reaction time: Compared to control trials, participants were faster on hedonic ($p < .001$) and eudaimonic ($p = .001$) trials, and they were faster on hedonic than eudaimonic trials ($p = .017$).

The Reward × Congruency interactions were significant for both hit rate, $F(2, 112) = 4.88$, $p = .009$, $\eta_p^2 = .08$, and reaction time, $F(2, 112) = 38.65$, $p < .001$, $\eta_p^2 = .41$. Hit rates were higher and reaction times were shorter for congruent versus incongruent trials for each reward type (*p*s < .001). Differences between congruent and incongruent trials were largest for eudaimonic reward trials compared to hedonic reward trials and control trials (hit rate: $d_{\text{Eudaimonic}} = 0.96$, $d_{\text{Hedonic}} = 0.84$, $d_{\text{Control}} = 0.84$; reaction time: $d_{\text{Eudaimonic}} = 2.38$, $d_{\text{Hedonic}} = 1.65$, $d_{\text{Control}} = 1.43$). We reported means and standard deviations in Table 4.1.

The Reward Positivity

We conducted a 3 (reward: hedonic, eudaimonic, control) × 2 (congruency: congruent, incongruent) repeated measures ANOVA on RewP amplitude. We obtained a significant main effect of reward, $F(2, 112) = 19.09$, $p < .001$, $\eta_p^2 = .25$. Post hoc analysis showed that the RewP was larger on hedonic ($M = 5.77$ μV, $SD = 4.66$ μV) than eudaimonic ($M = 4.76$ μV, $SD = 4.54$ μV, $p < 0.001$) trials, and higher than in the control ($M = 4.39$ μV, $SD = 4.50$ μV, $p < 0.001$) trials. However, there was no significant difference on RewP between eudaimonic trials and control trials ($p = .396$). In addition, consistent with the reward responsivity hypothesis of self-control (Kelley et al., 2019), the RewP was larger after self-control was exerted (i.e., incongruent trials, $M = 5.42$ μV, $SD = 4.66$ μV) compared to not exerted (i.e., congruent trials, $M = 4.53$ μV, $SD = 4.47$ μV), $F(1, 113) = 42.04$, $p < .001$, $\eta_p^2 = .27$. The interaction was not significant, $F(2, 112) = 2.25$, $p = .110$, $\eta_p^2 = .04$. We reported means and standard deviations of RewP amplitude in Table 4.1. We depicted grand average waveforms in Figure 4.3a and Figure S4.1, and the corresponding topographic maps in Figure 4.4a.

Next, we examined the modulation of RewP difference wave using a 2 (reward: hedonic, eudaimonic) × 2 (congruency: congruent, incongruent) repeated measures ANOVA. Consistent with past research (Luo et al., 2019), the RewP difference wave was larger on hedonic ($M = 1.42$ μV, $SD = 3.00$ μV) than eudaimonic ($M = 0.38$ μV, $SD = 3.05$ μV) trials, $F(1, 113) = 22.50$, $p < .001$, $\eta_p^2 = .17$. In addition, consistent with the reward responsivity hypothesis of self-control (Kelley et al., 2019), the RewP difference wave was larger after self-control was exerted (i.e., incongruent trials, $M = 1.17$ μV, $SD = 3.10$ μV) compared to not exerted (i.e., congruent trials, $M = 0.63$ μV, *SD*

= 2.95 μV), $F(1, 113) = 4.37$, $p = .039$, $\eta_p^2 = .04$. However, the interaction was not significant, $F(1, 113) = 0.61$, $p = .438$, $\eta_p^2 = .01$. We reported means and standard deviations of RewP difference wave in Table 4.1. We depicted grand average waveforms in Figure 4.3b and Figure S4.2, and the corresponding topographic maps in Figure 4.4b.

Table 4.1 *Means and Standard Deviations for Hit Rate, Reaction Time, RewP, and RewP Difference Wave*

Hit Rate (%)

	Eudaimonic	Hedonic	Control	Average
Incongruent	81.24 (11.21)	83.87 (10.52)	81.22 (12.29)	82.11 (11.34)
Congruent	92.82 (6.09)	93.11 (5.90)	90.22 (8.13)	88.49 (6.70)
Average	87.03 (8.65)	88.49 (8.21)	85.72 (10.21)	

Reaction Time (ms)

	Eudaimonic	Hedonic	Control	Average
Incongruent	674.04 (65.52)	658.55 (66.37)	669.66 (66.92)	667.42 (63.39)
Congruent	594.04 (54.78)	596.87 (54.47)	616.74 (57.39)	602.55 (52.97)
Average	634.04 (57.26)	627.71 (56.77)	643.20 (58.88)	

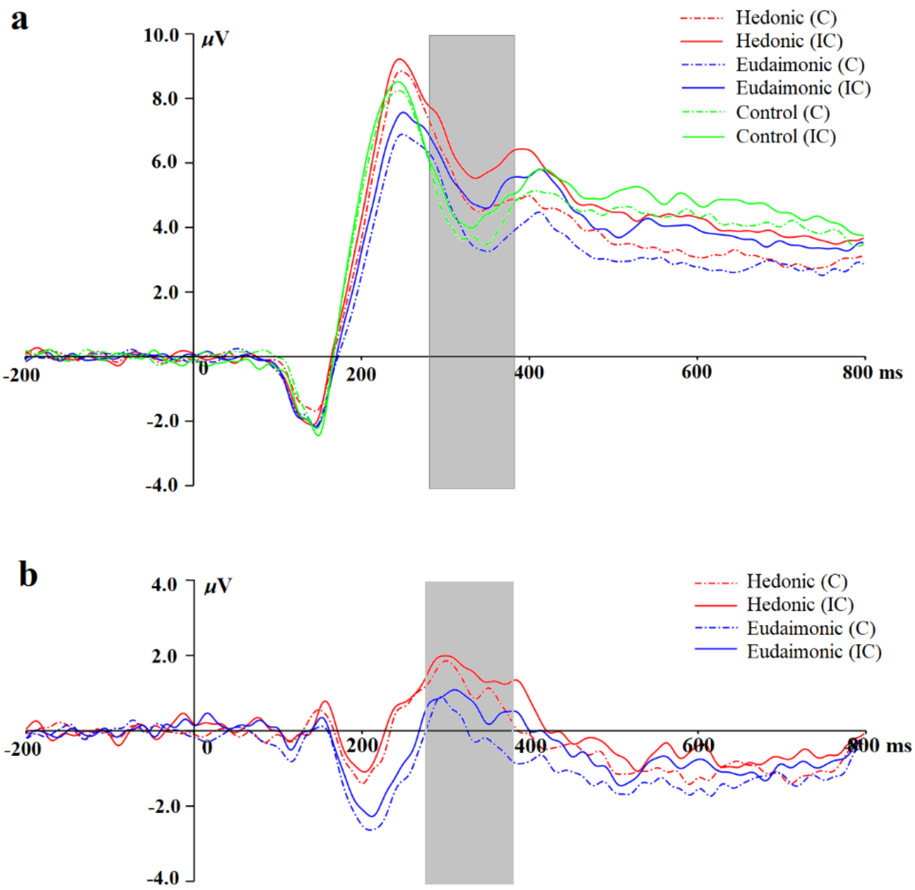
RewP (μV)

	Eudaimonic	Hedonic	Control	Average
Incongruent	5.35 (4.70)	6.24 (4.68)	4.66 (4.60)	5.42 (4.66)
Congruent	4.16 (4.38)	5.29 (4.63)	4.12 (4.40)	4.53 (4.47)
Average	4.76 (4.54)	5.77 (4.66)	4.39 (4.50)	

RewP Difference Wave (μV)

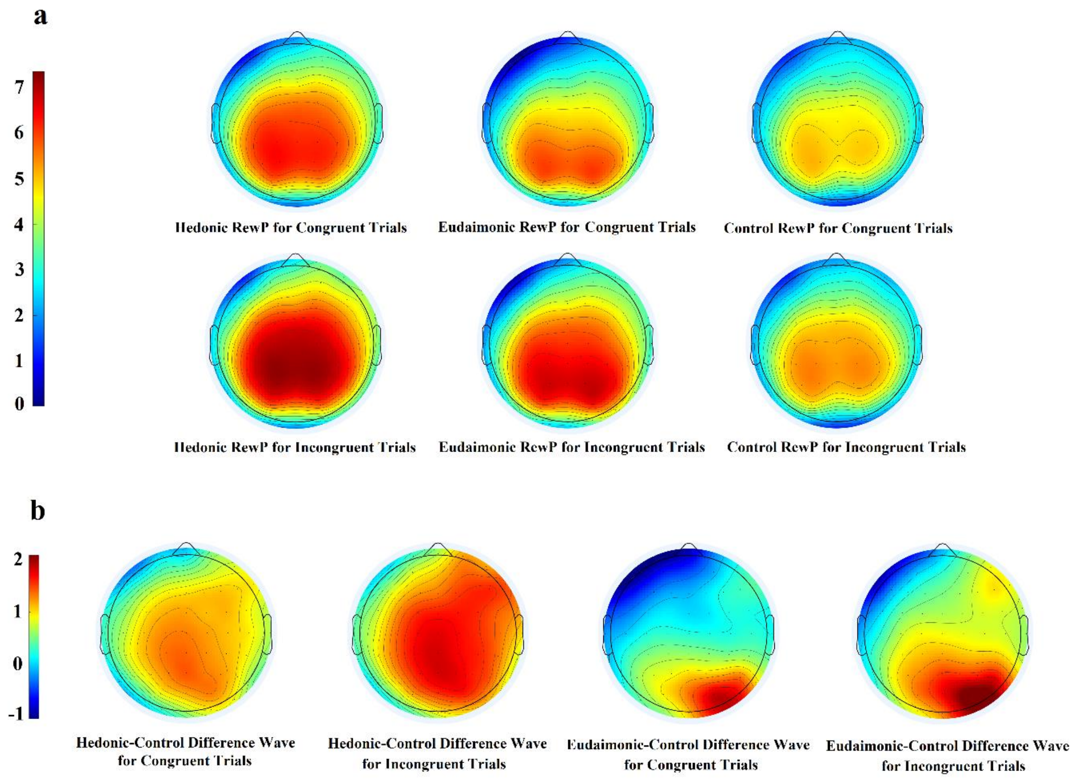
	Eudaimonic-Control	Hedonic-Control	Average
Incongruent	0.69 (2.94)	1.58 (2.96)	1.14 (2.58)
Congruent	0.05 (3.16)	1.17 (3.04)	0.61 (2.80)

Figure 4.3 *Grand Average Event-Related Potential Waveforms of the Reward Positivity as a Function of Self-control Exertion*



Note. (**a**) Grand averages of Reward Positivity for all conditions. (**b**) Reward Positivity difference waves (Reward - Control) in hedonic and eudaimonic conditions. The grand-averaged waveforms represent the grand average across 3 frontal-central sites (O1, O2, and Oz). The Reward Positivity measurement window (i.e., 280 ms – 380 ms) in shaded in light gray. C = congruent trials; IC = incongruent trials.

Figure 4.4 *Topographical Maps of the Reward Positivity as a Function of Self-control Exertion*



Note. (**a**) Topographical maps of the Reward Positivity for all conditions. (**b**) Topographical maps of the Reward Positivity difference waves (Reward - Control) in hedonic and eudaimonic conditions. The time window in the topographical maps is 280 ms – 380 ms.

4.1.4 General Discussion

We aimed to provide a rigorous test of the reward responsivity hypothesis of self-control (Kelley et al., 2019). Although this hypothesis is agnostic about how self-control exertion influences different types of rewards (hedonic vs. eudaimonic), it implicitly suggests that exercising self-control enhances reward responsivity generally. However, the majority of the literature on self-control and reward responsivity has focused on hedonic rewards such as responsivity to food (Haynes et al., 2016; Hoffman et al., 2007; Imhoff et al., 2014; Vohs et al., 2011, Study 3), drugs (Christiansen et al., 2012; Muraven et al., 2002; Schlauch et al., 2015; Shmueli & Prochaska, 2009), and money (Achtziger et al., 2015; Bruyneel et al., 2009; Osgood & Muraven, 2015; Schmeichel et al., 2010, Study 2b). Given this and recent evidence that exercising self-control increases meaning in life (Campbell et al., 2024), we sought to examine whether reward type (hedonic vs. eudaimonic) moderates the effect of self-control on reward responsivity. Consistent with the reward responsivity hypothesis of self-control, we showed that exercising self-control increases immediate neural responsivity to rewards (as indexed by RewP) in a domain-general fashion.

Theoretical Implications

The findings are consistent with theorizing in the self-control literature. According to the process model of self-control, exercising self-control causes shifts in attention and motivation toward rewards and gratification (Inzlicht et al., 2014; Inzlicht & Schmeichel, 2012). Insofar as the RewP is a reward prediction error linked to motivation and attention (Lange et al., 2012; Threadgill & Gable, 2016), its enhancement following self-control exertion is consistent with the central premise of the process model of self-control. Our results are also interpretable through the lens of the integrative self-control theory (Kotabe & Hofmann, 2015). According to it, conflict between immediate desires and long-term goals signals the need to mobilize self-control resources. When self-control resources are abundant (control > desire), behaviours in line with long term goals occur. However, when control resources are limited (control < desire), behaviours in line with immediate desires occur. Insofar as self-control attempts (i.e., incongruent Stroop trials) consume finite resources, they may tip the balance toward greater desire-driven reward-seeking behaviour reflected in the enhanced RewP. Moreover, the results are consistent with theoretical models that conceptualize self-control as a value-based choice (Berkman et al., 2017; Pfeifer & Berkman, 2018). According to them, exercising control shifts valued-based calculations in favour of more immediate options over (more effortful) options. Insofar as the RewP has been source localized to the striatum (Becker et al., 2014; Carlson et al., 2011; Foti et al., 2011; Gehring & Willoughby, 2002) and the striatum tracks subjective value (Knutson et al., 2009), the finding of an enhanced RewP after self-control exertion may reflect shifting valued-based calculations. Also, we note that the precise neural generators of the RewP remain uncertain (Cohen et al., 2011), and source localization of scalp-recorded ERPs is inherently challenging (Pizzagalli, 2007).

Moreover, cognitive dissonance (Aronson & Mills, 1959; Harmon-Jones & Mills, 1999) and psychological contrast (Zentall, 2010) accounts of effort suggest that aversive states elicited by the exertion of effort make the end-result or reward appear more valuable. In accord with these accounts, a greater subjective experience of effort is associated with a larger RewP in an effort justification paradigm (Harmon-Jones et al., 2020a), especially when perceptions of control are high (Harmon-Jones et al., 2024). To the extent that incongruent Stroop trials are effortful (Bouzidi & Gendolla, 2023), the current results are consistent with effort-based interpretations of enhanced reward responsivity.

Implications for Ego-Depletion and The Strength Model of Self-Control

The perspective advanced here adds conceptual and theoretical refinement to the resource model of self-control by identifying the specific circumstances under which exerting self-control influences subsequent behaviour: increased reward responsivity. The resource model of self-control has been generative, making self-control research a focal point in social psychology for about 25 years. However, this model has come under intense scrutiny and debate (Friese et al., 2019). Some researchers have suggested that the effects predicted by the resource model are smaller than once anticipated (Dang et al., 2017; Garrison et al., 2019), whereas others have suggested that these effects are negligible at best (Carter et al., 2015; Carter & McCullough, 2014; Hagger et al., 2016; Vohs et al., 2021). By contrast, the original authors have reaffirmed their commitment to the model (Baumeister et al., 2018; Baumeister & Vohs, 2016). The initial conceptualization of the resource model suggests that exercising self-control at Time 1 undermines the ability to exercise self-control at Time 2, resulting in decrement in performance on challenging tasks irrespective of task type. Stated otherwise, a domain-general, but finite, resource underlies all types of self-control (Baumeister et al., 1998; Muraven & Baumeister, 2000). The current findings challenge the notion of domain generality and suggest that the effects of self-control exertion on subsequent behaviour are specific to reward responsivity.

Limitations and Future Directions

Although we interpreted the significant main effect of congruency on RewP amplitudes as supportive evidence of the reward responsivity hypothesis of self-control (Kelley et al., 2019), some readers may remain unconvinced due to the non-significant Reward × Congruency interaction. The RewP increases as a function of reward magnitude and even zero magnitude wins (Meadows et al., 2016; Threadgill & Gable, 2018). Thus, the mere act of winning is rewarding even when it results in no monetary gain. Similarly, our findings indicated that exercising self-control increases reward responsivity even to no real rewards (win £0), suggesting that exercising self-control produces domain general increases in reward responsivity. Nonetheless, the RewP difference wave results indicate that, relative to these no monetary gain (win £0), participants were more sensitive to hedonic and eudemonic rewards, a pattern consistent with a domain general increase in reward responsivity following self-control exertion. Nonetheless, the RewP is a complex marker of reward responsivity that tracks reward linking (Angus et al., 2015; Brown et al., 2022; Huvermann et al., 2021; Jia et al., 2013; Peterburs et al., 2019), reward wanting (Angus et al., 2015; Banica et al., 2023; Huvermann et al., 2021; Threadgill & Gable, 2016), and reward learning (Cavanagh, 2015; Jackson & Cavanaugh, 2023). Given this complexity, future studies are needed to more precisely characterize how exercising self-control modulates the multidimensionality of the RewP, thus providing a rigorous test of the reward responsivity hypothesis of self-control.

Multiple theoretical perspectives indicate that eudaimonic processes (e.g., meaning) are central to psychological experience (Becker, 1971; Frankl, 1963; Greenberg et al., 2004; Heine et al., 2006; Martela et al., 2018; Pyszczynski et al., 2015; Sedikides & Wildschut, 2018; Steger et al., 2008b, 2009; Wong et al., 2013). In an effort to maintain homeostasis, the impulses of the self often need to be held in check. These efforts (i.e., self-control exertion) often come at the cost of eudaimonic processes that gives life meaning, that is, autonomy, volition, and choice. Based on this theorizing and recent research (Campbell et al., 2024), we would have expected that self-control exertion produced stronger responses to eudaimonic over hedonic rewards. We attain some evidence from the behavioural results that the eudaimonic condition showed the strongest Stroop effect in both hit rates and reaction times, suggesting the self-control was the most effortful in this condition. This finding aligns with prior research regarding to the relationship between self-control and eudaimonia. For example, self-control is positively associated with eudaimonic well-being (Li et al., 2022). Moreover, eudaimonic motives can promote well-being than hedonic motives because the former increase self-control, while the latter decrease it (Zeng & Chen, 2020). However, this pattern was not reflected in the results of the RewP, which instead showed a domain general reward response to both hedonic and eudaimonic rewards. This differentiation likely reflects that reward responsivity post-exertion could manifest differently from self-control exertion itself. Whilst the behavioural Stroop effects index *during-task* effort, the RewP reflects *post-task* response—consistent with theories distinguishing between the implementation of efforts and the evaluation of reward (Inzlicht et al., in press). Still other researchers even reported that the RewP to eudaimonic rewards is less sensitive to temporal decay than the RewP to hedonic rewards (Luo et al., 2022). In all, although self-control may not sensitize persons toward eudaimonic rewards in-the-moment, their weaker temporal decay may make eudaimonic rewards well-suited for countermanding the aversiveness of self-control over time. Indeed, a weaker temporal decay of the RewP to eudaimonic (vs. hedonic) rewards may help to explain longitudinal associations between self-control and positive life outcomes (Moffitt et al., 2011). Still other ERPs may be better suited to distinguish between hedonic and eudaimonic rewards after self-control exertion. For example, the late positive potential (LPP) is driven by stimulus significance above and beyond other factors (Hajcak & Foti, 2020). Insofar as eudaimonic (vs. hedonic) rewards are more psychologically enriching, they should modulate the late-positive potential after self-control exertion. Future studies could test these possibilities.

