Title: Are you as important as me? Self-Other discrimination within trait-adjective processing

Authors: Anna Hudson., McLennon J.G. Wilson., Emma S. Green., Roxane J. Itier., Heather A. Henderson

Corresponding author: Heather A. Henderson, 200 University Avenue West, Waterloo ON, Canada, N2L 3G1, PAS 4009; P: 519-888-4567 X31597; F: 519-746-8631 hhenderson@uwaterloo.ca

Affiliations: Department of Psychology**,** University of Waterloo

Abstract

Healthy adults typically display enhanced processing for self- (relative to other-) relevant and positive (relative to negative) information. However, it is unclear whether these two biases interact to form a self-positivity bias, whereby self-positive information receives prioritized processing. It is also unclear how a blocked versus mixed referent design impacts reference and valence processing. We addressed these questions using behavioral and electrophysiological indices across two studies using a Self-Referential Encoding Task, followed by surprise recall and recognition tasks. Early (P1) and late (LPP) event-related potentials were time-locked to a series of trait adjectives, encoded relative to oneself or a fictional character, with referent presented in a blocked (Exp. 1) or mixed (Exp. 2) trial design. Regardless of study design, participants recalled and recognized more self- than other-relevant adjectives, and recognized more positive than negative adjectives. Additionally, participants demonstrated larger LPP amplitudes for self-relevant and positive adjectives. The LPP self-relevance effect emerged earlier and persisted longer in the blocked (400-800ms) versus mixed design (600-800ms). The LPP valence effect was not apparent in the blocked design, but appeared late in the mixed design (600-1200ms). Critically, the interaction between self-relevance and valence appeared only behaviorally in the mixed design, suggesting that overall self-relevance and valence independently impact neural socio-cognitive processing.

1. Introduction

From early in development, humans preferentially attend to information relevant to themselves. This bias towards self-relevant information provides the foundation for the development of social cognition, wherein one learns about the self in relation to others (Gordon, 1986; Henderson & Mundy, 2013; Molnar-Szakacs & Uddin, 2013). Healthy individuals also preferentially process positively over negatively valenced self-referential information (Auerbach et al., 2016; Burrows, Usher, Mundy, & Henderson, 2017), a bias that may protect against internalizing disorders such as depression (e.g., Goldstein, Hayden, & Klein, 2015). These processing biases are reflected by enhanced neural encoding and memory performance for self-relevant (Knyazev, 2013; Macrae, Moran, Heatherton, Banfield, & Kelley, 2004; Pfeifer, Lieberman, & Dapretto, 2007) and positive information (Auerbach et al., 2015; Auerbach et al., 2016; Langeslag & van Strien, 2009; Shestyuk & Deldin, 2010). However, most studies have examined the positivity bias exclusively with respect to self-referential information. Consequently, it is unclear whether this privileged processing generalizes to other referents and, if so, whether self-relevance potentiates a more general positivity bias. Additionally, it is unclear how attentional demands that result from blocked versus mixed referent trial presentation impact referent and valence processing. To address these issues, the current study examined the effects of referent (self vs. other) and valence (positive vs. negative) simultaneously on healthy young adults' encoding and memory of trait adjectives using behavioral and electrocortical indices, when referent trials were blocked (Exp. 1) or mixed (Exp. 2).

The Self-Referential Encoding Task (SRET) is a well-established paradigm designed to identify biases in the implicit encoding of words into memory (Craik & Lockhart, 1972, Rogers, Kuiper, & Kirker, 1977). The SRET has reliably shown that information framed as relevant to the self (e.g., "Does this word describe *you*?") is encoded more deeply than information cued by structural, phonemic, and semantic properties of the text (for a meta-analytic review, see Symons & Johnson, 1997). This 'self-reference effect' is also apparent when self-referential information is contrasted with *other*-referential information (e.g., "Does this word describe [someone else]; Kuiper & Derry, 1982; Symons & Johnson, 1997). As such, this bias towards self-referential information appears to be a core feature of adaptive socio-cognitive functioning in healthy adults.

The SRET has also been used to examine the effect of word valence on information processing. In addition to exhibiting enhanced memory for self-relevant information, healthy individuals tend to endorse and recall more positive than negative self-relevant trait adjectives, a phenomenon known as the 'self-positivity effect' (Auerbach et al., 2016; Kuiper & Derry 1982; Shestyuk & Deldin, 2010; Symons & Johnson, 1997). This information processing bias may play an important role in maintaining positive self-regard and protecting against negative mental health outcomes (Rogers et al., 1977; Maddi, 1989; Fiske & Taylor, 1991). For example, in depressed individuals, the self-positivity effect is reduced or even reversed such that participants endorse and remember more negative than positive self-relevant trait adjectives (Auerbach, Stanton, Proudfit, & Pizzagalli, 2015; Goldstein et al., 2015; Kuiper & Derry, 1982). While a negativity bias for threat-related stimuli can be found in healthy populations during simple discrimination tasks (i.e., angry faces: Jackson, Wu, Linden, & Raymond, 2009; fearful faces: Righi, Marzi, Toscani, Baldassi, Ottonello, & Viggiano, 2012; negative images: Lto, Larsen, Smith, & Cacioppo, 1998), these paradigms are inherently non self-specific, and tap into simple attention capture rather than extended processing. Therefore, the prioritization of positive information specifically relevant to the self may be fundamental to adaptive social cognition (Mezulis, Abramson, Hyde, & Hankin, 2004). However, it is currently unclear how referent and valence cues within a SRET paradigm unfold at the time of information processing. This paper directly addresses this question using both behavioural and neural measures.

A promising means for exploring referent and valence biases at the neural level is to examine event-related potentials (ERPs). Past work suggests early electrocortical differentiation in response to referent (Shestyuk & Deldin, 2010) and valence (Auerbach et al., 2015; Auerbach et al., 2016; Shestyuk & Deldin, 2010) when processing trait adjectives. The P1 ERP component occurs around 100ms post-stimulus onset at parietal-occipital sites, and is thought to reflect both involuntary and voluntary attention capture (Luck, Woodman, & Vogel, 2000; Mangun, 1995). Focusing exclusively on the self-referential condition of the SRET, Auerbach et al. (2016) found enhanced P1 amplitudes across parietal and frontocentral sites during the encoding of positive relative to negative trait adjectives in healthy adolescent females, a pattern that was reversed in a sample of clinically depressed female adolescents (Auerbach et al., 2015). Similarly, Shestyuk and Deldin (2010) examined the effects of the SRET in current and remitted depression relative to a small non-depressed comparison group on behavioral and neural indices of self-referential processing. In their design, Shestyuk and Deldin recorded EEG during blocked self- and other-relevant encoding conditions (i.e., repeated trials with the same referent). Each referent condition was presented separately in 5 blocks of 18 trials, with an intentional recall task following each block. For healthy controls, Shestyuk and Deldin (2010) reported enhanced P2 amplitudes (200-300ms) in response to positive relative to negative items that were specific to the self-referential condition. In both aforementioned studies, the self-positivity effects carried through to the Late Positive Potential (LPP), an ERP component occurring between 400 and 1200 ms post-stimulus onset over parietal and frontocentral regions.

The LPP generally reflects sustained attentional processing of salient stimuli (Foti & Hajcak, 2008; Schupp et al., 2004; Wieser et al., 2014), and can be split into early (400-600 ms post stimulus onset) and late (600-1200ms) sub-components. The former is associated with encoding and processing of emotional information (Naumann, Batussek, Diedrich, & Laufer, 1992), while the latter is associated with memory storage and affective encoding (Ruchkin, Johnson, Mahaffey, Sutton, 1988). The enhanced LPP amplitudes specifically for self-positive items (Auerbach et al., 2016; Shestyuk & Deldin, 2010) suggest valence effects are unique to self-referential processing during a referent-blocked design, whereby self-relevant information potentiates the positivity bias in healthy adults. In contrast, in depressed samples of adolescent females (Auerbach et al., 2015) and depressed adults (Shestyuk & Deldin, 2010), the LPP was enhanced for negative relative to positive self-relevant adjectives. However, it is unclear whether the valence effects within self-referential processing were driven by the continual, same referent processing that occurs as a result of the referent-blocked experimental design. Thus, it is important to replicate and extend this work using a mixed referent trial design to determine whether the self-positivity bias holds across *intermittent* referent processing. Since Shestyuk and Deldin (2010) were the first to present an other-relevant condition, it is also critical to replicate their findings to understand the specificity of valence effects in self- and other-relevant processing, to examine whether these effects generalize across different trial designs (i.e. blocked relative to mixed relevance conditions), and to examine if these effects occur naturally in an incidental (relative to intentional) memory paradigm.

It is important to note that neural responses to valence vary across word-related ERP studies. In a design slightly differing from the original SRET, Fields and Kuperberg (2012) presented written social scenarios that were self- or other-directed, with ERPs time-locked to a critical word mid-sentence (e.g., A man knocks on *Sandra's/your* hotel room. *She/you* see(s) that he has a *gun/gift/tray* in his hand), and found enhanced LPP responses to unpleasant words (i.e., 'gun') relative to pleasant (i.e., 'gift') and neutral (i.e., 'tray') words regardless of referent. Additionally, Herbert, Herbert, Ethofer, and Pauli (2011) recorded ERPs to words depicting pleasant emotions (i.e., 'happiness'), unpleasant emotions (i.e. 'fear'), and neutral nouns (i.e. 'furniture') that were preceded with possessive self-relevant (i.e., 'my') or other-relevant (i.e., 'his') pronouns, or with 'the' indicating no referent. Reduced amplitudes for emotional compared to neutral nouns were found at parietal sites between 200-400ms, interpreted as an early emotion categorization effect. Critically, an interaction between referent and valence was found during the LPP window (350-550ms) at frontal sites, such that amplitudes for negative items in the self-relevant condition were more negative relative to negative-other, and negative-no referent conditions. Finally, at central-parietal sites, enhanced positive amplitudes were found for pleasant relative to neutral nouns regardless of referent, between 450-600ms of the LPP. However, at posterior-parietal sites, larger amplitudes for positive versus negative items were only found in the self-relevant condition. These valence differences seen on ERPs across studies may be a result of differences in ERP recording (reference site and/or electrode used for measurement), study design, or even arousal differences elicited by the experimental designs and stimulus choice. Indeed, arousal has been shown to impact ERPs, particularly the LPP (for review, see Olofsson, Nordin, Sequeira, & Polich, 2008; Speed & Hajcak, 2018) and is an important factor to control in ERP studies. However, the two studies described above (Fields and Kuperberg, 2012 and Herbert et al., 2011) did not focus on neural responses to adjectives characterizing people, making those mixed valence effects difficult to directly relate to those seen in classic SRET paradigms.

To the best of our knowledge, Shestyuk and Deldin (2010) were the only other paper in which ERPs in response to a self- and other-relevant conditions in the SRET were directly compared, with the majority of past behavioral and physiological work focusing exclusively on the effects of valence on self-referential encoding without an other-relevant condition for comparison (e.g., Auerbach et al., 2015; Auerbach et al., 2016; Dainer-Best, Trujillo, Schnyer, & Beevers, 2017; Quevedo et al., 2017; Zhang, Guan, Qi, & Yang, 2013). As a result, there are two plausible explanations for preferential processing of positively valenced self-referential information. It may be that the preferential processing of self-positive information simply reflects the existence of two independent processing biases working in tandem. Alternatively, there may exist a third bias that specifically prioritizes processing of positive, self-relevant information over and above the combined effects of these two independent biases. This unique potentiation would suggest the existence of a specific socio-cognitive mechanism underlying the self-positivity effect. Critically, in order to distinguish between these two explanations, an other-relevant condition is necessary to evaluate the specificity of the positivity effect to self-relevant information.

As the majority of past SRET studies have focused exclusively on the self when contrasting positive and negative word processing, all trials have typically been presented in a single referent block. Upon introducing a second referent (other), one must then consider how the blocking of trials might influence the attentional or processing demands of the task. For example, maintaining the same referent in mind across an entire block of trials might elicit an atypically strong referent effect relative to alternating randomly between referents within a block trial. It is possible that prolonged engagement with a single referent in a blocked design elicits an uncharacteristically strong referent effect which may reduce the valence effect. Differences between blocked and mixed designs have indeed been reported in past behavioural work although the direction of the effect is likely paradigm specific. For example, in a go/no-go task in which participants responded to either a line, or a cross, and withheld a response to a small ring, participants' reaction times did not differ between conditions when go-trials were blocked (i.e. line *or* cross as 'go' trials) but did when condition trials were mixed (i.e. line *and* cross as 'go' trials; Bruder, Ribeiro-do-Valle, 2009). This effect was interpreted as reflecting differences in strategy due to varying difficulty levels, where participants paid more attention in general to the mixed presentation type because three stimuli (as opposed to two) had to be discriminated. Some ERPs have also been shown to be modulated when stimulus types are blocked compared to mixed, such as the feedback-related negativity (FRN) peak in a feedback monitoring task (Pfabigan, Zeiler, Lamm & Sailer, 2014), or the P3 in a task manipulating difficulty (Wilson et al., 1998; although see Pastor et al., 2008, for null results on the LPP). Therefore, it may be that a mixed referent SRET design will improve valence processing, as participants must pay more attention in general to the intermittent referent cue. Given the scarcity of prior research examining trait adjective processing using an other-relevant condition, and given the above mentioned concern of potential attentional differences depending on the type of design used, two ERP experiments were conducted to assess self-relevance and valence processing when relevance type was blocked (Exp. 1) or mixed (Exp. 2).

