



Citation for published version:

Sui, J & Humphreys, GW 2017, 'The self survives extinction: Self-association biases attention in patients with visual extinction', *Cortex*, vol. 95, pp. 248-256. <https://doi.org/10.1016/j.cortex.2017.08.006>

DOI:

[10.1016/j.cortex.2017.08.006](https://doi.org/10.1016/j.cortex.2017.08.006)

Publication date:

2017

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

CC BY-NC-ND

University of Bath

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

The self survives extinction

Running head: The self survives extinction

**The self survives extinction: Self-association biases attention in patients with
visual extinction**

Jie Sui

Department of Experimental Psychology, University of Oxford, United Kingdom,

Department of Psychology, University of Bath, United Kingdom

Email: j.sui@bath.ac.uk

Glyn W. Humphreys

Department of Experimental Psychology, University of Oxford, United Kingdom

Abstract

People show biases to self-related information on a range of tasks. Key but controversial questions are whether self-related information is processed without attention, and whether self-related information determines what is attended. We examined this using patients showing visual extinction. We had patients associated shapes with themselves or their best friend prior to carrying out a shape identification task. We demonstrate that extinction was modulated by whether patients associated stimuli with themselves or their best friend. Notably, patients were biased to identify their own shape relative to the shape associated with their friend, when the two shapes were placed in competition. This occurred even when the self-associated shape fell in the contralesional field. The data indicate that self-relatedness can be computed pre-attentively and can cue attention to regions of space that would otherwise be ignored by neuropsychological patients.

Keywords: extinction; self; personal association

Introduction

Human cognition is biased towards self-related information in comparison with information relating to other people (Conway & Pleydell-Pearce, 2000; Sui, He & Humphreys, 2012), but it is controversial whether such biases operate without attention (Keyes & Dlugokencka, 2014; Humphreys & Sui, 2015a). For example in a classic study, Moray (1959) showed that participants tended to notice their own name more than other names when the stimuli were presented to an unattended ear, suggesting that self-reference is computed without attention. On the other hand, momentary shifts of attention to the unattended side could support the better identification of self-related items (Lachter, Forster, & Ruthruff, 2004). Stronger evidence would accrue if self-reference affected performance for stimuli that participants were otherwise unaware of. We assessed this by examining whether self-related information modulates extinction in neuropsychological patients.

Extinction patients can respond to a single stimulus on their affected side but fail to report the same item if another stimulus appears at the same time on the unimpaired side (Karnath, 1988; Driver & Vuilleumier, 2001). This can be attributed to the brain lesion biasing attention so that contralesional stimuli lose the competition for selection (Duncan, Humphreys, & Ward, 1997; Vuilleumier & Rafal, 2000). Although patients are typically unable to report the extinguished stimulus, there is evidence of stimulus processing – for example extinction reduces when contra- and ipsilesional stimuli group (Humphreys, 1998), when the stimuli are pictured to interact with one

The self survives extinction

another (Riddoch et al., 2003), when they have a common exemplar identity (Berti et al., 1992) and extinguished stimuli can still receive some residual on-line processing in order to enhance implicit memory (Vuilleumier et al., 2002a). Indeed when a salient object is on the contralesional side, extinction can even reverse so that patients report the contralesional not the ipsilesional stimulus (Riddoch et al., 2003).

Extinction in patients can then be used to examine whether contralesional stimuli are processed pre-attentively, indexed by the degree of extinction to these stimuli.

We evaluated self-reference effects by having patients associate themselves or a best friend with a shape and then asking patients to identify the shape(s) when they were presented as single items or pairs (Sui, Yankouskaya, & Humphreys, 2015a). Previous studies have shown that this self-association procedure induces strong biases for self-related items, even when factors such as the frequency, length and concreteness of the words has been controlled for, even when the self is pitted against a highly familiar other person (best friend, mother), and even when participants are just presented with the (formerly) neutral shapes (Sui et al., 2012). The effect is correlated with the psychological distance individuals feel to the people used for the other associations (Sui & Humphreys, 2015) and it shows stable trait-like properties in participants (Humphreys & Sui, 2015a and b). We assessed whether self-association to a shape modulates whether it is extinguished or consciously reported in a task requiring participants to judge whether a shape is associated to themselves, to their best friend or new. We tested performance in six patients, three of them showing right-side extinction after left-hemisphere damage and three of them showing left-side

The self survives extinction

extinction after right-hemisphere damage. Although left-side extinction after right hemisphere damage can be more severe than right-side extinction after left hemisphere lesion (Chechlacz et al., 2014), cases of right-side extinction are still prevalent. Furthermore there is evidence that self-biases may be associated with right hemisphere processing, at least with face stimuli (Keenan et al., 2001). By testing patients with either right or left hemisphere lesion here we sought to demonstrate a general effect of self-relatedness on extinction, to provide a proof-of-principle test for self-bias modulating pre-attentive processing. To optimise the effects, the patients who were selected also demonstrated self-bias effects on basic perceptual matching, a defining case to produce self-bias in subsequent shape identification (cf. Sui et al., 2012). To foreshadow the results we found that both self and friend-associated shapes were better reported than new shapes but, when the self-shape was paired with the shape for the friend, patients showed extinction of the friend shape and reported the self shape. This result occurred even when the self shape fell in the contralesional field, reversing the standard pattern of spatial extinction. The data indicate that self-related stimuli can be processed pre-attentively and modulate the subsequent allocation of attention.