The current study revealed a parietal-dominant RewP distribution, contrasting with the classic frontocentral pattern (Harmon-Jones et al., 2020b; Luo et al., 2022). This topographic shift likely reflects the increased task demands of my experimental paradigm. Specifically, the critical distinction between hedonic and eudaimonic reward processing likely engages parietal mechanisms involved in action valuation (Wisniewski et al., 2015). Future research should systematically manipulate reward complexity to test this interpretation.

4.1.5 Conclusion

Self-control has profound implications for a wide range of behaviours, with grave personal and societal costs. Consequently, experimental research on self-control has permeated many subfields of psychology. Although challenges to prominent models have damped enthusiasm, we provided evidence supporting the reward responsivity hypothesis, with an increased reward responsivity (Reward Positivity) following self-control exertion. This effect occurs regardless of the presence or type of reward, suggesting that self-control enhances reward responsivity in a domain-general manner. We hope our findings offer the conceptual and theoretical innovation necessary to renew interest and focus to the experimental study of self-control.

References

Achtziger, A., Alós-Ferrer, C., & Wagner, A. K. (2015). Money, depletion, and prosociality in the dictator game. *Journal of Neuroscience, Psychology, and Economics, 8*(1), 1–14. http://dx.doi.org/10.1037/npe0000031

Angus, D. J., Kemkes, K., Schutter, D. J. L. G., & Harmon-Jones, E. (2015). Anger is associated with reward-related electrocortical activity: Evidence from the reward positivity. *Psychophysiology, 52*(10), 1271–1280. https://doi.org/10.1111/psyp.12460

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, *59*(2), 177–181. https://doi.org/10.1037/h0047195

Banica, I., Allison, G., Racine, S. E., Foti, D., & Weinberg, A. (2023). All the Pringle ladies: Neural and behavioural responses to high-calorie food rewards in young adult women. *Psychophysiology*, *60*(3), e14188. https://doi.org/10.1111/psyp.14188

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Baumeister, R. F., Tice, D. M., & Bushman, B. J. (2023). A review of multisite replication projects in social psychology: is it viable to sustain any confidence in social psychology's knowledge base?. *Perspectives on Psychological Science*, *18*(4), 912–935. https://doi.org/10.1177/17456916221121815

Baumeister, R. F., Tice, D. M., & Vohs, K. D. (2018). The strength model of self-regulation: Conclusions from the second decade of willpower research. *Perspectives on Psychological Science*, *13*(2), 141-145. Conclusions from the second decade of willpower research. *Perspectives on Psychological Science*, *13*(2), 141–145. https://doi.org/10.1177/1745691617716946

Baumeister, R. F., & Vohs, K. D. (2016). Misguided effort with elusive implications. *Perspectives on Psychological Science*, *11*(4), 574–575. https://doi.org/10.1177/1745691616652878

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355. https://doi.org/10.1111/j.1467-8721.2007.00534.x

Becker, E. (1971). *The birth and death of meaning*. Free Press.

Becker, M. P., Nitsch, A. M., Miltner, W. H., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Bogdanov, M., Renault, H., LoParco, S., Weinberg, A., & Otto, A. R. (2022). Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cerebral Cortex*, *32*(19), 4255–4270. https://doi.org/10.1093/cercor/bhab480

Bouzidi, Y. S., & Gendolla, G. H. (2023). Action-orientation shields against primed cognitive conflict effects on effort-related cardiac response. *Psychophysiology, 60*(12), Article e14407. https://doi.org/10.1111/psyp.14407

Brown, D. R., Jackson, T. C., & Cavanagh, J. F. (2022). The reward positivity is sensitive to affective liking. *Cognitive, Affective, & Behavioural Neuroscience*, *22*, 258–267. https://doi.org/10.3758/s13415-021-00950-5

Bruyneel, S. D., Dewitte, S., Franses, P. H., & Dekimpe, M. G. (2009). I felt low and my purse feels light: Depleting mood regulation attempts affect risk decision making. *Journal of Behavioural Decision Making, 22*(2), 153–170. https://doi.org/10.1002/bdm.619

Cai, H., Wu, L., Shi, Y., Gu, R., & Sedikides, C. (2016). Self-enhancement among Westerners and Easterners: A cultural neuroscience approach. *Social Cognitive and Affective Neuroscience*, *11*(10), 1569–1578. https://doi.org/10.1093/scan/nsw072

Campbell, A. V., Wang, Y., & Inzlicht, M. (2024). Experimental evidence that exerting effort increases meaning. *PsyArXiv*. https://doi.org/10.31234/osf.io/z3g9c

Carlson, J. M., Foti, D., Harmon-Jones, E., & Proudfit, G. H. (2015). Midbrain volume predicts fMRI and ERP measures of reward reactivity. *Brain Structure and Function*, *220*(3), 1861–1866. https://doi.org/10.1007/s00429-014-0725-9

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *Neuroimage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796–815. https://doi.org/10.1037/xge0000083

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Frontiers in Psychology*, *5*, Article 823. https://doi.org/10.3389/fpsyg.2014.00823

Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioural adjustments, but at different times. *NeuroImage, 110,* 205–216. https://doi.org/10.1016/j.neuroimage.2015.02.007

Christiansen, P., Cole, J. C., & Field, M. (2012). Ego depletion increases ad-lib alcohol consumption: investigating cognitive mediators and moderators. *Experimental and Clinical Psychopharmacology, 20*(2), 118–128. https://doi.org/10.1037/a0026623

Cohen, M. X., Cavanagh, J. F., & Slagter, H. A. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity: Commentary. *Human brain mapping*, *32*(12), 2270-2271. https://doi.org/10.1002/hbm.21358

Dang, J. (2016). Testing the role of glucose in self-control: A meta-analysis. *Appetite, 107,* 222–230. https://doi.org/10.1016/j.appet.2016.07.021

Dang, J., Liu, Y., Liu, X., & Mao, L. (2017). The ego could be depleted, providing initial exertion is depleting. *Social Psychology*, *48*(4), 242–245. https://doi.org/10.1027/1864-9335/a000308

David, L., Vassena, E., & Bijleveld, E. (2024). The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect. *Psychological Bulletin.* Advance online publication. https://doi.org/10.1037/bul0000443

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. https://doi.org/10.1016/j.jneumeth.2003.10.009

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471–482. https://doi.org/10.1037/pas0000209

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behaviour Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Finley, A. J., & Schmeichel, B. J. (2019). Aftereffects of self-control on positive emotional reactivity. *Personality and Social Psychology Bulletin*, *45*(7), 1011–1027. https://doi.org/10.1177/0146167218802836

Foti, D., Carlson, J. M., Sauder, C. L., & Proudfit, G. H. (2014). Reward dysfunction in major depression: Multimodal neuroimaging evidence for refining the melancholic phenotype. *NeuroImage*, *101*, 50–58. https://doi.org/10.1016/j.neuroimage.2014.06.058

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping*, *32*(12), 2207–2216. https://doi.org/10.1002/hbm.21182

Frankl, V. E. (1963). *Man's search for meaning.* Washington Square Press.

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, *23*(2), 107–131. https://doi.org/10.1177/1088868318762183

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source:

Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*(2), 325–336. https://doi.org/10.1037/0022-3514.92.2.325

Gallagher, M. W., Lopez, S. J., & Preacher, K. J. (2009). The hierarchical structure of well-being. *Journal of Personality*, *77*(4), 1025–1050. https://doi.org/10.1111/j.1467-6494.2009.00573.x

Garrison, K. E., Finley, A. J., & Schmeichel, B. J. (2019). Ego depletion reduces attention control: Evidence from two high-powered preregistered experiments. *Personality and Social Psychology Bulletin*, *45*(5), 728–739. https://doi.org/10.1177/0146167218796473

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gendolla, G. H. E., & Wright, R. A. (2009). Effort. In D. Sander & K. R. Scherer (Eds.), *Oxford companion to emotion and the affective sciences* (pp. 134–135). Oxford University Press.

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Goodman, F. R., Disabato, D. J., Kashdan, T. B., & Kauffman, S. B. (2018). Measuring well-being: A comparison of subjective well-being and PERMA. *The Journal of Positive Psychology*, *13*(4), 321–332. https://doi.org/10.1080/17439760.2017.1388434

Greenberg, J., Koole, S. L., & Pyszczynski, T. A. (Eds.). (2004). *Handbook of experimental existential psychology*. Guilford Press.

Hagger, Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G., Bruyneel, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T., Crowell, A., De Ridder, D. T. D., Dewitte, S., ... Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*(4), 495–525. https://doi.org/10.1037/a0019486

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7), Article e13570. https://doi.org/10.1111/psyp.13570

Harmon-Jones, E., Clarke, D., Paul, K., & Harmon-Jones, C. (2020a). The effect of perceived effort on reward valuation: Taking the reward positivity (RewP) to dissonance theory. *Frontiers in Human Neuroscience*, *14,* Article 157. https://doi.org/10.3389/fnhum.2020.00157

Harmon-Jones, E., Matis, S., Angus, D. J., & Harmon-Jones, C. (2024). Does effort increase or decrease reward valuation? Considerations from cognitive dissonance theory. *Psychophysiology, 61*(6), Article e14536. https://doi.org/10.1111/psyp.14536

Harmon-Jones, E., & Mills, J. (1999). *Cognitive dissonance: Progress on a pivotal theory in social psychology.* American Psychological Association.

Harmon-Jones, E., Willoughby, C., Paul, K., & Harmon-Jones, C. (2020b). The effect of perceived effort and perceived control on reward valuation: Using the reward positivity to test a dissonance theory prediction. *Biological Psychology*, *154*, Article 107910. https://doi.org/10.1016/j.biopsycho.2020.107910

Haynes, A., Kemps, E., & Moffitt, R. (2016). Too depleted to try? Testing the process model of ego depletion in the context of unhealthy snack consumption. *Applied Psychology: Health and Well-Being, 8*(3), 386–404. https://doi.org/10.1111/aphw.12080

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review, 10*(2), 88–110. https://doi.org/10.1207/s15327957pspr1002_1

Henderson, L. W., Knight, T., & Richardson, B. (2013). An exploration of the well-being benefits of hedonic and eudaimonic behaviour. *The Journal of Positive Psychology*, *8*(4), 322–336. https://doi.org/10.1080/17439760.2013.803596

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behaviour. *Journal of Experimental Social Psychology, 43*(3), 497–504. https://doi.org/10.1016/j.jesp.2006.05.004

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Research*, *1105*(1), 93–101. https://doi.org/10.1016/j.brainres.2005.12.015

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *NeuroReport*, *22*(5), 249–252. https://doi.org/10.1097/WNR.0b013e328345441d

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies*, *15*(6), 1425–1456. https://doi.org/10.1007/s10902-013-9485-0

Huvermann, D. M., Bellebaum, C., & Peterburs, J. (2021). Selective devaluation affects the processing of preferred rewards. *Cognitive, Affective, & Behavioural Neuroscience*, *21*(5), 1010-1025. https://doi.org/10.3758/s13415-021-00904-x

Imhoff, R., Schmidt, A. F., & Gerstenberg, F. (2014). Exploring the interplay of trait self–control and ego depletion: Empirical evidence for ironic effects. *European Journal of Personality, 28*(5), 413–424. https://doi.org/10.1002/per.1899

Inzlicht, M., Campbell, A. V., Saunders, B., & Inzlicht, M. (in press). Effort Paradox Redux: Rethinking how effort shapes social behavior. *Advances in Experimental Social Psychology*.

Inzlicht, M., & Schmeichel, B. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, *7*(5), 450–463. https://doi.org/10.1177/1745691612454134

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*(3), 127–133. https://doi.org/10.1016/j.tics.2013.12.009

Jackson, T. C., & Cavanagh, J. F. (2023). Reduced positive affect alters reward learning via reduced information encoding in the Reward Positivity. *Psychophysiology*, *60*(8), e14276. https://doi.org/10.1111/psyp.14276

Jia, S., Zhang, W., Li, P., Feng, T., & Li, H. (2013). Attitude toward money modulates outcome processing: An ERP study. *Social Neuroscience*, *8*(1), 43-51. https://doi.org/10.1080/17470919.2012.713316

Joshanloo, M. (2016). Revisiting the empirical distinction between hedonic and eudaimonic aspects of well-being using exploratory structural equation modeling. *Journal of Happiness Studies*, *17*(5), 2023–2036. https://doi.org/10.1007/s10902-015-9683-z

Kashdan, T. B., Biswas-Diener, R., & King, L. A. (2008). Reconsidering happiness: The costs of distinguishing between hedonics and eudaimonia. *The Journal of Positive Psychology*, *3*(4), 219–233. https://doi.org/10.1080/17439760802303044

Kelley, N. J., Finley, A. J., & Schmeichel, B. J. (2019). After-effects of self-control: The reward responsivity hypothesis. *Cognitive, Affective, & Behavioural Neuroscience*, *19*(3), 600–618. https://doi.org/10.3758/s13415-019-00694-3

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, *21*(16), Article RC159. https://doi.org/10.1523/JNEUROSCI.21-16-j0002.2001

Knutson, B., Delgado, M. R., & Phillips, P. E. (2009). Representation of subjective value in the striatum. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 389–406). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-374176-9.00025-7

Kotabe, H. P., & Hofmann, W. (2015). On integrating the components of self-control. *Perspectives on Psychological Science*, *10*(5), 618–638. https://doi.org/10.1177/1745691615593382

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, *7*, 67–70. https://doi.org/10.1016/j.copsyc.2015.08.003

Lange, S., Leue, A., & Beauducel, A. (2012). Behavioural approach and reward processing: Results on feedback-related negativity and P3 component. *Biological Psychology, 89*(2), 416–425. https://doi.org/10.1016/j.biopsycho.2011.12.004

Li, Q., Xiang, G., Song, S., Huang, X., & Chen, H. (2022). Examining the associations of trait self-control with hedonic and eudaimonic well-being. *Journal of Happiness Studies, 23*(2), 667–687. https://doi.org/10.1007/s10902-021-00418-w

Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioural Reviews, 35*(5), 1219–1236. https://doi.org/10.1016/j.neubiorev.2010.12.012

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luo, Y., Jiang, H., Chen, X., Zhang, Y., & You, X. (2019). Temporal dynamics of hedonic and eudaimonic reward processing: An event-related potentials (ERPs) study. *International Journal of Psychophysiology*, *137*, 63–71. https://doi.org/10.1016/j.ijpsycho.2018.12.009

Luo, Y., Zhang, X., Jiang, H., & Chen, X. (2022). The neural habituation to hedonic and eudaimonic rewards: Evidence from reward positivity. *Psychophysiology*, *59*, Article e13977. https://doi.org/10.1111/psyp.13977

Ma, Q., Meng, L., Wang, L., & Shen, Q. (2014). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research*, *261*, 1–7. https://doi.org/10.1016/j.bbr.2013.11.045

Martela, F., Ryan, R. M., & Steger, M. F. (2018). Meaningfulness as satisfaction of autonomy, competence, relatedness, and beneficence: Comparing the four satisfactions and positive affect as predictors of meaning in life. *Journal of Happiness Studies, 19*, 1261–1282. https://doi.org/10.1007/s10902-017-9869-7

Meadows, C. C., Gable, P. A., Lohse, K. R., & Miller, M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology*, *118*, 154-160. https://doi.org/10.1016/j.biopsycho.2016.06.002

Morelli, S. A., Sacchet, M. D., & Zaki, J. (2015). Common and distinct neural correlates of personal and vicarious reward: A quantitative meta-analysis. *NeuroImage, 112,* 244–253. https://doi.org/10.1016/j.neuroimage.2014.12.056

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*(2), 247–259. https://doi.org/10.1037//0033-2909.126.2.247

Muraven, M., Collins, R. L., & Neinhaus, K. (2002). Self-control and alcohol restraint: an initial application of the self-control strength model. *Psychology of Addictive Behaviours, 16*(2), 113–120. https://doi.org/10.1037//0893-164X.16.2.113

Muraven, M., Shmueli, D., & Burkley, E. (2006). Conserving self-control strength. *Journal of personality and social psychology*, *91*(3), 524–537. https://doi.org/ 10.1037/0022-3514.91.3.524

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: regulatory depletion patterns. *Journal of personality and social psychology*, *74*(3), 774–789. https://doi.org/10.1037/0022-3514.74.3.774

Osgood, J. M., & Muraven, M. (2015). Self-control depletion does not diminish attitudes about being prosocial but does diminish prosocial behaviours. *Basic and Applied Social Psychology, 37*(1), 68–80. https://doi.org/10.1080/01973533.2014.996225

Pan, W., Lu, J., Wu, L., Kou, J., & Lei, Y. (2023). Expending effort may share neural responses with reward and evokes high subjective satisfaction. *Biological Psychology*, *177*, Article 108480. https://doi.org/10.1016/j.biopsycho.2022.108480

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peterburs, J., Sannemann, L., & Bellebaum, C. (2019). Subjective preferences differentially modulate the processing of rewards gained by own vs. observed choices. *Neuropsychologia*, *132*, 107-139. https://doi.org/10.1016/j.neuropsychologia.2019.107139

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184–187. https://doi.org/10.1016/0013-4694(89)90180-6

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Pizzagalli, D. A. (2007). Electroencephalography and high-density electrophysiology source localization. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 56–84). Cambridge University Press.

Pyszczynski T., Solomon S., Greenberg J. (2015). Thirty years of terror management theory: From genesis to revelation. *Advances in Experimental Social Psychology, 52*, 1–70. https://doi.org/10.1016/bs.aesp.2015.03.001

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166. https://doi.org/10.1146/annurev.psych.52.1.141

San Martín, R., Kwak, Y., Pearson, J. M., Woldorff, M. G., & Huettel, S. A. (2016). Altruistic traits are predicted by neural responses to monetary outcomes for self vs charity. *Social Cognitive and Affective Neuroscience*, *11*(6), 863–876. https://doi.org/10.1093/scan/nsw026

Schlauch, R. C., Christensen, R. L., Derrick, J. L., Crane, C. A., & Collins, R. L. (2015). Individual differences in approach and avoidance inclinations moderate the effect of self-control depletion on ad-lib drinking. *Alcoholism: Clinical and Experimental Research, 39*(12), 2480–2488. https://doi.org/10.1111/acer.12915

Schmeichel, B. J., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology, 99*(1), 162–173. https://doi.org/10.1037/a0019797

Sedikides, C., & Wildschut, T. (2018). Finding meaning in nostalgia. *Review of General Psychology, 22*(1), 48–61. https://doi.org/10.1037/gpr0000109

Sescousse, G., Caldú, X., Segura, B., & Dreher, J.-C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience and Biobehavioural Reviews, 37*(4), 681–696. https://doi.org/10.1016/j.neubiorev.2013.02.002

Shizgal, P. (1999). On the neural computation of utility: Implications from studies of brain stimulation reward. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 502–526). Russell Sage Foundation.

Shmueli, D., & Prochaska, J. J. (2009). Resisting tempting foods and smoking behaviour: implications from a self-control theory perspective. *Health Psychology, 28*(3), 300–306. https://doi.org/10.1037/a0013826

Smith, T., Panfil, K., Bailey, C., & Kirkpatrick, K. (2019). Cognitive and behavioural training interventions to promote self-control. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(3), 259–279. http://dx.doi.org/10.1037/xan0000208

Steger, M. F., Kashdan, T. B., & Oishi, S. (2008a). Being good by doing good: Daily eudaimonic activity and well-being. *Journal of Research in Personality*, *42*(1), 22–42. https://doi.org/10.1016/j.jrp.2007.03.004

Steger, M. F., Kashdan, T. B., Sullivan, B. A., & Lorentz, D. (2008b). Understanding the search for meaning in life: Personality, cognitive style, and the dynamic between seeking and experiencing meaning. *Journal of Personality, 76*(2), 199–228. https://doi.org/10.1111/j.1467-6494.2007.00484.x

Steger, M. F., Oishi, S., & Kashdan, T. B. (2009). Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. *The Journal of Positive Psychology, 4*(1), 43–52. https://doi.org/10.1080/17439760802303127

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galván, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proceedings of the National Academy of Sciences*, *111*(18), 6600–6605. https://doi.org/10.1073/pnas.1323014111

Threadgill, A. H., & Gable, P. A. (2016). Approach-motivated pregoal states enhance the reward positivity. *Psychophysiology, 53*, 733–738. https://doi.org/10.1111/psyp.12611

Threadgill, A. H., & Gable, P. A. (2018). The sweetness of successful goal pursuit: Approach-motivated pregoal states enhance the reward positivity during goal pursuit. *International Journal of Psychophysiology*, *132*, 277-286. https://doi.org/10.1016/j.ijpsycho.2017.12.010

Vohs, K. D., & Baumeister, R. F. (2016). *Handbook of self-regulation: Research, theory, and applications* (3rd Ed.). Guilford Press.