In summary, the current study sought to replicate and extend the results of Auerbach et al. (2015) and Shestyuk and Deldin (2010) through (1) the inclusion of an other-relevant condition and (2) using both a blocked and a mixed referent design, in a well characterized sample of healthy young adults. To directly examine the mechanism(s) underlying the self-positivity effect, the association between referent and valence in the SRET was assessed using both behavioral and electrophysiological indices. Behaviorally, we examined the effects of referent (self vs. other) and valence (positive vs. negative) on participants' rate of endorsement, spontaneous recall, recognition memory, and memory sensitivity for a series of trait adjectives. Following Auerbach et al. (2015), ERPs were time-locked to word onset and analyses focused on early (P1) and later (LPP; early and late) components. We used a within-subjects design, varying the referent (self vs. other) between trial blocks (Exp. 1), and within trial blocks (Exp. 2), and manipulating word valence (positive vs. negative) within blocks across both studies. Independent word lists were randomized across referent conditions to preclude any carry-over effects which may confound our memory results.

We hypothesized that participants would preferentially process self- (vs. other-) relevant and positive (vs. negative) trait adjectives, as indicated by improved memory and enhanced ERP amplitudes for these items across all components. We expected the effect of referent to be particularly strong in a blocked referent design (Exp. 1). Based on Shestyuk and Deldin (2010), we predicted an interaction between referent and valence in both experiments, such that participants would exhibit preferential processing specifically for self-positive information, with no effect of valence in the other condition

2. Common Methods

2.1 Stimuli

The SRET was designed and administered using Experiment Builder (version 1.11.0.1316). The experiment was run in a sound attenuated Faraday cage. Stimuli were presented on a screen set to a 1600x1200 resolution, with an 85Hz refresh rate, and responses were recorded on a computer keyboard. To reduce head movements during the task, participants used a chinrest throughout the experiment, positioned 70cm in front of the screen. Trait adjectives were presented in lowercase, size 50 Times New Roman font, with a vertical visual angle between .04 and .07 degrees, and a horizontal angle between 1.6 and 5.7 degrees. EEG data were continuously recorded throughout the endorsement phase of the SRET (see Electrophysiology section below for details).

Three word lists containing 64 words each (32 positive; 32 negative) were generated using a lexical database that includes normative ratings of valence, arousal, and dominance (ranging from submissive/weak to dominant/strong) for each word (Warriner, Kuperman & Brysbaert, 2013). Selected words were trait adjectives 3 to 8 characters long (e.g., amazing, awful), categorized as either positive or negative based on valence ratings that ranged from 1 (highly negative) to 9 (highly positive). We defined positive trait adjectives as having a mean affective valence score of 7 or higher, and negative trait adjectives as having a mean score of 3 or lower. Descriptive statistics for each word list are reported in

Table 1. The three word lists were statistically equivalent in terms of valence for both positive, $F(3,191) = .002$, $p = .998$, and negative trait adjectives, $F(3,191) = .007$, $p = .993$, as well as for arousal, $F(3,191) = .57$, $p = .567$, dominance, $F(3,191) = .070$, $p = .933$, and average word length, $F(3, 191) = 1.146$, $p = .320$. Within each list, positive and negative adjectives were not statistically different in length or arousal (ps < .05), although positive words were higher in dominance ($p < .001$). Two word lists were used to generate adjectives in SRET experimental trials, while the third list provided the distractors in the recognition task following the SRET (described below).

Task		Length		Valence		Arousal		Dominance	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
List 1									
	Positive	6.56	1.37	7.48	0.36	4.77	0.91	6.77	0.47
	Negative	6.31	1.45	2.60	0.27	4.78	0.71	3.89	0.71
List 2									
	Positive	6.47	1.37	7.49	0.36	4.93	0.96	6.70	0.41
	Negative	5.87	1.45	2.60	0.28	4.87	0.76	4.11	0.53
List 3									
	Positive	6.75	1.24	7.48	0.35	4.71	0.97	6.66	0.48
	Negative	6.17	1.42	2.60	0.27	4.87	0.76	3.97	0.65

Table 1. Descriptive statistics for each word list. Length is reported in number of characters, while valence, arousal, and dominance reflect the normative ratings taken from Warriner, Kuperman, and Brysbaert (2013). Within each list, positive and negative adjectives did not differ in length (List 1: $t(62) = 0.71$, $p = .480$; List 2: $t(62) = 1.68$, $p = .097$; List 3: $t(62) = 1.31$, $p = .195$) or in arousal (List 1: $t(62) = -0.07$, $p = .949$; List 2: $t(62) = 0.14$, $p = .890$; List 3: $t(62) = -0.71$, $p = .484$). For each list, positive adjectives were higher in dominance (List 1: $t(62) = -19.00$, $p < .001$; List 2: $t(62) = 21.88$, $p < .001$; List 3: $t(62) = 18.90$, $p < .001$).

2.2 Common Procedure

The SRET consisted of three consecutive tasks: (1) endorsement, (2) recall, and (3) recognition. Participants came to the lab under the pretense of completing the endorsement task, but were unaware that they would be asked to complete the recall and recognition tasks.

Endorsement Task. Participants were seated in front of the testing computer and asked to read the task instructions on screen, followed by a brief reiteration of the instructions by the experimenter to ensure participants' full understanding. Participants then completed a short practice block using non-characteristic descriptor words (i.e., blue, flat) followed by two blocks of experimental trials. The popular novel and film character Harry Potter was used as the subject of the Other condition as he is well-known to young adults, and has been used previously in self-referencing tasks (e.g., Burrows, Usher, Mundy, &

Henderson, 2016; Henderson et al., 2009; Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007; Schneider, Debbane, Lagioia, Salomon, d'Argembeau, & Eliez, 2012).

Recall Task. After completing the endorsement task, participants performed a distractor task (verbally counting backwards from 50), followed by a surprise recall task. Participants were handed a blank piece of paper and asked to write down all the adjectives they could recall from the endorsement task (from any condition). After three minutes (or after indicating that they could not remember any additional items), participants' response sheets were collected by the experimenter.

Recognition Task. Immediately following the recall task, participants were given a pencil-and-paper recognition task consisting of all 128 trait adjectives presented in the endorsement task, and an additional 64 distractor trait adjectives from the third list for a total of 192 trait adjectives. Participants were instructed to circle all words they recognized from the experiment. All trait adjectives were presented in a randomized order, with no more than two words per word list or valence being presented adjacently.

After completing the entire SRET, participants were debriefed and remunerated.

2.3 Electrophysiological Recording

EEG was continuously recorded during the endorsement task at 512 Hz by an Active-two BioSemi 64-channel system. The montage included a total of 72 recording sites: 66 channels in an electrode cap under the 10/20 system (64 standard locations plus custom PO9/PO10 sites), two additional pairs of electrodes situated under and beside the eyes to record the Electrooculography (EOG), and an additional pair of electrodes on the mastoids. EEG data were referenced through a Common Mode Sense (CMS) active electrode, and a Driven Right Leg (DRL) passive electrode, which also acted as a ground. Average-referencing occurred offline. Electrode direct current offsets were kept under 20mV.

2.4 Common data processing and analyses

2.4.1 Behavioural data

Endorsement task. Endorsement Scores were calculated as the number of 'Yes' responses by referent (Self vs. Other) and valence (positive vs. negative).

Recall task. A trait-adjective was considered correctly recalled if the root form of the word appearing in the word list was produced (e.g., admirable would be accepted for admired).

Recognition Task. Recognition scores were calculated as the number of words circled on the pen and paper administered recognition sheet. Distractor words that were circled were also recorded to compute *memory sensitivity* scores. The sensitivity metric *d'* was computed as the standardized probability of correctly recognizing a word, subtracted from the standardized probability of incorrectly recognizing a distractor word (Macmillan & Creelman, 2005). The calculation of *d'* requires non-zero values for hit rates (target trait adjectives recognized) and false alarm rates (distractor trait adjectives incorrectly recognized). To account for this (as per Macmillan & Creelman, 2005), the formula $\frac{1}{2N}$ was used to replace hit/false alarm rates of zero, resulting in a corrected value of .0132[1].

A series of 2 x 2 repeated-measures ANOVAs with Referent (Self, Other) and Valence (positive, negative) as within subject factors were conducted with endorsement, recall, recognition, and memory sensitivity (*d'*) scores.

2.4.2 Event-Related Potential Data

EEG and EOG data were pre-processed offline using EEGlab (Delorme & Makeig, 2004) and ERPlab (http://erpinfo.org/erplab) running under Matlab 2014b software (Mathworks, Inc.,2014). Raw data were first segmented into 1400ms long epochs, with a 200ms pre-stimulus baseline (-200ms, +1200ms) and were digitally band-pass filtered (0.01 – 30 Hz) using a two-way least-squares FIR filter. Trials contaminated with artifacts greater than ±70 μV were automatically rejected; any trial deemed too

[1] This correction was used for 6 participants in Exp. 1 (5 endorsing zero negative distractors, 1 endorsing zero positive distractors) and 4 participants in Exp. 2 (2 endorsing zero negative distractors, 2 endorsing zero positive distractors).

noisy after a second visual inspection was then manually rejected. Average waveforms time-locked to word onset were then computed for each participant and each of the four conditions (self-positive, self-negative, other-positive, other-negative). The P1 was averaged across the same electrodes as Auerbach et al. (2015): P1, P2, PO3, PO4, Pz and POz. In Exp. 1, the P1 was defined as the mean amplitude between 100-170ms post stimulus onset based on examining the grand average waveform. The early LPP (hereafter eLPP; mean amplitude between 400-600 ms) and late LPP (hereafter lLPP; mean amplitude between 600-1200 ms) were averaged across the Frontal Central Midline (FCM) sites Fz, FCz, and Cz. To further identify the temporal dynamics of referent and valence processing, the lLPP was further divided into three time windows (600-800 ms, 800-1000 ms, and 1000-1200 ms).

A series of 2 x 2 repeated measures ANOVAs with 2 Referent (Self, Other) and 2 Valence (positive, negative) were run independently for the P1 and eLPP amplitudes. For the lLPP amplitudes, a 2 (referent) x 2 (valence) x 3 (time window) ANOVA was computed to determine the stability of the effects across time. When an interaction with time window was found, follow-up 2 (referent) x 2 (valence) ANOVAs were computed separately for each time window using a Bonferroni corrected *p*-value threshold, such that only $p < 0.016$ were considered significant (0.05/3 comparisons).

2.4.3 Internal consistency

The reliability of the early and late LPP components was assessed using Cronbach's Alpha, a measure of internal consistency that typically describes the extent to which the items in a test measure the same construct (Tavakol & Dennick, 2011). In our analysis, following recommendations from recent ERP studies (Thigpen, Kappenman, Keil, 2017), the four condition-averaged ERPs (self-positive, self-negative, other-positive, other-negative) were entered as items for the computation of Cronbach's Alpha for each ERP component, in each Experiment. For the lLPP, we calculated one Cronbach's alpha for each of the three time windows. With regards to conventional benchmarks (George & Mallery, 2003), all reliability estimates were high, with alpha >.8. Cronbach's alphas for Exp. 1 (Panel A) and Exp. 2 (Panel B) are presented in Table 4.

3. Experiment 1

Exp. 1 Participants. Thirty-eight undergraduate and graduate students aged 18-25 years (Mean = 21.05 years; 20 females) were recruited from the University of Waterloo (UW) psychology participant pool. Participants provided written informed consent, and were compensated with either 2 credits towards an undergraduate course, or $20 upon completion of the study. Exclusion criteria included a history of neurological or psychiatric disease, brain lesions, and alcohol abuse or the regular use of drugs. Participants were required to speak and read English as their first and primary language. Additionally, participants had to have seen at least one Harry Potter movie or read one Harry Potter book to be deemed sufficiently familiar with the character in the Other condition of the SRET (see below). Participants also completed the Center for Epidemiologic Studies Depression Scale – Revised (CESD-R; Van Dam, Earleywine, 2011), in lab during the study. Due to the inherent difference in processing positively- and negatively-valenced items in depressed populations (i.e. Auerbach et al., 2016; Shestyuk & Deldin, 2010), participants scoring above the recommended cutoff of 16 on the CESD-R Depression screening were removed from analysis, leaving a final sample of 29 participants with behavioural data.

Exp. 1 (Blocked) Endorsement Task. Participants were cued to the referent at the beginning of each block. Each block subsequently corresponded to either the "self" or "other" condition. Trials began with the presentation of a central fixation cross, appearing on screen jittered between 1500-1700 ms. A trait adjective then appeared for 200ms, followed by a second fixation cross for 1800ms (see Figure 1 for sample trial progression). Following each adjective in the self condition, participants were prompted with the question "Does this word describe you?" and indicated 'yes' or 'no' via key press ('c' or 'm', mappings counterbalanced across participants). Prompts in the other condition read "Does this word describe Harry Potter?". Each condition presented all items from one of the two 64-word experimental lists. Condition order (self or other) was counterbalanced across participants, and word lists were counterbalanced across the conditions. Within each condition, words were semi-randomized such that no more than two words of the same valence were presented in a row.