Materials and methods

We first measured self-bias using a perceptual matching procedure introduced by Sui et al. (2012), to ensure that the patients were sensitive to self-related information. We subsequently evaluated whether shapes associated with the self or a best friend were

The self survives extinction

subject to extinction, when the patients were presented with shape pairs rather than single shapes.

Patients. The six patients (RR, PH, MH, SW, JB, and DT) were selected from the panel of neuropsychological volunteers at the Cognitive Neuropsychology Centre (CNC), University of Oxford. Patients were selected from (i) a continuous series in neuropsychological patients coming into the CNC (ii) whether they had extinction in the right or left visual field based on their unilateral lesion and (iii) whether they showed self-biases in perceptual matching. RR and PH had acquired left-hemisphere brain lesions following a stroke, and MH had suffered carbon monoxide poisoning. SW, JB and DT had acquired right-hemisphere brain lesions following a stroke. All patients were at a chronic stage (>12 months post-injury) and provided written informed consent in agreement with ethics protocols at the CNC. RR had lesions including left parietal, temporal cortex extending into inferior frontal and insula. PH's lesion extended across the left inferior frontal, parietal (angular gyrus) and superior temporal cortices, extending into the left caudate and adjacent subcortical regions. The main overlap between RR and PH's lesion was the left inferior parietal cortex extending to subcortical regions. MH had a grey matter lesion including the left parietal cortex, with an overlap with RR and PH in white matter underlying the parietal cortex. SW had lesions including the right parietal, supramarginal, temporal, and occipital cortex extending to hippocampus. JB had lesion including the right insular, putamen inferior frontal cortex extending to the parietal operculum and temporal cortex and subcortical regions. DT had lesions across the right parietal,

The self survives extinction

occipital, and temporal cortex extending to the middle frontal gyrus and subcortical regions and a lesion in the left inferior frontal gyrus and insula. The main overlap between SW, JB and DT's brain lesion was the right parietal cortex which to some degree was mirrored to the overlay with RR and PH (Figure 1a).

All patients had undertaken the Oxford Cognitive Screening (OCS) test (Demeyere et al., 2015; online available, <http://isis-innovation.com/outcome-measures/the-oxford-cognitive-screen-ocs/>) and the Birmingham Cognitive Screen (BCoS) (Bickerton et al., 2014; Humphreys et al., 2012) (online available, <http://www.bcos.bham.ac.uk/>) to provide a background neuropsychological profile (for details see Table 1).

Healthy controls for self-bias measure. Forty healthy controls (20 males and 20 females, mean of age \pm standard deviation = 39.95 ± 18.83 years, range 19-70) with no history of stroke, brain damage or neurological disorders were recruited as healthy controls to measure the magnitude of self-bias in the normal population and to provide cut-off scores for the patients.

The matching and extinction tasks. Each patient completed 4 sessions. In each session, the matching task was followed by the extinction task. Three geometric shapes (triangle, circle, and square) were selected. Patients learned that 2 shapes were associated with two labels representing two people (the patient self and their named best friend; the third shape as a neutral stimulus was used in the extinction task, see below). For example, patients were told that “the triangle represents you, the square

The self survives extinction

represents your best friend” (Sui et al., 2012). Following the association instruction, patients immediately conducted a shape-label matching task to judge whether a shape-label pair matched. A shape (covering $3.0^\circ \times 3.0^\circ$ of visual angle) was displayed above a white central fixation cross ($0.8^\circ \times 0.8^\circ$ visual angle). One of two labels (‘You’ or ‘Friend, $1.7^\circ/2.30^\circ \times 1.76^\circ$ of visual angle) was displayed below the white fixation cross. Each trial started a central fixation cross for 500 ms, followed by a shape-label pair at the center of the screen for 100 ms for RR and PH and 500 ms for MH. Half the pairings of the shape and label conformed to the instruction and were responded to as match trials; on the remaining trials the shapes and labels were re-paired to form mismatch trials. For mismatch trials, a shape was paired with each of the two possible labels (e.g., self shape with either the friend label). The next display was a 3000 ms blank field. Patients were encouraged to make a response as quickly and accurately as possible within this 3000 ms interval. A feedback message (correct or incorrect) was then given in the center of the screen for 500 ms. Patients were also informed of their overall accuracy at the end of each block. There were three blocks of 40 trials following 12 practice trials. Thus there were 30 trials for each match and mismatch condition.

