The experiential blink

Pincham, Hannah L; Bowman, H; Szucs, D

DOI:
10.1016/j.cortex.2016.04.007

The Experiential Blink: Mapping the Cost of Working Memory Encoding onto Conscious Perception in the Attentional Blink

H. L. Pincham, H. Bowman & D. Szucs

DOI: 10.1016/j.cortex.2016.04.007

Document Version: Early version of article, after reviews, also known as post-print.

Citation for published version (Harvard):
Pincham, HL, Bowman, H & Szucs, D 'The Experiential Blink: Mapping the Cost of Working Memory Encoding onto Conscious Perception in the Attentional Blink' Cortex, vol 8, pp 35-49. 2016.

The Experiential Blink:

Mapping the Cost of Working Memory Encoding onto Conscious Perception in the Attentional Blink

H. L. Pincham[1,2], H. Bowman[3,4*] & D. Szucs[1]

[1] Department of Psychology, University of Cambridge, UK

[2] Developmental Neuroscience Unit, Anna Freud Centre, London, UK

[3] Centre for Cognitive Neuroscience and Cognitive Systems and the School of Computing, University of Kent at Canterbury, UK

[4]School of Psychology, University of Birmingham, UK

* Corresponding author: Howard Bowman, Centre for Cognitive Neuroscience and Cognitive Systems and the School of Computing, University of Kent at Canterbury, Canterbury, Kent, CT2 7NF, United Kingdom, Tel. +44 1227 823815, Fax 01227 762811, E-mail: H.Bowman@kent.ac.uk

Abstract

The attentional blink (AB) represents a cognitive deficit in reporting the second of two targets (T2), when that second target appears 200-600msec after the first (T1). However, it is unclear how this paradigm impacts the subjective visibility (that is, the conscious perception) of T2, and whether the temporal profile of T2 report accuracy matches the temporal profile of subjective visibility. In order to compare report accuracy and subjective visibility, we asked participants to identify T1 and T2, and to rate the subjective visibility of T2 across two experiments. Event-related potentials were also measured. The results revealed different profiles for the report of T2 versus the subjective visibility of T2, particularly when T1 and T2 appeared within 200msec of one another. Specifically, T2 report accuracy was high but T2 visibility was low when the two targets appeared in close temporal succession, suggesting what we call the *Experiential Blink* is different from the classic AB. Electrophysiologically, at lag-1, the P3 component was modulated more by subjective visibility than by report accuracy. Collectively, the data indicate that the deficit in accurately reporting T2 is not the same as the deficit in subjectively experiencing T2. This suggests that traditional understandings of the AB may require adjustment and that, consistent with other findings, working memory encoding and conscious perception may not be synonymous.

Key words: Attention, consciousness, electroencephalography, P300, subjective visibility.

Introduction

A key objective in the study of mind and brain is to characterise the temporal dynamics of cognitive and perceptual functions. For example, researchers have sought to answer questions concerning how long attention has to be engaged on an item in order to be reported (Duncan, Ward, & Shapiro, 1994; Ward, Duncan, & Shapiro, 1996), and what the temporal profile of working memory (WM) encoding is. One phenomenon that has been frequently employed to study the temporal dynamics of cognition is the Attentional Blink (AB) (Raymond, Shapiro, & Arnell, 1992). The AB refers to a deficit in correctly reporting the second of two targets when that second target (T2) appears 200-600msec after the first (T1). Most major theories of the AB maintain that it indexes the temporal cost of encoding a stimulus into working memory (Bowman & Wyble, 2007; Chun & Potter, 1995; Olivers & Meeter, 2008). However, it is unclear whether the cost associated with encoding T1 is specific to *encoding* T2 into working memory, or whether it impacts other functions, such as the conscious perception of T2. The current study was designed to examine this.

This topic is particularly pertinent because it has the potential to throw light on whether WM encoding and conscious perception are synonymous. Within the AB domain, many researchers (including ourselves) have previously assumed that the correct report of T2 (which requires WM encoding) suggests that T2 was consciously perceived (Bowman & Wyble, 2007; Kranczioch, Debener, Maye, & Engel, 2007; Pincham & Szucs, 2012). The assumption that T2 is consciously perceived when (and, indeed, only when) it is correctly reported therefore implies that WM encoding and conscious perception are synonymous: what is consciously perceived enters WM, and everything that enters WM is consciously perceived. In other words, this position suggests that conscious perception is necessary and sufficient for entry into WM. Even though it might be intuitively plausible to view conscious perception and report accuracy as synonymous, the current study provides evidence to the contrary.

Only a small number of AB studies have distinguished between WM encoding and the conscious perception of T2. In those studies, the conscious perception of T2 has been operationalized using subjective visibility measures, and WM encoding of T2 has been operationalized using T2 identity (report) accuracy (Nieuwenhuis & de Kleijn, 2011; Sergent & Dehaene, 2004). For example, Sergent and Dehaene (2004) collected subjective visibility measures of T2 and suggested that the distribution from non-conscious to conscious perception is bimodal. In that study, it appeared that T2 was either ‘seen’ (high subjective visibility rating) or ‘not seen’ (low subjective visibility rating). By contrast, Nieuwenhuis and de Kleijn (2011) collected subjective visibility and report accuracy measures for T2 and revealed that the conscious perception of T2 could be a more gradual distribution between low and high subjective visibility ratings. Importantly, neither of these existing studies asked whether conscious perception and accuracy are synonymous in the AB. Further empirical work is needed to uncover the relationship between WM encoding and subjective visibility in the AB.

Outside of the AB, the relationship between conscious perception and WM encoding has been frequently debated in terms of the notion of phenomenological awareness (Block, 2007). This debate really considers whether conscious perception is *sufficient* to ensure WM encoding, and the existence of phenomenological awareness would suggest it is not. Although the current investigation is related to Block’s work, it is distinct from that body of literature because we focus on the dual concept – whether conscious perception is *necessary* for WM encoding, and we will argue our findings suggest it is not.

To examine the relationship between WM encoding and conscious perception in the AB, we presented two targets in a Rapid Serial Visual Presentation (RSVP) stream, and asked participants to report the identities of T1 and T2. Participants were also asked to provide a subjective visibility rating for T2. Across lags, these data can generate two temporal profile curves: report accuracy across lags versus subjective visibility across lags. The accuracy profile represents the (classic) AB curve. We argue that if WM encoding and conscious perception are equivalent in the AB, then the report accuracy curve and the subjective visibility curve would have the same shape.

In addition to examining behavioural data, the current study employed the temporal resolution of electroencephalography (EEG) to help contrast the temporal profiles of WM encoding and conscious perception in the AB. The P3 event-related potential component has been frequently viewed as an electrophysiological correlate of WM encoding (Polich, 2007; Vogel, Luck, & Shapiro, 1998). Studies have consistently found that P3 amplitude is reduced or even absent altogether on trials where T2 is reported incorrectly or not at all (Craston, Wyble, Chennu, & Bowman, 2009; Kranczioch, Debener, & Engel, 2003; Kranczioch, et al., 2007; Martens, Elmallah, London, & Johnson, 2006; Pincham & Szucs, 2012; Robitaille, Jolicoeur, Dell'Acqua, & Sessa, 2007; Vogel, et al., 1998). More importantly, the timing of the P3 component has been taken to be the temporal profile of the AB deficit. A demonstration of this was provided by McArthur, Budd and Michie (1999), who matched the temporal profile of the T2-P3 to the temporal profile of the T2 report accuracy deficit observed in the AB.

Given that previous AB work has rarely separated subjective visibility from report accuracy, it is not clear whether, in this context, the P3 indexes conscious perception, WM encoding or both. To that end, we examined how the amplitude and topography of the P3 is modulated by subjective visibility versus report accuracy. In a related study, Lamy, Salti and Bar-Haim (2008) measured ERPs in the context of a backwards masking paradigm to investigate the role of consciousness in online responding. To examine the neural correlate of awareness, Lamy et al. contrasted P3 amplitudes across high subjective visibility and low subjective visibility trials, while holding report accuracy constant. To examine the neural correlate of unconscious perception, P3 amplitude was contrasted across accurate and inaccurate trials, while holding subjective awareness constant (low visibility). In that study, both report accuracy and subjective visibility were shown to modulate P3 amplitude, but the impact of subjective visibility on the P3 was larger and topographically more widespread. Whereas Lamy et al. examined online responding, our study was designed to uncover the relationship between conscious perception and WM *encoding* (where responding is later and offline).