Vohs, K., Baumeister, R., Mead, N., Ramanathan, S., & Schmeichel, B. (2011). Engaging in self-control intensifies desires and feelings. In Ahluwalia, R., Chartrand, T. L., & Ratner, R. K. (2012). *Building connections: Proceedings of the Association for Consumer Research Annual Conference.* Association for Consumer Research.

Vohs, K. D., Schmeichel, B. J., Lohmann, S., Gronau, Q. F., Finley, A. J., Ainsworth, S. E., Alquist, J. L., Baker, M. D., Brizi, A., Bunyi, A., Butschek, G. J., Campbell, C., Capaldi, J., Cau, C., Chambers, H., Chatzisarantis, N. L. D., Christensen, W. J., Clay, S. L., Curtis, J., ... Albarracín, D. (2021). A multisite preregistered paradigmatic test of the ego-depletion effect. *Psychological Science*, *32*(10), 1566–1581. https://doi.org/10.1177/0956797621989733

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioural choice. *Neuroscience & Biobehavioural Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Wisniewski, D., Reverberi, C., Momennejad, I., Kahnt, T., & Haynes, J. D. (2015). The role of the parietal cortex in the representation of task–reward associations. *Journal of Neuroscience, 35*(36), 12355-12365. https://doi.org/10.1523/JNEUROSCI.4882-14.2015

Wong, P. T. (2013). *The human quest for meaning: Theories, research, and applications*. Routledge.

Zentall, T. R. (2010). Justification of effort by humans and pigeons: Cognitive dissonance or contrast? *Current Directions in Psychological Science*, *19*(5), 196–300. https://doi.org/10.1177/0963721410383381

Zeng, Z., & Chen, H. (2020). Distinct associations of hedonic and eudaimonic motives with well-being: Mediating role of self-control. *International Journal of Environmental Research and Public Health, 17*(15), Article 5547. https://doi.org/10.3390/ijerph17155547

Zhang, Y., Rong, Y., & Wei, P. (2023). Mothers exhibit higher neural activity in gaining rewards for their children than for themselves. *Social Cognitive and Affective Neuroscience*, *18*(1), 1–14. https://doi.org/10.1093/scan/nsad048

4.2 Neural Evidence for the Reward Responsivity Hypothesis of Self-control: Domain-general Effects and the Role of Trait Authenticity

Abstract

Exercising self-control increased reward responsivity comparably for hedonic and eudaimonic rewards. I re-examined this process by examining whether trait authenticity—a facilitator of self-control and one indicator of eudaemonic well-being—moderates the effect of self-control exertion on reward responsivity. A sample of 114 participants completed a speeded reaction time task where they exercised self-control (incongruent Stroop trials) or not (congruent Stroop trials), followed by opportunities to win money for themselves (hedonic rewards) or a charity (eudaimonic rewards), with trait authenticity measured. The results revealed that trait authenticity did not interact with self-control exertion to influence the reward positivity (RewP) component. Nevertheless, participants with low authenticity in the non-self-control condition and those with high authenticity exhibited a stronger RewP response to hedonic than eudaimonic rewards. These findings indicate that the enhancement of reward responsivity following self-control exertion was independent of trait authenticity, and suggest a potential influence of authenticity on the resilience of self-control resources.

Keywords: authenticity, self-control, reward positivity, hedonic rewards, eudaimonic rewards,

4.2.1 Introduction

In my previous investigation (see Chapter 4 Part A), I found that exercising self-control increased reward responsivity (i.e., Reward Positivity amplitudes), and this aftereffect of self-control manifested in a domain-general manner. In other words, self-control exertion increased reward responsivity comparably for hedonic rewards and eudaimonic rewards.

One accepted method for elucidating a process underlying an experimental effect is to examine the role of individual difference variables that influence the tendency to engage the proposed process (Gohn & Clore, 2000; Underwood, 1975). Examining whether such individual differences moderate (i.e., qualify) an experimental effect is a way to test assumptions about the processes underlying the effect. Thus, in Chapter 4 Part B I will reexamine the possibility that eudaemonic processes drive the aftereffects of self-control exertion (Campbell et al., 2024) by examining how one such indicator of eudaemonic well-being – trait authenticity – moderates the effect of self-control exertion on reward responsivity.

Authenticity is defined as the perception of being one's true self (Kernis & Goldman, 2006). Previous researches indicated that there is a positive relationship between authenticity and self-control. The true-self-as-guide theory of authenticity posits that true self should guide behaviour (Rivera et al., 2019). From this perspective, acting in alignment with one's authentic self serves as an internal norm (e.g., follow who you real are), promoting congruence between individual choices and the true self, navigating conflict situations that require self-control, thereby contributing to well-being. The self-concordance model similarly suggests that pursuing self-concordant goals (i.e., goals consistent with one's intrinsic interests and core values) enables individuals to invest sustained effort in achieving them (Sheldon & Elliot, 1999). Empirical evidence provides preliminary and indirect support for the notion that a clear and consistent self-concept enables individuals to identify and prioritize self-initiated and personally valued goals, thereby enhancing effective self-control in goal pursuit. For example, people with higher self-concept clarity—possessing a clear and coherent self-view—are more likely to exhibit effective self-control (Jiang et al., 2023). In contract, those with low self-concept clarity tend to rely on external factors to maintain a coherent identity, making them more dependent on, and susceptible to, external influences (Campbell, 1990). Similarly, self-concept clarity and grit—defined as perseverance and passion for long-term goals (Duckworth et al., 2007), a construct closely related to self-control—have been shown to positively reinforce on each other over time (Wong & Vallacher, 2018). Recently, Li et al. (2023) found that authenticity predicted increases in self-control over time, and vice versa, providing direct evidence for the predictive effect of authenticity on self-control. In summary, although research on this topic is still limited, initial findings suggest that authenticity may indeed foster self-control.

Furthermore, as a eudaemonic construct, authenticity may also shape the eudemonic reward responsivity of self-control exertion. Eudaimonia calls for individuals to live authentically, aligning their lives with their true selves to achieve personal growth (Disabato et al., 2016; Ryan & Deci, 2001; Waterman, 1993). As one of the core elements of eudaimonia (Huta & Waterman, 2014; Smallenbroek et al., 2017), authenticity is positively related to the experience of meaning in life (Lutz et al., 2023; Schlegel et al., 2009; Schlegel & Hicks, 2011). Moreover, this meaning-making function of authenticity distinguishes from other forms of well-being (i.e., hedonia) (Schlegel et al., 2013; Smallenbroek et al., 2017). Consequently, individuals who live authentically are more likely to experience eudaimonia, in contrast to hedonia, which is focused primarily on immediate pleasure and satisfaction. In the context of reward processing, such relationship between authenticity and eudaimonia may also shape reward responsivity following self-control exertion, potentially differing based on reward types (e.g., hedonic versus eudaimonic rewards).

Taken together, I hypothesized that individuals with high authenticity may possess greater self-control resources, thus exerting self-control would have minimal impact on their already ample resources, leading to a relatively blunted reward sensitivity following self-control exertion; in contrast, individuals with low authenticity may have fewer self-control resources, thus would experience a significant depletion of these limited resources during self-control exertion, resulting in heightened reward sensitivity (i.e., Congruency × Authenticity interaction). Moreover, this effect would be stronger for eudaimonic (vs. hedonic) rewards (i.e., Congruency × Authenticity × Reward interaction). To investigate these hypotheses, participants exerted self-control (incongruent Stroop trials) or not (congruent Stroop trials) in a speeded reaction time task where they had the opportunity to win money for themselves (a hedonic reward) or a charity of their choosing (a eudaimonic reward), with trait authenticity measured.

4.2.2 Method

The methods and participants were as same as those reported in Chapter 4 Part A.

In addition, I also measured trait authenticity in this part. I assessed authenticity with the Southampton Authenticity Scale (SAS; Kelley et al., 2022), which consisted of four items, e.g., “In general, I feel authentic”. The SAS provides a psychometrically robust and concise assessment of authenticity as a unidimensional construct, exhibiting excellent reliability (α = 0.91) and strong convergent validity with the Authentic Living subscale (r = 0.62) (Kelley et al., 2022). Participants indicated the extent to which each item descripted them (1 = *strongly disagree*, 7 = *strongly agree*; M = 5.58, SD = 0.95, α = 0.88). I averaged scores across all items to formulate an average index representing trait authenticity, with higher values indicating higher trait authenticity.

In order to control for individual differences that are consistent within participants but vary between individuals—such as baseline reaction time and RewP—I applied a subtraction method, independently calculating the differences between the control group and each of the two experimental conditions (Tucker-Drob, 2011). This approach resulted in two distinct reward conditions: the hedonic condition (hedonic minus control) and the eudaimonic condition (eudaimonic minus control). By isolating the effects of each reward type relative to the control, it mitigates the confounding impact of baseline variability.

4.2.3 Results

I conducted a 2 (Reward: hedonic, eudaimonic) × 2 (Congruency: congruent trials, incongruent trials) × 2 (Authenticity: high, low) mixed Analysis of Variance (ANOVA) on hit rate, reaction time, and RewP (Figure 4.5). We reported means and standard deviations in Table 4.2.

Hit Rate and Reaction Time

The main effects of congruency were significant, as participants had a non-significantly higher hit rate, $F(1, 112) = 3.49$, $p = .064$, $\eta_p^2 = .03$, and were faster, $F(1, 112) = 40.62$, $p < .001$, $\eta_p^2 = .27$, on congruent than incongruent trials. The main effects of Reward were significant, as participants had a higher hit rate, $F(1, 112) = 8.50$, $p = .004$, $\eta_p^2 = .07$, and were faster, $F(1, 112) = 6.99$, $p = .009$, $\eta_p^2 = .06$, on hedonic than eudaimonic trials.

The Reward × Congruency interactions were significant for both hit rate, $F(1, 112) = 6.09$, $p = .015$, $\eta_p^2 = .05$, and reaction time, $F(1, 112) = 35.22$, $p < .001$, $\eta_p^2 = .24$. Hit rates were higher and reaction times were shorter for hedonic versus eudaimonic on incongruent trials (hit rate: $p = .002$, 95% *CI* = [0.92%, 4.05%]; reaction time: $p < .001$, 95% *CI* = [-20.55, -9.26]), but hit rates and reaction times were equal for hedonic versus eudaimonic on congruent trials (hit rate: $p = .557$, 95% *CI* = [-0.66%, 1.23%]; reaction time: $p = .253$, 95% *CI* = [-2.14, 8.05]).

However, neither the main effect of authenticity nor its interactions were significant for both hit rate and reaction time, *ps* > .05, indicating that trait authenticity did not impact reaction times and hit rates. No other main effects or interactions was significant, *ps* > .05.

RewP

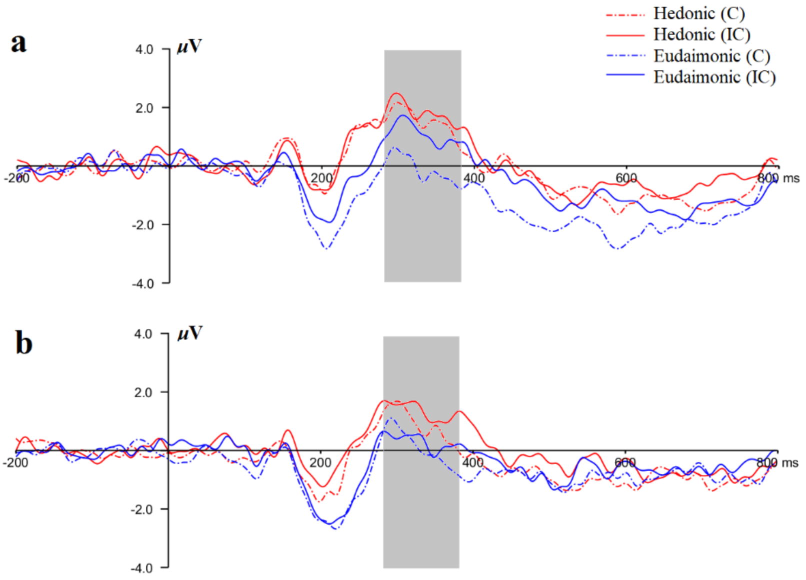
The Reward × Congruency × Authenticity interaction was significant, $F(1, 112) = 4.87$, $p = .029$, $\eta_p^2 = .04$. Simple effects tests showed that for low-level authenticity participants, the RewP was stronger in hedonic condition than eudaimonic condition on congruent trials, $p < .001$, 95% *CI* = [0.92, 2.40], but there was no significant difference between hedonic condition and eudaimonic condition on incongruent trials, p = .074, 95% *CI* = [-0.07, 1.54]. For high-level authenticity participants, the RewP was stronger in hedonic condition than eudaimonic condition on incongruent trials, p = .005, 95% *CI* = [0.32, 1.74], and the RewP was also stronger in hedonic condition than eudaimonic condition on congruent trials, p = .035, 95% *CI* = [0.05, 1.36]. I will return to these findings in the discussion section. However, the Congruency × Authenticity interaction was not significant, $p > .05$, indicating that trait authenticity did not impact the reward responsivity whether exerting self-control or not.

The main effect of Reward was significant, $F(1, 112) = 22.50$, $p < .001$, $\eta_p^2 = .17$. Post hoc tests showed that the RewP was stronger in hedonic condition than eudaimonic condition. The main effect of Consistency was significant, $F(1, 112) = 4.37$, $p = .039$, $\eta_p^2 = .04$. Post hoc tests showed that the RewP was stronger on incongruent trials than congruent trials, providing direct evidence for the reward responsivity hypothesis of self-control. However, the Congruency × Authenticity interaction was not significant, $F(1, 112) = 0.43$, $p = .516$, $\eta_p^2 = .004$, suggesting no significant differences in RewP among individuals high and low in authenticity between congruent trials and incongruent trials. I further examined Congruency × Authenticity interactions separately for two reward conditions, and found that there were no significant Congruency × Authenticity interactions for both reward conditions (Hedonic rewards: $F(1, 112) = 0.24$, $p = .624$, $\eta_p^2 = .002$; Eudaimonic rewards: $F(1, 112) = 2.53$, $p = .115$, $\eta_p^2 = .02$). No other main effects or interactions was significant, *p*s > .05.

Table 4.2 *Means and Standard Deviations for Hit Rate, Reaction Time, and RewP*

	Hit Rate (%)			
	Low Authenticity (*N* = 50)		High Authenticity (*N* = 64)	
	Hedonic	Eudaimonic	Hedonic	Eudaimonic
Congruent	1.85 (5.03)	1.67 (5.41)	3.70 (7.60)	3.33 (7.28)
Incongruent	1.30 (7.35)	0.00 (6.15)	3.70 (11.43)	0.03 (10.96)
	Reaction Time (ms)			
	Low Authenticity (*N* = 50)		High Authenticity (*N* = 64)	
	Hedonic	Eudaimonic	Hedonic	Eudaimonic
Congruent	-20.52 (35.15)	-24.58 (31.39)	-19.37 (31.72)	-21.23 (22.11)
Incongruent	-4.87 (40.37)	5.23 (39.55)	-15.99 (30.93)	3.72 (29.49)
	RewP (μV)			
	Low Authenticity (*N* = 50)		High Authenticity (*N* = 64)	
	Hedonic	Eudaimonic	Hedonic	Eudaimonic
Congruent	1.59 (2.97)	-0.07 (3.01)	0.85 (3.07)	0.14 (3.30)
Incongruent	1.84 (2.82)	1.11 (2.78)	1.38 (3.07)	0.36 (3.04)

Figure 4.5 *Grand Average Event-Related Potential Waveforms of the Reward Positivity as a Function of Self-control Exertion Among Individual in Low and High Authenticity*



Note. (**a**) Grand averages of Reward Positivity for individuals in low authenticity. (**b**) Grand averages of Reward Positivity for individuals in high authenticity. The grand-averaged waveforms represent the grand average across 3 frontal-central sites (O1, O2, and Oz). The Reward Positivity measurement window (i.e., 280 ms – 380 ms) in shaded in light gray. C = congruent trials; IC = incongruent trials.

4.2.4 Discussion

The primary goal of Chapter 4 Part B was to re-examine the possibility that eudaimonic processes shape the aftereffects of self-control exertion on reward responsivity with an individual differences approach (Gohn & Clore, 2000; Underwood, 1975). To do so, I re-analysed the data from Chapter 4 Part A and considered the moderating role of trait authenticity because it is a eudaemonic construct (Smallenbroek et al., 2017) and previous research has shown that exercising self-control increases eudaemonic processes (e.g., meaning in life; Campbell et al., 2024).

The results indicated that there were no significant differences in RewP among individuals high and low in authenticity between exerting self-control or not. While authenticity influences self-control (Li et al., 2023) and self-control exertion can enhance reward responsivity, the direct impact of authenticity on reward responsivity might not be as pronounced. In this case, RewP—an index of reward sensitivity—may remain relatively stable across groups, irrespective of self-control exertion. On the other hand, researchers emphasized authenticity as one's core self, an underlying, and potentially invisible aspect of the self (Harter, 2002; Kernis & Goldman, 2006; Lenton et al., 2013; Sedikides et al., 2017, 2019; Strohminger et al., 2017). Stated otherwise, authenticity is a deeper self beneath the veneer of social trappings and is not often or easily expressed to others. Therefore, although authenticity may not sensitize persons toward immediate neural response to rewards after self-control exertion, it is possible that authenticity may interact with self-control in a more delayed manner, influencing downstream processes such as subjective satisfaction, or emotional well-being. Taken together, future researches could explore the potential effect of authenticity on reward responsivity after self-control exertion in a more fined way.

The current study did not observe a stronger reward responsivity for individuals with a higher trait authenticity after self-control exertion. Unexpectedly, it was found that the RewP was stronger to hedonic rewards than eudaimonic rewards for people with high authenticity, no matter of exerting self-control or not. Meanwhile, the current study also found that the RewP was also stronger for people with low authenticity in no self-control exertion condition, whereas the RewP was comparable to hedonic rewards and eudaimonic rewards in the self-control exertion condition among individuals low in authenticity. Authenticity is positively related to self-control (Li et al., 2023; Rivera et al., 2019) and serves as a resilience factor against struggling event, such as stressful events (Reed et al., 2021; Ryan et al., 2005). In this vein, individuals high in authenticity are likely to be equipped with more self-control resources and are better able to recover efficiently from the cognitive and emotional demands of self-control, whereas individuals low in authenticity are likely to possess less self-control resources and may struggle to effectively recover from the cognitive and emotional demands of self-control, specifically in the self-control exertion condition, thus resulting in a blunted overall reward sensitivity. Future researches could further validate this possibility.

4.2.5 Conclusion

This study indicates that the enhancement of reward responsivity following self-control exertion was independent of trait authenticity. However, the results suggest that authenticity may play a role in the resilience of self-control resources. These insights contribute to a clearer understanding of the complex relationship between authenticity and self-control, offering a foundation for future research in this domain.