After the matching task, patients immediately undertook the extinction task, where a learned and/or a new shape of $3.0^\circ \times 3.0^\circ$ appeared in the left and/or right visual field with white central fixation cross of $0.8^\circ \times 0.8^\circ$. The outer edge of shape to the center of the screen was 7.8° . There were three types of single item – a self-associated shape, friend-associated shape, and a new shape, where they equally

The self survives extinction

displayed in the ipsilesional and contralesional visual field. There were also three types of two-item pairs (self-neutral, friend-neutral, and self-friend) where they equally displayed either in the left-right and right-left visual fields. The patients had to orally report how many shape(s) they saw and who the shape(s) represented (themselves, their friend, a new shape, or they were aware of the presence of a stimulus even if they could not identify it). Each type of single item trial was repeated three times in a block and each type of two-item trial was repeated six times in a block. Each session included a minimum of three blocks which created 54 single item trials in 6 conditions (3 shape: self, friend, or neutral by 2 visual field: ipsilesional vs. contralesional), 9 trials per condition, and 108 two-item trials in 6 conditions (3 pairs [self-neutral, friend-neutral, or self-friend] by 2 visual fields [the more familiar stimulus on the ipsilesional or contralesional side]). In total across sessions, there were in 36 trials for each type of single item condition and 72 trials for each type of two-item trial. The exposure durations differed across patients with the aim of achieving approximately 90% correct identification of single items on the contralesional side (50 ms for RR, 100 ms for PH, SW, JB and DT, and 200 ms for MH), which was tested in a pilot session. Each trial was preceded by a fixation cross for 1000 ms. Trials were self-paced. Each session took about one hour.

All the stimuli were displayed on a grey background. E-prime software (version 2.0) was used to present the stimuli and to record responses. The experiment was run on a PC with a 22-in monitor (1920 × 1080 pixels) at 60 Hz. There was identical assignment of shapes to the self and friend in the first two sessions, then the

The self survives extinction

assignment of shapes and persons was swapped in the other two sessions to rule out the effect of shapes. By using geometric shape stimuli, equated for self and friend in the associative learning task, we equated the effects of familiarity for the self and friend stimuli. The interval of two sessions took place over one-four weeks.

The cutoff score in self-bias. The self-bias effect for the matching task in healthy controls was used to create cutoff scores for the patients (Humphreys & Sui, in press). Self-bias was indexed by the relative difference in the reaction time between the friend and the self multiplied by 100, divided by the sum of the two conditions. The controls showed reliable self-advantage effects, consistent with evidence from prior studies (Humphreys & Sui, 2015a, b). The self-bias scores were used to define the self-bias measure for individual patients, which was defined by subtracting the mean for the controls from that of each patient and dividing by the standard deviation for the controls. The cut-off to classify patients as impaired was based on them having a mean level of self-advantage effect either less or more than 2.5 SDs from the control mean (defining respectively a hypo- or a hyper-SFP deficit) (Sui et al., 2013b, 2015b).

Extinction. There was a two-response stage in the extinction task. The patients were asked to first report the number of items present (detection task) and then they were asked to identify the shapes (identification task). All six patients showed high performance in reporting the number of items (the range of correct responses in all conditions, RR: 98%-99%, MH: 97-99%, PH: 96%-97%, SW: 95-97%, JB: 97-99%, DT: 96-99%), suggesting that they showed relatively mild extinction. As a

The self survives extinction

consequence we report identification performance only in the Results section based on the contralesional and ipsilesional visual field respectively.

Assessments of attentional-related tasks. In addition to the experimental tasks we also measured patients' attentional abilities using standardised neuropsychological tests of spatial and selective attention (the Apples cancellation and the Auditory Attention task from the BCoS battery; Bickerton et al., 2014; Humphreys et al., 2012, Table 1). For the Apples task the patient had to cancel drawings of complete apples while ignoring distractor apples that had a section missing on either the left or right side. Targets and distractors were randomly positioned on the sheet with the proviso that there were equal numbers of each type of item within each five columns across the page. The test provides separate measures of 'egocentric' neglect (where items are missed according to their position on the sheet) and 'allocentric neglect' (based on false positive responses to distractors to whether participants respond incorrectly to distractors with a gap on the left or right of individual stimuli irrespective of the positions of the stimuli on the sheet). For the selective attention task, patients were instructed to respond to 3 high frequency, auditorily-presented target words (no, goodbye, please) and to ignore (not respond to) 3 high frequency, related distractors (yes, hello, thanks). The index for the selective attention was the ability to respond to the targets while ignoring the distractors. Cutoff scores for these tasks were extracted from 100 age-matched non-lesioned healthy controls. The six patients had no neglect deficits, and only PH showed a deficit in the selective attention test (Table 1).

Results

Perceptual matching. We evaluated matching and shape identification performance in the six neuropsychological patients (RR, PH, MH, SW, JB, and DT). RR, PH, SW and JB showed a hyper self-bias effect (greater self-bias than the control participants); MH and DT showed a normal self-bias effect, compared to the controls (Figure 1b). The larger self-bias effects, in RR and PH compared with the controls, may reflect a lack executive control in these patients (Sui et al., 2015b). Irrespective of this, the results demonstrate that the patients were sensitive to self-reference in perceptual matching.

Extinction was assessed by comparing identification performance on one-item and two-item trials, averaging across the two-item conditions with a neutral shape, using a mixed-design ANOVA. The number of items (one vs. two) and shape (self/self-neutral vs. friend/friend-neutral) were within-subjects factors and patient was a between-subjects factor (to demonstrate generalization across patients). Each test session was treated as a participant nested within the patient factor. There was a significant main effect of item number, $F(1,11) = 316.13, p < .001, \eta^2 = .97$; responses were more accurate on one than two items trials – an extinction effect. There was also a significant interaction between item number and patient, $F(5,11) = 10.75, p = .001, \eta^2 = .83$. The degree of extinction effect varied across patients but was reliable in all cases ($\chi^2 > 4.55, p < .02$ for patient JB, who showed the weakest effect; **Figure 1c**). There was no main effect of shape (self vs. friend) and no interactions involving this

The self survives extinction

factor ($F < 0.75$, $p > .41$). Overall there were no differences in reporting self and friend shapes when they were presented alone.