To summarise, the current study contrasted the temporal profile of WM encoding and conscious perception in the AB. To achieve this, two experiments were conducted. In Experiment 1, we aimed to sample the entire AB curve. Behavioural data were collected while T2 appeared at lags 1, 2, 3, 4, 6 or 8. In Experiment 2, behavioural and EEG data were collected, while T2 predominantly appeared at lags 1 and 3 (only two lags were used to enhance EEG signal strength). Following Lamy et al. (2008), we hypothesised that both report accuracy and subjective visibility would contribute to P3 amplitudes, although with subjective visibility possibly more so. Given that the current study is the first to compare report accuracy curves and subjective visibility curves in the AB, we did not make specific predictions about whether the shapes of those curves would be similar or dissimilar.

Materials & Methods

Participants

Initially, twenty-one young adults took part in the study. One participant was removed due to an inability to achieve 50% accuracy for T1. Two more participants were removed because more than 50% of the epochs extracted from their EEG data were rejected through the artifact detection criteria. Data from 18 participants (15 females) were therefore analysed. Participants were 19-28 years old (mean age: 21.67 years, SD = 2.93 years). Participants provided informed, written consent, had normal or corrected-to-normal vision and were fluent in English. The study was approved by the Psychology Research Ethics Committee at the University of Cambridge, UK.

Stimuli and Procedure

Stimuli were presented on a Sony Graphics Display CRT monitor with a 100Hz refresh rate. Targets were the uppercase letters excluding I, M, O, Q, W. These letters were excluded because of their physical similarity to digits (I, O and Q) or because their physical size meant that they were not adequately masked by digits (M and W). Each trial contained one or two targets – T1 occurred on every trial and was always presented in red, and T2 (if it occurred) was presented in white. Distractors were single digits excluding 0 and 1, presented in white. The rationale for presenting T1 in red and all other items in white was so that the visibility question (that is, "How visible was the white letter?") would clearly refer to T2 and not T1. All alphanumeric stimuli appeared on a black screen. Stimuli subtended visual angles of 3.8˚ vertically and 2.9˚ horizontally, assuming a viewing distance of 57cm. On each trial, a fixation cue (a cross shape subtending 2˚×2˚) was presented in the centre of the monitor for 200msec. The RSVP stream began 1000msec after the onset of the fixation cross. Each RSVP stream contained 15 items that were presented one after the other in the centre of the monitor (see Figure 1). The identities of the target letters and the digit distractors were randomly assigned on each trial with the restriction that successive items were not the same. Distractors were presented for 90msec with no ISI. T1 randomly appeared as the fourth, fifth or sixth item in the RSVP stream.

At the end of each RSVP stream, participants were asked to rate the subjective visibility of T2 using a self-report scale: "On a scale of 1-6, please indicate how well you have seen the white letter." This prompt appeared immediately after the RSVP stream finished, which corresponded to 540-900msec after the onset of T2 (depending on T2 lag, and the position of T1 in the stream: fourth, fifth or sixth item in the stream). The numbers 1 2 3 4 5 6 were presented in a horizontal line on the screen, with the description "not seen" presented beneath the number 1 and the description "maximal visibility" presented beneath the number 6 (see Figure 1). Participants used the number keys (1-6) on the keyboard to indicate their subjective visibility ratings. Although participants were not time-restricted in their subjective visibility ratings, participants typically responded very rapidly to the visibility prompt, suggesting that they were not replying on a decayed episodic memory of T2's subjective visibility. Participants then reported the identity of T1 and T2 (even if a second target did not occur) using the keyboard letter keys. Participants were required to guess if they were unsure of the target identities."

All participants completed two experiments, spaced at least one week apart. Experiment 1 exclusively collected behavioural data and Experiment 2 collected both behavioural and EEG data. Experiment 1 consisted of four blocks, each with a different target/mask duration combination. The mask, if it occurred, was always the hash (#) symbol. In Block 1, targets appeared for 90msec with no mask. In Blocks 2, 3 and 4, the target/mask durations were 70msec/20msec, 60msec/30msec and 50msec/40msec respectively. In Experiment 1, T2 appeared at lags 1, 2, 3, 4, 6 or 8 with equal frequency. Experiment 1 deliberately sampled a large number of lags in order to examine the relationship between T2 accuracy and subjective visibility across the entire AB curve. Trials that did not present a second target (no-T2 trials) were also included with equal frequency (hence, one in seven trials did not contain a second target). The no-T2 trials provide an index of the T2 guessing rate. Experiment 1 contained 4 blocks of 49 trials, totalling 196 experimental trials.

For each participant, data from Experiment 1 were analysed to determine which of the four target/mask durations resulted in T2 being correctly reported on approximately 50% of lag 3 trials. Each participant's optimal target/mask duration was then employed in Experiment 2. As a result, 28% of participants received the 70msec/20msec target/mask duration in Experiment 2, 50% of participants received the 60msec/30msec duration and the remaining participants received the 50msec/40msec duration. Experiment 2 contained 5 blocks of 100 trials, totalling 500 trials. To maximise ERP signal strength in Experiment 2, T2 appeared at lag 1 on 200 trials, at lag 3 on 200 trials, at lag 6 on 50 trials and was absent on 50 trials.

In Experiments 1 and 2, a distractor appeared in the place of T2 on no-T2 trials. However, the experimental program still assigned a target identity to T2, and participants were asked to report the subjective visibility and identity of T2 – even when a second target did not appear. In this manner, T2 'accuracy' on no-T2 trials (trials where T2 was correctly guessed by chance) could be calculated. The no-T2 trials were included for two reasons. First, subjective visibility for T2 could be determined for trials where the second target was not present (this is the guess rate). It was therefore possible to confirm that participants were accurately using the visibility scale, because subjective visibility should be very low on trials where T2 did not occur. Second, T2 report accuracy on no-T2 trials could be calculated and compared with theoretical (chance) levels of correct T2 report accuracy.

The order of the trials within each block was randomised. Participants could take short breaks between blocks. Testing occurred individually in an acoustically and electrically shielded booth.

[INSERT FIGURE 1 ABOUT HERE]

EEG Acquisition and Pre-processing

EEG was recorded using the Electrical Geodesics Inc. system and a 129-channel hydrocel geodesic sensor net. The sampling rate was 500Hz. An anti-aliasing lowpass filter of 100Hz was applied during data acquisition. Offline, the data were bandpass filtered between 0.01–30Hz and recomputed to an average reference. The continuous EEG was segmented into epochs between -200 to 1000msec relative to the onset of T1. Spline interpolation was carried out on individual channels if required. The mean percentage of interpolated channels was 4.60% (range: 0–8.59%). Epoched data were smoothed using a 50msec Gaussian filter.

Epochs were excluded from analysis if they met any of the following artifact rejection criteria: voltage deviations exceeded ±100µV relative to baseline, the maximum gradient exceeded 50µV, or activity was lower than 0.5µV. Across participants, 78.02% of trials were retained after artifact rejection.

Data Analysis

The behavioural data were analysed using two different approaches. Note that T2 report accuracy was conditional on T1 being correctly reported (T2|T1), such that trials where T1 was incorrectly reported were not included in calculation of T2 accuracy. Data were first analysed using the standard approach, which compares T1 report accuracy and T2 report accuracy in an omnibus ANOVA and then looks at lag effects for T1 and T2, where the latter of these is the classic way to demonstrate an attentional blink. This approach collapsed the data across the subjective visibility ratings. The second analysis considered the interaction between measure type (T2 accuracy and mean T2 subjective visibility rating) and lag (again, only for trials where T1 was correctly reported).

The data were analysed in these two ways because the first analysis reflects the standard analysis conducted in the AB literature, enabling comparisons with extant research. The second analysis is a novel approach that allowed us to compare the AB accuracy curve with the subjective visibility curve. In Experiments 1 and 2, data from the no-T2 trials were not analysed. Tukey's adjustment procedure was used to correct for multiple comparisons in all post-hoc contrast analyses.

Behavioural Data: Experiment 1

T1 vs. T2 Accuracy: Target accuracy for T1 and T2 were analysed using a block (1 vs. 2 vs. 3 vs. 4) × target (T1 vs. T2) × lag (1 vs. 2 vs. 3 vs. 4 vs. 6 vs. 8) repeated measures ANOVA. We also burrowed into this ANOVA to analyse simple effects of T1 by lag and T2 by lag separately. This was justified since the interaction of target and lag was significant. Indeed, one might argue for running planned comparisons across lags for T2, since the Attentional Blink is a well attested phenomenon. However, our effects were extremely strong, even without this.