References

Bogdanov, M., Renault, H., LoParco, S., Weinberg, A., & Otto, A. R. (2022). Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cerebral Cortex, 32*(19), 4255–4270. https://doi.org/10.1093/cercor/bhab480

Campbell, A. V., Wang, Y., & Inzlicht, M. (2024). Experimental evidence that exerting effort increases meaning. *PsyArXiv*. https://doi.org/10.31234/osf.io/z3g9c

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, *59*(3), 538–549. https://doi.org/10.1037/0022-3514.59.3.538

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471–482. https://doi.org/10.1037/pas0000209

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087. https://doi.org/10.1037/0022-3514.92.6.1087

Gendolla, G. H. E., & Wright, R. A. (2009). Effort. In D. Sander & K. R. Scherer (Eds.), *Oxford companion to emotion and the affective sciences* (pp. 134–135). Oxford University Press.

Gohm, C. L., & Clore, G. L. (2000). Individual differences in emotional experience: Mapping available scales to processes. *Personality and Social Psychology Bulletin, 26*, 679-697.

Harmon-Jones, E., Matis, S., Angus, D. J., & Harmon-Jones, C. (2024). Does effort increase or decrease reward valuation? Considerations from cognitive dissonance theory. *Psychophysiology, 61*(6), Article e14536. https://doi.org/10.1111/psyp.14536

Harter, S. (2002). Authenticity. In C. S. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 382–394). New York: Oxford University Press.

Heintzelman, S. J. (2018). Eudaimonia in the contemporary science of subjective well-being: Psychological well-being, self-determination, and meaning in life. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being* (pp. 167–180). Def Publishers Salt Lake City, UT.

Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies*, *15*(6), 1425–1456. https://doi.org/10.1007/s10902-013-9485-0

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. https://doi.org/10.1177/01461672221109664

Kelley, N. J., Davis, W. E., Dang, J., Liu, L., Wildschut, T., & Sedikides, C. (2022). Nostalgia confers psychological wellbeing by increasing authenticity. *Journal of Experimental Social Psychology, 102*, 104379. https://doi.org/10.1016/j.jesp.2022.104379

Kelley, N. J., Finley, A. J., & Schmeichel, B. J. (2019). After-effects of self-control: The reward responsivity hypothesis. *Cognitive, Affective, & Behavioural Neuroscience, 19*(3), 600–618. https://doi.org/10.3758/s13415-019-00694-3

Kernis, M. H., & Goldman, B. M. (2005). From thought and experience to behaviour and interpersonal relationships: A multicomponent conceptualization of authenticity. In *Building, defending and regulating the self: A psychological perspective* (pp. 31–52). Psychology Press. https://doi.org/10.4324/9780203998052

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 283–357). Academic Press. https://doi.org/10.1016/S0065-2601(06)38006-9

Ma, Q., Meng, L., Wang, L., & Shen, Q. (2014). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research, 261*, 1–7. https://doi.org/10.1016/j.bbr.2013.11.045

Lenton, A. P., Bruder, M., Slabu, L., & Sedikides, C. (2013). How does “being real” feel? The experience of state authenticity: state authenticity. *Journal of Personality*, *81*(3), 276–289. https://doi.org/10.1111/j.1467-6494.2012.00805.x

Li, Q., Ren, X., Zhou, Z., & Wang, J. (2023). Reciprocal relationships between self-control and self-authenticity: a two-wave study. *Frontiers in Psychology*, *14*, Article 1207230.

Lutz, P. K., Newman, D. B., Schlegel, R. J., & Wirtz, D. (2023). Authenticity, meaning in life, and life satisfaction: A multicomponent investigation of relationships at the trait and state levels. *Journal of Personality*, *91*(3), 541–555. https://doi.org/10.1111/jopy.12753

Pan, W., Lu, J., Wu, L., Kou, J., & Lei, Y. (2023). Expending effort may share neural responses with reward and evokes high subjective satisfaction. *Biological Psychology, 177*, Article 108480. https://doi.org/10.1016/j.biopsycho.2022.108480

Reed, D. E., Lehinger, E., Cobos, B., Vail III, K. E., Nabity, P. S., Helm, P. J., Galgali, M. S., & McGeary, D. D. (2021). Authenticity as a resilience factor against CV-19 threat among those with chronic pain and posttraumatic stress disorder. *Frontiers in Psychology*, *12*, 643869.

Rivera, G. N., Christy, A. G., Kim, J., Vess, M., Hicks, J. A., & Schlegel, R. J. (2019). Understanding the relationship between perceived authenticity and well-being. *Review of General Psychology*, *23*(1), 113–126. https://doi.org/10.1037/gpr0000161

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166. https://doi.org/10.1146/annurev.psych.52.1.141

Ryan, R. M., LaGuardia, J. G., & Rawsthorne, L. J. (2005). Self-complexity and the authenticity of self-aspects: Effects on well being and resilience to stressful events. *North American Journal of Psychology*, *7*(3).

Schlegel, R. J., & Hicks, J. A. (2011). The true self and psychological health: Emerging evidence and future directions. *Social and Personality Psychology Compass*, *5*(12), 989–1003. https://doi.org/10.1111/j.1751-9004.2011.00401.x

Schlegel, R. J., Hicks, J. A., Arndt, J., & King, L. A. (2009). Thine own self: True self-concept accessibility and meaning in life. *Journal of Personality and Social Psychology*, *96*(2), 473–490. https://doi.org/10.1037/a0014060

Schlegel, R. J., Hirsch, K. A., & Smith, C. M. (2013). The importance of who you really are: The role of the true self in eudaimonia. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonic functioning* (pp. 207–226). Washington, DC: American Psychological Association

Sedikides, C., Lenton, A. P., Slabu, L., & Thomaes, S. (2019). Sketching the contours of state authenticity. *Review of General Psychology*, *23*(1), 73–88. https://doi.org/10.1037/gpr0000156

Sedikides, C., Slabu, L., Lenton, A., & Thomaes, S. (2017). State authenticity. *Current Directions in Psychological Science*, *26*(6), 521–525. https://doi.org/10.1177/0963721417713296

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*(3), 482–497.

Shizgal, P. (1999). On the neural computation of utility: Implications from studies of brain stimulation reward. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 502–526). Russell Sage Foundation.

Smallenbroek, O., Zelenski, J. M., & Whelan, D. C. (2017). Authenticity as a eudaimonic construct: The relationships among authenticity, values, and valence. *The Journal of Positive Psychology, 12*(2), 197-209. https://doi.org/10.1080/17439760.2016.1187198

Steger, M. F., Kashdan, T. B., & Oishi, S. (2008a). Being good by doing good: Daily eudaimonic activity and well-being. *Journal of Research in Personality, 42*(1), 22–42. https://doi.org/10.1016/j.jrp.2007.03.004

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science, 12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galván, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proceedings of the National Academy of Sciences, 111*(18), 6600–6605. https://doi.org/10.1073/pnas.1323014111

Tucker-Drob, E. M. (2011). Individual differences methods for randomized experiments. *Psychological Methods*, *16*(3), 298–318. https://doi.org/10.1037/a0023349

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30*, 128-134.

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, *64*(4), 678–691.

Wong, A. E., & Vallacher, R. R. (2018). Reciprocal feedback between self-concept and goal pursuit in daily life. *Journal of Personality*, *86*(3), 543–554. https://doi.org/10.1111/jopy.12334

Chapter 5 General Discussion

The purpose of this thesis was to examine authenticity, including its nature (self-enhancement, self-accuracy, and self-consistency views) by comparing the authentic self against the presented self (Empirical Paper I and Empirical Paper II), and its influence on reward processing for different types of rewards (Empirical Paper III). In this chapter, I will summarise the key findings of all studies. Next, I will discuss theoretical and empirical implications of the findings. I will then consider strengths of this thesis followed by limitations and recommendations for future research. I will end the chapter with a conclusion.

5.1 Summary of Key Findings

In Chapter 2, I presented two experiments testing, via the SR-valence task, the relative strength of self-positivity for the authentic and presented selves. I offered two competing hypotheses. First, in line with the self-enhancement view, I hypothesized that the strength of self-positivity would be comparable for the authentic and presented selves. Alternatively, in line with the self-accuracy and self-consistency views, I hypothesized that the strength of self-positivity would be weaker for the authentic compared to the presented self. The results of Experiments 1 and 2 showed that participants ascribed less positivity to the authentic (than the presented) self: They endorsed more negative traits and were faster to admit having them; also, they endorsed fewer positive traits and were slower to admit having them. Neurally, in Experiment 2, participants manifested preferential processing of threatening information (P1), followed by preferential processing of favourable information (N170), about the presented self, indicating its brittleness. At a later stage (LPP), participants engaged in more elaborate processing of both threatening and favourable information about the authentic self, indicating its subjective importance. Authenticity, albeit mostly positive, allows room for negativity.

In Chapter 3, I evaluated the replicability of the findings from Empirical Paper I by testing how the authentic and presented selves are influenced by emotional self-descriptiveness through a self-referent emotional Stroop task. Based on the emotional Stroop effect—pronounced slowdown in responding to the ink colour of negative (vs. positive or neutral) words—I offered two competing hypotheses. First, in line with the self-enhancement view, the interference of negative (vs. positive) self-evaluations on the authentic self would be on par with the effect observed for the presented self (i.e., an amplified emotional Stroop effect). Second, in line with the self-accuracy and self-consistency views, the interference of negative (vs. positive) self-evaluations would be less pronounced for the authentic self than the presented self (i.e., an attenuated emotional Stroop effect). The results showed that colour naming latency was slower to negative (vs. positive) traits for the presented-self and control condition. In contract, colour naming latency was faster to negative (vs. positive) traits for the authentic self. ERPSs indicated that negative and positive traits for the authentic self did not differ at either early (P1) and later (P3) stages of attentional processing. However, P1 was larger for negative, and P3 was larger for positive, traits for the presented self. These findings suggest that the presented self is highly motivated by self-enhancement, rendering it vulnerable to negative self-descriptors, whereas the authentic self is also guided by self-consistency and self-accuracy motivation, integrating both positive and negative self-aspects with comparable weight.

In Chapter 4, I expanded the reward responsivity hypothesis of self-control by testing the extent to which exercising self-control influences the reward system differently for hedonic versus eudaimonic rewards, and examined how this relationship is influenced by trait authenticity using a modified monetary incentive delay task. I offered three competing hypotheses regarding the influence of self-control exertion on reward responsivity. First, exercising self-control would increase hedonic reward responsivity. Second, exercising self-control would increase eudaimonic reward responsivity. Third, exercising self-control would increase reward responsivity similarly for hedonic and eudaimonic rewards (i.e., in a domain-general manner). Also, based on the established links between authenticity and both self-control and eudaimonia, I hypothesized that individuals high on authenticity would exhibit a relatively blunted reward sensitivity following self-control exertion, whereas individuals low on authenticity would manifest a heightened reward sensitivity following self-control exertion, and this effect would be stronger for eudaimonic (vs. hedonic) rewards. Participants evinced a larger RewP after exercising (vs. not exercising) self-control, in accordance with the reward responsivity hypothesis. Self-control and reward type did not interactively modulate RewP, suggesting that self-control increases the reward responsivity in a domain-general manner. In addition, trait authenticity did not interact with self-control exertion to influence reward positivity, suggesting that the increased reward responsivity after self-control exertion was not influenced by trait authenticity. These findings offer a neurophysiological basis for the reward responsivity hypothesis and underscore the role of authenticity in self-control dynamics.

5.2 Implications

The thesis helps to clarify the three theoretical views on authenticity. Some researchers conceptualized authenticity through the lens of self-accuracy, the candid and unbiased processing of identity relevant information (Kernis & Goldman, 2006; Lakey et al., 2008). In part because self-accuracy is difficult to empirically verify (Vazire & Wilson, 2012), others conceptualized authenticity as self-consistency, the alignment of one's behaviour with internal standards, goals, or values (Kernis & Goldman, 2006; Wood et al., 2008). Still, other researchers considered authenticity as self-enhancement (Bailey & Iyengar, 2023; Bench et al., 2015; Guenther et al., 2024; Strohminger et al., 2017). In Chapters 2 and 3, I subjected the three views of authenticity to empirical scrutiny by comparing the authentic self to the presented self. The results of Chapters 2 and 3 indicate that the authentic self is more than self-enhancement, challenging prior evidence favouring the self-enhancement view (Sedikides & Schlegel, 2024). Specifically, although participants endorsed more positive than negative traits as self-descriptive and demonstrated faster endorsement of positive than negative traits as self-descriptive when making decisions for the authentic self, these effects were weaker in reference to the presented self (Chapter 2). Moreover, the authentic self exhibited attenuated emotional Stroop effect (e.g., a threat-driven slowdown caused by automatic attention allocation to threatening stimuli at the expense of concurrent task demands; Algom et al., 2004; Öhman, 1993; Öhman et al., 2001), whereas the presented self exhibited attenuated emotional Stroop effect (Chapter 3). Therefore, while motived by self-enhancement, the authentic self—compared to the presented self—demonstrates two critical functional characteristics: (1) a greater capacity to acknowledge negative self-relevant traits (Chapter 2), and (2) superior resistance to negativity interference effects (Chapter 3). These results also align with both self-accuracy (e.g., its tolerance for negative yet veridical self-appraisal) and self-consistency (e.g., its resistance to external influences) views, suggesting that authenticity achieves a dynamic equilibrium between enhancement and realism. In fact, although the current thesis provides a clear empirical distinction between the self-consistency/self-accuracy and self-enhancement views by comparing the strength of self-positivity of the authentic self and the presented self, these three pathways to authenticity may not be mutually exclusive. Self-enhancement is as nourishing to the psychological self as food is for the physical self (Sedikides, 2021; Sedikides & Gregg, 2008), thus it is be integral to one's core self. Although, the findings are compatible with the self-accuracy and self-consistency views, they opened the possibility that these motivations (i.e., to seek the authentic accurately, consistently, and positively) might jointly contribute to the experience of authenticity.

This thesis also contributes a novel perspective to the literature on self-enhancement as it pertains to authentic self by incorporating the processing of negativity. Self-enhancement is thought to operate broadly, manifesting across self-representations (Alicke & Sedikides, 2009; Sedikides, 2020, 2021), influencing substantially the authentic self (Sedikides & Schlegel, 2024). However, as mentioned in the last paragraph, the authentic self is more than self-enhancement. Although prior research has predominantly emphasized the connection between authenticity and the endorsement of positivity (Guenther & Sedikides, in press), this thesis suggests that authenticity is not merely about amplifying positive self-aspects but also involves an engagement with negative self-aspects. The results strengthen a humanistic perspective (Kernis & Goldman, 2005, 2006), stating that authenticity involves integration of both positivity and negativity. This interplay parallels other phenomena in which balanced experiences yield favourable outcomes. For example, nostalgia, considered as a bittersweet emotion, has been shown to enhance psychological well-being (Sedikides & Wildschut, 2016; Wildschut et al., 2006; Wildschut & Sedikides, 2020). Similarly, emodiversity—the variety and relative abundance of the emotions people experience, encompassing both positive and negative emotions (Quoidbach et al., 2014)—can improve both mental and physical health outcomes (Ong et al., 2018; Quoidbach et al., 2018). In this vein, the dual nature of authenticity—embracing both positivity and negativity—may serve as a core mechanism for maintaining psychological equilibrium. Specifically, by reconciling conflicting self-aspects, the authentic self demonstrates an acceptance of all facets of one's identity (Rogers, 1961), fostering a pathway toward self-actualization (Maslow, 1943) and, ultimately, improving well-being.

The thesis additionally offers insights into the role of authenticity in shaping self-control and reward processing. Although previous research has indicated that authenticity can enhance self-control (Li et al., 2023; Rivera et al., 2019) and that self-control exertion can enhance reward responsivity (Bogdanov et al., 2022; Harmon-Jones et al., 2024), the present findings demonstrate that the effect of self-control exertion on reward responsivity is not influenced by trait authenticity. Although unexpected, this result contributes to refining theoretical models by delineating where the effect of authenticity may be limited, thus offering a more subtle understanding of the boundaries of its influence in self-control processes and reward dynamics. Specifically, the results revealed that the RewP was stronger for hedonic (vs. eudaimonic) rewards among participants with high authenticity, as well as among those with low authenticity in the absence of self-control exertion. However, this pattern did not hold under the self-control exertion condition among participants with low authenticity. These findings suggest that the influence of authenticity may be more pronounced in the resilience of self-control resources (Reed et al., 2021; Ryan et al., 2005)— for example, participants with low authenticity exhibited greater difficulty in replenishing self-control resources following exertion—rather than in generalized neurophysiological processes such as reward responsivity. Along these lines, authenticity may facilitate recovery from resource depletion and support subsequent self-control efforts, but it does not necessarily modulate the fundamental neural sensitivity to rewards following self-control exertion. From another perspective, authenticity is conceptualized as an underlying aspect of the self (Harter, 2002; Kernis & Goldman, 2006; Lenton et al., 2013; Sedikides et al., 2017, 2019; Strohminger et al., 2017). Thus, whereas authenticity does not enhance immediate neural responses to rewards after self-control exertion, it may exert its influence in a more delayed or indirect manner, potentially affecting downstream processes such as emotional well-being, or long-term goal alignment. In summary, this thesis clarifies the nature of authenticity within the framework of three theoretical views and delineates a boundary for its influence, contributing to a more nuanced understanding of authenticity and its role in the context of self-control and reward processes.

Leveraging ERPs, the thesis elucidated the intricate neural mechanism underlying authenticity and provided direct evidence for the reward responsivity hypothesis of self-control. At the initial processing stage, the authentic (vs. presented) self exhibited attenuated neural responses to negative information. Specifically, the P1 was diminished when participants endorsed negative traits as descriptive of the authentic than presented self (Chapter 2). Similarly, the P1 was larger when identifying the colour of negative than positive traits for the presented self, whereas no such difference was observed for the authentic self (Chapter 3). P1 is sensitive to emotional stimuli (Mueller et al., 2013; Schindler & Bublatzky, 2020), and can further differentiate between non-threatening and potentially threatening information (Zhang et al., 2014). Therefore, the P1 findings in Chapters 2 and 3 suggested preferential processing of negative information referring to the presented (vs. authentic) self. In contrast, the authentic self exhibited reduced sensitivity to potentially threatening information at the very initial stage of processing.

This pattern was reversed at the subsequent processing stage: the presented self (vs. the authentic self) exhibited heightened neural responses to positive information. In particular, the N170 was strengthened when participants endorsed positive traits as descriptive of the presented than authentic self (Chapter 2). Similarly, the P3 was larger for positive than negative traits in the context of the presented self, whereas no difference was observed for the authentic self (Chapter 3). Taken together, the findings reveal preferential processing toward positive information related to the presented (vs. authentic) self. In contrast, positive information holds less significance for the authentic self.

This asymmetrical neural patterns during the early and subsequent processing stages aligns with the mobilization-minimization hypothesis (Taylor, 1991; also see Sedikides et al., 2016), which posits that negative or threatening information elicits 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 this thesis, negative self-descriptive information initially received preferential processing initially (mobilization; P1), followed by a shift towards positive self-descriptive information (minimization; N170, P3). Notably, this mobilization-minimization dynamic was more pronounced for the presented than authentic self, suggesting that it requires greater defence mechanisms due to its malleable nature, whereas the authentic self remains more stable and less susceptible to such fluctuations.