We also conducted an ANOVA on the accuracy of performance on two-item trials with stimulus pair (self-neutral, friend-neutral, or self-friend) as a within-subjects factor and patient as a between-subjects factor. Neither the main effect of stimulus pair nor the interaction between stimulus pairs and patients was significant, $F < 0.67$, $p > .74$. The results indicate that there was no overall difference in correct responses across the three types of pairs. The overall probability of correct responses to the three types of two-item trials is illustrated in **Figure 2**.

Incorrect responses to self-neutral and friend-neutral pairs were evaluated by comparing the probability of correctly reporting self or friend vs. neutral shapes, on trials where only one item was identified. Stimulus pair (self-neutral vs. friend-neutral) and shape (self or friend vs. neutral) were within-subjects factors and patient a between-subjects factor. There were no effects on report of the ipsilesional item, $F < 3.35$, $p > .10$ (**Figure 3a**). In contrast, there was a significant effect of shape on reporting contralesional stimuli, $F(1,11) = 5.09$, $p < .02$, $\eta^2 = .32$; patients reported more self and friend shapes than neutral shapes (**Figure 3b**). There was no difference between self and friend shapes ($F < 0.11$). This bias, to report the self and friend shapes over the neutral shape, could reflect the personal familiarity of the self and friend or it could reflect the effect of familiarising participants to these stimuli (in the matching task). Given that there was no effect of shape on the report of an ipsilesional item, the

The self survives extinction

effect on contralesional stimuli suggests that, for items on that side, attention was more likely to be attracted if the item was personally known or more familiar.

Incorrect responses to self-friend pairs were analyzed when participants correctly reported just one self or friend shape (i.e., on trials where extinction occurred). There were two within-subjects factors, shape (self vs. friend) and stimulus pair (field of stimulus: self-friend vs. friend-self) and patient was a between-subjects factor.

Patients reported more self than friend shapes, $F(1,11) = 9.53, p = .01, \eta^2 = .54$. There was also a significant interaction between shape and stimulus pair, $F(1,11) = 6.47, p = .02, \eta^2 = .37$ (**Figure 3c**) but no interaction involving patient. Self-associated shapes were reported more than friend-associated shapes on the contralesional side, $t(16) = 3.33, p = .004$. There was a weak effect of shape on the report of ipsilesional stimuli, $t(16) = 2.81, p = .01$.

When self and friend stimuli competed for selection there was a bias to report the contralesional item when it was a self-related shape compared with when it was associated with a friend. In contrast, there were no differences in the probability of reporting a single ipsilesional self- or friend-related shape on two-item trials. In the latter case it can be argued that items in the ipsilesional (attended) field tended to attract attention irrespective of whether they were associated with the self or a friend. In contrast, self-related items on the contralesional side were more likely to attract attention than friend-related items on that side, and hence were reported more often on extinction trials. The effect of self is unlikely to reflect the ipsilesional stimulus.

The self survives extinction

Notably, there were no significant differences in reporting self- and friend-related items in the ipsilesional visual field, both in the self-friend conditions and the self-neutral and friend-neutral conditions (see above).

Discussion

The data indicate a self bias effect on whether self- or friend-associated shapes are extinguished, when placed in competition for report. When extinction occurred on stimulus identification, self shapes were more likely to be reported than friend shapes, when the shapes appeared in the contralesional field. Note that, on extinction trials where the contralesional item was reported, the patients were unable to make a response to the ipsilesional stimulus; that is, the standard pattern of spatial extinction (ipsilesional item > contralesional item) was reversed.

These results on self-bias are complemented by a more general benefit for self and friend shapes over neutral shapes, and the magnitude of the benefit did not differ. In this case, the benefits for both self and friend shapes may reflect several factors – for example, the personal relevance of these stimuli, the fact that only these stimuli had specific identities associated with them or the familiarity gained in the initial matching task. Each of these factors could have increased the saliency of the self and friend shapes. In comparisons of report relative to the neutral shape, a self advantage may not emerge over the friend condition because both the friend- and the self-related shapes had sufficient saliency to sometimes be selected instead of the neutral item on the ipsilesional side. It is only when the self and friend shapes are placed in

The self survives extinction

competition that the self advantage emerges, when the relative increase in saliency for the self shape (over the friend) is critical. The drawing of attention to the contralesional side, by a self-related shape, can then lead to patients failing to attend to the ipsilesional item, generating extinction of the ipsilesional (e.g., friend-related) stimulus.

The results indicate that both general personal relevance/specific identity/familiarity (for friend and self stimuli) and self-specific reference can be computed pre-attentively, even when a patient carries a chronic attentional bias to the ipsilesional field. Both factors can then modify whether a contralesional stimulus that would otherwise be subject to extinction can attract attention and be reported. Notably, though, the pre-attentive computation of self-relevance, and the subsequent biasing of attention (Moray, 1959), operate over and above more general effects of personal relevance, association to a specific identity and/or stimulus familiarity, found with friend-related stimuli. The data indicate that the pre-attentive computation of self-relevance can provide an important cue for visual selection, biasing attention to items that are of significant social importance to the individual.