T2 Accuracy vs. T2 Visibility: To uncover the relationship between subjective visibility and report accuracy, mean T2 accuracy and mean subjective visibility ratings were calculated for each participant at each lag. For comparability across the two dependent measures, subjective visibility ratings were transformed into scores out of 100. Only trials where T1 was correctly identified were included. An ANOVA was employed with T2 measurement type (T2 Accuracy vs. T2 visibility) and lag (1 vs. 2 vs. 3 vs. 4 vs. 6 vs. 8) as within-subjects factors. Although the main effect of T2 measurement type is not meaningful here (because two different dependent measures are being contrasted), a factorial ANOVA was included in order to examine the all-important interaction effect. An interaction between measurement type and lag would suggest that lag-related changes in T2 visibility differ from lag-related changes in T2 accuracy. In other words, a significant interaction would suggest that the shapes of the report accuracy curve and subjective visibility curve differ, such that the AB paradigm does not equally impact T2 WM encoding and T2 subjective visibility.

Behavioural Data: Experiment 2

In Experiment 2, T2 appeared at lag 1 or lag 3 on the majority of trials. Behavioural and EEG analyses for Experiment 2 are therefore restricted to these lags. Behavioural data from lag 6 are included in the appropriate figures for illustrative purposes.

T1 vs. T2 Accuracy: Target accuracy scores for T1 and T2 were compared using a target (T1 vs. T2) × lag (1 vs. 3) within-subjects ANOVA, and again, since the interactions were significant, we looked at the simple effect of lag for T1 and of lag for T2 separately. Again, a planned comparison was not performed, since the post hoc test was already highly significant.

T2 Accuracy vs. T2 Visibility: A 2 × 2 ANOVA was employed with measurement type (T2 accuracy vs. T2 visibility) and lag (1 vs. 3) as within-subjects factors. Pearson's correlations between T2 report accuracy (% correct) and T2 subjective visibility were also calculated separately for lag 1, lag 3, lag 6 and no-T2 trials. The correlations were employed to track the extent to which report accuracy and subjective visibility are linearly related at each lag.

ERP Data: Experiment 2

For every participant and every lag, the proportion of T2 visibility ratings falling into each of the six rating options was calculated. To pre-empt the empirical findings, participants' subjective visibility responses suggested that they were reluctant to use the upper visibility ratings. More than 50% of trials were quite low in subjective visibility. It therefore made statistical sense to create two visibility bins, creating – as far as possible – an approximate median split of the data (see Supplementary Material A). The two lowest visibility ratings formed the low visibility bin, and the remaining four visibility ratings formed the high visibility bin. The formation of two visibility bins therefore enhanced ERP signal strength (more epochs falling within each bin). The distribution of visibility ratings across the lags is shown in Supplementary Material A.

Epochs were time-locked to T1-onset, so that the zero time point reflects T1-onset, 90msec represents T2 onset at lag 1, and 270msec reflects T2 onset at lag 3. Post-stimulus P3 effects were analysed at Pz (electrode 62 in the EGI net) because the P3 component is maximal at centro-parietal locations (Polich, 2007). Epochs were categorised according to lag (1 vs. 3), T2 report accuracy (T2-correct vs. T2-incorrect) and subjective visibility (low vs. high).

Consecutive time windows of two-hundred millisecond durations were employed to capture activity associated with the T1-P3 and the T2-P3 components. Specifically, mean T1-P3 amplitude was examined between 300-500msec after T1 onset, T2-P3 amplitude at lag 1 was analysed 500-700msec after T1 onset (which corresponds to 410-610msec after T2 onset at lag 1) and mean T2-P3 amplitude at lag 3 was analysed 700-900msec after T1 onset (which corresponds to 340-540msec after T2 onset at lag 3). The choice of window placements is further justified in Supplementary Material B. Mean amplitudes were calculated rather than peak activity because previous work has shown that two distinct peaks (for T1 and for T2) are not apparent when T2 is presented at lag 1 (for example, Craston et al., 2009). Further, calculating mean amplitude is generally preferable because peaks are biased by uneven trial numbers across conditions (Luck, 2005). For all analyses, ERPs were only included if T1 was correctly reported.

T1 vs. T2 Accuracy and Visibility: A target (T1 vs. T2) × lag (1 vs. 3) × T2 accuracy (T2-correct vs. T2-incorrect) repeated measures ANOVA was used to compare P3 mean amplitudes. This analysis ignored subjective visibility ratings. Recall that T2 report accuracy was only calculated on trials where T1 was correctly detected. We re-analysed the ERP data based on subjective visibility rating, rather than report accuracy. Therefore, T2 accuracy was ignored and a target (T1 vs. T2) × lag (1 vs. 3) × visibility (low vs. high) repeated measures ANOVA was employed.

T2 Accuracy vs. T2 Visibility: Two ANOVAs were conducted to make comparisons between report accuracy and subjective visibility (see also Lamy et al. (2008) for similar statistical comparisons). In the first, only T2-correct trials were considered and compared across the two visibility ratings using a repeated-measures ANOVA with target (T1 vs. T2), lag (1 vs. 3) and visibility (low vs. high) as factors. The second only considered cases of low visibility, and compared across T2 accuracy: target (T1 vs. T2) × lag (1 vs. 3) × accuracy (T2-correct vs. T2-incorrect).

Results

Behavioural Data: Experiment 1

T1 vs. T2 Accuracy: This analysis examined report accuracy according to block, target and lag. Across blocks, T2 report accuracy decreased with decreasing T2 duration ($F(3,51)=55.482$, $p<.001$, $\eta^2=.765$). Apart from a main effect of block, the pattern of findings (as described below) was the same across the blocks. Data were therefore collapsed across the four blocks for ease of interpretation. As shown in Figure 2, T1 accuracy was significantly higher than T2 accuracy ($F(1,17)=173.936$, $p<.001$, $\eta^2=.911$). Increased accuracy for T1 (even at lags 1 and 8) is likely due to the colour marking of T1, effectively rendering T1 a singleton. Consistent with the traditional AB deficit, accuracy differed across lags ($F(5,85)=4.703$, $p=.001$, $\eta^2=.217$). An interaction between target and lag also emerged ($F(5,85)=7.120$, $p<.001$, $\eta^2=.295$), and accordingly we looked separately at T1 and T2 report accuracy across lags. In the T1 analysis, the main effect of lag was significant ($F(1,17)=3.344$, $p=.008$, $\eta^2=.164$). Tukey corrected pairwise post-hoc contrasts indicated that this effect was due to reduced T1 accuracy at lag 1 versus lag 3 ($p=.002$). No other pairwise comparisons between lags were significant. In the T2 analysis, the main effect of lag was also significant, indicating that T2|T1 report accuracy differed according to lag ($F(1,17)=6.042$, $p<.001$, $\eta^2=.262$), i.e. there was a clear attentional blink. On trials where T2 was not presented, T1 accuracy was 94.25% and the identity of T2 was correctly guessed on 8.57% of trials on which T1 was correctly reported.

T2 Accuracy vs. T2 Visibility: Figure 2 displays data from the T2 measurement type × lag ANOVA. This analysis revealed a main effect of lag ($F(5,85)=14.073$, $p<.001$, $\eta^2=.453$). Importantly, the main effects were modulated by an interaction between the type of measure and lag ($F(5,85)=14.633$, $p<.001$, $\eta^2=.463$). As is clear in Figure 2, accuracy and subjective visibility appeared to track one another at most lags, but the absolute difference between T2 report accuracy and T2 visibility ratings was greatest at earlier lags and largest at lag 1. Post-hoc contrasts confirmed that the difference between T2 report accuracy and T2 visibility was significantly larger at lag 1 than at other lags (lag 1 vs. lag 2: $p=.072$, lag 1 vs. lags 3, 4, 6, 8: all $p<.002$). The difference between T2 report accuracy and T2 visibility were not statistically different across all other lags. Further analyses (used to assess the robustness of these findings, especially the validity of comparing two different measures: report accuracy and subjective visibility) are included in Supplementary Material C.