Furthermore, these findings underscore the complexity of self-referential processing, particularly when intertwined with emotional content. Previous studies have highlighted the multi-stage nature of self-related processing in emotional contexts. For example, Herbert et al. (2011) demonstrated that emotional content rapidly captured attention (reflected in augmented early posterior negativity for unpleasant and pleasant nouns vs. neutral ones), followed by self-referential processing (manifested as augmented LPPs for unpleasant nouns only when preceded by personal pronouns). However, self-referential processing may occur earlier than emotional processing, with self-other discrimination emerging as early as the P1, and interactions between self-reference and emotional valence appearing later, manifested in the LPP (Zhou et al., 2017). Despite variations in prioritizing self-referential versus emotional processing, the current thesis extends this body of work by examining these processes within the framework of different self-representations (e.g., the authentic self, the presented self), offering a fine-grained neural deconstruction of authenticity.

Finally, the results concerning RewP provides robust empirical evidence for the reward responsivity hypothesis of self-control (Kelley et al., 2019). Although prior evidence has indicated that effort—an essential component of self-control (Kotabe & Hofmann, 2015)—can increase immediate neural responsivity to rewards (as indexed by RewP), relatively few studies have directly examined the specific effects of self-control exertion on RewP. Therefore, the present findings address this gap by offering direct empirical evidence of the aftereffects of self-control exertion on reward processing. Moreover, although the majority of studies examining the effects of self-control exertion on reward responsivity have focused on hedonic rewards (Kelley et al., 2019), here exercising self-control increased reward responsivity comparably for hedonic and eudaimonic rewards, indicating that the exercising self-control produces domain general increases in reward responsivity.

In summary, by empirically examining the multi-stage processing of authenticity via emotion-related and self-related ERPs, this thesis advances understanding of the neurophysiological mechanisms underlying authenticity. In addition, by empirically investigating the reward responsivity after self-control exertion via RewP, the thesis provided direct evidence for the reward responsivity hypothesis of self-control.

5.3 Strengths, Limitations and Future Directions

5.3.1 Strengths

Theoretical Strengths

I was concerned how two complementary pathways—the pursuit of positivity (and acknowledgment of negativity) and the resistance to negativity—contribute to authenticity-seeking. Specifically, in Chapter 2, I tested the three theoretical views on authenticity by examining the endorsement, and reaction time thereof, of positive traits and rejection of negative traits, whereas, in Chapter 3, I addressed the same issue by assessing resistance to the interference of negative information. The dual pathways of embracing positivity and managing negativity may be central to explaining the complexity of self-related processes, specifically for authenticity. Whereas the self is often marked by relentless pursuit of positivity, the capacity to fearlessly acknowledge, manage, and mitigate the influence of negative self-aspects is likely a contributing factor in sustaining psychological well-being and social functioning (Carson & Langer, 2006; Neff, 2003; Rogers, 1961). Therefore, authenticity emerges not merely as an amplifier of positive self-perceptions but also as a regulator of negative self-representations. By integrating these dual pathways, this thesis offers a more comprehensive framework for understanding authenticity, expanding the literature to encompass its roles in both positive and negative self-dynamics.

Methodological Strengths

Beyond traditional assessments using self-reports (e.g., the Authenticity Scale; Wood et al., 2008) and behavioural tasks (e.g., SR-valence tasks; Cai et al., 2016) - which risk response distortions due to defensive biases or introspective limitations (Koole, 2003) - this thesis additionally introduces novel electrophysiological measures (i.e., ERPs) to investigate authenticity. ERPs, which index fluctuations in neural activity with exceptional temporal resolution across time, are suitable for capturing rapid neural dynamics in complex processing (e.g., emotional processing, self-related processing, and reward processing) that manifest within temporally proximal substages (Luck, 2014), especially for complex concept such as authenticity. Therefore, the current thesis combined behavioural and neuroscientific approach to examine both the behavioural mechanism and electrophysiological underpinnings underlying self by recording participants' electroencephalography signal while they carried out the behavioural task. In Chapters 2 and 3, I used ERPs relevant to early-stage selective attention (P1) and attention allocation (N170) in response to emotionally evocative stimuli to test the self-enhancement view of the self, as well as earlier (P2) and later (P3) stages of self-relevant information processing and later elaborate processing and stimulus significance (LPP) to compare the difference between the authentic self and the presented self. In Chapter 4, I applied RewP—a sensitive index to feedback signalling the outcome of an action—to test the reward responsivity hypothesis of self-control and the potential influence of trait authenticity. These multi-stage ERP components serve as covert measures independently of behavioural responses, compensating for the defect that singular behavioural measurement may not be able to capture the complexity of authenticity. By mapping the above processes to specific ERP components, I capture the real-time neural dynamics of authenticity processing across distinct temporal stages.

5.3.2 Limitations and Future Directions

Individuals can simultaneously possess both authentic and presented selves (Sedikides & Gregg, 2003). One future research direction is to identify specific contextual factors that enhance the salience of one self-representation over the other. Existential threats such as mortality salience may serve as one of the potential moderators of this relative prominence. Distal death priming (e.g., reflecting on death in the distant future) tends to enhance meaning maintenance and construction (Vess et al., 2017), which is related with authentic self-pursuing. However, proximal priming (e.g., thinking about imminent death) can induce defensive self-enhancement strategies, as evidenced by increased preference for high-status products like luxury watches people (Mandel & Heine, 1999), which is more associated with presented self. To elucidate these mechanisms, future research could employ within-subjects death priming paradigms (varying prime proximity) to test the effect of death priming on self-representation.

In Experiment 2 of Empirical Paper I, trial numbers were imbalanced across conditions (see Table S2.2). This disparity could potentially introduce frequency-related biases in neural adaptation. For example, components like P2 and P3 are sensitive to infrequent stimuli and exhibit enhanced amplitudes to infrequent stimuli (Glazer & Nusslock, 2022; Luck & Hillyard, 1994; Polich, 2012; Pritchard, 1981). Such disparity could also reduce reliability in amplitude quantification due to unequal signal-to-noise ratios (Luck, 2014). Future designs could mitigate this issue by (a) having participants select self-descriptive traits prior to the experiment (Schlegel et al., 2009) to ensure balanced and sufficient trial counts for stable EEG analysis; (b) implementing multivariate analysis techniques (e.g., Multivariate Pattern Analysis, MVPA), which is less sensitive to trial-count differences (Grootswagers et al., 2017).

The current thesis focused exclusively on Western undergraduate samples and this reliance on Western, Educated, Industrialized, Rich, and Democratic (WEIRD) samples may constrain the generalizability of the results. Specifically, this reliance may overlook critical cultural variations in the conceptualization and experience of authenticity. Cultural frameworks appear to diverge in their emphasis on self-construal: Western cultures prioritize autonomy and promote independent self-construal, whereas Eastern cultures emphasize hierarchy and encourage interdependent self-construal (Cross et al., 2011; Markus & Kitayama, 2010, 2014). Therefore, in Western contexts, the authentic self is primarily driven by autonomous motivation, with perceived authenticity often judged based on internal needs and personal motivations; in contract, in Eastern contexts, the authentic self is more frequently driven by relational dynamics, such as hierarchy, proximity, and contextual harmony (Liang & Xie, 2021). In such settings, the authentic self balances personal needs with the requirements of relational contexts, integrating the authentic self into social environments in a way that satisfies both individual and collective needs (Liang & Xie, 2021; Robinson et al., 2013). In this vein, individuals with interdependent self-construal (characteristic of Eastern cultures) would demonstrate stronger self-enhancement when evaluating authenticity, as their authentic self-concept may incorporate socially presented aspects to maintain relational harmony, thus creating a positivity priority in self-evaluation. Conversely, those with independent self-construal (characteristic of Western cultures) should show greater self-consistency/self-accuracy, as they prioritize internal consistency over social approval in authenticity judgments. On the other hand, although there is some common ground regarding authenticity across cultural groups, such as a shared belief in the "goodness" of the authentic self (De Freitas et al., 2018) and a positive relationship between authenticity and psychological well-being (English & Chen, 2011; Rathi & Lee, 2021), people from independent and interdependent cultural contexts may express authenticity differently due to self-construal or thinking style (English & Chen, 2011; Slabu et al., 2014). To address these cultural nuances and enhance the cross-cultural validity of the findings, future research should replicate and extend this work in non-Western contexts, particularly in cultures characterized by interdependent self-construal. Such investigations will provide a more comprehensive understanding of the essential of authenticity and its role in social functioning across diverse cultural landscapes.

An additional limitation of the current research pertains to neuroscience methodology. The neural evidence relied on EEG, which has excellent temporal resolution but poor spatial resolution (Cohen, 2017). Functional magnetic resonance imaging (fMRI) studies of reward-relevant brain regions may complement the findings. Reward-related brain regions like the striatum are critical to self-processing (Berridge & Kringelbach, 2013; Delgado, 2007). Thinking about the self feels good and activates parts of the striatum (Enzi et al., 2009). Evidence of decreased striatal activation when making judgments about the authentic (vs. presented) self would bolster the findings. However, increased striatal activation when making judgments about the authentic (vs. presented) self would support the self-enhancement view of authenticity. Other neuroscientific studies link self-enhancement to both structural (Chavez & Heatherton, 2015; Chester et al., 2016) and functional (Chavez & Heatherton, 2015) connectivity between self-relevant (medial prefrontal cortex) and reward relevant (striatum) brain regions. If self-enhancement underlies the authentic self, one would anticipate especially strong connectivity between medial prefrontal cortex and the striatum when participants make judgments about the authentic self. However, based on the current findings, one would obtain weaker connectivity between medial prefrontal cortex and the striatum when participants make judgements about the authentic self. Furthermore, reward-related fMRI studies could provide valuable insights into the neural mechanisms underlying the relationship between self-control, reward processing, and authenticity. Although I did not observe a direct influence of authenticity on reward responsivity after self-control exertion, examining whether individuals high versus low on authenticity show distinct strengths of functional connectivity in response to rewards following self-control exertion could clarify how authenticity modulates reward sensitivity via self-control.

A further limitation of this thesis may be the inability to delineate the role of authenticity in reward responsivity following self-control exertion. On the one hand, there is a close connection between authenticity and eudaimonia. Authenticity is one of the core elements of eudaimonia (Huta & Waterman, 2014; Smallenbroek et al., 2017). Moreover, eudaimonia emphasizes living in alignment with one's true self to foster personal growth (Disabato et al., 2016; Ryan & Deci, 2001; Waterman, 1993). On the other hand, previous research has demonstrated a positive relationship between authenticity and self-control, with individuals high on authenticity exhibiting greater self-control capabilities (Li et al., 2023; Rivera et al., 2019). In the context of reward processing, the relationship between authenticity and both self-control and eudaimonia may indicate an intricate dynamic of how individuals respond to different types of rewards (e.g., hedonic vs. eudaimonic) following self-control exertion. Specifically, authenticity may simultaneously influence self-control and reward processing, shaping reward responsivity to these rewards. For instance, individuals low on authenticity may exhibit diminished reward responsivity to eudaimonic rewards, potentially due to their limited self-control resource or an inherent lower responsiveness to eudaimonic rewards. This dual influence introduces additional complexity in the relationship among authenticity, self-control, and reward processing. Therefore, future research is needed to adopt sophisticated designs that would disentangle these dynamics. For example, longitudinal designs could explore how authenticity influences self-control and reward processing dynamically, shedding light on the temporal dynamics of these relationships.

An additional limitation of the thesis is the limited attention to the nuanced neural dynamics of eudaimonic processes. Multiple theoretical perspectives indicate that eudaimonic processes (e.g., the pursuit of meaning) are central to psychological experience (Becker, 1971; Frankl, 1963; Greenberg et al., 2004; Heine et al., 2006; Martela et al., 2018; Pyszczynski et al., 2015; Sedikides & Wildschut, 2018; Steger et al., 2008, 2009; Wong et al., 2013). Despite their theoretical prominence, the present findings revealed that reward responsivity, as indexed by the RewP, was weaker for eudaimonic rewards compared to hedonic rewards. This could be contributed its weak temporal decay—the RewP to eudaimonic rewards is less sensitive to temporal decay than the RewP to hedonic rewards (Luo et al., 2022). Although self-control may not sensitize persons toward eudaimonic rewards in-the-moment, their weaker temporal decay may make eudaimonic rewards well-suited for countermanding the aversiveness of self-control over time. Indeed, a weaker temporal decay of the RewP to eudaimonic (vs. hedonic) rewards may help to explain longitudinal associations between self-control and positive life outcomes, such as well-being and achievement (Moffitt et al., 2011). Therefore, given the centrality of eudaimonic processes to human life and their resistance to temporal decay, further studies should explore strategies to maximize the impact of eudaimonic rewards in enhancing social functioning and long-term well-being. Additionally, beyond the RewP, other ERPs may provide more differentiated insights into the differential processing of hedonic and eudaimonic rewards following self-control exertion. For example, the late-positive potential (LPP) is driven by stimulus significance above and beyond other factors (Hajcak & Foti, 2020). Insofar as eudaimonic (vs. hedonic) rewards are more psychologically enriching, they should modulate the LPP after self-control exertion.

The thesis did not identify differential effects of self-control exertion on the neural processing of hedonic versus eudaimonic rewards. That is, I did not observe a differential effect of self-control exertion on reward responsivity enhancement for hedonic verses eudaimonic rewards. However, previous studies have highlighted relative differentiation between eudaimonic rewards and hedonic rewards. For example, neural activity associated with eudaimonic rewards predicts increases in well-being, whereas neural activity associated with hedonic rewards predicts decreases in well-being (Luo et al., 2019, 2022; Telzer et al., 2014). More relevant, the RewP to eudaimonic rewards is less sensitive to temporal decay than the RewP to hedonic rewards (Luo et al., 2022). Thus, although self-control exertion strengthens reward responsivity similarly for both hedonic and eudaimonic rewards, the underlying neural mechanisms may differ. For example, eudaimonic rewards may engage more stable and enduring motivational processes, whereas hedonic rewards may operate through transient affective pathways. Future research could address these distinctions by examining the temporal dynamics of reward processing following self-control exertion, providing insights into how different reward types are distinguished within the self-control.

Finally, when re-examining whether trait authenticity moderates the effect of self-control exertion on reward responsivity in Chapter 4 Part B. I employed a median split to categorize individuals into high versus low trait authenticity groups. Although this approach has certain advantages (MacCallum et al., 2002), literature indicates that dividing continuous data using a median split (or similar grouping methods) may reduce statistical power and increase likelihood of producing spurious significant results (Irwin & McClelland, 2003; Fitzsimons, 2008). Future research could consider utilizing Linear Mixed Models (LMM) to analyse repeated measures data, thereby providing a more comprehensive understanding of the study's findings.

5.4 General Conclusion

Authenticity has long captivated the interest of scholars, researchers, and the public alike. This thesis sought to uncover the essence of authenticity and examine its influence on social functions (e.g., the aftereffects of self-control). Although authenticity is predominantly self-enhancing, it also encompasses a willingness to confront less favourable aspects of the self, such as acknowledging personal weaknesses and resisting interference from negative information—dimensions that reflect self-accuracy and self-consistency. Moreover, authenticity has the potential to bolster resilience in self-control resources following exertion. In all, the thesis offers behavioural and neurophysiological evidence supporting the nuanced balance between self-enhancement and self-accuracy/self-consistency motivations, while highlighting its pivotal role in the dynamics of self-control.

References

Algom, D., Chajut, E., & Lev, S. (2004). A Rational Look at the Emotional Stroop Phenomenon: A Generic Slowdown, Not a Stroop Effect. *Journal of Experimental Psychology: General*, *133*(3), 323–338. https://doi.org/10.1037/0096-3445.133.3.323

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*(1), 1–48. https://doi.org/10.1080/10463280802613866

Bailey, E. R., & Iyengar, S. S. (2023). Positive—More than unbiased—Self-perceptions increase subjective authenticity. *Journal of Personality and Social Psychology*, *125*(6), 1351–1372. https://doi.org/10.1037/pspa0000359

Becker, E. (1971). *The birth and death of meaning.* Free Press.

Bench, S. W., Schlegel, R. J., Davis, W. E., & Vess, M. (2015). Thinking about change in the self and others: The role of self-discovery metaphors and the true self. *Social Cognition*, *33*(3), 169–185. https://doi.org/10.1521/soco.2015.33.3.2

Berridge, K. C., & Kringelbach, M. L. (2013). Neuroscience of affect: Brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology*, *23*(3), 294–303. https://doi.org/10.1016/j.conb.2013.01.017

Bogdanov, M., Renault, H., LoParco, S., Weinberg, A., & Otto, A. R. (2022). Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cerebral Cortex*, *32*, 4255–4270. https://doi.org/10.1093/cercor/bhab480

Cai, H., Wu, L., Shi, Y., Gu, R., & Sedikides, C. (2016). Self-enhancement among Westerners and Easterners: A cultural neuroscience approach. *Social Cognitive and Affective Neuroscience*, *11*(10), 1569–1578. https://doi.org/10.1093/scan/nsw072

Carson, S. H., & Langer, E. J. (2006). Mindfulness and self-acceptance. *Journal of Rational-Emotive & Cognitive-Behaviour Therapy*, *24*(1), 29–43. https://doi.org/10.1007/s10942-006-0022-5

Chavez, R. S., & Heatherton, T. F. (2015). Multimodal frontostriatal connectivity underlies individual differences in self-esteem. *Social Cognitive and Affective Neuroscience*, *10*(3), 364–370. https://doi.org/10.1093/scan/nsu063

Chester, D. S., Lynam, D. R., Powell, D. K., & DeWall, C. N. (2016). Narcissism is associated with weakened frontostriatal connectivity: A DTI study. *Social Cognitive and Affective Neuroscience*, *11*(7), 1036–1040. https://doi.org/10.1093/scan/nsv069

Cohen, M. X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences*, *40*(4), 208–218. https://doi.org/10.1016/j.tins.2017.02.004

Cross, S. E., Hardin, E. E., & Gercek-Swing, B. (2011). The what, how, why, and where of self-construal. *Personality and Social Psychology Review*, *15*(2), 142–179. https://doi.org/10.1177/1088868310373752

De Freitas, J., Sarkissian, H., Newman, G. E., Grossmann, I., De Brigard, F., Luco, A., & Knobe, J. (2018). Consistent belief in a good true self in misanthropes and three interdependent cultures. *Cognitive Science*, *42*(S1), 134–160. https://doi.org/10.1111/cogs.12505

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*(1), 70–88. https://doi.org/10.1196/annals.1390.002

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471–482. https://doi.org/10.1037/pas0000209

English, T., & Chen, S. (2011). Self-concept consistency and culture: The differential impact of two forms of consistency. *Personality and Social Psychology Bulletin*, *37*(6), 838–849. https://doi.org/10.1177/0146167211400621

Enzi, B., Greck, M. de, Prösch, U., Tempelmann, C., & Northoff, G. (2009). Is our self nothing but reward? Neuronal overlap and distinction between reward and personal relevance and its relation to human personality. *PLOS ONE*, *4*(12), e8429. https://doi.org/10.1371/journal.pone.0008429

Fitzsimons, G. J. (2008). Death to dichotomizing. *Journal of Consumer Research, 35*(1), 5–8. https://doi.org/10.1086/589561

Frankl, V. E. (1963). Man's search for meaning. Washington Square Press.

Goldman, B. M. (2002). The role of authenticity in healthy psychological functioning and subjective well-being. *Annals of the American Psychotherapy Association*, *5*(6), 18–20.

Greenberg, J., Koole, S. L., & Pyszczynski, T. A. (Eds.). (2004). *Handbook of experimental existential psychology.* Guilford Press.

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of Cognitive Neuroscience, 29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Guenther, C. L., & Sedikides, C. (in press). Authenticity as self-enhancement. In M. Vess, B. Schlegel, & J. Hicks (Eds.), *The Oxford handbook of authenticity*. Oxford University Press.