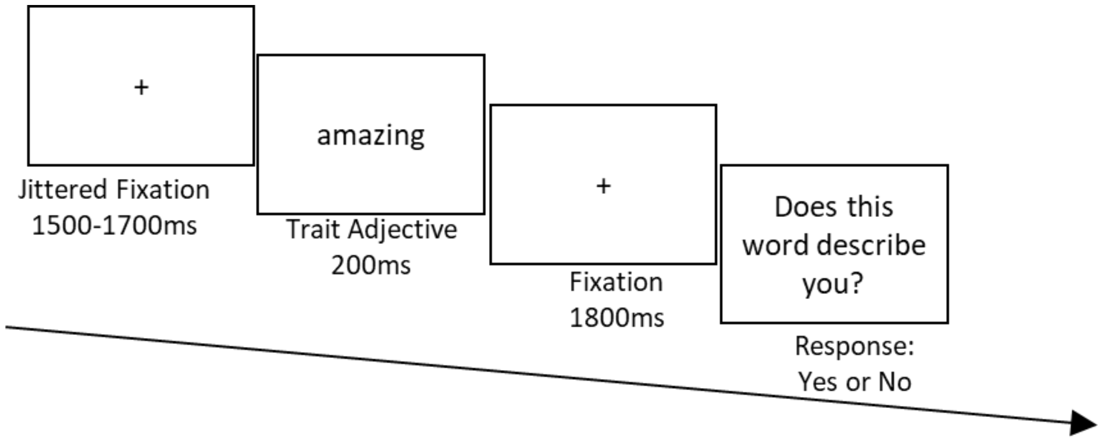

Figure 1. Exp. 1 trial example demonstrating a positive trait adjective in the self-relevant condition. Participants completed two blocked conditions (Self, Other) and each condition included 64 trials (32 positive, 32 negative) with the following progression. First, a jittered fixation cross was presented on

Exp. 1 ERP Data processing. After artifact rejection, 2 participants who had fewer than 20 trials per condition were rejected, leaving 27 participants for the ERP analyses (18 female). Participants had an average of 25.2 (SD = 3.12) trials per condition.

3.1 Exp. 1 Results

3.1.1 Behavioural Data. Descriptive statistics are reported in Table 2. Endorsement data for one participant were lost due to a technical error, resulting in N = 28 for all behavioural analyses involving endorsement.

Task		Self		Other	
		Mean	SD	Mean	SD
Endorsement					
	Positive	23.35	4.33	22.04	4.75
	Negative	4.12	2.64	3.88	3.15
Recall					
	Positive	4.41	2.33	3.96	2.08
	Negative	2.26	1.68	2.11	1.65
Recognition					
	Positive	22.10	5.67	19.28	5.08
	Negative	19.83	5.85	17.38	5.01
Memory sensitivity (*d'*)					
	Positive	1.82	.564	1.53	.557
	Negative	1.95	.577	1.70	.537

Table 2. Means and standard deviations for each variable of interest and each of the four conditions in Exp. 1 (blocked design; self-positive, self-negative, other-positive, other-negative). The maximum number of words in each condition was 32.

Word endorsement. As hypothesized, there was a significant effect of valence, $F(1,27) = 117.47$, $p < .001$, $\eta_p^2 = .81$, such that participants endorsed more positive than negative trait adjectives. However, the effect of referent was not significant, $F(1,27) = 2.85$, $p = .103$, $\eta_p^2 = .10$, nor was the referent-by-valence interaction, $F(1,27) = 0.16$, $p = .691$, $\eta_p^2 = .01$ (Figure 2; Panel A).

Free recall. There was a significant effect of valence, $F(1,28) = 63.64$, $p < .001$, $\eta_p^2 = .69$, such that participants recalled more positive than negative trait adjectives. The main effect of referent was not significant $F(1,28) = 2.04$, $p = .164$, $\eta_p^2 = .07$, nor was the referent-by-valence interaction, $F(1,28) = 0.80$, $p = .384$, $\eta_p^2 = .03$ (Figure 3; Panel A).

Recognition. There was a significant effect of referent, $F(1,28) = 6.96$, $p = .013$, $\eta_p^2 = .20$, such that participants recognized more self-relevant than other-relevant trait adjectives. There was also a significant effect of valence, $F(1,28) = 13.46$, $p = .001$, $\eta_p^2 = .33$, such that participants recognized more positive than negative trait adjectives. The referent-by-valence interaction was not significant, $F(1,28) = .23$, $p = .639$, $\eta_p^2 = .01$ (Figure 4; Panel A).

Memory sensitivity (*d*'). There was a significant main effect of referent, $F(1, 28) = 7.49$, $p = .011$, $\eta_p^2 = .21$. In line with recognition scores, participants were more accurate in their recognition of self- versus other-relevant trait adjectives. The main effect of valence, $F(1,28) = 2.17$, $p = .152$, $\eta_p^2 = .07$, and the referent-by-valence interaction, $F(1,28) = 0.78$, $p = .384$, $\eta_p^2 = .03$, were not significant (Figure 5; Panel A).

3.1.2 Exp. 1 ERP data

Descriptive statistics for each ERP component by processing condition are reported in Table 3. Waveforms showing LPP component are displayed in Figure 6, Panel A. Topographic maps (Figure 7, Panel A) further illustrate the different neural processing patterns across recording epochs for both the referent and valence conditions.

		Condition							
		Self-Positive		Self-Negative		Other-Positive		Other-Negative	
ERP		Mean	SD	Mean	SD	Mean	SD	Mean	SD
P1									
	100-170ms	.700	1.90	.671	1.90	.301	1.75	.322	1.64
eLPP									
	400-600ms	1.60	2.19	1.46	2.12	.476	1.65	.353	2.02
lLPP									
	600-800ms	1.66	2.05	.994	2.24	.270	2.10	-.217	2.30
	800-1000ms	1.32	2.18	.655	2.41	.441	2.59	.084	2.49
	1000-1200ms	.755	2.40	.229	2.53	.60	2.96	-.280	2.80

Table 3. Mean amplitudes (μV) and Standard Deviations averaged across parieto-occipital sites (P1, P2, Pz, POz, PO3, PO4) for the P1 and across the Frontal Central Midline (FCM; Fz, FCz, Cz) for the eLPP and lLPP, for each of the four conditions of Exp.1 (block design).

Cronbach's Alpha			
		Exp. 1	Exp. 2
	Time		
eLPP	400-600ms	.873	.914
lLPP	600-800ms	.819	.893
	800-1000ms	.848	.890
	1000-1200ms	.841	.931

Table 4. Cronbach's Alpha for each of the ERP components analyzed, using the four conditions as "items", in Exp. 1 (A) and Exp. 2 (B).

P1. Analysis of the P1 amplitude (mean amplitude between 100-170 ms post stimulus) yielded no effect of referent, $F(1,26) = 3.80$, $p = .062$, $\eta_p^2 = .127$, no effect of valence $F(1,26) < .001$, $p = .985$, $\eta_p^2 < .001$, nor an interaction between valence and referent $F(1,26) = .002$, $p = .888$, $\eta_p^2 = .001$.

eLPP. There was a main effect of referent, with an increased eLPP amplitude for self-relevant compared to other-relevant items, $F(1, 26) = 17.04$, $p < .001$, $\eta_p^2 = .39$, see Figure 6, Panel A. There was no effect of valence $F(1,26) = .292$, $p = .594$, $\eta_p^2 = .011$, nor a referent-by-valence interaction $F(1,26) = .002$, $p = .963$, $\eta_p^2 < .001$.

lLPP. The main effect of time window was significant $F(1.236, 32.85) = 6.58$, $p = .010$, $\eta_p^2 = .202$. There was no main effect of valence, $F(1, 26) = 3.95$, $p = .058$, $\eta_p^2 = .132$, nor a valence by time window interaction $F(2, 52) = .376$, $p = .688$, $\eta_p^2 = .014$. The main effect of referent was significant, $F(1, 26) = 5.16$, $p = .032$, $\eta_p^2 = .166$, but qualified by a referent by time window interaction, $F(2, 52) = 7.21$, $p = .002$, $\eta_p^2 = .217$. Thus, each time window (600-800ms; 800-1000ms; 1000-1200ms) was subjected to a 2 (referent) by 2 (valence) ANOVA, with a Bonferroni correction such that significance was set at $p < .016$.

Analyses of the three sub-windows of the lLPP demonstrated an effect of referent between 600-800ms, $F(1, 26) = 13.47$, $p = .001$, $\eta_p^2 = .341$, such that amplitudes for self-relevant items were enhanced relative to other-relevant items. The effect of referent was no longer significant between 800-1200ms (800-1000ms: $F(1, 26) = 3.58$, $p = .070$, $\eta_p^2 = .121$; 1000-1200ms $F(1,31) = 2.25$, p = .143), see Figure 6, Panel A. Moreover, in the omnibus ANOVA, no interaction between referent and valence was found

$F(1,26) = .258$, $p = .616$, $\eta_p^2 = .010$, nor a three-way interaction between referent, valence and time window $F(2,52) = .245$, $p = .784$, $\eta_p^2 = .009$.

Topographic Maps. Figure 7, Panel A demonstrates reference effects across frontal electrodes, where amplitudes for self-relevant items were enhanced relative to other-relevant items. This effect appeared between 400-600ms, and was maintained between 600-800ms. Alternatively, valence effects were not statistically significant in any time window, as reflected by noisier maps.

3.2 Exp. 1 Discussion

The goal of Exp. 1 was to examine whether the self-positivity bias reflects the existence of two independent processing biases working in concert or a unique bias specifically prioritizing the processing of self-relevant, positively-valenced information. Replicating past work using a referent-blocked design, we found evidence for both self-referencing and positivity effects at the behavioural level (Auerbach et al., 2015; Kuiper & Derry 1982; Symons & Johnson, 1997; Shestyuk & Deldin, 2010), although only referent effects at the ERP level. Participants recognized more self- than other-relevant words with corresponding early (eLPP) and late (lLPP) neural enhancements. Participants also endorsed, recalled, and recognized more positive relative to negative words. However, contrary to our hypothesis, there was no interaction between referent and valence at either the behavioral or neural level, failing to support the existence of a unique socio-cognitive mechanism for processing self-positive information in a referent blocked paradigm.

The early differentiation of referent on the eLPP through to the beginning stage of the lLPP in the current study may be a result of the blocked presentation of referent conditions. That is, participants were informed of the referent condition at the beginning of each block. In this design, participants are instructed to think of the same individual across the entire block, which required no task switching between trials. Consequently, the processing of each word was strongly primed by the referent of each block, which may have enhanced the referent effect, particularly at the neural level. To address the effect of long-term, continual, same-referent processing relative to inconsistent and variable referent-processing, Exp. 2 employed a mixed trial design in which self-relevant and other-relevant trait adjectives were presented randomly within a block, changing at a trial by trial level.

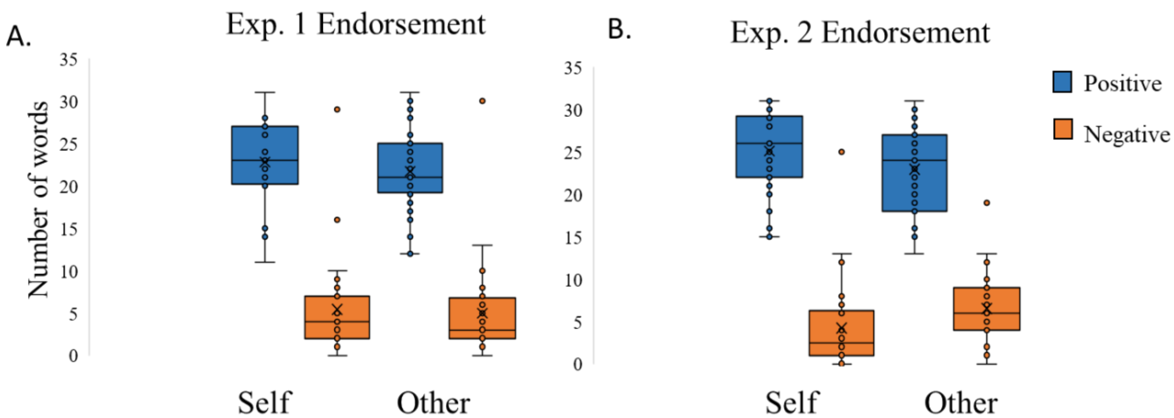

Figure 2. Endorsement rates for Exp. 1 (N = 28; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* Inter Quartile Range (IQR; minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle. The maximum number of words in each condition was 32.

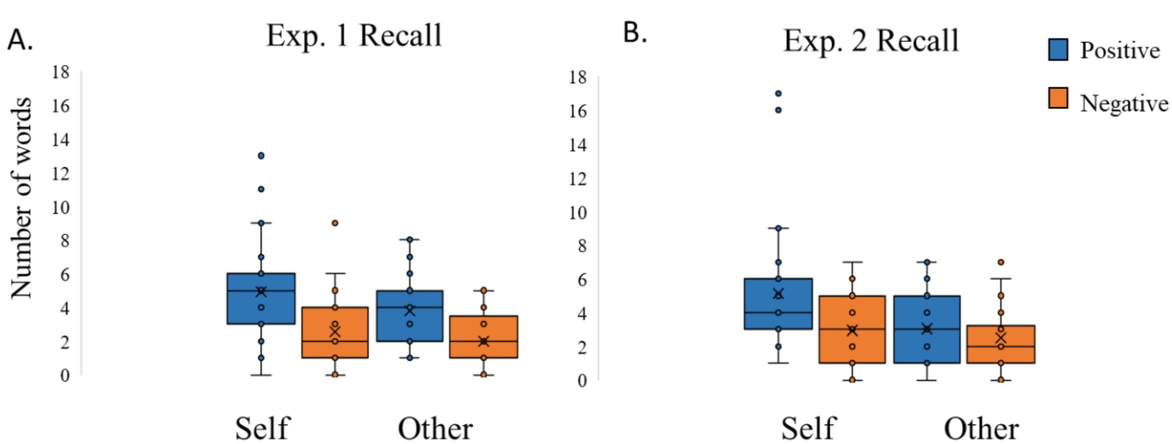

Figure 3. Recall rates for Exp. 1 (N = 29; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* IQR (minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle. The maximum number of words in each condition was 32.