The exact factors driving this self advantage effect are not set out here. For example, possibilities are that self-related stimuli carry a higher intrinsic reward value than other stimuli, and that this reward value attracts attention (see Northoff & Hayes, 2011 for an argument in relation to the self; see Anderson, Laurent, & Yantis, 2011 for evidence on attentional capture by reward). On the other hand, in other work we

The self survives extinction

have shown some dissociations between self-bias effects and the effects of reward (Sui & Humphreys, 2015; Sui, Yankouskaya & Humphreys, 2015a), arguing against reward being critical. Another possibility is that self-associated items have a high positive valence, and this leads to attention being directed to the self-related shape (Stolte et al., 2017). Previous work indicates that emotional stimuli can reduce visual neglect and extinction on the contralesional side (e.g. Domínguez-Borràs, et al., 2012; Tamietto et al., 2015; Vuilleumier et al., 2002b; Tamietto et al., 2015; Vuilleumier & Schwartz, 2001). Ma and Han (2010) reported that the self-advantage effect in face perception was reduced when healthy participants were instructed to evaluate negative personality traits in relation to themselves compared with neutral conditions. We (Sui, Ohrling, & Humphreys, 2016) have found that self-biases in perceptual matching with neutral shapes reduce when a low mood is induced. A question for future work is whether such manipulations of the emotional valence of stimuli may modulate the advantage for self-stimuli under extinction conditions.

Alternatively, the self-advantage effect may reflect enhanced visual awareness for self-related shapes after personal tagging and this facilitates the recovery of extinguished items. Previous neuropsychological studies have shown that there is a dissociation between processing of sensory input and visual awareness in patients with unilateral neglect (Vallar, 1998). For example, Berti and Rizzolatti (1998) had patients discriminate targets following primes in a categorization task in which there were different pairs of primes and targets (congruent vs. incongruent trials). The researchers found that the responses to targets in the contralesional visual field were

The self survives extinction

facilitated following primes in the congruent condition compared to incongruent condition. The data indicate that neglect patients are able to process stimuli presented in the impaired visual field although they are unaware of them. The dissociation between visual awareness and processing of sensory input has also been reported in other types of patients, such as dyslexia and anosognosia (Heilman, Barrett, & Adair, 1998; Warrington & Shallice, 1979). The neuropsychological evidence has additionally been supported by neuroimaging studies showing that sensory processing and visual awareness are associated with different neural patterns of activity, but they interact in terms of the levels of self-awareness (Goldberg, Harel, & Malach, 2006). The present study tested the relationship between visual awareness and stimulus processing by manipulating self-awareness using the self-association task, as it has been reported that self-relevance prioritizes access to visual awareness (Macrae, Visokomogilski, Cunningham, & Sahraie, 2017; but the effect is subject to tasks where identification is required, see Stein, Siebold, & van Zoest, 2016). Here the patients with extinction reported more self-related shapes in the impaired visual field than friend-related shapes, indicating privileged awareness towards self-related shapes after shape-personal label tagging.

Prior work has shown that the self-biases in matching are associated with a specific neural circuit involving the ventromedial prefrontal cortex (vmPFC) and the left posterior superior temporal sulcus (LpSTS). The vmPFC is associated with internal self-representation (Northoff & Bermpohl, 2004; Sui & Humphreys, 2015; Sui, 2016) and the LpSTS is thought to reflect the social salience of external stimuli

The self survives extinction

(Sui, Rotshtein, & Humphreys, 2013a). The strength of coupling between these two regions mediates self-bias in behaviour. In addition, the processing other-related stimuli (e.g., associated to a friend or a stranger) recruited the dorsal executive control network (Corbetta & Shulman, 2002), and there was negative correlation in activity between the vmPFC and the executive control network (e.g., left lateral dorsal prefrontal cortex (LDPFC)). The results indicate that self-relevant stimuli have enhanced saliency and attract greater attentional resources than stimuli associated to other people, and this leads to biased responses to self-stimuli compared to stimuli related to others. In contrast, the processing other-related stimuli recruits an executive control network in order to support the more difficult matching conditions (Humphreys & Sui, 2015a). Here the patients with extinction had lesions in the executive control network including the parietal cortex extending to the inferior frontal cortex whereas the vmPFC was intact. We speculate that the effects of lesions to the attentional control network could be reduced when there was strong enhancement of activity from the vmPFC, resulting in attention being shifted to self-related items (Humphreys & Sui, 2015a). This speculation needs to be tested by assessing the interaction between the self-network and the attentional control network using fMRI and/or TMS/tDCS techniques.

Although the present evidence indicates that there is pre-attentional processing of the self-relatedness of stimuli, we should not conclude that self-related processing is normal (see Humphreys & Sui, 2015a). The present results indicate only that there is sufficient processing of extinguished items for them to modulate how attention is

The self survives extinction

allocated. On the other hand, the report of contralesional items remained impaired, even when self stimuli fell in the contralesional field; this demonstrates that some processing constraints remain. In addition we should note that the present effects emerged here on stimulus identification, while the patients were typically able to detect both items on two item trials. The effects on identification were not due to guessing, since the patients made minimal false reports of the second item when extinction occurred and there were also few reports of two items on single item trials. Nevertheless, it may be argued that the results reflect a late-acting effect of self-bias, occurring after stimulus detection but before stimulus identification takes place. However, this proposal cannot be distinguished here from the fact that extinction was not severe in the present cases, minimising the ‘space’ for effects to be found on detection trials. Here it would be useful to explore effects of self-bias in patients with greater impairments. This argument notwithstanding, the current data provide proof-of-principle that self-relatedness can modulate attention and hence contact with self-representations is made pre-attentively.