[INSERT FIGURE 2 ABOUT HERE]

Behavioural Data: Experiment 2

T1 vs. T2 Accuracy: The target × lag ANOVA indicated that T1 report performance significantly exceeded T2 performance ($F(1,17)=311.040$, $p<.001$, $\eta^2=.948$). As expected, target report accuracy was higher at lag 1 than at lag 3 ($F(1,17)=18.446$, $p<.001$, $\eta^2=.520$). Report accuracy was also influenced by the interaction between target and lag, as the impact of lag differed for T2 accuracy versus T1 accuracy ($F(1,17)=93.872$, $p<.001$, $\eta^2=.847$). We then looked inside this significant target × lag interaction separately for T1 and T2 report accuracy across lags. In the T1 analysis, the main effect of lag was significant, as T1 report accuracy was reduced at lag 1 versus lag 3 ($F(1,17)=32.200$, $p<.001$, $\eta^2=.654$). In the T2 analysis, the main effect of lag was also significant, as T2|T1 report accuracy was higher at lag 1 than at lag 3 ($F(1,17)=51.676$, $p<.001$, $\eta^2=.752$).

Similarly to Experiment 1, on no-T2 trials, T1 accuracy was high (93.73%) and T2 accuracy was at approximately the level expected due to chance (3.20%). Importantly, the calibration procedure used in Experiment 1 to equate the number of T2-correct and T2-incorrect trials at lag 3 was successful here because T2 was correctly reported on 42.3% of trials at lag 3. These data are shown in Figure 3.

T2 Accuracy vs. T2 Visibility: The comparison of T2 report accuracy and mean subjective visibility ratings is shown in Figure 3. This analysis revealed a main effect of lag ($F(1,17)=29.266$, $p<.001$, $\eta^2=.633$). There was also an interaction between the type of measure and lag ($F(1,17)=87.052$, $p<.001$, $\eta^2=.837$). Tukey post-hoc contrasts showed that while subjective visibility ratings did not differ across lags 1 and 3 ($p=.423$), target accuracy scores did ($p<.001$).

[INSERT FIGURE 3 ABOUT HERE]

Confirming the results of the ANOVA, T2 report accuracy was significantly positively correlated with subjective visibility at lag 3 and at lag 6 suggesting that higher rates of T2 report accuracy tended to be associated with higher subjective visibility ratings at those lags (lag 3: $r=.516$, $p=.028$; lag 6: $r=.759$; $p<.001$). The correlation was not, though, significant at lag 1 ($r=.460$; $p=.055$), although there was a trend towards a significant effect. On no-T2 trials, there was no relationship between T2 report accuracy and subjective visibility ($r=.021$, $p=.936$). Figure 4 displays the data used in the correlational analyses.

[INSERT FIGURE 4 ABOUT HERE]

ERP Data: Experiment 2

T1 vs. T2 Accuracy: In this analysis, a target × lag × T2 report accuracy ANOVA was used to examine P3 mean amplitudes. As shown in Figure 5A, mean T1-P3 amplitude (mean=2.003; SE=0.572) was larger than mean T2-P3 amplitude (mean=0.697; SE=0.699) ($F(1,17)=11.125$, $p=.004$, $\eta^2=.396$). P3 amplitudes were larger on T2-correct trials (mean=1.693; SE=0.647) than on T2-incorrect trials (mean=1.007; SE=0.591) ($F(1,17)=8.584$, $p=.009$, $\eta^2=.336$). A significant target × T2 report accuracy effect indicated that the difference between T2-correct and T2-incorrect trials was more pronounced for the T2-P3 than the T1-P3 ($F(1,17)=6.988$, $p=.017$, $\eta^2=.291$). Tukey post-hoc pairwise comparisons confirmed that the T2-P3 was significantly larger on T2-correct trials (mean=1.324; SE=0.587) compared with T2-incorrect trials (mean=0.694; SE=0.634; $p<.001$). Mean T1-P3 amplitude did not differ significantly across T2-correct (mean=2.061; SE=0.778) and T2-incorrect trials (mean=1.320; SE=0.652; $p=.269$). The lag main effect approached significance, suggesting that P3 mean amplitude on lag 1 trials (mean=1.691; SE=0.665) tended to be higher than P3 mean amplitude on lag 3 trials (mean=1.009; SE=0.600) ($F(1,17)=3.741$, $p=.070$, $\eta^2=.180$). No other effects were statistically significant (largest F=1.718). Difference topographies (shown in Figures 5B – 5E) reveal that the post-stimulus effects were strongest in parietal scalp locations, as expected. This was particularly the case for the T2-P3 difference topographies (Figures 5C and 5E).

T1 vs. T2 Visibility: This analysis compared P3 amplitudes using a target × lag × visibility repeated measures ANOVA. As in the previous Analysis, mean T1-P3 amplitude (mean=2.067; SE=0.621) was larger than mean T2-P3 amplitude (mean=0.849; SE=0.708) ($F(1,17)=9.641$, $p=.006$, $\eta^2=.362$; see Figure 5F). Further, P3 amplitude was significantly larger for trials rated as high in subjective visibility (mean=1.992; SE=0.743) than on trials rated as low visibility (mean=0.924; SE=0.578) ($F(1,17)=7.525$, $p=.014$, $\eta^2=.307$). A target × lag interaction indicated that mean amplitude differences between lag 1 and lag 3 trials were more pronounced for the T2-P3 than the T1-P3 ($F(1,17)=6.850$, $p=.018$, $\eta^2=.287$). Tukey post-hoc comparisons confirmed that there was no statistical difference between lag 1 and lag 3 for the T1-P3 (Lag 1 Mean=2.219; SE=0.664 versus Lag 3 Mean=1.915; SE=0.655; p=.691). However, the lag 1 T2-P3 was significantly larger than the lag 3 T2-P3 (Lag 1 Mean=1.511; SE=0.681 versus Lag 3 Mean=0.187; SE=0.842; $p<.001$). Difference topographies confirmed that the P3 effect was strongest around Pz. This was especially clear in the T2-P3 difference topographies (Figures 5H and 5J) and for the T1-P3 difference topography at lag 1 (Figure 5G). No other effects attained statistical significance (largest F=2.873).

[INSERT FIGURE 5 ABOUT HERE]

T2 Accuracy vs. T2 Visibility: To further explore what the P3 is most representative of: encoding into WM or conscious perception, we sought to keep one variable fixed (e.g. visibility), and examine the extent to which the P3 varied with the other variable (e.g. report accuracy). The relevant ERPs are shown in Figure 6. Statistically, we first considered only T2-correct trials and examined P3 amplitudes using an ANOVA with target, lag and visibility as factors. Here, T1-P3 amplitude (mean=2.001; SE=0.591) was significantly larger than T2-P3 amplitude (mean=0.829; SE=0.702) ($F(1,17)=7.499$, $p=.014$, $\eta^2=.306$). Further, mean P3 amplitudes were significantly higher under high visibility (mean=2.166; SE=0.670) than under low visibility (mean=0.664; SE=0.653) ($F(1,17)=9.075$, $p=.008$, $\eta^2=.348$). Finally, the target × lag interaction was significant, indicating that the differences between lags 1 and 3 were driven by the T2-P3 and not the T1-P3 ($F(1,17)=5.323$, $p=.032$, $\eta^2=.238$).

Next, we only considered cases of low visibility, and compared P3 amplitudes using a target × lag × accuracy ANOVA. This analysis confirmed that the T1-P3 (mean=1.632; SE=0.539) was larger than the T2-P3 (mean=0.036; SE=0.662) ($F(1,17)=16.149$, $p=.001$, $\eta^2=.487$). The target × lag interaction also approached significance ($F(1,17)=3.511$, $p=.078$, $\eta^2=.171$). Importantly, the main effect of report accuracy was not significant ($F<1$), nor were any other interactions with report accuracy (largest $F=1.798$). The data from this analysis are shown in Figure 6.

As seen in the analyses above, subjective visibility significantly impacted P3 amplitudes, whereas T2 accuracy did not significantly impact P3 amplitudes. In order to confirm that visibility was a more powerful mediator of P3 amplitudes than accuracy, it was necessary to directly contrast accuracy and visibility, in terms of their effects on P3 amplitude. We therefore calculated visibility difference scores (high visibility/T2-correct minus low visibility/T2-correct),) and accuracy difference scores (low visibility/T2-correct minus low visibility/T2-incorrect) for each participant. Based on Lamy et al. (2008), we hypothesized that subjective visibility would be a more powerful determinant of P3 amplitude than accuracy. Paired samples one-tailed t-tests confirmed that visibility difference scores were larger (mean=1.892; SE=0.801) than accuracy difference scores at lag 1 (mean=-0.581; SE=0.904) ($t(17)=1.756$, $p=.048$, Cohen's *d*=.647), but not at lag 3 (visibility difference score mean=1.114, SD=0.833 versus accuracy difference score mean=-0.101, SE=0.604) ($t(17)=0.910$, $p=.188$, Cohen's *d*=.386). Collectively, these data suggest that subjective visibility plays a greater role than accuracy in determining P3 amplitudes, at least at lag 1[1].