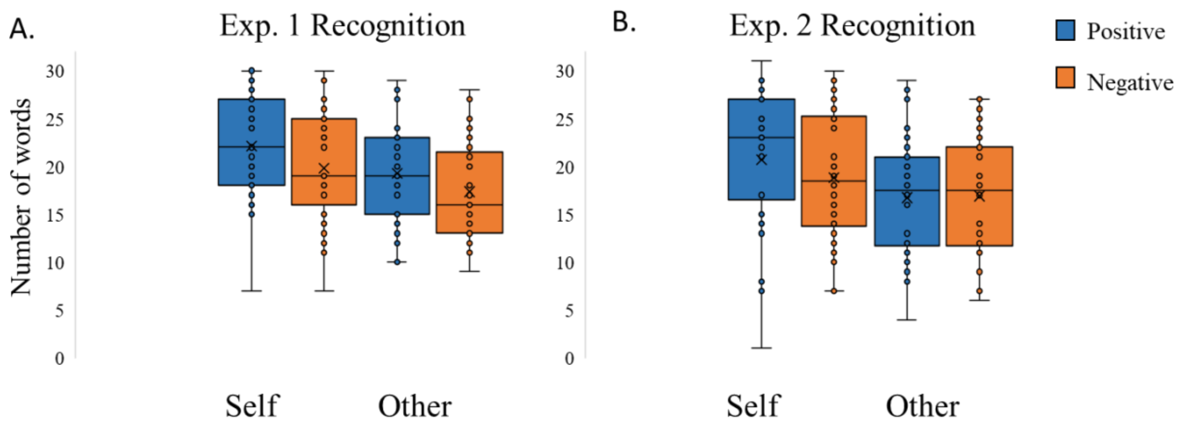

Figure 4. Recognition rates for Exp. 1 (N = 29; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* IQR (minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle.

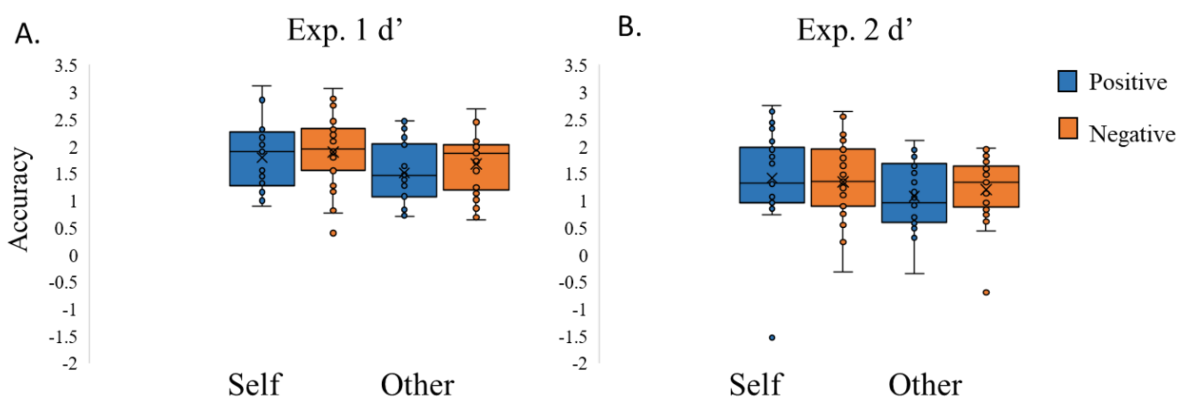

Figure 5. D Prime (accuracy) score for Exp. 1 (N = 29; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* IQR (minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle.

4. Experiment 2

Exp. 2 Participants. A fully independent (from Exp. 1) sample of forty-four UW undergraduate and graduate participants aged 18-25 years (Mean = 20.12; 22 female) were recruited from the psychology participant pool. Participants provided written informed consent and were compensated with either 2 credits towards an undergraduate course, or $20 upon completion of the study. Exclusion criteria were identical to Exp. 1, with the additional criterion that they were not to have participated in the first version of the study. Participants scoring above the recommended score of 16 on the CESD-R were not included in analyses, leaving 29 participants with behavioural data.

Exp. 2 Mixed Trials Endorsement Task. Each trial began with a prompt that said either "you" or "Harry Potter" for 1000ms, and participants were instructed to think about that person for the remainder of that trial. Following the prompt, a central fixation cross appeared on the screen, jittered between 1500-1700ms, following by the trait adjective for 200ms. A second central fixation cross appeared for 1800ms before asking participants to make a behavioural response. Participants were asked "Does this describe you?" if they were prompted with 'You" at the beginning of the trial, or asked "Does this word describe Harry Potter?" if they were prompted with "Harry Potter" at the beginning of the trial. Participants indicated a 'yes' or 'no' response via key press ("c' or 'm', counterbalanced). Participants were required to respond at the end of each trial as to whether or not the word presented described the person prompted at the beginning of each trial (see Figure 8 for trial design).

Exp. 2 ERP Data processing. After trial rejection, 7 participants who had fewer than 20 trials per condition were rejected, leaving 23 participants for the ERP analyses (16 female). Participants had an average of 22.39 (SD = 3.36) trials per condition. The P1 was defined as the mean amplitude between 90-170ms post stimulus onset based on examining the grand average waveform. The LPP was measured across the same electrodes and time windows identified in Exp. 1.

4.2. Exp. 2 Results

4.2.1 Behaviour Data. Descriptive statistics are reported in Table 5.

	Self		Other	
Task	Mean	SD	Mean	SD
Endorsement				
Positive	24.79	4.80	23.21	4.82
Negative	3.54	3.84	5.89	3.37
Recall				
Positive	4.20	2.0	2.85	1.93
Negative	2.62	1.80	2.23	1.50
Recognition				
Positive	21.18	6.40	17.00	6.32
Negative	18.89	6.72	17.12	6.09
Memory sensitivity (*d'*)				
Positive	1.48	.591	1.10	.600
Negative	1.37	.625	1.25	.452

Table 5. Means and standard deviations for each variable of interest and each of the four conditions in Exp. 2 (mixed design; self-positive, self-negative, other-positive, other-negative) in the mixed SRET design. The maximum number of words in each condition was 32.

Word endorsement. The main effect of valence was significant, $F(1,28) = 206.35$, $p < .001$, $\eta_p^2 = .88$, such that participants endorsed more positive than negative trait adjectives. The main effect of referent was not significant, but there was a referent by valence interaction, $F(1,28) = 15.99$, $p < .001$, $\eta_p^2 = .36$, whereby the difference between positive and negative word endorsement was greater in the self-relevant condition than in the other-relevant condition (Figure 2; Panel B).

Free recall. There was a main effect of referent, $F(1,28) = 10.39$, $p = .003$, $\eta_p^2 = .27$ in which participants recalled more self-relevant than other-relevant words. There was also a main effect of valence, $F(1,28) = 17.81$, $p < .001$, $\eta_p^2 = .39$ in which more positive words were recalled. The referent by valence interaction was also significant, $F(1,28) = 6.16$, $p = .019$, $\eta_p^2 = .18$ such that positive words in the self-relevant condition were better remembered over all other trial types[2] (Figure 3; Panel B).

[2] Free recall scores for two participants in the self-positive condition were determined to be statistically extreme outliers in this sample (scores of 16 and 17 out of a possible 32 words), as classified by the SPSS 'explore' function to highlight outliers. This function identifies extreme outliers as 3 times the inter-quartile range. They remained in the analyses as they did not differ on other measures, and simply performed well on this task. However, when they were removed from the sample, the interaction became insignificant, $F(1,26) = 3.73$, $p = .065$, $\eta_p^2 = .13$.

Recognition. The main effect of referent was significant, $F(1,28) = 17.42$, $p < .001$, $\eta_p^2 = .38$ such that participants recognized more self-relevant items. However this effect was modulated by a significant referent by valence interaction, $F(1,28) = 4.75$, $p = .038$, $\eta_p^2 = .15$, in which there was an enhanced recognition for self-positive over self-negative items, but no difference between positive and negative items in the other-relevant condition. The valence effect was not significant (Figure 4; Panel B)

Memory sensitivity (*d'*). The main effect of referent was significant such that items presented in the self condition were more accurately recognized, $F(1, 28) = 15.52$, $p < .001$, $\eta_p^2 = .357$. The critical interaction between valence within each referent were not significant, however a referent by valence interaction emerged, $F(1, 28) = 4.41$, $p = .045$, $\eta_p^2 = .136$, such that self-positive and self-negative items were more accurately recognized (with fewer false alarms) than other-positive items, and marginally different than other-negative items. However, there was no difference between self-positive and self-negative, nor between other-positive and other-negative items. The main effect of valence was not significant, $F(1, 28) = .075$, $p = .786$, $\eta_p^2 = .003$ (Figure 5; Panel B)

4.2.2 Exp. 2 ERP data.

Participants had an average of 22.69 trials per condition (S.D = 3.08). Descriptive statistics for each ERP component are reported in Table 6. The LPP waveform is displayed in Figure 6, Panel B. Topographic maps (Figure 7, Panel B) further illustrate the different neural processing patterns across recording epochs for both the referent and valence conditions.

		Condition							
		Self-Positive		Self-Negative		Other-Positive		Other-Negative	
ERP		Mean	SD	Mean	SD	Mean	SD	Mean	SD
P1									
	90-170ms	.060	1.43	.353	.995	.821	.996	.380	1.09
eLPP									
	400-600ms	1.84	1.86	1.49	2.18	1.40	2.39	.816	1.92
lLPP									
	600-800ms	1.02	1.60	.572	1.83	.360	2.22	-.549	1.86
	800-1000ms	.382	1.76	-.220	2.40	.044	2.11	-1.00	1.53
	1000-1200ms	-.323	1.85	-.830	2.49	-.422	2.20	-1.58	1.78

Table 6. Mean amplitudes (µV) and Standard Deviations averaged across parieto-occipital sites (P1, P2, Pz, POz, PO3, PO4) for the P1 and across the Frontal Central Midline (FCM; Fz, FCz, Cz) for the eLPP and lLPP, for each of the four conditions in Exp.2 (mixed design).

P1. The effect of referent was not significant, $F(1,22) = 3.03$, $p = .096$, $\eta_p^2 = .12$, nor were the effects of valence $F(1,22) = .25$, $p = .620$, $\eta_p^2 = .01$, or the referent by valence interaction, $F(1,22) = 3.45$, $p = .077$, $\eta_p^2 = .14$.

eLPP. The main effect of referent was not significant, $F(1,22) = 3.22$, $p = .086$, $\eta_p^2 = .13$, nor was the effect of valence, $F(1,22) = 3.68$, $p = .068$, $\eta_p^2 = .14$, or the interaction, $F(1,22) = .38$, $p = .543$, $\eta_p^2 = .02$ (see Figure 6; Panel B).

lLPP. The main effect of time was significant $F(1.62, 25.56) = 16.95$, $p < .001$, $\eta_p^2 = .435$. The main effect of valence was significant, $F(1, 22) = 4.84$, $p = .039$, $\eta_p^2 = .180$, but did not interact with time $F(2, 44) = .855$, $p = .432$, $\eta_p^2 = .037$. Across the waveform, positive items elicited more positive amplitude than negative items (Figure 6, Panel B). There was no main effect of referent, but the referent by time window interaction was significant, $F(2, 44) = 6.50$, $p = .003$, $\eta_p^2 = .228$. Thus, each time window (600-800ms; 800-1000ms; 1000-1200ms) was subjected to a 2 (referent) by 2 (valence) ANOVA, with a Bonferroni correction such that significance was set at $p < .016$. Analysis of the three lLPP sub-windows demonstrated a main effect of referent only during the 600-800ms window, $F(1,22) = 7.53$, $p = .012$, $\eta_p^2 = .26$ such that self-relevant items elicited more positive amplitudes than other-relevant items.

In the omnibus ANOVA, no valence by referent interaction was found, nor a valence by reference by time window interaction.