Acknowledgements: This paper is dedicated to the memory of Glyn W. Humphreys (1954-2016). This work was supported by a grant from the Economic and Social Research Councils (ES/J001597/1 and ES/K013434/1).

References

The self survives extinction

- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 10367-10371.
- Berti, A., Allport, A., Driver, J., Dienes, Z., Oxbury, J., & Oxbury, S. (1992). Level of processing for stimuli in an “extinguished” visual field. *Neuropsychologia* *30*, 403-415.
- Berti, A., & Rizzolatti, G. (1992). Visual processing without awareness: evidence from unilateral neglect. *Journal of Cognitive Neuroscience*, *4*, 345-351.
- Bickerton, W-L., Demeyere, N., Francis, D., Kumar, V., Remoundou, M., Balani, A., Harris, L., Williamson, J., Lau, J.K., Samson, D., Riddoch, M.J. & Humphreys, G.W. (2014) The BCoS cognitive profile screen: Utility and predictive value for stroke. *Neuropsychology*, *29*. 638-648.
- Chechlacz, M., Rotshtein, P., Demeyere, N., Bickerton, W-L. & Humphreys, G.W. (2014) The frequency and severity of extinction after stroke affecting different vascular territories. *Neuropsychologia*, *54*, 11-17.
- Conway, M.A, & Pleydell-Pearce, C.W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*. *107*, 261-288.
- Corbetta, M., & Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201-215.
- Demeyere, N., Riddoch, M.J., Slavkova, E.D., Bickerton, W.L., & Humphreys, G.W. (2015). The Oxford Cognitive Screen (OCS): Validation of a stroke-specific short cognitive screening tool. *Psychological Assessment*. *27*, 883-894.

The self survives extinction

Domínguez-Borràs, J., Saj, A., Armony, J. L., & Vuilleumier P. (2012). Emotional processing and its impact on unilateral neglect and extinction.

Neuropsychologia, 50, 1054-1071.

Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition*, 79, 39-88.

Duncan, J., Humphreys, G.W., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion Neurobiology*, 7, 255–261.

Goldberg, I. I., Harel, M., & Malach R. (2006). When the Brain loses its self: prefrontal inactivation during sensorimotor processing. *Neuron*, 50, 329-339.

Heilman, K. M., Barrett, A. M., & Adair, J. C. (1998). Possible mechanisms of anosognosia: a defect in self-awareness. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 353, 1903-1909.

Humphreys, G.W. (1998). Neural representation of objects in space A dual coding account. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353, 1341–1351.

Humphreys, G.W., Bickerton, W.L., Samson, D., & Riddoch, M.J. (2012). The Birmingham Cognitive Screen (BCoS). (London: Psychology Press).

Humphreys, G. W., & Sui, J. (2015a). Attentional control and the self: the self-attention network (SAN). *Cognitive Neuroscience*,

DOI:10.1080/17588928.2015.1044427.

Humphreys, G.W. & Sui, J. (2015b) The salient self: Social saliency effects based on self bias. *Journal of Cognitive Psychology*, 27, 129-140.

The self survives extinction

- Karnath, H. O. (1988). Deficits of attention in acute and recovered visual hemi-neglect. *Neuropsychologia*, **26**, 27-43.
- Keenan J, Nelson A, O'Connor M, & Pascual-Leone A. (2001). Self-recognition and the right hemisphere. *Nature*, 409, 305.
- Keyes, H., & Dlugokencka, A. (2014). Do I have my attention? Speed of processing advantages for the self-face are not driven by automatic attention capture. PLOS ONE. doi: 10.1371/journal.pone.0110792.
- Lachter, J., Forster, K., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): still no identification without attention. *Psychological Review*, *111*, 880-913
- Macrae, C. N., Visokomogilski, A., Cunningham, W., & Sahraie, A. (2017). Self-relevance prioritizes access to visual awareness. *Journal of Experimental Psychology, Human Perception and Performance*, *43*, 438-443.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*, 56–60.
- Northoff, G. & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*. *8*, 102-107.
- Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*. *69*, 1019-1025.
- Riddoch, M.J., Humphreys, G.W., Edwards, S., Baker, T., & Willson, K. (2003). Seeing the action: Neuropsychological evidence for action-based effects on object selection. *Nature Neuroscience*, *6*, 82-89.

The self survives extinction

Stein, T, Siebold, A, & van Zoest, W. (2016). Testing the idea of privileged awareness of self-relevant information. *Journal of Experimental Psychology, Human Perception and Performance*, 42, 303-307.

Stolte, M., Humphreys, G. W., Yankouskaya, A. & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, 70, 1011-1022.

Sui, J. (2016). Self-reference acts as a golden thread in binding. *Trends in Cognitive Sciences*. 20, 482-483.