[INSERT FIGURE 6 ABOUT HERE]

Discussion

The current study was designed to contrast T2 report accuracy and T2 subjective visibility in the AB. Critically, the behavioural findings indicate that, across lags, the shapes of the accuracy curve and the subjective visibility curve were not identical. Rather, the data suggest that the subjective visibility curve and the AB curve differ at early lags: a little at lag 2 and a lot at lag 1. Despite high T2 report accuracy during early lags, subjective visibility of T2 was seen to be particularly low. Our behavioural results therefore suggest that WM encoding and conscious perception are not synonymous in the AB. Further, the data suggest that the classic AB deficit is distinct from, what we name, the 'experiential blink'. The

[1] A more complete analysis would have involved a full two-by-two crossing of visibility and accuracy. However, this 2 X 2 ANOVA on P3 amplitude could not be formed with this data set, since the response T2-incorrect/ High Visibility was almost never made. This prevents us from being able to interrogate the visibility by report accuracy interaction, which would have been the most compelling demonstration of our thesis.

experiential blink is defined as a deficit in subjective visibility for T2 (as opposed to T2 report accuracy), which results from the temporal proximity of T1 and T2. The EEG data also support a distinction between WM encoding and conscious perception. At lag-1, subjective visibility appeared to be a more powerful determinant of P3 amplitude than report accuracy (see especially Figure 6), suggesting that the conscious perception of T2 (as indexed by subjective visibility) is not synonymous with the encoding of T2 into WM (as indexed by report accuracy).

Behavioural Data

In Experiments 1 and 2, T2 report accuracy and subjective visibility tracked one another between lags 4 – 8. This pattern may have started to diverge at lag-3, but then certainly did at lag 2, with the greatest difference between T2 report accuracy and visibility observed at lag 1; see Figures 2 and 3. The increased difference between accuracy and subjective visibility at early lags suggests that WM encoding is more effective at these lags than would be expected by the relatively poor conscious perception of T2. These central findings were evident across both experiments, and were confirmed with supplementary analyses (see Supplementary Material C). If it is assumed that T2 must be encoded into WM in order to be correctly reported and that subjective visibility ratings reflect conscious perception, then the current data highlight a potential disparity between WM encoding and conscious perception. Additionally, we suggest that the traditional *attentional* blink be viewed as distinct from the *experiential* blink. At the very least, the data undermine the (albeit implicit) assumption common in the AB literature that the conscious perception of T2 is synonymous with accurate report of T2.

It is surely the case that, in most everyday contexts, conscious perception is synonymous with WM encoding. Indeed, under most lags employed here, conscious perception closely tracked WM encoding. However, conscious perception has been dissociated from the visuomotor system in blindsight patients where accuracy can be above chance, despite poor subjective visibility (Weiskrantz, 1986). Our data may therefore be indicative of a form of WM 'blindsight' in non-patient (control) populations. Importantly, the current study should be distinguished from related phenomena such as implicit learning and subliminal priming. During implicit learning, the rule or relationship that is implicitly learned is, by definition, not reportable (Lovibond & Shanks, 2002). During subliminal priming, the prime's identity is also inaccessible (Bar & Biederman, 1998; Cheesman & Merikle, 1984). Both of these situations are in contrast to what we have here, whereby T2's identity was accurately reported at the earliest lag, despite no/low subjective visibility of that item. Thus, our difference is between conscious perception and *explicit recall* of identity, which must come from a stored and maintained representation (that is, a representation from WM). Indeed, one might call the behavioural phenomenon presented here *sight-blink recall*, since the item is not perceptually "seen", but it is recalled.

The correlational data presented in Experiment 2 support the idea that consciously perceiving a target is not necessary for that target to be encoded in WM. At lag 6, participants appeared to be most accurate in predicting (with subjective visibility) whether T2's identity was correctly reported. However, WM encoding and conscious perception were most poorly related at lag 1. These data therefore additionally suggest that metacognitive (or introspective) capabilities monotonically increase with increases in target onset asynchrony between T1 and T2. This is in stark contrast to classic AB report accuracy, which robustly exhibits a U-shaped pattern, strictly formalised as an inverted gamma pattern across lags (Su, Bowman, & Barnard, 2011).

Lau and Passingham (2006) present a related finding to ours. Specifically, using metacontrast masking they were able to compare two target-mask SOAs that exhibited the same stimulus discrimination performance, but different Seen/Not-seen visibility judgements. This finding, and its temporal profile, in which participants reported lower visibility at shorter SOAs, is consistent with our observation that subjective visibility drops relative to report accuracy as the T1 and T2 become temporally closer, with lag-1 being the closest. This raises the interesting possibility that the same perceptual mechanism may underlie both findings, which is further discussed under *Theoretical Interpretation* later in this Section.

However, a number of differences do exist between Lau and Passingham's (2006) study and ours. For example, their study is more directly differentiating objective and subjective perception, while, as previously discussed, our findings are more about differentiating encoding into WM from subjective perceptual experience. This is reflected in our task requiring identification and encoding into WM amongst a large number of alternatives – most of the letters of the alphabet. Further, Lau and Passingham identify fMRI correlates of their finding, whereas our use of EEG enables us to consider a response time-course, with a fine temporal resolution.

The current behavioural findings also nicely complement those identifying WM maintenance without conscious awareness (Soto & Silvanto, 2014) and particularly with data reported by Bergstrom and Eriksson (2014) who demonstrated such an effect for the AB. Our finding might be considered as the identification of a WM encoding mechanism without awareness, from which extended pre-conscious WM maintenance traces could arise.

ERP Data

Confirming existing work, P3 amplitude was increased when T2 was correctly reported versus incorrectly reported (Craston, et al., 2009; Pincham & Szucs, 2012; Vogel, et al., 1998). P3 amplitude was also enhanced under high compared with low visibility (Del Cul, Baillet, & Dehaene, 2007; Sergent, Baillet, & Dehaene, 2005). The more interesting finding, however, was that the P3 component was differentially affected by changes in report accuracy versus subjective visibility. When T2 report accuracy was held constant (that is, only T2-correct trials were analysed), P3 amplitude was more positive on high visibility trials and seemed to have been eliminated under low visibility (see Figure 6A). By contrast, when visibility was equalised (that is, only low visibility trials were analysed), there was no apparent difference between T2-corrrect and T2-incorrect trials (see Figure 6B). Moreover, these two differences, (1) between High and Low visibility (with report accuracy constant) and (2) between correct and incorrect (with visibility constant), were themselves statistically different. Subjective visibility of T2 therefore appeared to be a more powerful determinant of P3 size than T2 report accuracy. For example, the T2-correct/low visibility ERP was much more similar to the T2-incorrect/low visibility ERP than to the T2-correct/high visibility ERP (see Figure 6), suggesting that changes in subjective visibility are what most significantly modulate the P3 in our RSVP context. Additionally, if one views the P3 as likely to be generated by a fronto-parietal network, these ERP findings would be somewhat consistent with Frassle and colleagues' (2014) association of such an fMRI network with introspection of subjective experience.

Lamy et al. (2008) also found that subjective visibility more strongly impacted P3 amplitudes than accuracy did. However, making this inference in Lamy et al.'s work was to some extent confounded by the fact that guesses might bias the correct/low visibility ERP. Lamy et al. estimated the form of that ERP with guessing corrected for, but acknowledged that this was only an estimate. Because our behavioral task required letter identification, the probability of guessing a correct response was negligible (one in twenty), and much smaller than that in the Lamy et al. study (where it was one in four). It is also important to emphasise that the task employed by Lamy et al. (2008) is very different to ours. Specifically, in their study, an online task was performed, for which the involvement of WM was unclear. Our task explicitly engages WM. Thus, Lamy et al. explored the differentiation of subjective visibility and perceptual discrimination, while we explore the differentiation of subjective visibility and WM encoding.