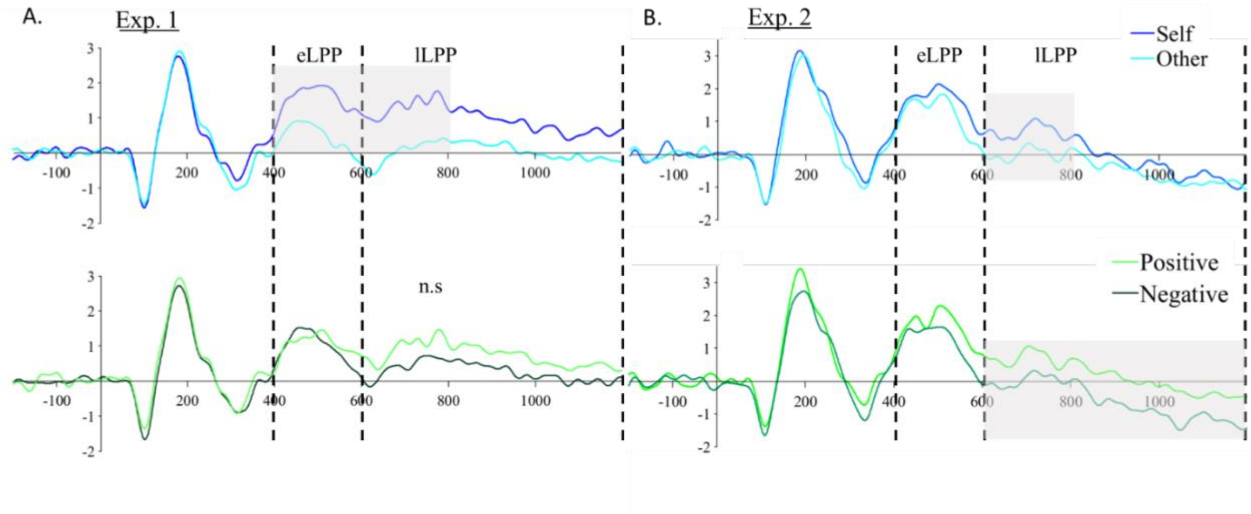

Figure 6. The entire 1200ms ERP waveform averaged across electrodes Fz, FCz, and Cz to create the Frontal Central Midline (FCM). ERPs were time locked to the onset of word presentation during the encoding phase of the task. The significant effects of referent (top) and valence (bottom) for Exp. 1 (Panel A) and Exp. 2 (Panel B) are highlighted in grey.

Topographic Maps. Figure 7, Panel B demonstrates reference effects across frontal electrodes, where amplitudes for self-relevant items were enhanced relative to other-relevant items. This effect appeared to begin early in the epoch, peaking between 600-800ms, before becoming weaker. Alternatively, valence effects did not appear until about 600ms, where positive items elicited an enhanced amplitude relative to negative items. This valence effect persisted until the end of the epoch.

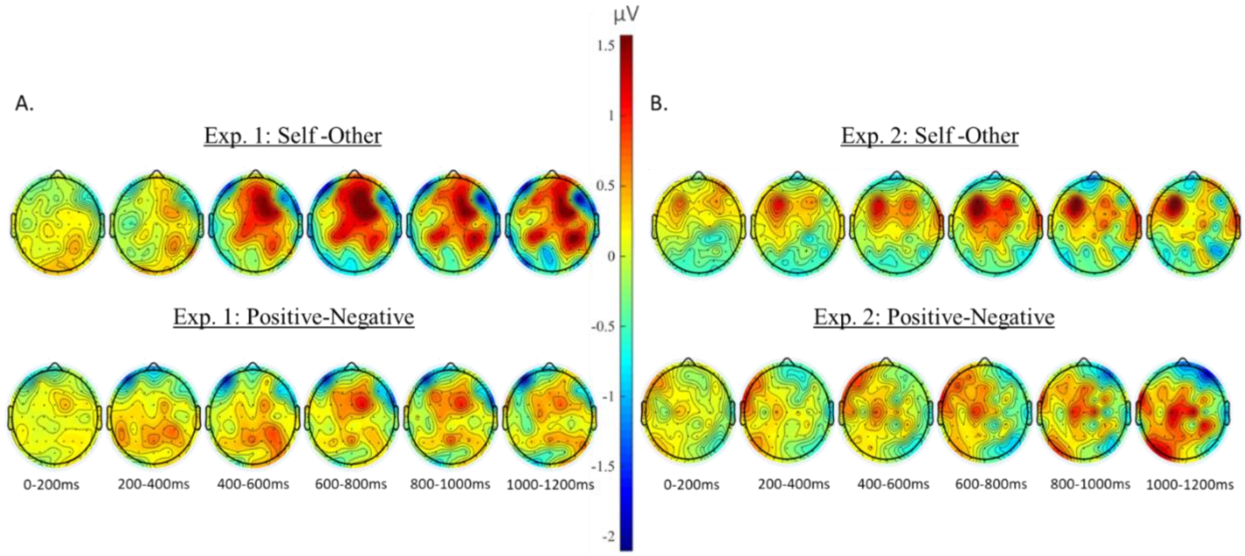

Figure 7. Topographic maps across the 1200ms waveform. Average Self-Other differences are displayed across the top for both Exp. 1 (Panel A) and Exp. 2 (Panel B). Warmer tones reflect more positive ERP amplitudes for Self-relevant trials relative to Other-relevant trial types. Maps across the bottom reflect average Positive-Negative differences, with warmer tones reflecting more positive amplitudes for Positive trials relative to Negative trial types.

4.3 Exp. 2 Discussion

The goal of Exp. 2 was to examine the influence of a mixed referent design in the processing of referent and valence cues. Critically, at the behavioural level, there was an interaction between referent and valence on endorsement, free recall, recognition, and memory sensitivity (*d'*), such that the difference between positive and negative items was enhanced in the Self condition, but was reduced in the Other condition (endorsement) or not apparent at all (recognition, *d'*). This finding is in contrast to Exp.1, where participants did not differ in the way they processed positive and negative information within the self-referential condition. This suggests that intermittent referent processing, relative to consistent referent processing, modulates memory. Interestingly, a comparable interaction was not apparent at the neural level, where, like in Exp. 1, referent and valence information were encoded separately, although the onset and duration of these effects differed. This lack of interaction at the neural level suggests that, regardless of study design, referent and valence information were encoded separately, at least up until 1200ms.

Importantly, participants did not make a response until 1800ms, however our epoch of interest only extended to 1200ms. It is possible that a neural interaction may exist after 1200ms.

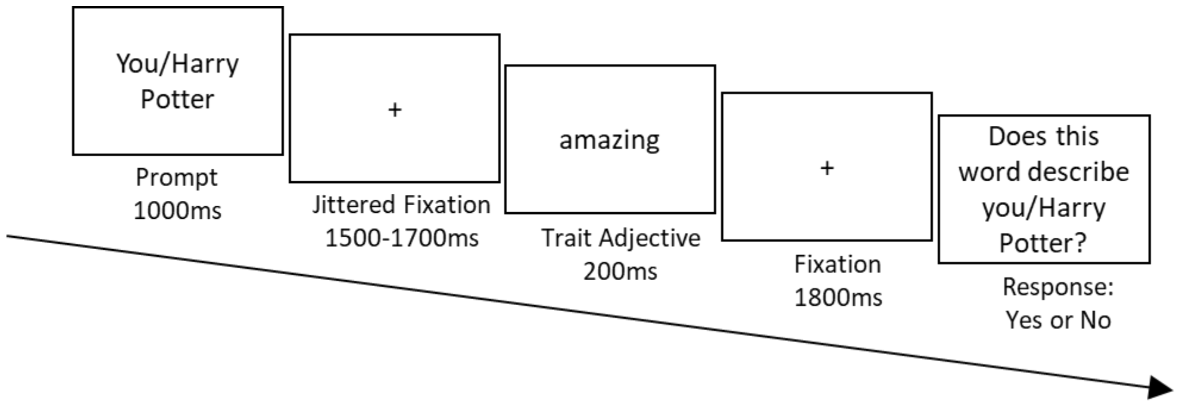

Figure 8. Exp. 2 trial sample in which each trial begins with a prompt indicating who the participant should think about for the remainder of the trial.

5. General Discussion

The self-reference effect was apparent across recognition and memory sensitivity (*d'*) scores in both Exp. 1 and Exp. 2, such that memory for self-relevant items was enhanced relative to other-relevant items. Given that equal numbers of traits were endorsed in the Self and Other conditions, these findings demonstrate that adjective words related to the self were better encoded than words related to a fictional character. Likewise, across both experiments, the positivity effect was apparent in rates of endorsement, where participants endorsed more positive than negative words regardless of the referent condition. The positivity effect was also found in recall scores, such that participants remembered significantly more positive than negative trait adjectives. Given that the word lists were comprised of carefully controlled (valence, word length, dominance, and arousal) words, these results are a particularly strong replication and extension of past behavioral findings. Moreover, unlike past studies that used the same word list in both referent conditions (e.g., Fields & Kuperberg, 2012; 2015; Shestyuk & Deldin, 2010), the current study employed fully independent word lists across referent blocks (counter-balanced across participants), precluding carry-over effects for previously encoded words.

Interestingly, in contrast to Exp.1, enhanced recall for items that were self-relevant was found in Exp.2, as well as improved recognition and accuracy specifically for self-positive items. As the only difference between studies was the blocked relative to the mixed referent design, these findings suggest there may be different cognitive processes involved in consolidating referent and valence information following intermittent rather than prolonged referent processing. The blocked referent design did not require participants to task-switch between referent cues at the beginning of each trial in the way the mixed referent design did. Consequently, participants could remain less vigilant to each word presented within each block during Exp.1, as they could systematically process referent information in a similar way across all trials. That is, participants did not need to update their referent cue at each trial, allowing them to perhaps drift or mind-wander through the blocked referent design. The constant referent switching in Exp.2 may have resulted in enhanced attention to every trial in order to complete the task, which may have facilitated participants' memory. Previous work assessing the effects of learning across trial presentation types suggest blocked designs highlight the similarities within a category, while mixed designs highlight the differences between categories (Carvalho & Goldstone; 2014a; 2014b). This explanation may account for the increased recognition specifically for self-positive items in our mixed referent design. Since a mixed design highlights differences between categories, an enhanced positivity bias for themselves compared to Harry Potter may be more noticeable to participants, relative to the blocked referent design. Additionally, there was no referent effect found in recall scores during the blocked referent design, but there was in the mixed referent design. This again may be a result of mixed designs highlighting between-category differences, as well as the potentially enhanced vigilance across trials. However, to further establish the effect of blocking and intermixing trial types on referent and valence processing, future studies should consider blocking based on valence.

It is important to note that the interaction between referent and valence for participants' recall scores was reduced to non-significance in Exp. 2 when two statistical outliers were removed. Consequently, it remains unclear how participants consolidate referent and valence cues. It is possible that the large stimulus set, which was necessary to ensure enough trials to analyze ERPs, may have contributed to participants' overall low recall rates. As such, it is possible that self-positive adjectives *did* receive prioritized processing, but that the effect was masked due to the overall poor recall performance. Previous ERP SRET paradigms included far fewer target words in the endorsement task. For example, Auerbach et al., (2015) and Auerbach et al., (2016) had participants complete an endorsement task of 80 target words and each reported about 17 total words recalled, that is, roughly 21% of the presented words. Shestyuk and Deldin (2010) presented just 44 words in total, repeated for both self and other-referent blocks but split into 5 blocks of 18 words with a free recall after each block. Participants recalled roughly 30 of the 44 words or about 68% of the words. This is in stark contrast to our study, in which 128 target words were presented in total, with participants recalling on average a total of 13 words in Exp. 1 (~10%), and about 11 words in Exp. 2 (~9%). Future studies should consider presenting the memory tasks following a shorter endorsement period to determine whether an interaction for self-positive items exists when fewer items are held in memory, which would support the hypothesis that self-positive items receive prioritized processing in healthy adults.

Not only did trial presentation impact memory results, it also modulated ERPs during encoding. At the neural level, the blocked referent design resulted in an earlier onset of the referent effect, which also persisted for longer relative to the mixed referent design (400-800ms in Exp.1; 600-800ms in Exp.2, see Fig.6). Moreover, valence effects only occurred in the mixed referent design (600-1200ms, see Fig 6). The mixed referent design required constant referent updating at the beginning of each trial, requiring participants to remain vigilant throughout each block, knowing the referent could change at each trial. The LPP has been reported to be impacted by attention modulations (for review, see Schupp, Flaisch, Stockburger, & Junghofer., 2006; however, see Pastor et al., 2008 for no effects), as well as priming in lexical decision-making tasks (Brown, Hagoort, & Chwilla., 2000). The blocked referent design strongly primed participants with the referent cue, and participants could remain less vigilant across each block as they could process each item with a systematic referent cue. Consequently, the blocked referent design enhanced the neural response to referent during encoding due to strong priming, but potentially reduced valence processing due to overall reduced vigilance. In contrast, in the mixed referent design, participants had to remain more vigilant to each trial, as the referent cue could change at any time, resulting in a reduced priming effect by referent, but enhanced neural processing to valence due to overall increased vigilance.

The enhanced processing of positive words in the SRET is in contrast to the LPP literature with other, non-word and non-self-relevant stimuli, where the LPP is enhanced for negative faces (Schupp et al., 2004; Fruhholz, Jellinghaus, & Herrmann, 2011; Rellecke, Sommer, & Schacht, 2012) and unpleasant images (Schupp et al., 2003; Weinberg & Hajcak, 2010). This parallel line of literature suggests that self-relevant social information is processed uniquely, when compared to non-self visual information that may speak to a threat-based evolutionary theory. While the neural onsets of valence effects in Exp. 2 suggests valence was processed beginning around 600ms, it is important to note that we did not report an exhaustive measure of all electrodes and time points. It is possible that valence effects may be seen earlier than 600ms at different electrode sites. Additionally, this epoch did not extend to the full 1800ms after which point a behavioral response was made. Consequently, an interaction between referent and valence may be seen beyond 1200ms. For this reason, mass univariate analyses, which provide a more exhaustive approach to interpreting EEG data, would be beneficial in future studies. However, mass univariate toolboxes (i.e. the Hierarchical Linear Modeling of Electroencephalographic Data; LIMO; Pernet et al., 2015; and the Factorial Mass Univariate ERP Toolbox; fMUT; Fields, 2017) require considerably more trials per condition to establish reliable results, which as mentioned above likely impacts memory performance on the SRET. The importance of a full scalp analysis approach is highlighted in the differences seen in the topographic distributions between Exp. 1 and Exp. 2 (Figure 7). These heat maps showed differences between the experiments as to where the neural effects were strongest using the group grand averages. For example, panel A (Fig. 7) highlighted strong referent effects beginning around 400ms for Exp.1, which appeared to be maintained across the epoch. Panel B (Fig. 7) highlighted a similar temporal referent effect, however much weaker in Exp. 2. Further analyses of these topographic maps may provide further insight to the differences in neural encoding as a result of trial type. Regardless, the results across both studies illustrate distinguishable timelines in the processing of referent and valence with respect to incoming social information.