Guenther, C. L., Zhang, Y., & Sedikides, C. (2024). The authentic self is the self-enhancing self: A self-enhancement framework of authenticity. *Personality and Social Psychology Bulletin*, *50*(8), 1182–1196. https://doi.org/10.1177/01461672231160653

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7), e13570. https://doi.org/10.1111/psyp.13570

Harmon-Jones, E., Matis, S., Angus, D. J., & Harmon-Jones, C. (2024). Does effort increase or decrease reward valuation? Considerations from cognitive dissonance theory. *Psychophysiology*, *61*(6), e14536. https://doi.org/10.1111/psyp.14536

Harter, S. (2002). Authenticity. In *Handbook of positive psychology*. Oxford University Press.

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review, 10(*2), 88–110. https://doi.org/10.1207/s15327957pspr1002_1

Herbert, C., Pauli, P., & Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Social Cognitive and Affective Neuroscience*, *6*(5), 653–661. https://doi.org/10.1093/scan/nsq082

Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies*, *15*(6), 1425–1456. https://doi.org/10.1007/s10902-013-9485-0

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research, 40*(3), 366–371. https://doi.org/10.1509/jmkr.40.3.366.19237

Kelley, N. J., Finley, A. J., & Schmeichel, B. J. (2019). After-effects of self-control: The reward responsivity hypothesis. *Cognitive, Affective, & Behavioural Neuroscience*, *19*(3), 600–618. https://doi.org/10.3758/s13415-019-00694-3

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, *14*(1), 1–26. https://doi.org/10.1207/S15327965PLI1401_01

Kernis, M. H., & Goldman, B. M. (2005). From thought and experience to behaviour and interpersonal relationships: A multicomponent conceptualization of authenticity. In *Building, defending and regulating the self: A psychological perspective* (pp. 31–52). Psychology Press. https://doi.org/10.4324/9780203998052

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 283–357). Academic Press. https://doi.org/10.1016/S0065-2601(06)38006-9

Koole, S. L. (2003). In search of the real self: A functional perspective on optimal self-esteem and authenticity. *Psychological Inquiry*, *14*(1), 43–48.

Kotabe, H. P., & Hofmann, W. (2015). On integrating the components of self-control. *Perspectives on Psychological Science*, *10*(5), 618–638. https://doi.org/10.1177/1745691615593382

Lakey, C. E., Kernis, M. H., Heppner, W. L., & Lance, C. E. (2008). Individual differences in authenticity and mindfulness as predictors of verbal defensiveness. *Journal of Research in Personality*, *42*(1), 230–238. https://doi.org/10.1016/j.jrp.2007.05.002

Lenton, A. P., Bruder, M., Slabu, L., & Sedikides, C. (2013). How does “being real” feel? The experience of state authenticity: state authenticity. *Journal of Personality*, *81*(3), 276–289. https://doi.org/10.1111/j.1467-6494.2012.00805.x

Li, Q., Ren, X., Zhou, Z., & Wang, J. (2023). Reciprocal relationships between self-control and self-authenticity: A two-wave study. *Frontiers in Psychology*, *14*, 1207230. https://doi.org/10.3389/fpsyg.2023.1207230

Liang Y., & Xie T. (2021). True self in east and west from a Guanxi perspective. *Advances in Psychological Science*, *29*(5), 894–905. https://doi.org/10.3724/SP.J.1042.2021.00894

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed). MIT Press.

Luo, Y., Jiang, H., Chen, X., Zhang, Y., & You, X. (2019). Temporal dynamics of hedonic and eudaimonic reward processing: An event-related potentials (ERPs) study. *International Journal of Psychophysiology*, *137*, 63–71. https://doi.org/10.1016/j.ijpsycho.2018.12.009

Luo, Y., Zhang, X., Jiang, H., & Chen, X. (2022). The neural habituation to hedonic and eudaimonic rewards: Evidence from reward positivity. *Psychophysiology*, *59*, Airtle e13977. https://doi.org/10.1111/psyp.13977

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19–40. https://doi.org/10.1037//1082-989X.7.1.19

Mandel, N., & Heine, S. J. (1999). Terror management and marketing: He who dies with the most toys wins. *Advances in Consumer Research, 26*, 527–532.

Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, *5*(4), 420–430. https://doi.org/10.1177/1745691610375557

Markus, H. R., & Kitayama, S. (2014). Culture and the self: Implications for cognition, emotion, and motivation. In *College student development and academic life* (pp. 264–293). Routledge.

Martela, F., Ryan, R. M., & Steger, M. F. (2018). Meaningfulness as satisfaction of autonomy, competence, relatedness, and beneficence: Comparing the four satisfactions and positive affect as predictors of meaning in life. *Journal of Happiness Studies, 19*, 1261–1282. https://doi.org/10.1007/s10902-017-9869-7

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. https://doi.org/10.1037/h0054346

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Mueller, A. K., Falter, C. M., & Tucha, O. (2013). Distinct attentional mechanisms for social information processing as revealed by event-related potential studies. *International Journal of Brain and Cognitive Sciences*, *2*(2), 23–37. https://doi.org/10.5923/j.ijbcs.20130202.03

Neff, K. (2003). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, *2*(2), 85–101. https://doi.org/10.1080/15298860309032

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In *Handbook of emotions* (pp. 511–536). Guilford Press.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478. https://doi.org/10.1037/AXJ96-3445.130.3.466

Ong, A. D., Benson, L., Zautra, A. J., & Ram, N. (2018). Emodiversity and biomarkers of inflammation. *Emotion*, *18*(1), 3–14. https://doi.org/10.1037/emo0000343

Pyszczynski T., Solomon S., Greenberg J. (2015). Thirty years of terror management theory: From genesis to revelation. *Advances in Experimental Social Psychology, 52*, 1–70. https://doi.org/10.1016/bs.aesp.2015.03.001

Quoidbach, J., Gruber, J., Mikolajczak, M., Kogan, A., Kotsou, I., & Norton, M. I. (2014). Emodiversity and the emotional ecosystem. *Journal of Experimental Psychology: General*, *143*(6), 2057–2066. https://doi.org/10.1037/a0038025

Quoidbach, J., Mikolajczak, M., Gruber, J., Kotsou, I., Kogan, A., & Norton, M. I. (2018). *Robust, replicable, and theoretically-grounded: A response to Brown and Coyne's (2017) commentary on the relationship between emodiversity and health.* https://psycnet.apa.org/fulltext/2018-06787-002.html

Rathi, N., & Lee, K. (2021). Does it pay to be authentic? Implications of authenticity for life satisfaction and psychological well-being in a collectivist culture. *Journal of Happiness Studies*, *22*(1), 147–161. https://doi.org/10.1007/s10902-020-00223-x

Reed, D. E., Lehinger, E., Cobos, B., Vail III, K. E., Nabity, P. S., Helm, P. J., Galgali, M. S., & McGeary, D. D. (2021). Authenticity as a resilience factor against CV-19 threat among those with chronic pain and posttraumatic stress disorder. *Frontiers in Psychology*, *12*, 643869.

Rivera, G. N., Christy, A. G., Kim, J., Vess, M., Hicks, J. A., & Schlegel, R. J. (2019). Understanding the relationship between perceived authenticity and well-being. *Review of General Psychology*, *23*(1), 113–126. https://doi.org/10.1037/gpr0000161

Robinson, O. C., Lopez, F. G., Ramos, K., & Nartova-Bochaver, S. (2013). Authenticity, social context, and well-being in the United States, England, and Russia: A three country comparative analysis. *Journal of Cross-Cultural Psychology*, *44*(5), 719–737. https://doi.org/10.1177/0022022112465672

Rogers, C. R. (1961). *On becoming a person*. Houghton Mifflin Harcourt.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166. https://doi.org/10.1146/annurev.psych.52.1.141

Ryan, R. M., LaGuardia, J. G., & Rawsthorne, L. J. (2005). Self-complexity and the authenticity of self-aspects: Effects on well being and resilience to stressful events. *North American Journal of Psychology*, *7*(3).

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, *130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Sedikides, C. (2020). On the doggedness of self-enhancement and self-protection: How constraining are reality constraints? *Self and Identity, 19*(3), 251–271. https://doi.org/10.1080/15298868.2018.1562961

Sedikides, C. (2021). Self-construction, self-protection, and self-enhancement: A homeostatic model of identity protection. *Psychological Inquiry*, *32*(4), 197–221. https://doi.org/10.1080/1047840X.2021.2004812

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one’s faults. *European Review of Social Psychology*, *27*(1), 1–62. https://doi.org/10.1080/10463283.2016.1183913

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science*, *3*(2), 102–116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

Sedikides, C., Lenton, A. P., Slabu, L., & Thomaes, S. (2019). Sketching the contours of state authenticity. *Review of General Psychology*, *23*(1), 73–88. https://doi.org/10.1037/gpr0000156

Sedikides, C., & Schlegel, R. J. (2024). Distilling the concept of authenticity. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-024-00323-y

Sedikides, C., Slabu, L., Lenton, A., & Thomaes, S. (2017). State authenticity. *Current Directions in Psychological Science*, *26*(6), 521–525. https://doi.org/10.1177/0963721417713296

Sedikides, C., & Wildschut, T. (2016). Nostalgia: A bittersweet emotion that confers psychological health benefits. In A. M. Wood & J. Johnson (Eds.), *The Wiley Handbook of Positive Clinical Psychology* (pp. 125–136). Wiley.

Sedikides, C., & Wildschut, T. (2018). Finding meaning in nostalgia. *Review of General Psychology, 22*(1), 48–61. https://doi.org/10.1037/gpr0000109

Slabu, L., Lenton, A. P., Sedikides, C., & Bruder, M. (2014). Trait and state authenticity across cultures. *Journal of Cross-Cultural Psychology*, *45*(9), 1347–1373. https://doi.org/10.1177/0022022114543520

Smallenbroek, O., Zelenski, J. M., & Whelan, D. C. (2017). Authenticity as a eudaimonic construct: The relationships among authenticity, values, and valence. *The Journal of Positive Psychology*, *12*(2), 197–209. https://doi.org/10.1080/17439760.2016.1187198

Steger, M. F., Kashdan, T. B., Sullivan, B. A., & Lorentz, D. (2008). Understanding the search for meaning in life: Personality, cognitive style, and the dynamic between seeking and experiencing meaning. *Journal of Personality, 76*(2), 199–228. https://doi.org/10.1111/j.1467-6494.2007.00484.x

Steger, M. F., Oishi, S., & Kashdan, T. B. (2009). Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. *The Journal of Positive Psychology, 4*(1), 43–52. https://doi.org/10.1080/17439760802303127

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, *12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*(1), 67–85. https://doi.org/10.1037/0033-2909.110.1.67

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galván, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proceedings of the National Academy of Sciences*, *111*(18), 6600–6605. https://doi.org/10.1073/pnas.1323014111

Vazire, S., & Wilson, T. D. (2012). *Handbook of self-knowledge*. Guilford Press.

Vess, M., Rogers, R., Routledge, C., & Hicks, J. A. (2017). When being far away is good: Exploring how mortality salience, regulatory mode, and goal progress affect judgments of meaning in life. *European Journal of Social Psychology, 47*(1), 82–91. https://doi.org/10.1002/ejsp.2192

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, *64*(4), 678–691.

Wildschut, T., & Sedikides, C. (2020). The psychology of nostalgia: Delineating the emotion’s nature and functions. In *Nostalgia now* (pp. 47–65). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780429287602-2/psychology-nostalgia-tim-wildschut-constantine-sedikides

Wildschut, T., Sedikides, C., Arndt, J., & Routledge, C. (2006). Nostalgia: Content, triggers, functions. *Journal of Personality and Social Psychology*, *91*(5), 975–993. https://doi.org/10.1037/0022-3514.91.5.975

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the Authenticity Scale. *Journal of Counseling Psychology*, *55*(3), 385–399. https://doi.org/10.1037/0022-0167.55.3.385

Wong, P. T. (2013). *The human quest for meaning: Theories, research, and applications*. Routledge.

Zhang, D., He, W., Wang, T., Luo, W., Zhu, X., Gu, R., Li, H., & Luo, Y. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, *9*(12), 1897–1903. https://doi.org/10.1093/scan/nst188

Zhou, H., Guo, J., Ma, X., Zhang, M., Liu, L., Feng, L., Yang, J., Wang, Z., Wang, G., & Zhong, N. (2017). Self-reference emerges earlier than emotion during an implicit self-referential emotion processing task: Event-related potential evidence. *Frontiers in Human Neuroscience*, *11*, Article 451.

Appendix A Supplementary Material for Chapter 2

The appendix includes:

Supplementary Results for Experiment 1 and 2

Supplementary Study S1 to S3

References

Tables S2.1 to S2.12

Figure S2.1 to S2.3

Appendix A

Experiment 1

Results

Trait Endorsement

Self-Positivity for the Authentic Self and Presented Self. The Self × Valence × Endorsement interaction was significant, $F(1, 221) = 47.85$, $p < 0.001$, $\eta_p^2 = 0.18$. We examined the Valence × Endorsement interaction separately for the authentic self and the presented self, asking whether self-positivity was evident for both kinds of self. In the case of the authentic self, participants endorsed more positive traits (65.48 ± 9.74) than negative traits (18.09 ± 9.63) as descriptive, $t(221) = 44.70$, $p < 0.001$, 95% *CI* = [45.30, 49.48], *Cohen's d* = 3.00, but judged more negative traits (63.78 ± 10.83) than positive traits (16.16 ± 8.04) as non-descriptive, $t(221) = 45.59$, $p < 0.001$, 95% *CI* = [45.56, 49.68], *Cohen's d* = 3.06. Similarly, in the case of the presented self, participants endorsed more positive traits (66.71 ± 8.76) than negative traits (12.58 ± 8.06) as descriptive, $t(221) = 57.95$, $p < 0.001$, 95% *CI* = [52.29, 55.97], *Cohen's d* = 3.89, but judged more negative traits (69.79 ± 8.86) than positive traits (15.40 ± 7.81) as nondescriptive, $t(221) = 58.23$, $p < 0.001$, 95% *CI* = [52.55, 56.23], *Cohen's d* = 3.91. Self-positivity was evident for both the authentic self and the presented self.

Reaction Time

Model Settings. We aimed to test the difference in self-positivity, as reflected in RT, between the authentic self and the presented self; thus, the dependent variable was RT for each valid trial. Three variables and their interactions (i.e., self, valence, endorsement, Self × Valence, Self × Endorsement, Endorsement × Valence, Self × Valence × Endorsement) were level-1 predictors (i.e., fixed effects); for self: authentic self = 1, presented self = -1; for valence: positive traits = 1, negative traits = -1; for endorsement: yes (self-descriptive) = 1, no (non-self-descriptive) = -1. Following Volpert-Esmond et al.'s (2018, 2021) model specification procedures, we also calculated each participant's mean RT based on the average of all relevant trials for the level-2 factor and treated it as a grouping variable ("random factor"); given individual differences in RTs, we estimated random intercepts of mean RT for each participant; also, given that the effect of mean RT on the outcome varied across participants, we estimated a random slope for each participant's mean RT.

Self-Positivity for the Authentic and Presented Self. The Self × Valence × Endorsement interaction was significant, $\beta = 8.76$, $t_{72627} = 2.48$, $p = 0.013$. We examined the Valence × Endorsement interaction separately for the authentic self and the presented self, asking whether self-positivity was evident for both kinds of self. In the case of the authentic self, participants responded faster to endorse positive traits (1283.66 ± 317.32) than negative traits

(1518.55 ± 404.04) as descriptive, $\gamma = -105.20$, $z = -15.73$, $p < 0.001$, but responded faster to judge negative traits (1349.25 ± 290.88) than positive traits (1639.01 ± 454.81) as non-descriptive, $\gamma = -129.27$, $z = -18.48$, $p < 0.001$. Likewise, in the case of the presented self, participants responded faster to endorse positive traits (1261.61 ± 280.48) than negative traits (1592.39 ± 446.82) as descriptive, $\gamma = -152.52$, $z = -19.98$, $p < 0.001$, but responded faster to judge negative traits (1332.29 ± 282.78) than positive traits (1597.69 ± 472.48) as non-descriptive, $\gamma = -116.88$, $z = -16.77$, $p < 0.001$. We observed self-positivity both for the authentic self and the presented self.

Experiment 2

Results

Trait Endorsement

Self-Positivity for the Authentic and Presented Self. The Self × Valence × Endorsement interaction was significant, $F(1, 120) = 57.50$, $p < 0.001$, $\eta_p^2 = 0.32$. We examined the Valence × Endorsement interaction separately for the authentic self and the presented self, asking whether self-positivity was evident for both kinds of self. In the case of the authentic self, participants endorsed more positive traits (141.83 ± 25.01) than negative traits (41.35 ± 23.02) as descriptive, $t(120) = 25.72$, $p < 0.001$, 95% *CI* = [92.75, 108.22], *Cohen's d* = 2.34, but judged more negative traits (154.81 ± 24.90) than positive traits (53.14 ± 24.72) as nondescriptive, $t(120) = 25.58$, $p < 0.001$, 95% *CI* = [93.80, 109.54], *Cohen's d* = 2.33. Similarly, in the case of the presented self, participants endorsed more positive traits (150.61 ± 24.03) than negative traits (27.88 ± 17.66) as descriptive, $t(120) = 37.63$, $p < 0.001$, 95% *CI* = [116.27, 129.19], *Cohen's d* = 3.42, but judged more negative traits (167.08 ± 19.27) than positive traits (43.83 ± 22.31) as nondescriptive, $t(120) = 39.26$, $p < 0.001$, 95% *CI* = [117.04, 129.47], *Cohen's d* = 3.57. Self-positivity manifested itself both for the authentic self and presented self.

Reaction Time

Model Settings. The model settings are the same as Experiment 1.

Self-Positivity for the Authentic and Presented Self. The Self × Valence × Endorsement interaction was significant, $\beta = 21.75$, $t_{94329} = 7.74$, $p < 0.001$. We examined the Valence × Endorsement interaction separately for the authentic self and the presented self, asking whether self-positivity was evident for both kinds of self. With regard to the authentic self, participants responded faster to endorse positive traits (1326.18 ± 337.71) than negative traits (1624.60 ± 423.74) as descriptive, $\gamma = -116.86$, $z = -20.98$, $p < 0.001$, but responded faster to judge negative traits (1297.50 ± 327.94) than positive traits (1683.54 ± 413.92) as non-descriptive, $\gamma = -170.33$, $z = -34.08$, $p < 0.001$. In a similar vein, with regard to the presented self, participants responded faster to endorse positive traits (1304.22 ± 353.80) than negative traits (1726.73 ± 456.19) as descriptive, $\gamma = -165.70$, $z = -26.07$, $p < 0.001$, but responded faster to judge negative traits (1287.29 ± 364.70) than positive traits (1734.97 ± 445.94) as non-descriptive, $\gamma = -192.90$, $z = -36.93$, $p < 0.001$. Self-positivity characterized both the authentic self and the presented self.

P1

Model Settings. We aimed to test the difference in self-positivity, as reflected in P1 amplitudes, between the authentic self and the presented self. We examined in particular how P1 amplitudes varied as function of each variable (participants, electrodes, self, valence, endorsement). The last three variables and their interactions (Self × Valence, Self × Endorsement, Endorsement × Valence, Self × Valence × Endorsement) were level-1 predictors (i.e., fixed effects)—self: authentic self = 1, presented self = -1; valence: positive = 1, negative = -1; endorsement: endorse = 1, reject = -1. We also calculated each participant's mean P1 amplitude, indicated by the average of all trials for that participant, as a level-2 factor. Finally, participants and electrodes were grouping variables or random factors. Following Volpert-Esmond et al.'s (2018, 2021) model specification procedures, we specified participants and electrodes as independent factors (i.e., a cross-classified model), allowed the intercept and slope of mean amplitude to vary by participants (i.e., random intercept, random slope), and allowed the intercept to vary by electrode (i.e., random intercept).