Within the AB literature, the P3 component has been primarily viewed as an index of WM encoding and cognitive resource allocation (for example, Pincham & Szucs, 2012; Shapiro, Schmitz, Martens, Hommel, & Schnitzler, 2006; Vogel, et al., 1998). However, outside of the AB, the P3 is linked to various cognitive mechanisms including task difficulty/demands (Courchesne, 1978), stimulus categorisation (Mecklinger & Ullsperger, 1993), decision making (Nieuwenhuis, Aston-Jones, & Cohen, 2005) and even subjective visibility (Del Cul, et al., 2007; Melloni, Schwiedrzik, Müller, Rodriguez, & Singer, 2011; Sergent, et al., 2005). This variety of characterisations of the P3, seems to stand against the possibility that there is a single unifying cognitive characterisation of the component. Thus, it is important to consider the generality of our P3 findings.

In this respect, we found little, if any, evidence of the P3 responding to report accuracy in our experiment. For example, the higher amplitudes for Correct conditions in Figure 5A, could be explained by the correlation between report accuracy and subjective visibility (compared with Figure 4). Further, Figure 6 (which is the clearest contrast for this question) gives very little evidence that the P3 is modulated by report accuracy in our experimental paradigm.

Despite this, we are not in a position to definitively infer that the P3 does not respond to report accuracy in the attentional blink. In particular, due to a lack of trials in the T2-incorrect/High Visibility condition, we were unable to consider whether T2-correct generates a larger P3 pattern than T2-incorrect, when subjective visibility is high,. Furthermore, it was not possible to analyse the P3 at lag-6, due to a limited number of trials. In addition, the AB conditions we have considered might be a special case, where the perceptual demands are so hard that subjective experience and report accuracy decouple. Importantly, this would be sufficient to support the basic claim we are making, viz that conscious experience and encoding into WM are not exactly the same. Strictly, a single counterexample suffices to disprove an assertion, and we present that here.

One inconsistent set of findings in the literature are a number of impressive studies that ascribe a post-perceptual locus for the P3 (Pitts, Padwal, Fennelly, Martínez, & Hillyard, 2014; Shafto & Pitts, 2015; Squires, Hillyard, & Lindsay, 1973). Our experiment though has a number of differences to these previous studies, which presumably explain the difference in findings. We discuss these in turn.

1) The stimuli that were being assessed for perceptual experience in these previous studies were typically drawn from a very small set, e.g. a detection task in Squires et al. and two shapes in Pitts et al.. Consequently, the perceptual judgement / encoding into WM processes were much more informationally-rich in our study (i.e. identify a letter).
2) The elegant demonstration by Pitts and colleagues (Pitts, et al., 2014; Shafto & Pitts, 2015) that the P3 in their experiment was modulated by task set, does not naturally carry over to our setting. This is because task set, certainly in respect of instruction, is constant throughout our experiment. What modulates the P3 in the current study is behaviour (specifically High vs Low visibility).
3) Pitts et al. (2014) highlighted a number of confounds associated with studies of conscious experience that mean that post-perceptual processes are not equated across conditions. These confounds are, at least to a large extent, resolved in our experiment. In particular, our key P3 comparison contrasts subjective visibility levels, while report accuracy is controlled, i.e. the “same” identity report (although not, of course, the

same subjective visibility report) is made for both High and Low subjective visibilities (see Figure 6).

Additionally, the P3 in RSVP may present differently to the P3 in paradigms without a rapid sequence of repeated onsets. Importantly, the key P3 finding in RSVP previous to this paper was that a P3 is present when an item breaks through into awareness, and is reported. But to all intents and purposes, the P3 is not present at all when it is not correctly reported in the current study. Our specific claim in this paper is about the P3 as it manifests in this "fringe of awareness" context.

Furthermore, we do not definitively know whether the P3 is the earliest component that responds to either subjective visibility or report accuracy in RSVP. Indeed, it could be that there is an earlier process in the timecourse, which we do not see in our EEG experiments. Possibilities for such a component are the modulation of the P2 (Vogel, et al., 1998) or the N2 (Sergent, et al., 2005), which might be related to the Visual Awareness Negativity (Shafto & Pitts, 2015). This said, we present empirical evidence that the P3 was modulated by subjective visibility in our experimental setting – this is our key finding. Thus, while it might be that the P3 is not the earliest component that responds to conscious experience in our paradigm, it would seem that, at the least, it is driven by such a component.

Integration

In the current study, we seek to attribute the increased difference between report accuracy and subjective visibility that we observe at very early lags (and particularly lag-1) to a divergence between encoding into working memory and conscious perception. The case for this inference may be somewhat complicated by current theories of the lag-1 data point. Specifically, many have argued that lag-1 is a special case, in which the T1 and T2 are sometimes processed together, even as a single integrated percept (Akyürek et al., 2012;

Bowman & Wyble, 2007; Hommel & Akyürek, 2005; Wyble, Bowman, & Nieuwenstein, 2009).

This then raises the possibility that the reduction in subjective visibility ratings at lag-1 are not specifically a reduction in conscious experience, but more *confusion about* the conscious experience. That is, that conscious experience of T2 at lag-1 is not exactly reduced in strength, but rather yields an unfamiliar conscious experience leading participants to report a low subjective visibility because of what one may call a "loss of confidence" in their experience. Put in other terms, it could be that in some sense the perceptual image of T2 at lag-1 is in fact strong, but the co-presence of a T1 in that perceptual experience means one cannot accurately rate the T2's perceptual strength. This is a subtle distinction, since our hypothesis is fundamentally about the subjective and confidence is part of that very subjectivity. However, even if we entertain the possibility that it would be a confound, there are a number of reasons to believe that this confusion due to order explanation cannot account for our findings.

First, colour marking was incorporated in the experiment to ensure that T1 was coloured distinctly from the rest of the stream, and particularly from T2. One reason for doing this was to mitigate against the possibility of obtaining integrated percepts at lag-1, and there is evidence that this worked. Specifically, in letters-in-digits tasks, of which the experiment here is an example, the cardinal indicator of integrated percepts is order errors. That is, if T1 and T2 are encoded as a single undifferentiated whole, information about the order in which they occurred would be lost, and in fact, conjunction information between the T1, the T2 and the colour in which T2 occurred would be lost. Importantly though, order errors were very rare in our experiments: they were around 10% compared to 30% in classic, un-colour marked letters-in-digits tasks (e.g., Chun & Potter, 1995). This suggests that the percentage of trials in which no order information was present was no more than 2 x 10%, since on such trials if participants would guess the order, 50% of the time they would get it right. This then implies that on at least 80% (100% minus 20%) of the trials, the T2 was successfully differentiated as the second target, i.e. the letter that was *not* colour-marked. This stands against the suggestion that T1-T2 integrations are prevalent at lag-1 and underlie the drop in relative subjective visibility at early lags in our experiments.

Second, it is important to note that the reduction in relative subjective visibility can also be observed at lag-2, and perhaps also the beginning of the effect at lag-3, see Figure 2. The integration argument is though classically ascribed specifically to lag-1 and not later lags, in which there are intervening distractors.

Third and most significantly, we have endeavoured to perform an analysis uncontaminated by integration trials. This is presented as the fourth analysis in supplementary material C. This has been performed by (1) excluding all order error trials, and (2) excluding an equivalent number of order correct trials, under a conservativeness assumption, which ignores the possibility of prior entry in an order error and assumes that on integration trials, participants will randomly guess the order of the targets. Thus, there will be an equal number of correct order as incorrect order trials arising from integration. We do not though know which of the correct order trials were the integration ones. So, we discard correct order trials that will make it hardest for us to demonstrate the effect we seek to show. If the effect is still significant after such a procedure, we know the effect is robust, although we are likely to have overestimated the p-value (i.e. underestimated the size of the effect). Under this analysis, the interaction effect between measure type (report accuracy and subjective visibility) and lag remained significant. This is strong evidence that our findings are not explainable by integrated percepts at lag-1.

Further, the ERP findings do not obviously fit with the integrated percept explanation. In this context, the critical condition is T2 correct/ low visibility, since this is the disparity case (T2 was correctly reported, but was poorly perceived consciously). The prevalence of these trials at lag-1 is what causes the relative reduction in subjective visibility at that serial position. If perceptual integration were characteristic of this phenomenon at lag-1, we would expect to see an ERP reflecting this integration. The obvious candidate pattern for this situation is a relatively short, but relatively high amplitude P3, and indeed, this is the pattern normally observed for T1 & T2 correct *versus* T1 & not T2 correct at lag-1 (see, for example Craston, et al., 2009, Figure 7). The slightly higher amplitude arises because T1 and T2 are being encoded at the same time, and the relatively short positive deflection (not much longer than for T1 & not T2) arises from the simultaneity of encoding, i.e. T2 does not have to wait for T1 to be encoded, which would generate a very long positive deflection. Importantly, the ERP we observe for T2-correct/ low visibility, see Figure 6, panel A, does not follow this classic integration pattern. In particular, the initial positive (P3) deflection for T2 correct/ low visibility is very low amplitude, and indeed, smaller than for the other conditions in Figure 6, panel A. This does not look like a characteristic integration pattern. Overall, this set of arguments would seem to counter the possibility that an integrated percept at lag-1 explains our findings.