In contrast to previous reports of an enhanced P1 for positive relative to negative trait adjectives in the SRET (Auerbach et al., 2015; Auerbach et al., 2016) we did not find an effect of word valence on the P1 in either experiment. The lack of early differentiation for valence in the current set of studies highlights the controversial P1 modulations. This component is highly sensitive to low level features of a stimulus (Mangun, 1995), such as contrast and pixel intensity. Consequently, early neural differentiation found (or not found) on this component for either variable (referent or valence) may occur due to low level factors such as pixel intensity, contrast, or even the differences in word length and shape, rather than by attentional effects linked to the actual semantic differences between the words. Indeed, several other studies have found no effect of word valence on the P1 (Dainer-Best, Trujillo, Schnyer, & Beevers, 2017; Fields and Kuperberg, 2012; Watson, Dritschel, Obonsawin, & Jentzsch, 2007). Future replications with controlled stimuli are needed to address the valence ambiguities on this early neural component.

Critically, valence and referent did not interact for any ERP components up until 1200ms, suggesting that attention was captured and allocated preferentially to self-relevant (regardless of valence) and positively-valenced (regardless of referent) information. That is, there was no evidence, in either experiment, for privileged neural processing of self-positive information within this epoch during encoding. This is particularly relevant to the interpretation of past studies that have only tested the role of valence in the context of self-relevant information processing (e.g., Auerbach et al., 2015; Auerbach et al., 2016; Dainer-Best, Trujillo, Schnyer, & Beevers, 2017; Quevedo, et al., 2017; Zhang, Guan, Qi, & Yang, 2013). In the absence of an other-referent condition, the positivity bias cannot be interpreted as specific to the self. We found an interaction only in the recall and recognition scores in Exp. 2, which suggests self-positive information may undergo privileged processing, but this would be during later consolidation or retrieval, and not during initial encoding given our ERP findings.

Beyond trait-adjective processing, self-relevance reliably impacts the neural processing of faces (Schwartz et al., 2012; Wieser et al., 2014; McCrackin & Itier, 2018), as well as self-relevant sentences (Fields & Kuperberg, 2012; 2015). This general impact of referent on early cortical activity across various stimuli speaks to the fundamental role of self-other discrimination in the processing of information, a distinction that influences subsequent cognition and behavior. The salience of the self as measured using comparable SRET tasks is evident behaviorally by middle childhood with healthy children aged 8-16 demonstrating the same self-reference effect as adults in referent-blocked experimental designs (Burrows et al., 2017; Cunningham, Brebner, Quinn, & Turk, 2014; Henderson et al., 2009). This self-focused interpretative lens beginning in childhood may be central to the development of critical socio-cognitive skills over the course of the lifespan. Specifically, processing social information in relation to the self allows children to develop the ability to mentalize about others based on how they themselves would feel in a given situation (Ross & Nisbett, 1991). It will be important in future work to examine the cortical underpinnings of these processing biases in early childhood, using blocked and mixed designs with more trials, to run more robust mass univariate analyses across the entire epoch.

There were some limitations to the current set of studies. The first limitation concerns the choice of the character imagined in the Other condition. As only Harry Potter was used in the Other condition, traits specific to this character may have influenced participants' performance. While Harry Potter was selected based on previous studies (i.e. Burrows, Usher, Mundy & Henderson, 2016; Knyazev, 2013) who believed he was a neutral character that was not idolized as "good", it may be that Harry Potter is in fact not always seen as 'neutral'. Lombardo et al. (2007) had 30 participants rate Harry Potter on a likeability scale from 1 (not very likeable) to 6 (very likeable), with a mean rating of 4.07 (SD 1.34). On their scale, a "neutral" rating would be a 3, suggesting participants were skewed to rating Harry Potter as more likeable than a truly neutral character would be. Similarly, our endorsement data suggest that participants viewed Harry Potter in a positive light. Consequently, the general positivity bias seen particularly in the endorsement data, but also perhaps in the memory and ERP data, may have occurred due to Harry Potter being viewed as positively as participants viewed themselves. Shestyuk and Deldin (2010) found similar valence effects in their Other condition (Bill Clinton) on endorsement data, but no valence effects on recall data or ERPs. This suggests that healthy participants may be positively biased in their assessments of others who are not personally known to them, however uncontrolled factors of the protagonists may impact the depth of encoding. Future studies could include multiple "other" referents to test the generalizability of these effects beyond the specific character of Harry Potter, as well as examining how the reputation, and familiarity of the other character, impacts the SRET results. However, even if our choice of "other" is inherently positive, our results suggest that valence and referent information might be processed independently during encoding at the neural level, as each effect was distinguishable.

A second limitation concerns the choice of the words. While the positive and negative trait adjectives used in this study were matched on many attributes, positive words drawn from a previously validated word corpus were significantly more dominant (defined as the extent to which a word indicates something as being strong or weak; Warriner, Kuperman & Brysbaert, 2013) within each word list relative to the negative words. It may be that an effect of dominance could be in fact an effect of arousal, which can impact the LPP (Olofson et al., 2008; Speed and Hajcak 2018). It would be necessary to develop a larger corpus of standardized words to have perfectly matched word lists. The effect of dominance may have enhanced the positivity effect seen on the behavioural and physiological measures; however, since there was no consistent interaction with referent, we again propose that valence and referent were encoded independently.

In conclusion, the results of these studies highlight the central but possibly independent roles of referent and valence in guiding the encoding and prolonged processing of incoming social information.

These biases are invoked frequently in everyday life and their implications are numerous and broad in scope. We believe the existence of these biases provide adaptive lenses that support socio-cognitive development and emotional well-being across the lifespan. The results also suggest that a mixed referent design may be a more reliable, and ecologically-valid method for assessing self-referenced processing. The strategy participants use in a mixed referent design may be more similar to a real-world scenario, where individuals very rarely encode information about a single individual for a long period of time (as in the referent-blocked design). It is important that future research keep in mind the attentional differences of their experimental designs during the SRET, as design appears to impact behavioural and neural responses.

Funding

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant #418431), the Canada Foundation for Innovation (CFI, #213322), and the Canada Research Chair program (CRC #230407) awarded to RJI, and by the Social Science and Humanities Research Council (SSHERC Insight grant #435-2016-0494) awarded to HAH.

References

Auerbach, R. P., Bondy, E., Stanton, C. H., Webb, C. A., Shankman, S. A., & Pizzagalli, D. A. (2016). Self-referential processing in adolescents: Stability of behavioral and ERP markers. *Psychophysiology, 53*(9), 1398-1406. doi:10.1111/psyp.12686

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. doi: 10.1037/abn0000023

Brown, C. M., Hagoort, P., & Chwilla, D. J. (2000). An event-related brain potential analysis of visual word priming effects. Brain and Language, 72, 158-190. doi:10.1006/brln.1999.2284.

Burrows, C. A., Usher, L. V., Mundy, P. C., & Henderson, H. A. (2017). The salience of the self: Self-referential processing and internalizing problems in children and adolescents with autism spectrum disorder. A*utism Research, 10*(5), 949-960. doi:10.1002/aur.1727

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684. doi:10.1016/S0022-5371(72)80001-X

George, D., & Mallery, P, (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn and Bacon.

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. doi:10.1111/cdev.12144

Dainer-Best, J., Trujillo, L. T., Schnyer, D. M., & Beevers, C. G. (2017). Sustained engagement of attention is associated with increased negative self-referent processing in major depressive disorder. *Biological psychology*, *129*, 231-241.

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. doi:10.1016/j.jneumeth.2003.10.009

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. doi:10.1016/j.neuroimage.2012.05.003

Fields, E.C., & Kuperberg, G.R. (2015). Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Social Cognitive and Affective Neuroscience*, *10*(9), 1202-1209. https://doi.org/10.1093/scan/nsv004

Fields, E.C., & Kuperberg, G.R. (2019), Having your cake and eating it too: Flexibility and power with mass univariates statistics for ERP data. *Psychophysiology, 57*(2). doi: 10.1111/psyp.13468

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed. ed.). New York, NY: McGraw-Hill Book Company.

Foti, D. & Hajcak, G. (2008). Deconstructing reappraisal: descriptions preceding arousing pictures modulate the subsequent neural response. *Journal of Cognitive Neuroscience, 20*(6):977-88. doi: 10.1162/jocn.2008.20066.

Fruhholz, S., Jellinghaus, A., & Herrmann, M. (2011). Time course of implicit processing and explicit processing of emotional faces and emotional words. *Biological Psychology, 87*(2), 265-274. doi:10.1016/j.biopsycho.2011.03.008

Goldstein, B. L., Hayden, E. P., & Klein, D. N. (2015). Stability of self-referent encoding task performance and associations with change in depressive symptoms from early to middle childhood. *Cognition and Emotion, 29*(8), 1445. doi:10.1080/02699931.2014.990358

Gordon, R.M. (1986). Folk psychology as simulation. *Mind & Language*, *1*(2), 158-171. doi:10.1111/j.1468-0017.1986.tb00324.x

Henderson, H. A., & Mundy, P. C. (2013). The integration of self and other in the development of self-regulation: Typical and atypical processes. In K. Caplovitz Barrett, N.A. Fox, G.A. Morgan, D. Fidler, & L. Daunhauer (Eds.), *Handbook of Self-Regulatory Processes in Development: New Directions and International Perspectives* (pp. 113-134). New York: Routledge.

Henderson, H. A., Zahka, N. E., Kojkowski, N. M., Inge, A. P., Schwartz, C. B., Hileman, C. M., . . . Mundy, P. C. (2009). Self-referenced memory, social cognition, and symptom presentation in autism. *Journal of Child Psychology and Psychiatry*, *50*(7), 853-861. doi:10.1111/j.1469-7610.2008.02059.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. doi: 10.1080/17470919.2010.523543

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology, 45*(3), 487-498. doi:10.1111/j.1469-8986.2007.00638.x

Jackson, M. C., Wu, C.-Y., Linden, D. E. J., & Raymond, J. E. (2009). Enhanced visual short-term memory for angry faces. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 363–374. https://doi.org/10.1037/a0013895

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

Kuiper, N., & Derry, P. (1982). Depressed and nondepressed content self-reference in mild depressives. *Journal of Personality*, *50*(1), 67-80. doi:10.1111/j.1467-6494.1982.tb00746.x

Langeslag, S. J. E., & van Strien, J. W. (2009). Aging and emotional memory: The co-occurrence of neurophysiological and behavioral positivity effects. *Emotion, 9*(3), 369-377. doi:10.1037/a0015356

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *Plos One, 2*(9) doi:10.1371/journal.pone.0000883

Lto, T.A., Larsen, J.T., Smith K.N., & Cacioppo, J.T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*(4), 887-900. doi: 10.1037/0022-3514.75.4.887

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Science, 4*(2000) 432-440. doi: 10.1016/S1364-6613(00)01545-X

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.) Psychology Press. doi:10.4324/9781410611147.

Macrae, C. N., Moran J, M., Heatherton, T.,F., Banfield, J.,F., & Kelley, W.,M. (2004). Medial prefrontal activity predicts memory for self. Cerebral Cortex, 14(6), 647-654. doi:10.1093/cercor/bhh025

Maddi, S.R. (1989). *Personality theories: A comparative analysis (5th ed.)* Homewood, IL, US: Dorsey Press.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32*(1), 4-18. doi:10.1111/j.1469-8986.1995.tb03400.x

MATLAB and Statistical Toolbox Release 2014b, The MathWorks, Inc., Natick, Massachusetts, United States.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin, 130*(5), 711-747. doi:10.1037/0033-2909.130.5.711

Molnar-Szakacs, I., & Uddin, L. Q. (2013). Self-processing and the default mode network: Interactions with the mirror neuron system. *Frontiers in Human Neuroscience, 7* doi:10.3389/fnhum.2013.00571

Naumann, E., Bartussek, D., Diedrich, O., & Laufer, M. E. (1992). Assessing cognitive and affective information processing functions of the brain by means of the late positive complex of the event-related potential. *Journal of Psychophysiology, 6*(4), 285-298.

Olofsson, J.K., Nordin, S., Sequeira, H., and Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology, 77*(3), 247-265. doi: 10.1016/j.biopsycho.2007.11.006.

Pernet, C.R., Chauveau, N., Gaspar, C., and Rousselet, G.A. (2011). LIMO EEG: A toolbox for heriarchical linear modeling of electroencephalographic data. *Computational Intelligence and Neuroscience.* doi.org/10.1155/2011/831409

Pfeifer, J. H., Lieberman, M. D., & Dapretto, M. (2007). "I know you are but what am I?!": Neural bases of self- and social knowledge retrieval in children and adults. *Journal of Cognitive Neuroscience, 19*(8), 1323.