Sui, J., Chechlacz, M., & Rotshtein, P., & Humphreys, G. W. (2013b) Lesion-symptom mapping of self-prioritization in explicit face categorization: Distinguishing hypo- and hyper-self bias. *Cerebral Cortex*, **25**, 374-383.

Sui, J., Enock, F., Ralph, J., & Humphreys, G. W. (2015b). Dissociating hyper- and hypo-self biases to a core self-representation. *Cortex*, 70, 202-212.

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

Sui, J., & Humphreys, G. W. (2015). The interaction between self-bias and reward: Evidence for common and distinct processes. *Quarterly Journal of Experimental Psychology*. 68, 1952-1964.

The self survives extinction

- Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences*, 19, 719-728.
- Sui, J., Ohrling, E., & Humphreys, G. W. (2016). Negative music disrupts self and reward biases in perception. *Quarterly Journal of Experimental Psychology*, 69, 1438-1448.
- Sui, J., Rotshtein, P., & Humphreys, G. W. (2013a). Coupling social attention to the self forms a network for personal significance. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 7607-7612.
- Sui, J., Yankouskaya, A. & Humphreys, G. W. (2015a). Super-capacity and violations of race independence for self- but not for reward-associated stimuli. *Journal of Experimental Psychology, Human Perception and Performance*, 41, 441-452.
- Tamietto, M., Cauda, F., Celegnin, A., Diano, M., Costa, T., Cossa, F.M., Sacco, K., Duca, S., Geminiani, G.C., & de Gelder, B. (2015). Once you feel it, you see it: insula and sensory-motor contribution to visual awareness for fearful bodies in parietal neglect. *Cortex*, 62, 56-72.
- Vallar, G. (1998). Spatial hemineglect in humans. *Trends in Cognitive Sciences*, 2, 87-97.
- Vuilleumier, P., Armony, J.L., Clarke, K., Husain, M., & Driver, J., & Dolan, R.J. (2002b). Neural response to emotional faces with and without awareness: event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia*, 40, 2156-2166.

The self survives extinction

Vuilleumier, P.O., Rafal, R.D. (2000). A systematic study of visual extinction.

Between- and within-field deficits of attention in hemispatial neglect. *Brain*, 123, 1263-1279.

Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture.

Neurology, 23, 153-158.

Vuilleumier, P., Schwartz, S., Clarke, K., Husain, M., & Driver, J. (2002a). Testing

memory for unseen visual stimuli in patients with extinction and spatial neglect. *Journal of Cognitive science*, 14, 875-886.

Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain*, 102, 43-

63.

Table 1. Patients' demographic, clinical details and cognitive scores.

	<i>RR</i>	<i>PH</i>	<i>MH</i>	<i>SW</i>	<i>JB</i>	<i>DT</i>
Age in years	33	41	60	74	67	64
Gender	Male	Male	Male	Female	Male	Male
Aetiology	Stroke	Stroke	CM ¹	Stroke	Stroke	Stroke
Handedness	Right	Right	Right	Right	Right	Right
Post-lesion in years	2	15	6	13	2	2
Self-bias in matching task	Hyper self-bias	Hyper self-bias	Normal self-bias	Hyper self-bias	Hyper self-bias	Normal self-bias
Extinction*	Yes	Yes	Yes	Yes	Yes	Yes
Selective attention deficit*	No	Yes	No	No	No	No
Spatial neglect*	No	No	No	No	No	No
Object neglect*	No	No	Yes	No	No	No

Footnotes. ¹carbon monoxide poisoning. The cognitive profile for the patients was extracted from the scores on the OCS (Demeyere et al., 2015) and BCoS tests (Bickerton et al., 2014; Humphreys et al., 2012). The patients were classified

The self survives extinction

according to whether a clinical deficit was present or absent relative to age-matched control data on attention-related tasks from the BCoS.

Figure legends

Figure 1. (a) Brain lesions across the three patients. (b) The self-bias effect on perceptual matching in patients compared to healthy controls. RR, PH, SW, and JB show a hyper self-bias effect (effect size outside the control population) while MH and DT shows a self-bias effect falling in the range of the healthy controls. (c) The extinction effect on single- vs. two-item trials. The error bars represent one standard error.

Figure 2. Proportion of correction responses to two-item trials as a function of stimulus pairs (self-neutral, friend-neutral, or self-friend). The error bars represent one standard error.

Figure 3. (a) The report of the ipsilesional item as a function of stimulus pair (self-neutral vs. friend-neutral) and shape (self or friend vs. neutral). (b) The report of the contralesional item as a function of stimulus pair (self-neutral vs. friend-neutral) and shape (self or friend vs. neutral). (c) The report of single ipsilesional and contralesional stimuli on two-item trials as a function of shape (self vs. friend) and stimulus pair (field of stimulus: self-friend vs. friend-self). The error bars represent one standard error.