Subjective Visibility Measure

In this paper, we are taking subjective visibility ratings as proxies for conscious experience. There is clearly an inferential step here, which we now seek to justify more fully. Importantly, in respect of subjective visibility, we are following what is now a large body of literature that probes the strength of conscious experience using such measures (e.g., Lamy, et al., 2008; Overgaard, Rote, Mouridsen, & Ramsoy, 2006; Sergent & Dehaene, 2004) including carefully considered comparisons of different scales (Sandberg, Timmermans, Overgaard, & Cleeremans, 2010). Of most significance, participants seem to use the scales consistently, with little evidence of confusion on their part. In particular, uncertainty regarding the interpretation of subjective visibility would manifest as increased variability in the measure, which would reduce statistical power and effect sizes. However, the statistical power of our key behavioural effects was large. Specifically, the degrees of freedom in the two experiments were pretty standard for attentional blink studies. Nonetheless, the interactions we observed were highly significant.

In particular, the interaction between measure type and lag had an F-value of $F(5,85) = 14.633$ for Experiment 1 and of $F(1,17) = 87.052$ for Experiment 2, with p-values as small as possible within the normal reporting range, i.e. $p < .001$ in both cases. Furthermore, the inclusion of a replication in the same paper, which also proves very significant, is highly unlikely for a weak finding hampered by high error variability. Finally, if degrees of freedom are taken out of the equation and raw effect sizes are considered, the statistical power associated with this finding is further emphasized. Specifically, the interaction between measure type and lag generates an eta-squared of 0.463 for experiment one and 0.837 for experiment two. These are large effect sizes.

Overall, then, the treatment of the subjective visibility scale seems to be extremely consistent across participants and experiments. This stands against there being confusion amongst participants with regard to how to interpret this measure, suggesting that participants responded according to the strength of the visual experience, or in other words "how well [they] have seen the [T2]", the terms in which the question was formulated.

On this issue, it should also be emphasized that the kind of uncertainty over measure employed is intrinsic to the question being considered. That is, we are endeavouring to characterise conscious experience; this is a fundamentally subjective phenomenon. Thus, there is no clear alternative to probing participants' *subjective* experience, and within this context, the approach we have employed here seems close to as good as one can do.

Process Impurity

Additionally, it is important to note that there is no sense to which our two measures – report accuracy and subjective visibility – are *process pure*. In particular, they certainly do not reflect distinct classes of information, Report accuracy is surely not independent of subjective visibility, or in other words one would (normally) expect report accuracy to increase with the vividness of conscious experience, and this relationship is apparent in our data from lag 3 upwards, i.e. higher levels of report accuracy coincide with higher levels of subjective visibility. But, the critical point for us is that this relationship of strong dependence breaks at short lags.

Our argument could be explicitly framed as that the extent to which report accuracy indexes conscious experience specifically changes at early lags, as demonstrated by reduction in subjective visibility relative to report accuracy. This suggests that encoding into working memory (which report accuracy should index) becomes decoupled from conscious experience, at these lags. We would argue that it is only possible to arrive at such a decoupling, if working memory encoding and conscious experience are not synonymous. This is the essential line of argument of this paper.

It is also important to note that our key index of conscious experience – subjective visibility – is an indirect measure. Indeed, both report accuracy and subjective visibility are assessed at the end of the RSVP stream. Consequently, both measures depend upon memory to bridge the gap from perception to report. In this respect, for our argument, we need that this memory (which would also be expected to involve working memory) provides a relatively accurate storage and retrieval of the quality of a conscious experience. The points we have made above re. the reliability and precision of subjective visibility report would seem to justify this claim. This enables us to use subjective visibility as an (offline) proxy for an (online) “in the moment” conscious experience.

To assess working memory encoding, we rely upon the broadly accepted perspective that the core AB deficit is at the point of encoding, suggesting that the level of T2|T1 reflects success at encoding into WM, with the necessity to discriminate amongst a large set of letter alternatives ensuring encoding is a demanding process. Moreover, it does not seem a concern that representations of subjective experience may be being simultaneously maintained in WM, with representations of letter identity. In particular, there is no sense to which the demand associated with holding such an experiential representation would change with lag.

To reiterate, the key point for us is that the memory of experiential quality diverges from the memory of identity at short lags. They are indeed both memories, but they are memories of different things. A stronger claim would be that WM encoding and conscious perception are *dissociable*, in the strict sense given to that term in cognitive psychology (Dunn & Kirsner, 1988). This would imply that, at some stage of processing, WM encoding and conscious perception are implemented by separate modules. Note that in the presence of nonlinearities, which are certainly there in the brain, it is possible to obtain different behavioural patterns for two measures even though they are implemented by a single module (Dunn & Kirsner, 1988). Thus, while we argue we have demonstrated that WM encoding and conscious perception are not synonymous, a further step is required to show that they are not just different “read-outs” from the same module. This question awaits further work.

Theoretical Interpretation

Further to the question of the nature of the subjective visibility measure, it is important to note that interpretation of it as an index of conscious experience is supported by a clear and intuitive theoretical interpretation of our findings. In particular, lag-1 is exactly the data point where simultaneity is most pronounced – the T1 and T2 are presented to the perceptual system almost concurrently. Accordingly, one could explain our findings in terms of a differential ability to simultaneously process; that is, we are, to a large extent, able to simultaneously *encode* two items, e.g. T1 and T2 here, but to a much lesser extent to simultaneously *perceive* multiple items consciously. In other words, there may be a sense to which we really can "only perceive one thing at a time".

Clearly, more work is required before such a strong claim can be confirmed, but our results are suggestive in this respect, and particularly if a further striking pattern in our ERP data is considered. Specifically, if one compares the *lag 1/T2 correct* ERP in Figure 5, panel A, with the *lag 1/T2 Correct/High visibility* ERP in Figure 6, panel A, we see that the latter is extremely extended in time.

This long lag 1/T2 Correct/High visibility ERP is suggestive of serial processing, whereby processing of T1 has to complete before T2 can start. In other words, when a short (combined) P3 is observed at lag-1, the T1 and T2 are encoded simultaneously, and this is sufficient to consolidate the T2 into working memory (the lag1/ T2 correct ERP), but not to obtain a strong conscious experience of the T2; that requires a serial process (the lag-1/T2 Correct/High visibility ERP).

Such seriality resonates with a recent proposal concerning the AB (Marti, Sigman, & Dehaene, 2012). Our position, though, is not exactly the same as Marti et al's. Specifically, we ascribe to the basic explanation of the report-accuracy AB from the Simultaneous Type Serial Token (STST) model (Bowman & Wyble, 2007; Wyble, et al., 2009). This has similarities to Marti et al.'s serial processing explanation, but goes further, since, for example, a specific mechanism for generating the seriality is posited, viz unavailability of the attentional enhancement (STST's blaster). This unavailability of attention during the blink would induce a subjective visibility blink similar to that for report-accuracy. However, the basic STST model explains (report-accuracy) lag-1 sparing in terms of joint encoding of the T1 and T2, which arises because both items benefit from a single attentional enhancement. It is the subjective visibility analogue of this joint encoding that we are arguing does not obtain. That is, even though two items can be simultaneously encoded, our findings suggest that they may not be able to be simultaneously consciously perceived. It is not clear what mechanism might induce this particular seriality of perception. In particular, we are not at this stage attributing this to the with-holding of attention, i.e. the same mechanism as that which generates the (report-accuracy) AB in STST.

Conclusion

Overall, Experiments 1 and 2 reveal that the process of encoding T1 differentially impacts the accuracy of reporting T2's identity versus the conscious perception of having seen T2. Behaviourally, the AB accuracy curve differs from the subjective visibility curve – especially at early lags. Electrophysiologically, the P3 component we observe is most substantially modulated by subjective visibility. Collectively, these data suggest that the classic AB deficit is not the same as the experiential blink. As a result, conscious perception is not synonymous with WM encoding in the AB context.