Quevedo, K., Ng, R., Scott, H., Smyda, G., Pfeifer, J. H., & Malone, S. (2017). Neurobiology of self-processing in abused depressed adolescents. *Development and Psychopathology*, *29*, 1057-1073. doi:10.1017/S0954579416001024

Rellecke, J., Sommer, W., & Schacht, A. (2012). Does processing of emotional facial expressions depend on intention? time-resolved evidence from event-related brain potentials. *Biological Psychology*, *90*(1), 23-32. doi:10.1016/j.biopsycho.2012.02.002

Righi, S., Marzi, T., Toscani, M., Baldassi, S., Ottonello, S., & Viggiano, M.P. (2010). Fearful expressions enhance recognition memory: Electrophysiological evidence. *Acta Pscyhological*, *139*(1), 7-18. doi.org/10.1016/j.actpsy.2011.09.015

Ross, L., & Nisbett, R.E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY, England: Mcgraw-Hill Book Company.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*(9), 677-688. doi:10.1037/0022-3514.35.9.677

Ruchkin, D.S.,Johnson,R., Mahaffey, D., & Sutton, S. (1988). Toward a functional categorization of slow waves. *Psychophysiology, 25*(3), 339-353. doi:10.1111/j.1469-8986.1988.tb01253.x

Schupp, H. T., Flaisch, T., Stockburger, J., & Jungho¨ fer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in Brain Research*, *156*, 31-51. doi: 10.1016/S0079-6123(06)56002-9.

Schupp, H. T., Jungho¨ fer, M., Weike, A., & Hamm, A. (2003). Emotional facilitation of sensory processing in the visual cortex. Psychological Science, 14, 7–13. doi: 10.1111/1467-9280.01411

Schupp, H. T., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., & Hamm, A. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion*, *4*(2), 189-200. doi:10.1037/1528-3542.4.2.189

Shestyuk, A. Y., & Deldin, P. J. (2010). Automatic and strategic representation of the self in major depression: Trait and state abnormalities. *Am.J.Psychiatry, 167*(5), 536-544. doi:10.1176/appi.ajp.2009.060919444

Speed and Hajcak. (2018). *Event-Related potentials and emotion dysregulation* (Theodore P. Beauchaine and Sheila E. Corwell, Ed).. The Oxford Handbook of Emotion Dysregulation. doi. 10.1093/oxfordhb/9780190689285.013.13

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

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371-394. doi:10.1037/0033-2909.121.3.371

Tavakol, M., and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55. doi: 10.5116/ijme.4dfb.8dfd

Thigpen, N., Kappenman, E., & Keli, A. (2017). Assessing the internal consistency of the event-related potential: An example analysis. *Psychophysiology*, *54*(1), 123-138. doi:10.111/psyp.12629.

Van Dam, T.N., & Earleywine, M. (2011). Validation of the Center for Epidemiologic Studies Depression Scale—Revised (CESD-R): Pragmatic depression assessment in the general population. *Psychiatry Research, 186*(1), 128-132. doi: https://doi.org/10.1016/j.psychres.2010.08.018

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior Research Methods, 45*(4), 1191-1207. doi:10.3758/s13428-012-0314-x.

Watson, L.A., Dritschel, B., Obonsawin, M.C., & Jentzsch, I. (2007). Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research*, 1152, 106-110. doi: 10.1016/j.brainres.2007.03.049

Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: The time-course of neural activity elicited by specific picture content. *Emotion, 10*(6), 767-782. doi: 10.1037/a0020242

Wieser, M. J., Gerdes, A. B. M., 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 15*(92), 74-82. doi:https://doi.org/10.1016/j.neuroimage.2014.01.022

Zhang, H., Guan, L., Qi M., & Yang, J. (2013). Self-esteem modulates the time course of self-positive bias in explicit self-evalutation. *PLoS ONE*, *8*(12). doi:https://doi.org/10.1371/journal.pone.0081169

Task		Length		Valence		Arousal		Dominance	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
List 1									
	Positive	6.56	1.37	7.48	0.36	4.77	0.91	6.77	0.47
	Negative	6.31	1.45	2.60	0.27	4.78	0.71	3.89	0.71
List 2									
	Positive	6.47	1.37	7.49	0.36	4.93	0.96	6.70	0.41
	Negative	5.87	1.45	2.60	0.28	4.87	0.76	4.11	0.53
List 3									
	Positive	6.75	1.24	7.48	0.35	4.71	0.97	6.66	0.48
	Negative	6.17	1.42	2.60	0.27	4.87	0.76	3.97	0.65

Table 1. Descriptive statistics for each word list. Length is reported in number of characters, while valence, arousal, and dominance reflect the normative ratings taken from Warriner, Kuperman, and Brysbaert (2013). Within each list, positive and negative adjectives did not differ in length (List 1: $t(62) = 0.71$, $p = .480$; List 2: $t(62) = 1.68$, $p = .097$; List 3: $t(62) = 1.31$, $p = .195$) or in arousal (List 1: $t(62) = -0.07$, $p = .949$; List 2: $t(62) = 0.14$, $p = .890$; List 3: $t(62) = -0.71$, $p = .484$). For each list, positive adjectives were higher in dominance (List 1: $t(62) = -19.00$, $p < .001$; List 2: $t(62) = 21.88$, $p < .001$; List 3: $t(62) = 18.90$, $p < .001$).

Task		Self		Other	
		Mean	SD	Mean	SD
Endorsement					
	Positive	23.35	4.33	22.04	4.75
	Negative	4.12	2.64	3.88	3.15
Recall					
	Positive	4.41	2.33	3.96	2.08
	Negative	2.26	1.68	2.11	1.65
Recognition					
	Positive	22.10	5.67	19.28	5.08
	Negative	19.83	5.85	17.38	5.01
Memory sensitivity (*d'*)					
	Positive	1.82	.564	1.53	.557
	Negative	1.95	.577	1.70	.537

Table 2. Means and standard deviations for each variable of interest and each of the four conditions in Exp. 1 (blocked design; self-positive, self-negative, other-positive, other-negative). The maximum number of words in each condition was 32.

		Condition							
		Self-Positive		Self-Negative		Other-Positive		Other-Negative	
ERP		Mean	SD	Mean	SD	Mean	SD	Mean	SD
P1									
	100-170ms	.700	1.90	.671	1.90	.301	1.75	.322	1.64
eLPP									
	400-600ms	1.60	2.19	1.46	2.12	.476	1.65	.353	2.02
lLPP									
	600-800ms	1.66	2.05	.994	2.24	.270	2.10	-.217	2.30
	800-1000ms	1.32	2.18	.655	2.41	.441	2.59	.084	2.49
	1000-1200ms	.755	2.40	.229	2.53	.60	2.96	-.280	2.80

Table 3. Mean amplitudes (µV) and Standard Deviations averaged across parieto-occipital sites (P1, P2, Pz, POz, PO3, PO4) for the P1 and across the Frontal Central Midline (FCM; Fz, FCz, Cz) for the eLPP and lLPP, for each of the four conditions of Exp.1 (block design).

Cronbach's Alpha			
		Exp. 1	**Exp. 2**
	Time		
eLPP	400-600ms	.873	.914
lLPP	600-800ms	.819	.893
	800-1000ms	.848	.890
	1000-1200ms	.841	.931

Table 4. Cronbach's Alpha for each of the ERP components analyzed, using the four conditions as "items", in Exp. 1 (A) and Exp. 2 (B).

Task		Self		Other	
		Mean	SD	Mean	SD
Endorsement					
	Positive	24.79	4.80	23.21	4.82
	Negative	3.54	3.84	5.89	3.37
Recall					
	Positive	4.20	2.0	2.85	1.93
	Negative	2.62	1.80	2.23	1.50
Recognition					
	Positive	21.18	6.40	17.00	6.32
	Negative	18.89	6.72	17.12	6.09
Memory sensitivity (*d'*)					
	Positive	1.48	.591	1.10	.600
	Negative	1.37	.625	1.25	.452

Table 5. Means and standard deviations for each variable of interest and each of the four conditions in Exp. 2 (mixed design; self-positive, self-negative, other-positive, other-negative) in the mixed SRET design. The maximum number of words in each condition was 32.

		Condition							
		Self-Positive		Self-Negative		Other-Positive		Other-Negative	
ERP		Mean	SD	Mean	SD	Mean	SD	Mean	SD
P1									
	90-170ms	.060	1.43	.353	.995	.821	.996	.380	1.09
eLPP									
	400-600ms	1.84	1.86	1.49	2.18	1.40	2.39	.816	1.92
lLPP									
	600-800ms	1.02	1.60	.572	1.83	.360	2.22	-.549	1.86
	800-1000ms	.382	1.76	-.220	2.40	.044	2.11	-1.00	1.53
	1000-1200ms	-.323	1.85	-.830	2.49	-.422	2.20	-1.58	1.78

Table 6. Mean amplitudes (μV) and Standard Deviations averaged across parieto-occipital sites (P1, P2, Pz, POz, PO3, PO4) for the P1 and across the Frontal Central Midline (FCM; Fz, FCz, Cz) for the eLPP and lLPP, for each of the four conditions in Exp.2 (mixed design).

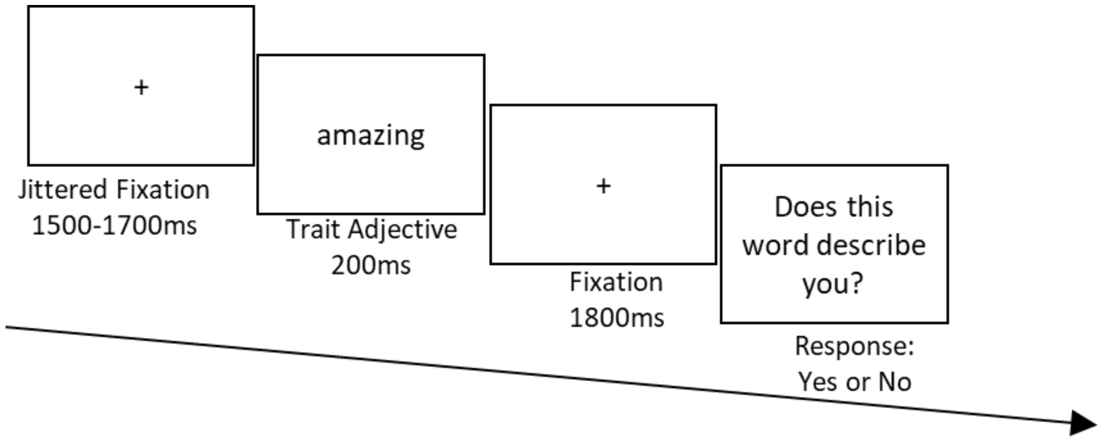

Figure 2. Exp. 1 trial example demonstrating a positive trait adjective in the self-relevant condition. Participants completed two blocked conditions (Self, Other) and each condition included 64 trials (32 positive, 32 negative) with the following progression. First, a jittered fixation cross was presented on

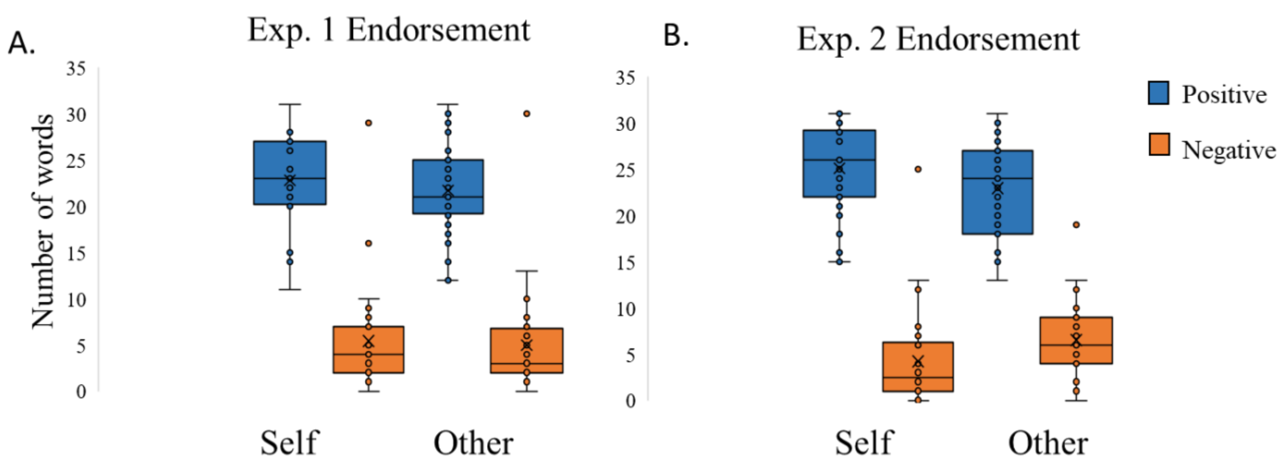

Figure 2. Endorsement rates for Exp. 1 (N = 28; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* Inter Quartile Range (IQR; minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle. The maximum number of words in each condition was 32. ***Use Color in print**