P1 in Judging the Non-Self-Descriptiveness of Positive and Negative Traits. The model revealed a significant Self × Valence × Endorsement interaction, $\beta = 0.06$, $t_{7616} = 3.93$, $p < 0.001$ (Figure 2.3a). We proceeded with simple slope tests to test whether the simple slopes of P1 on self and valence were significant in the endorsement (YES) or rejection (NO) conditions, respectively. We found a significant simple slope only in regard to endorsement of negative traits: the elicited P1 was larger when endorsing negative traits as descriptive of the presented self than the authentic self ($\gamma = 0.26$, $z = 9.20$, $p < 0.001$). The elicited P1 was equivalent when judging negative traits as non-descriptive of the presented self and the authentic self ($\gamma = -0.03$, $z = -0.93$, $p = 0.350$), was equivalent when endorsing positive traits as descriptive of the presented self and the authentic self ($\gamma = -0.004$, $z = -0.14$, $p = 0.882$), and was equivalent when judging positive traits as non-descriptive of the presented self and the authentic self ($\gamma = 0.01$, $z = 0.27$, $p = 0.791$). We display in Table S2.7 detailed results for fixed effects. The results were comparable for *N*s of 107 and 61 (Table S2.7 and Table S2.8).

N170

Model Settings. The model settings are the same as P1.

N170 in Judging the Non-Self-Descriptiveness of Positive and Negative Traits. The model revealed a significant Self × Valence × Endorsement interaction, $\beta = 0.03$, $t_{13544} = 2.94$, $p = 0.003$ (Figure 2.3b). We proceeded with simple slope tests to test whether the simple slopes of N170 on self and valence were significant in the endorsement (YES) or rejection (NO) conditions, respectively. We found a significant simple slope only in regard to endorsement of positive traits: the elicited N170 was larger when endorsing positive traits as descriptive of the presented self than the authentic self ($\gamma = 0.06$, $z = 3.35$, $p < 0.001$), whereas the elicited N170 tended to be larger when endorsing negative traits as descriptive of the authentic self than the presented self ($\gamma = 0.03$, $z = 1.72$, $p = 0.085$). However, the elicited N170 was equivalent when judging positive traits as non-descriptive of the presented self and the authentic self ($\gamma = 0.03$, $z = 1.59$, $p = 0.113$), and the elicited N170 was equivalent when judging negative traits as non-descriptive of the presented self and the authentic self ($\gamma = 0.01$, $z = 0.78$, $p = 0.435$). We display in Table S2.9 detailed results for the fixed effects. The results were comparable for *N*s of 107 and 61 (Table S2.9 and Table S2.10).

LPP

Model Settings. The model settings are the same as P1.

LPP in Judging the Non-Self-Descriptiveness of Positive and Negative Traits. The model revealed a significant Self × Valence × Endorsement interaction, $\beta = 0.04$, $t_{12698} = 2.54$, $p = 0.011$ (Figure 2.3c). We proceeded with simple slope tests to test whether the simple slopes of LPP on self and valence were significant in the endorsement (YES) or rejection (NO) conditions, respectively. We found significant simple slopes in the endorsement condition: the elicited LPP was larger when judging positive traits as descriptive of the authentic self than the presented self ($\gamma = 0.14$, $z = 4.72$, $p < 0.001$), and the elicited LPP was larger when judging negative traits as descriptive of the authentic self than the presented self ($\gamma = 0.19$, $z = 6.35$, $p < 0.001$). These results are consistent with a stimulus significance view of the LPP (Hajcak & Foti, 2020). In the rejection condition, the elicited LPP was larger when judging negative traits as non-descriptive of the authentic self than the presented self ($\gamma = 0.16$, $z = 5.21$, $p < 0.001$), but the elicited LPP was equal when judging positive traits as non-descriptive of the presented self and the authentic self ($\gamma = -0.05$, $z = -1.49$, $p = 0.136$). We display in Table S2.11 the detailed results of fixed effects. The results were comparable for *N*s of 107 and 61 (Table S2.11 and Table S2.12).

Appendix A

Study S1: Is the Presented Self More Sanitized Than the Authentic Self?

In Study S1, we examined how sanitized participants consider their presented and authentic selves. We hypothesized that they would consider their presented self as more sanitized than their authentic self. We operationalized “sanitized” in terms of the trait adjectives polished, refined, sleek, glossy, and smooth.

To test our hypothesis, we recruited 52 [NATIONALITY MASKED] Prolific workers ranging in age from 18 to 69 years (*M* = 39.88, *SD* = 14.80). Thirty-four of them (65.38%) identified as female, 17 (32.69%) as male, and 1 (1.92%) as non-binary/third gender. Participants’ ethnic background was as follows: White (*n* = 48, 92.31%), Black (*n* = 1, 1.92%), Asian (*n* = 2, 3.85%), Mixed (*n* = 1, 1.92%).

Participants learned that they would make judgments about two different versions of the self. One version, the presented self, was defined as “the self you present to others,” whereas the other version, the authentic self, was defined as your “true, real, genuine self.” Next, participants rated how polished, refined, sleek, glossy, and smooth they regarded their presented self ($\alpha = .95$) and their authentic self ($\alpha = .94$) and on a scale ranging from 1 (*not at all*) to 7 (*very much*). In line with our hypothesis, participants considered their presented self (*M* = 3.43, *SD* = 1.45) as more sanitized than their authentic self (*M* = 2.86, *SD* = 1.29), *t*(51) = 3.13, *p* = .003, Cohen’s *d* = -0.44, 95%CI [-0.72, -0.15]. Sensitivity analysis in G*Power indicated that 52 participants gave us 80% power to detect effects as small as *d* = ± 0.40.

Study S2: Is the Authentic Self More Robust Than the Presented Self?

In Study S2, we examined how robust participants regarded their authentic and presented selves. We hypothesized that they would regard their authentic self as more robust than their presented self. We operationalized "robustness" in terms of the trait adjectives variable, changeable, malleable, fluctuating, and shifting.

To test this hypothesis, we recruited 59 [NATIONALITY MASKED] Prolific workers ranging in age between 20 and 77 years (*M* = 41.86, *SD* = 14.50). Forty of them (67.80%) identified as female, 16 (27.12%) as male 2 (3.39%) as non-binary/third gender, and 1 (1.69%) preferred not to disclose their gender. Participants' ethnic background was as follows: White (*n* = 47, 79.66%), Black (*n* = 1, 1.69%), Asian (*n* = 9, 15.25%), Mixed (*n* = 1, 1.69%), Prefer not to say (*n* = 1, 1.69%).

Participants were informed that that they would make judgements about two different versions of the self. One version, the presented self, was defined as "the self you present to others." The other version, the authentic self, was defined as your "true, real, genuine self." Subsequently, participants rated how variable, changeable, malleable, fluctuating, and shifting they regarded the authentic self ($\alpha = .91$) and presented self ($\alpha = .95$) on a scale ranging from 1 (*not at all*) to 7 (*very much*). Higher scores on this composite reflect less robustness, whereas lower scores reflect more robustness. In accord with our hypothesis, participants rated their authentic self (*M* = 3.30, *SD* = 1.24) as more robust than their presented self (*M* = 4.58, *SD* = 1.41), $t(58) = -5.81$, $p < .001$, Cohen's $d = -0.76$, 95%CI [-1.04, -0.46]. Sensitivity analysis in G*Power indicated that 59 participants gave us 80% power to detect effects as small as $d = \pm 0.37$.

Appendix A

Study S3: Is the Authentic Self More Significant Than the Presented Self?

In Study S3, we examined how significant participants deemed their authentic and presented selves. We hypothesized that they would deem their authentic self as more significant than their presented self. We operationalized significance in terms of the trait adjectives important, significant, and valuable.

To test our hypothesis. We recruited 50 [NATIONALITY MASKED] Prolific workers aged between 18 and 26 years ($M = 21.22$, $SD = 1.87$). Twenty-five of them (50%) identified as male, 23 (46.00%) as female, and 2 (4.00%) as non-binary/third gender. Participants' ethnic background was as follows: White ($n = 32$, 64.00%), Black ($n = 4$, 8.00%), Asian ($n = 9$, 18.00%), Mixed ($n = 2$, 4.00%), Other ($n = 2$, 4.00%), Prefer not to say ($n = 1$, 2.00%).

Participants were instructed that they would make judgements about two different versions of the self. One version, the authentic self, was defined as your "true, real, genuine self." The other version, the presented self, was defined as "the self you present to others." Participants then rated how important, significant, and valuable they found the authentic self ($\alpha = .85$) and presented self ($\alpha = .95$) on a scale ranging from 1 (*not at all*) to 7 (*very much*). Congruent with our hypothesis, participants deemed their authentic self ($M = 5.77$, $SD = 1.04$) as more significant than their presented self ($M = 4.67$, $SD = 1.59$), $t(49) = 4.57$, $p < .001$, Cohen's $d = 0.65$, 95%CI [0.34, 0.95]. Sensitivity analysis in G*Power indicated that 50 participants gave us 80% power to detect effects as small as $d = \pm 0.40$.

Appendix A

References

Volpert-Esmond, H. I., Merkle, E. C., Levsen, M. P., Ito, T. A., & Bartholow, B. D. (2018). Using trial-level data and multilevel modeling to investigate within-task change in event-related potentials. *Psychophysiology, 55*(5), Article e13044. https://doi.org/10.1111/psyp.13044

Volpert-Esmond, H. I., Page-Gould, E., & Bartholow, B. D. (2021). Using multilevel models for the analysis of event-related potentials. *International Journal of Psychophysiology, 162*, 145–156. https://doi.org/10.1016/j.ijpsycho.2021.02.006

Appendix A

Supplementary Tables

Table S2. 1

Coefficient Estimation of Self, Valence, Endorsement, and Their Interactions on Reaction Time in Experiment 1

Effect	Estimate	*SE*	*t*	*p*
Intercept	1427.15	18.59	76.78	**< .001**
Self ^{a}	0.09	3.52	0.03	.979
Valence ^{b}	-3.29	3.52	-0.94	.350
Endorse ^{c}	-31.75	3.54	-8.97	**< .001**
Self ^{a} × Valence ^{b}	15.21	3.52	4.33	**< .001**
Self ^{a} ×Endorsement ^{c}	-14.18	3.52	-4.03	**< .001**
Valence × Endorsement ^{c}	-126.53	3.58	-35.40	**< .001**
Self ^{a} × Valence ^{b} × Endorsement ^{c}	8.76	3.53	2.48	**.013**

Note. We display unstandardized betas. We used Satterthwaite approximations to estimate degrees of freedom for calculating *p*-values.

^{a} authentic self = 1, presented self = −1. ^{b} positive = 1, negative = −1. ^{c} yes = 1, no = −1.

Table S2. 2

Number of Retained EEG Trials in Each Condition in Experiment 2

Condition	*Mean*	*Max*	*Min*
Authentic self, positive traits / Yes	133.59	194	52
Authentic self, positive traits / No	50.82	133	6
Authentic self, negative traits / Yes	39.28	108	5
Authentic self, negative traits / No	146.93	190	54
Presented self, positive traits / Yes	141.72	191	71
Presented self, positive traits / No	41.74	97	5
Presented self, negative traits / Yes	26.83	75	5
Presented self, negative traits / No	158.63	193	103

Note. Yes = self-descriptive; No = non-self-descriptive

Table S2. 3

Effects of Self, Valence, Endorsement, and Their Interactions on Trait Endorsement for Ns of 121, 107, and 61 in Experiment 2

	N = 121			*N* = 107			*N* = 61		
Variable	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Self	1.95	.165	0.02	1.45	.231	0.01	1.41	.241	0.02
Valence	3.09	.081	0.03	2.69	.104	0.03	2.51	.118	0.04
Endorsement	66.54	**< .001**	0.36	51.42	**< .001**	0.33	28.18	**< .001**	0.32
Self × Valence	0.74	.391	0.01	0.22	.644	0.002	0.07	.790	0.01
Self ×Endorsement	5.24	**.024**	0.04	6.67	**.011**	0.06	3.15	.081	0.05
Valence × Endorsement	1172.22	**< .001**	0.91	1044.20	**< .001**	0.91	488.14	**< .001**	0.89
Self × Valence × Endorsement	57.50	**< .001**	0.32	49.08	**< .001**	0.32	24.07	**< .001**	0.29

Table S2. 4

Pairwise Comparisons of the Authentic Self and Presented Self in the Context of the Three-Way Interaction for Ns of 121, 107, and 61 in Experiment 2

Variable	*Mean Difference (SE)*	*p*	*95% CI*
	121 participants		
1	-8.78 (1.75)	< **.001**	[-12.24, -5.31]
2	9.31 (1.75)	< **.001**	[5.84, 12.24]
3	13.46 (1.57)	< **.001**	[10.36, 16.57]
4	-12.27 (1.75)	< **.001**	[-15.74, -8.81]
	107 participants		
1	-8.61 (1.91)	< **.001**	[-12.40, -4.81]
2	9.23 (1.92)	< **.001**	[5.43, 13.04]
3	14.02 (1.73)	< **.001**	[10.58, 17.46]
4	-13.01 (1.92)	< **.001**	[-16.82, -9.20]
	61 participants		
1	-8.71 (2.48)	< **.001**	[-13.67, -3.74]
2	9.51 (2.45)	< **.001**	[4.60, 14.42]
3	13.48 (2.49)	< **.001**	[8.50, 18.45]
4	-12.49 (2.71)	< **.001**	[-17.92, -7.06]

1 = authentic-self, positive descriptive endorsements minus presented-self, positive descriptive endorsements
2 = authentic-self, negative descriptive endorsements minus presented-self, negative descriptive endorsement
3 = authentic-self, positive non-descriptive endorsements minus presented-self, positive non-descriptive endorsements
4 = authentic-self, negative non-descriptive endorsements minus presented-self, negative non-descriptive endorsement

Table S2. 5

Coefficient Estimation of Self, Valence, Endorsement, and Their Interactions on Reaction Time for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Effect	*Coefficient (SE)*	*t*	*p*	*Coefficient (SE)*	*t*	*p*	*Coefficient (SE)*	*t*	*p*
Intercept	1460.55 (26.99)	54.12	**< .001**	1458.11 (27.65)	52.74	**< .001**	1460.92 (39.97)	60.19	**< .001**
Self [a]	-13.58 (2.80)	-4.85	**< .001**	-14.71 (2.91)	-5.05	**< .001**	-6.35 (3.52)	-1.81	.071
Valence [b]	20.88 (2.81)	7.45	**< .001**	17.06 (2.91)	5.85	**< .001**	8.62 (3.52)	2.45	**.014**
Endorsement [c]	-10.63 (2.82)	-3.77	**< .001**	-12.35 (2.93)	-4.22	**< .001**	-23.80 (3.53)	-6.97	**< .001**
Self [a] × Valence [b]	4.29 (2.80)	1.53	.125	3.26 (2.91)	1.12	.263	3.73 (3.51)	1.06	.289
Self [a] ×Endorsement [c]	2.19 (2.80)	0.78	.434	0.40 (2.91)	0.14	.891	0.41 (3.52)	0.12	.907
Valence × Endorsement [c]	-166.94 (2.87)	-58.22	**< .001**	-156.92 (2.97)	-52.87	**< .001**	-126.63 (3.57)	-35.50	**< .001**
Self [a] × Valence [b] × Endorsement [c]	21.75 (2.81)	7.74	**< .001**	21.17 (2.92)	7.25	**< .001**	16.93 (3.53)	4.80	**< .001**

Note. We display unstandardized betas. We used Satterthwaite approximations to estimate degrees of freedom for calculating *p*-values.

[a] authentic self = 1, presented self = −1. [b] positive = 1, negative = −1. [c] yes = 1, no = −1.

Table S2. 6

Simple Slopes (γ) of the Authentic self and The Presented Self Regarding the Three-Way Interaction on Reaction Time for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Variable	*Mean (SD)*	*z*	*p*	*Mean (SD)*	*z*	*p*	*Mean (SD)*	*z*	*p*
1	14.65 (3.66)	4.00	**< .001**	10.12 (3.95)	2.56	**.010**	14.72 (5.22)	2.82	**.005**
2	-37.43 (7.69)	-4.87	**< .001**	-38.73 (2.91)	-4.93	**< .001**	-26.60 (9.01)	-2.80	**.003**
3	-33.23 (6.40)	-5.19	**< .001**	-33.02 (6.66)	-4.96	**< .001**	-19.96 (8.03)	-2.49	**.013**
4	1.69 (3.49)	0.48	.629	2.81 (3.77)	0.75	.456	6.45 (5.01)	1.29	.198

Note. We display unstandardized betas. We effect coded Self (authentic self = 1, presented self = −1).

1 = authentic self, positive traits, self-descriptiveness versus presented self, positive traits, self-descriptiveness

2 = authentic self, negative traits, self-descriptiveness versus presented self, negative traits, self-descriptiveness

3 = authentic self, positive traits, non-descriptiveness versus presented self, positive traits, non-self-descriptiveness

4 = authentic self, negative traits, non-descriptiveness versus presented self, negative traits, non-self-descriptiveness

Table S2. 7

Coefficient Estimation of Self, Valence, Endorsement, and Their Interactions on P1 for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Effect	Estimate *(SE)*	*t*	*p*	Estimate *(SE)*	*t*	*p*	Estimate *(SE)*	*t*	*p*
Intercept	1.43 (0.13)	11.42	**< .001**	1.20 (0.14)	8.77	**< .001**	0.19 (0.18)	1.02	.307
Self [a]	-0.08 (0.01)	-5.01	**< .001**	-0.07 (0.01)	-4.82	**< .001**	-0.15 (0.02)	-9.30	**< .001**
Valence [b]	0.11 (0.01)	7.31	**< .001**	0.14 (0.01)	9.64	**< .001**	0.07 (0.02)	4.22	**< .001**
Endorsement [c]	0.02 (0.01)	1.72	.086	-0.01 (0.01)	-0.68	.497	0.03 (0.02)	1.68	.094
Self [a] × Valence [b]	0.07 (0.01)	5.13	**< .001**	0.05 (0.01)	3.49	**< .001**	0.06 (0.02)	3.86	**< .001**
Self [a] ×Endorsement [c]	0.06 (0.01)	-4.34	**< .001**	0.04 (0.01)	2.53	**.011**	-0.05 (0.02)	-3.24	**.001**
Valence × Endorsement [c]	-0.04 (0.01)	3.12	**.002**	-0.02 (0.01)	-1.74	.081	-0.11 (0.02)	-6.61	**< .001**
Self [a] × Valence [b] × Endorsement [c]	0.06 (0.01)	3.93	**< .001**	0.03 (0.01)	2.17	**.030**	0.05 (0.02)	2.96	**.003**

Note. We display unstandardized betas. We used Satterthwaite approximations to estimate degrees of freedom for calculating *p*-values.

[a] authentic self = 1, presented self = −1. [b] positive = 1, negative = −1. [c] yes = 1, no = −1.

Table S2. 8

Simple Slopes (γ) of the Authentic Self and The Presented Self Regarding the Three-Way Interaction on P1 for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Variable	*γ*	*z*	*p*	*γ*	*z*	*p*	*γ*	*Z*	*p*
1	-0.004 (0.03)	-0.14	.882	-0.02 (0.03)	-0.84	.399	-0.09 (0.03)	-2.86	**.004**
2	-0.26 (0.03)	-9.20	**< .001**	-0.19 (0.03)	-6.50	**< .001**	-0.31 (0.03)	-9.68	**< .001**
3	0.01 (0.03)	0.27	.791	-0.01 (0.03)	-0.49	.625	-0.08 (0.03)	-2.58	**.010**
4	-0.03 (0.03)	-0.93	.350	-0.05 (0.03)	-1.80	.072	-0.11 (0.03)	-3.48	**< .001**

Note. We used unstandardized betas. We effect coded Self (authentic self = 1, presented self = −1).