The self survives extinction

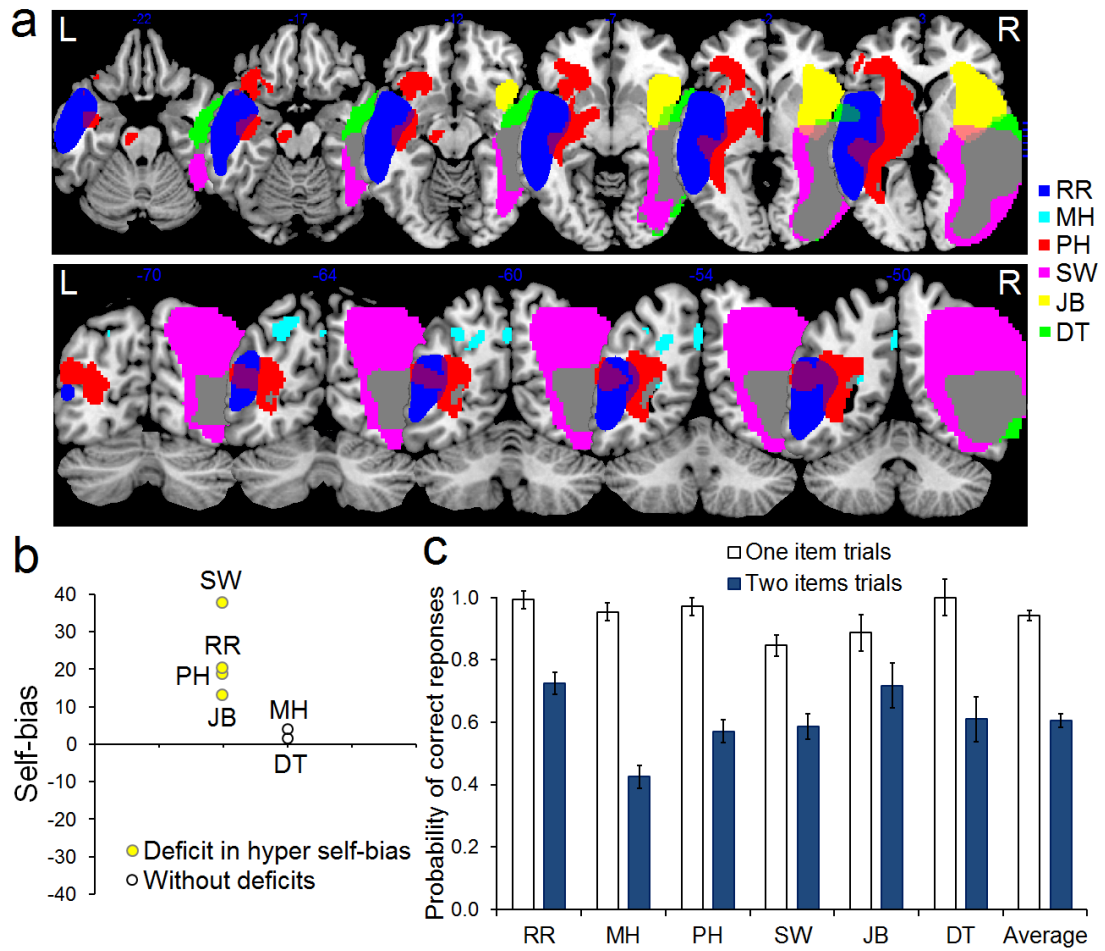


Figure 1

The self survives extinction

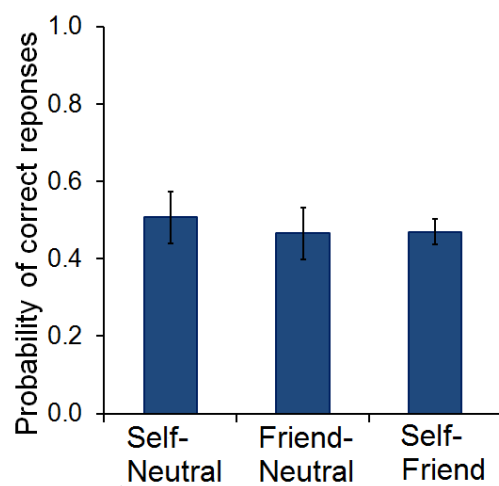


Figure 2

The self survives extinction

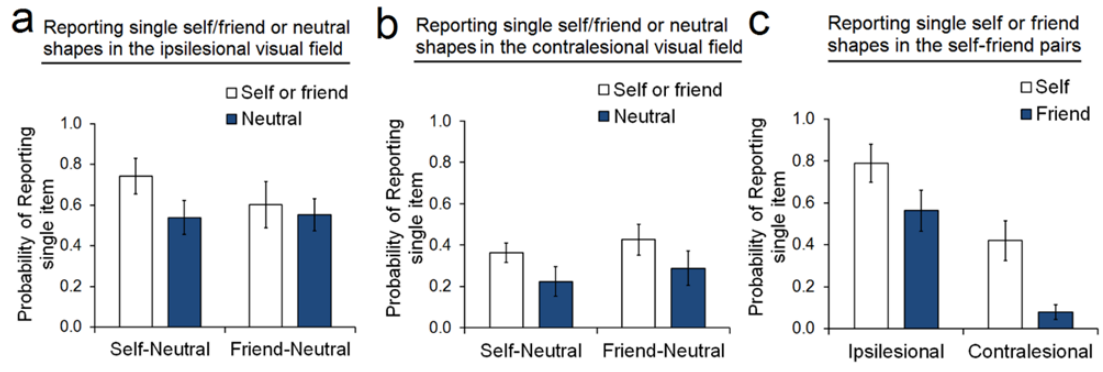


Figure 3