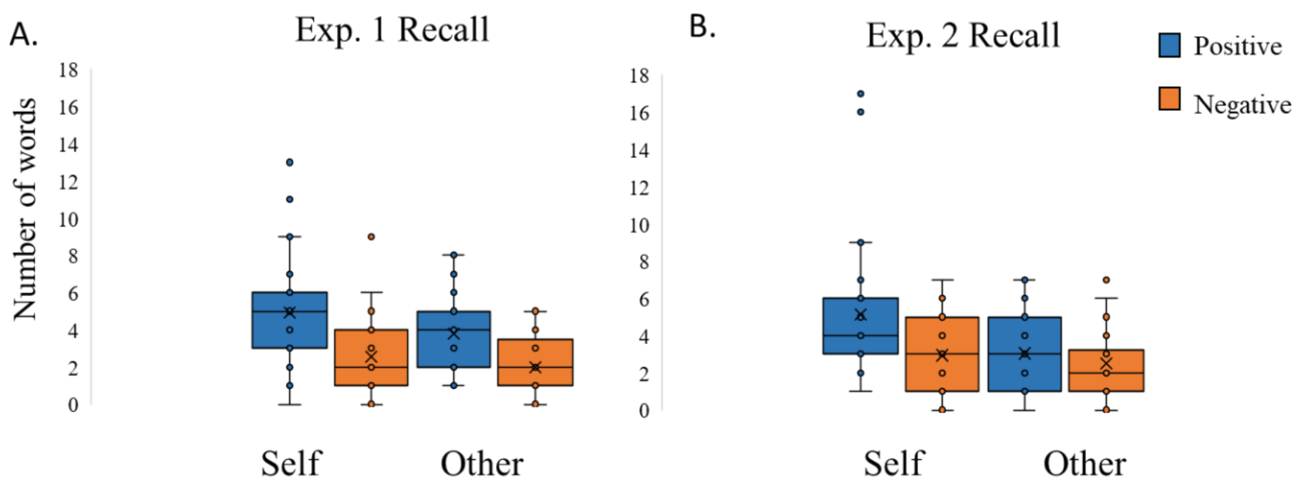

Figure 3. Recall rates for Exp. 1 (N = 29; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* IQR (minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle. The maximum number of words in each condition was 32. ***Use Color in print**

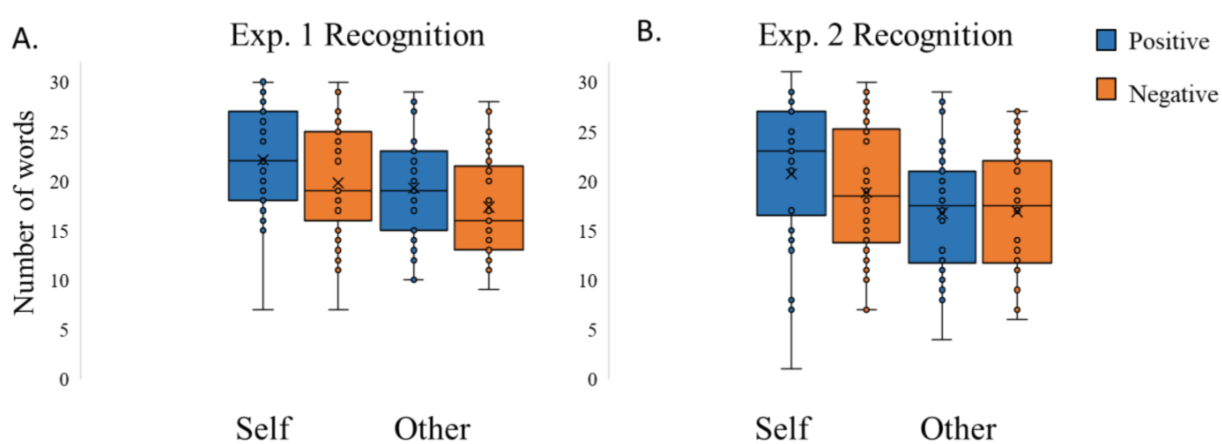

Figure 4. Recognition rates for Exp. 1 (N = 29; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* IQR (minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle. ***Use Color in print**

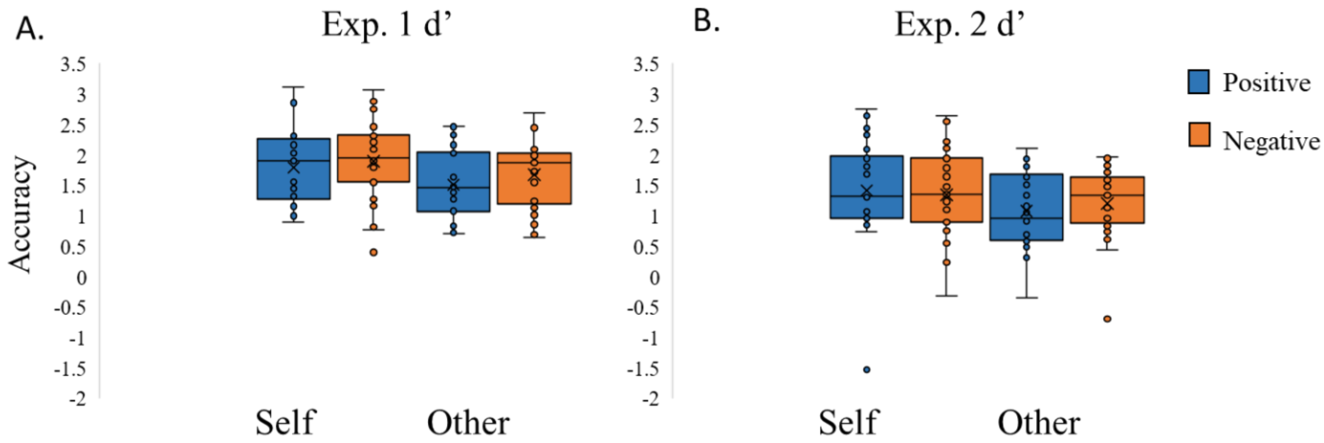

Figure 5. D Prime (accuracy) score for Exp. 1 (N = 29; Panel A) and Exp. 2 (N = 29; Panel B). The 25th and 75th percentile are represented by the upper and lower limits of each box plot, with the median indicated by a central horizontal line. An "X" marks each Mean within the boxplot. Minimum and maximum values on the lower and upper whiskers are calculated as Q1-1.5* IQR (minimum) and Q3+1.5*IQR (maximum) with potential extreme values marked with an external circle. ***Use Color in print**

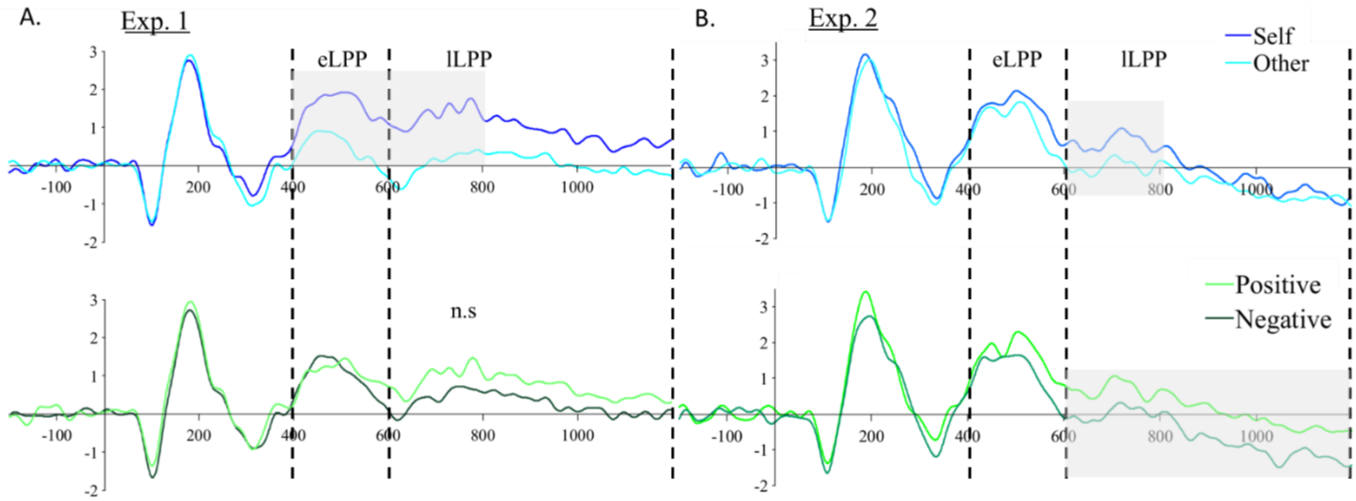

Figure 6. The entire 1200ms ERP waveform averaged across electrodes Fz, FCz, and Cz to create the Frontal Central Midline (FCM). ERPs were time locked to the onset of word presentation during the encoding phase of the task. The significant effects of referent (top) and valence (bottom) for Exp. 1 (Panel A) and Exp. 2 (Panel B) are highlighted in grey. ***Use Color in print**

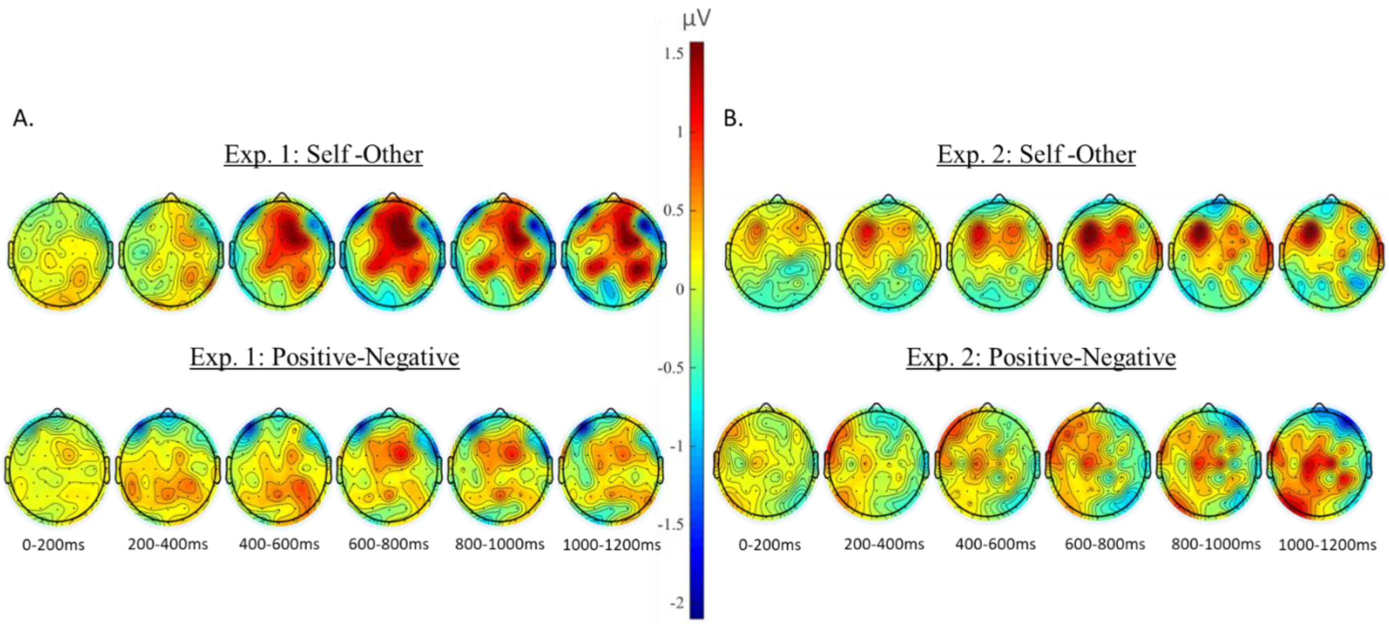

Figure 7. Topographic maps across the 1200ms waveform. Average Self-Other differences are displayed across the top for both Exp. 1 (Panel A) and Exp. 2 (Panel B). Warmer tones reflect more positive ERP amplitudes for Self-relevant trials relative to Other-relevant trial types. Maps across the bottom reflect average Positive-Negative differences, with warmer tones reflecting more positive amplitudes for Positive trials relative to Negative trial types. ***Use Color in print**

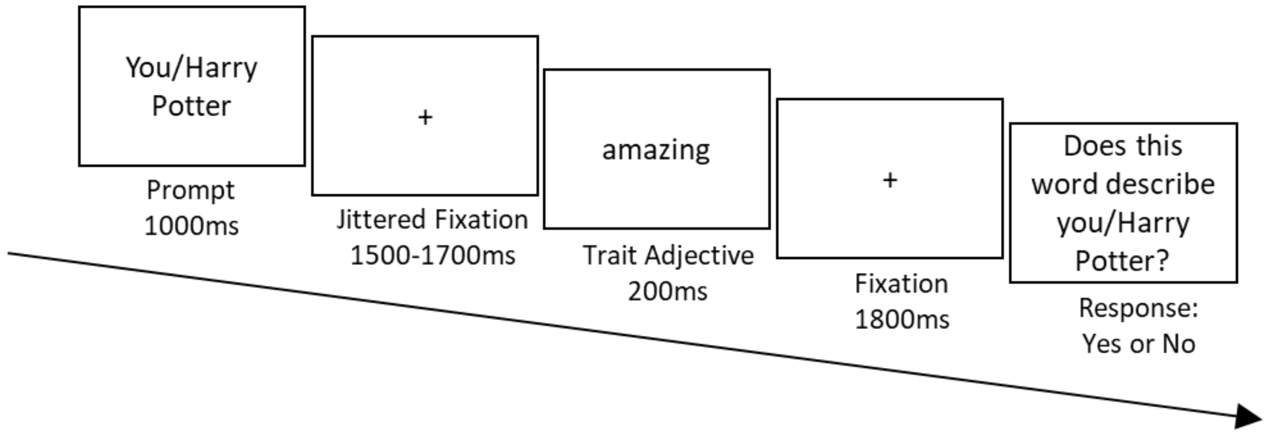

Figure 8. Exp. 2 trial sample in which each trial begins with a prompt indicating who the participant should think about for the remainder of the trial.