1 = authentic self, positive traits, self-descriptiveness versus presented self, positive traits, self-descriptiveness

2 = authentic self, negative traits, self-descriptiveness versus presented self, negative traits, self-descriptiveness

3 = authentic self, positive traits, non-descriptiveness versus presented self, positive traits, non-self-descriptiveness

4 = authentic self, negative traits, non-descriptiveness versus presented self, negative traits, non-self-descriptiveness

Table S2. 9

Coefficient Estimation of Self, Valence, Endorsement, and Their Interactions on N170 for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Effect	Estimate *(SE)*	*t*	*p*	Estimate *(SE)*	*t*	*p*	Estimate *(SE)*	*t*	*p*
Intercept	2.36 (0.15)	15.55	**< .001**	2.31 (0.16)	14.40	**< .001**	2.48 (0.14)	17.83	**< .001**
Self [a]	-0.01 (0.01)	-0.37	.714	-0.01 (0.01)	-1.27	.204	-0.02 (0.01)	-2.18	**.030**
Valence [b]	0.10 (0.01)	11.11	**< .001**	0.11 (0.01)	12.15	**< .001**	0.05 (0.01)	4.55	**< .001**
Endorsement [c]	-0.05 (0.01)	-4.90	**< .001**	-0.05 (0.01)	-5.59	**< .001**	0.01 (0.01)	0.22	.822
Self [a] × Valence [b]	0.02 (0.01)	2.14	**.033**	0.02 (0.01)	1.80	.072	0.02 (0.01)	2.05	**.041**
Self [a] ×Endorsement [c]	0.02 (0.01)	2.00	**.045**	0.02 (0.01)	1.95	**.050**	-0.01 (0.01)	-0.20	.844
Valence × Endorsement [c]	-0.01 (0.01)	-0.40	.693	-0.01 (0.01)	-1.49	.136	-0.07 (0.01)	-6.20	**< .001**
Self [a] × Valence [b] × Endorsement [c]	0.03 (0.01)	2.94	**.003**	0.03 (0.01)	3.36	**< .001**	0.03 (0.01)	2.56	**.011**

Note. We presented unstandardized betas. We used Satterthwaite approximations to estimate degrees of freedom for calculating *p*-values.

[a] authentic self = 1, presented self = −1. [b] positive = 1, negative = −1. [c] yes = 1, no = −1.

Table S2. 10

Simple Slopes (γ) of the Authentic self and The Presented Self Regarding the Three-Way Interaction on N170 for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Variable	*γ*	*z*	*p*	*γ*	*z*	*p*	*γ*	*z*	*p*
1	0.06 (0.02)	3.35	**< .001**	0.06 (0.02)	2.39	**.017**	0.03 (0.02)	1.12	.264
2	-0.03 (0.02)	-1.72	.085	-0.04 (0.01)	-3.17	**.002**	-0.08 (0.02)	-3.49	**< .001**
3	-0.06 (0.02)	-1.59	.113	-0.05 (0.01)	-3.38	**< .001**	-0.03 (0.02)	-1.25	.213
4	-0.01 (0.02)	-0.78	.435	-0.02 (0.02)	-0.68	.497	-0.02 (0.02)	-0.73	.463

Note. We presented unstandardized betas. We effect coded Self (authentic self = 1, presented self = −1).

1 = authentic self, positive traits, self-descriptiveness versus presented self, positive traits, self-descriptiveness

2 = authentic self, negative traits, self-descriptiveness versus presented self, negative traits, self-descriptiveness

3 = authentic self, positive traits, non-descriptiveness versus presented self, positive traits, non-self-descriptiveness

4 = authentic self, negative traits, non-descriptiveness versus presented self, negative traits, non-self-descriptiveness

Table S2. 11

Coefficient Estimation of Self, Valence, Endorsement, and Their Interactions on LPP for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Effect	Estimate *(SE)*	*t*	*p*	Estimate *(SE)*	*t*	*p*	Estimate *(SE)*	*t*	*p*
Intercept	3.05 (0.12)	25.59	**< .001**	2.49 (0.13)	19.23	**< .001**	2.36 (0.13)	18.74	**< .001**
Self [a]	0.11 (0.02)	7.39	**< .001**	0.08 (0.02)	5.58	**< .001**	0.07 (0.02)	4.54	**< .001**
Valence [b]	0.14 (0.02)	9.56	**< .001**	0.19 (0.02)	12.15	**< .001**	0.13 (0.02)	8.79	**< .001**
Endorsement [c]	0.26 (0.02)	17.15	**< .001**	0.25 (0.02)	18.73	**< .001**	0.33 (0.02)	20.35	**< .001**
Self [a] × Valence [b]	-0.06 (0.02)	-4.16	**< .001**	-0.04 (0.02)	-1.05	**0.004**	-0.06 (0.02)	-0.76	**< .001**
Self [a] ×Endorsement [c]	0.06 (0.02)	3.68	**< .001**	0.05 (0.02)	2.38	**< .001**	0.02 (0.02)	1.58	.194
Valence × Endorsement [c]	-0.31 (0.02)	-20.57	**< .001**	-0.25 (0.02)	-16.38	**< .001**	-0.21 (0.02)	-11.10	**< .001**
Self [a] × Valence [b] × Endorsement [c]	0.04 (0.02)	2.54	**.011**	0.04 (0.02)	2.80	**0.016**	0.08 (0.02)	5.67	**< .001**

Note. We presented unstandardized betas are presented. We used Satterthwaite approximations to estimate degrees of freedom for calculating *p*-values.

[a] authentic self = 1, presented self = −1. [b] positive = 1, negative = −1. [c] yes = 1, no = −1.

Table S2. 12

Simple Slopes (γ) of the Authentic self and The Presented Self Regarding the Three-Way Interaction on LPP for Ns of 121, 107, and 61 in Experiment 2

	121 participants			107 participants			61 participants		
Variable	*γ*	*Z*	*p*	*γ*	*z*	*p*	*γ*	*z*	*p*
1	0.14 (0.03)	4.72	**< .001**	0.13 (0.03)	4.19	**< .001**	0.11 (0.04)	3.06	**.002**
2	0.19 (0.03)	6.35	**< .001**	0.14 (0.03)	4.65	**< .001**	0.07 (0.04)	1.86	.063
3	-0.05 (0.03)	-1.49	.136	-0.05 (0.03)	-1.62	.105	-0.10(0.04)	-2.74	**.006**
4	0.16 (0.03)	5.21	**< .001**	0.11(0.03)	3.69	**< .001**	0.18 (0.04)	5.06	**< .001**

Note. We presented unstandardized betas. We effect coded Self (authentic self = 1, presented self = −1).

1 = authentic self, positive traits, self-descriptiveness versus presented self, positive traits, self-descriptiveness

2 = authentic self, negative traits, self-descriptiveness versus presented self, negative traits, self-descriptiveness

3 = authentic self, positive traits, non-descriptiveness versus presented self, positive traits, non-self-descriptiveness

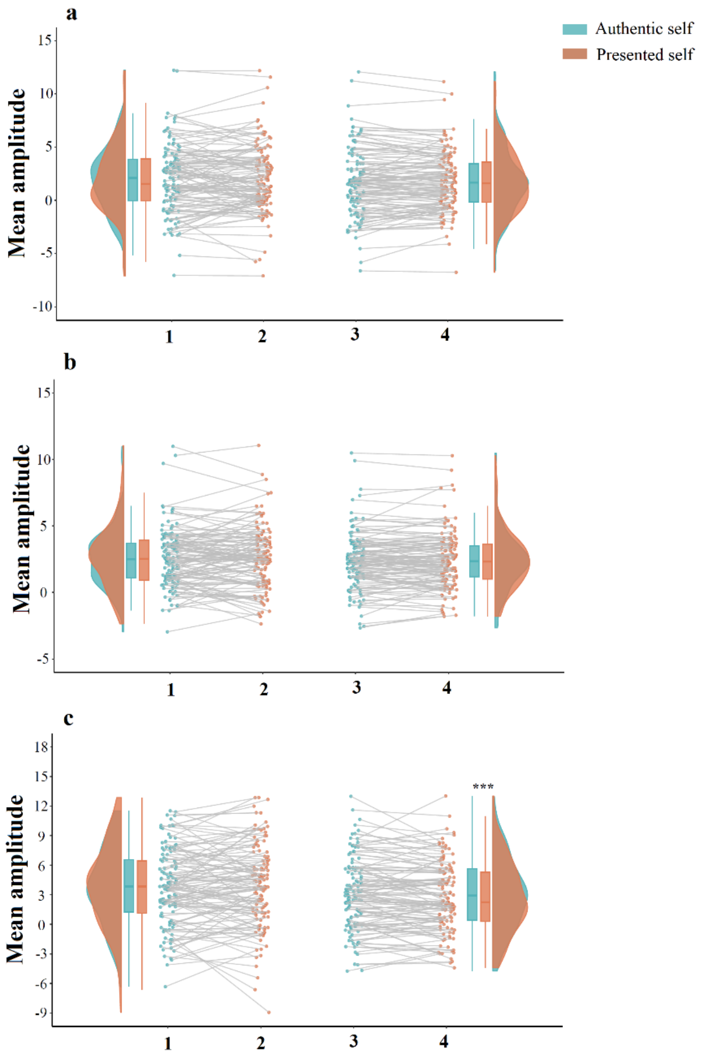
4 = authentic self, negative traits, non-descriptiveness versus presented self, negative traits, non-self-descriptiveness

Appendix A

Supplementary Figures

Figure S2. 1

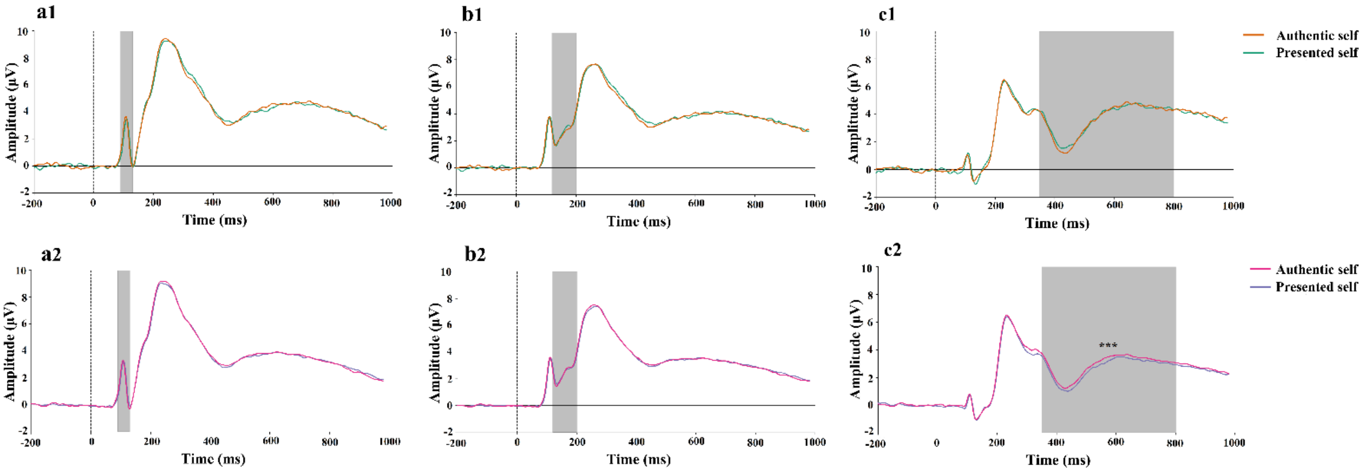
Neural Manifestations of Self-positivity for the Authentic Self and Presented Self in Experiment 2



Note. (**a**) P1 mean amplitude for the authentic self and the presented self in judging the non-self-descriptiveness of positive and negative traits. (**b**) N170 mean amplitude for the authentic self and the presented self in judging the non-self-descriptiveness of positive and negative traits. (**c**) LPP mean amplitude for the authentic self and the presented self in judging the non-self-descriptiveness of positive and negative traits (see Table S2.7 to S2.12 for fixed effects of self, valence, endorse, and their interactions on P1, N170, and LPP). 1 = authentic self, positive traits, non-self-descriptiveness; 2 = presented self, positive traits, non-self-descriptiveness; 3 = authentic self, negative traits, non-self-descriptiveness; 4 = presented self, negative traits, non-self-descriptiveness. $^{***}p < .001$.

Figure S2. 2

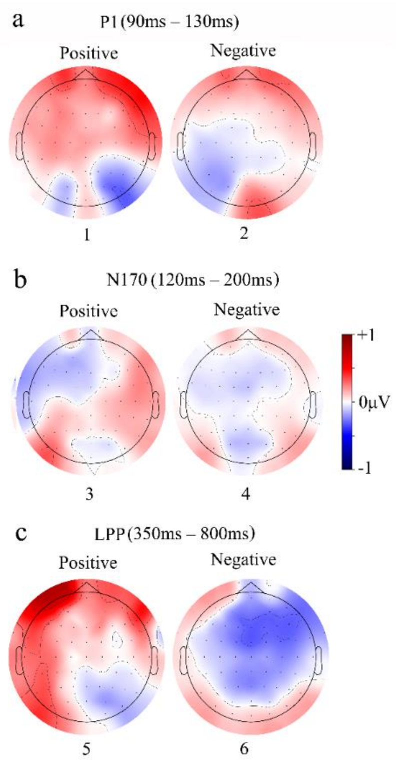
Grand Averages for the ERPs of Self-positivity for the Authentic Self and Presented Self in Experiment 2



Note. (**a1**) Grand averages of P1 for the authentic self and presented self in judging the non-self-descriptiveness of positive traits. (**a2**) Grand averages of P1 for the authentic self and presented self in judging the non-self-descriptiveness of negative traits. (**b1**) Grand averages of N170 for the authentic self and presented self in judging the non-self-descriptiveness of positive traits. (**b2**) Grand averages of N170 for the authentic self and presented self in judging the non-self-descriptiveness of negative traits. (**c1**) Grand averages of LPP for the authentic self and presented self in judging the non-self-descriptiveness of positive traits. (**c2**) Grand averages of LPP for the authentic self and presented self in judging the non-self-descriptiveness of negative traits. The grand-averaged waveforms represent the grand average across component-specific electrode clusters: for the P1 component, signals were averaged across nine posterior sites (P3, P4, Pz, PO3, PO4, POz, O1, O2, and Oz); for the N170 component, across 16 temporo-parieto-occipital sites (TP7, TP8, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, O1, and O2); and for the LPP component, across 15 fronto-centro-parietal sites (F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4, CP3 CPZ, CP4, P3, P4, and Pz). $^{***}p < .001$.

Figure S2. 3

Topological Maps of Self-positivity for the Authentic Self and Presented Self in Experiment 2



Note. (**a**) P1 amplitude difference between the presented self and authentic self in judging the non-self-descriptiveness of positive and negative traits. The time window of P1 is 90 ms – 130 ms. (**b**) N170 amplitude difference between the presented self and authentic self in judging the non-self-descriptiveness of positive and negative traits. The time window of N170 is 120 ms – 200 ms. (**c**) LPP amplitude difference between the presented self and authentic self in judging the non-self-descriptiveness of positive and negative traits. The time window of LPP is 350 ms – 800 ms. 1 = P1 amplitude of non-descriptiveness judgments on positive traits for the presented self minus P1 amplitude of non-descriptiveness judgments on positive traits for the authentic self; 2 = P1 amplitude of non-descriptiveness judgments on negative traits for the presented self minus P1 amplitude of non-descriptiveness judgments on negative traits for the authentic self; 3 = N170 amplitude of non-descriptiveness judgments on positive traits for the presented self minus N170 amplitude of non-descriptiveness judgments on positive traits for the authentic self; 4 = N170 amplitude of non-descriptiveness judgments on negative traits for the presented self minus N170 amplitude of non-descriptiveness judgments on negative traits for the authentic self; 5 = LPP amplitude of non-descriptiveness judgments on positive traits for the presented self minus LPP amplitude of non-descriptiveness judgments on positive traits for the authentic self; 6 = LPP amplitude of non-descriptiveness judgments on negative traits for the presented self minus LPP amplitude of non-descriptiveness judgments on negative traits for the authentic self.

Appendix B Supplementary Material for Chapter 4

The appendix includes:

Figure S4.1 to S4.2 (Part A)

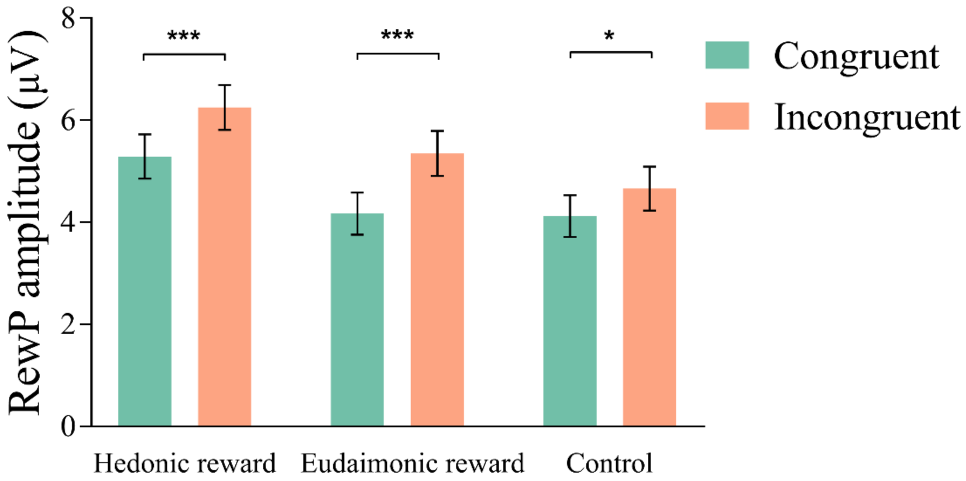
Southampton Authenticity Scale (Part B)

Appendix B

Supplementary Figures (Part A)

Figure S4. 1

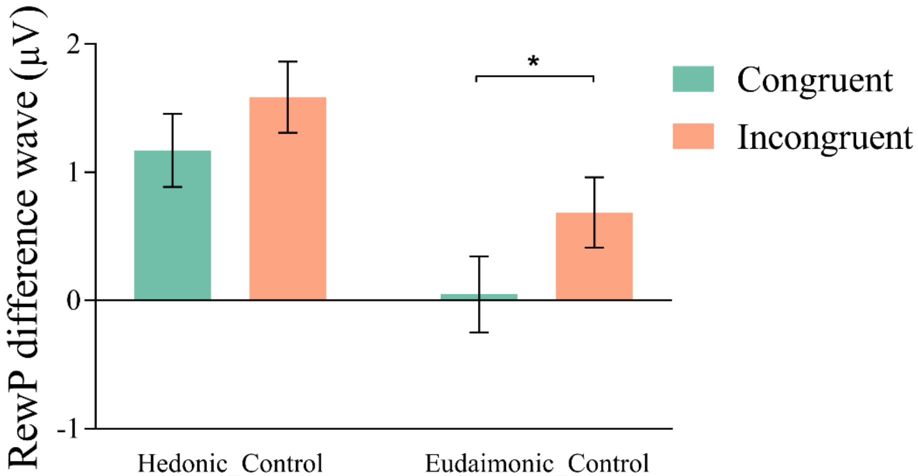
Grand Averages of RewP Amplitude as a Function of Reward and Congruency



Note. Error bars represent SEM; $^{***}p$ < .001. $^{**}p$ < .01. $^{*}p$ < .05.

Figure S4. 2

Grand Averages of RewP Difference Wave as a Function of Reward and Congruency



Note. Error bars represent SEM; $^{*}p < .05$.

Appendix B

Southampton Authenticity Scale (Part B)

Please indicate your level of agreement with each of the statements below.

1 = Strongly disagree, 2 = Mostly disagree, 3 = Slightly disagree, 4 = Neither Agree nor Disagree, 5 = Slightly agree, 6 = Mostly agree, 7 = Strongly agree

1. In general, I feel authentic.
2. In general, I feel true to myself.
3. In general, I feel like the real me.
4. In general, I feel genuine.