Acknowledgements

We acknowledge the contribution of Brad Wyble and Sander Nieuwenhuis in early discussions from which the presented research arose. HLP was funded by a studentship from the Gates Cambridge Trust. We would like to also thank Zoltan Dienes and two anonymous reviewers for very valuable comments on an early version of this article. Zoltan Dienes' insights on dissociability of components were particularly valuable.

References

Akyürek, E. G., Eshuis, S. A., Nieuwenstein, M. R., Saija, J. D., Başkent, D., & Hommel, B. (2012). Temporal target integration underlies performance at lag 1 in the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1448.

Bar, M., & Biederman, I. (1998). Subliminal visual priming. *Psychological Science, 9*(6), 464-468.

Bennett, C. M., Miller, M., & Wolford, G. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. *Neuroimage, 47*(1), S125.

Bergström, F., & Eriksson, J. (2014). Maintenance of non-consciously presented information engages the prefrontal cortex. *Frontiers in Human Neuroscience, 8*, 938.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*(5), 481.

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review, 114*(1), 38-70.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics, 36*(4), 387-395.

Chun, M. M., & Potter, M. C. (1995). A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 109-127.

Courchesne, E. (1978). Changes in P3 waves with event repetition: Long-term effects on scalp distribution and amplitude. *Electroencephalography and Clinical Neurophysiology, 45*(6), 754-766.

Craston, P., Wyble, B., Chennu, S., & Bowman, H. (2009). The attentional blink reveals serial working memory encoding: Evidence from virtual and human event-related potentials. *Journal of Cognitive Neuroscience, 21*(3), 550-566.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5*(10), e260.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369*(6478), 313-315.

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: the principle of reversed association. *Psychological Review, 95*(1), 91-101.

Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *The Journal of Neuroscience, 34*(5), 1738-1747.

Hommel, B., & Akyürek, E. G. (2005). Lag-1 sparing in the attentional blink: Benefits and costs of integrating two events into a single episode. *The Quarterly Journal of Experimental Psychology Section A, 58*(8), 1415-1433.

Kilner, J. M. (2013). Bias in a common EEG and MEG statistical analysis and how to avoid it. *Clinical Neurophysiology, 124*(10), 2062-2063.

Kranczioch, C., Debener, S., & Engel, A. K. (2003). Event-related potential correlates of the attentional blink phenomenon. *Cognitive Brain Research, 17*(1), 177-187.

Kranczioch, C., Debener, S., Maye, A., & Engel, A. K. (2007). Temporal dynamics of access to consciousness in the attentional blink. *Neuroimage, 37*(3), 947-955.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 12*(5), 535-540.

Lamy, D., Salti, M., & Bar-Haim, Y. (2008). Neural Correlates of Subjective Awareness and Unconscious Processing: An ERP Study. *Journal of Cognitive Neuroscience, 21*(7), 1435-1446.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences, 103*(49), 18763-18768.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28*(1), 3.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: The MIT Press.

Martens, S., Elmallah, K., London, R., & Johnson, A. (2006). Cuing and stimulus probability effects on the P3 and the AB. *Acta Psychologica, 123*(3), 204-218.

Marti, S., Sigman, M., & Dehaene, S. (2012). A shared cortical bottleneck underlying Attentional Blink and Psychological Refractory Period. *Neuroimage, 59*(3), 2883-2898.

McArthur, G., Budd, T., & Michie, P. (1999). The attentional blink and P300. *Neuroreport, 10*, 3691-3695.

Mecklinger, A., & Ullsperger, P. (1993). P3 varies with stimulus categorization rather than probability. *Electroencephalography and Clinical Neurophysiology, 86*(6), 395-407.

Melloni, L., Schwiedrzik, C. M., Müller, N., Rodriguez, E., & Singer, W. (2011). Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *Journal of Neuroscience, 31*(4), 1386-1396.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological Bulletin, 131*(4), 510-532.

Nieuwenhuis, S., & de Kleijn, R. (2011). Consciousness of targets during the attentional blink: a gradual or all-or-none dimension? *Attention, Perception, and Psychophysics, 73*, 364-373.

Olivers, C. N. L., & Meeter, M. (2008). A Boost and Bounce Theory of Temporal Attention. *Psychological Review, 115*(4), 836-863.

Overgaard, M., Rote, J., Mouridsen, K., & Ramsoy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition, 15*(4), 700-708.

Pincham, H. L., & Szucs, D. (2012). Conscious Access Is Linked to Ongoing Brain State: Electrophysiological Evidence from the Attentional Blink. *Cerebral Cortex, 22*(10), 2346-2353.

Pitts, M. A., Padwal, J., Fennelly, D., Martínez, A., & Hillyard, S. A. (2014). Gamma band activity and the P3 reflect post-perceptual processes, not visual awareness. *NeuroImage, 101*, 337-350.

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

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849-860.

Robitaille, N., Jolicoeur, P., Dell'Acqua, R., & Sessa, P. (2007). Short-term consolidation of visual patterns interferes with visuo-spatial attention: Converging evidence from human electrophysiology. *Brain Research, 1185*, 158-169.

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Consciousness and Cognition, 19*(4), 1069-1078.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience, 8*(10), 1391-1400.

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science, 15*, 720-728.

Shafto, J. P., & Pitts, M. A. (2015). Neural signatures of conscious face perception in an inattentional blindness paradigm. *The Journal of Neuroscience, 35*(31), 10940-10948.

Shapiro, K. L., Schmitz, F., Martens, S., Hommel, B., & Schnitzler, A. (2006). Resource sharing in the attentional blink. *Neuroreport, 17*(2), 163-166.

Soto, D., & Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends in Cognitive Sciences, 18*(10), 520-525.

Squires, K. C., Hillyard, S. A., & Lindsay, P. H. (1973). Cortical potentials evoked by confirming and disconfirming feedback following an auditory discrimination. *Perception & Psychophysics, 13*(1), 25-31.

Su, L., Bowman, H., & Barnard, P. (2011). Glancing and then looking: on the role of body, affect and meaning in cognitive control. *Frontiers in Psychology, 2*, 348.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological Evidence for a Postperceptual Locus of Suppression During the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1636-1674.

Ward, R., Duncan, J., & Shapiro, K. L. (1996). The Slow Time-Course of Visual Attention. *Cognitive Psychology, 30*(1), 79-109.

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford: Oxford University Press.

Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 787.

Figure Captions

Figure 1: The experimental design. (A) shows that T1 and T2 were letters presented between white digit distractors. T1 was always a red letter and T2 (if it appeared) was a white letter. All stimuli were presented on a black monitor. In this figure, T2 appears at lag 3. In Experiment 2, T2 appeared at lag 1 (45% of trials), lag 3 (45% of trials), lag 6 (5% of trials) or did not appear (5% of trials). On no-T2 trials, T2 was replaced by a digit distractor. Although the entire stream length is not shown, each stream contained 15 stimuli. Items appeared for 90msec duration, with no ISI. A hash (#) mask could appear in the interstimulus interval following target stimuli, with varying durations (as described above). This was used to adjust the visual strength of targets according to an individual's perceptual threshold. (B) shows the phrasing and screen layout of the subjective visibility question. Importantly, the subjective visibility question asked about the white letter and therefore referred to T2 and not T1.

Figure 2: T1 accuracy, T2|T1 accuracy and T2 mean subjective visibility ratings in Experiment 1 (note, only trials in which T1 was correctly reported are included in the T2 subjective visibility rating). Mean ratings of subjective visibility appeared to track T2 accuracy at the later lags, but not at lag 1 and probably also not at lag-2.

Figure 3: T1 accuracy, T2|T1 accuracy and T2 mean subjective visibility rating in Experiment 2 (note, only trials in which T1 was correctly reported are included in the T2 subjective visibility rating). Mean ratings of subjective visibility appeared to track T2 accuracy at lags 3 and 6, but not at lag 1. Data for lag 6 is shown for illustrative purposes only, as these conditions were not included in the main statistical analyses (due to low trial counts).

Figure 4: Correlations between T2 report accuracy and mean subjective visibility ratings. (A) shows the correlation between T2 report accuracy and subjective visibility at lag 1. Although a trend was present, this relationship was not statistically significant. (B) shows the correlation between T2 report accuracy and subjective visibility at lag 3. This positive linear relationship was statistically significant, indicating that an increase in T2 report accuracy was associated with an increase in subjective visibility. (C) shows the correlation between T2 report accuracy and subjective visibility at lag 6. This positive linear relationship was statistically significant, indicating that an increase in T2 report accuracy was associated with an increase in subjective visibility. (D) shows the correlation between T2 "accuracy" and subjective visibility when T2 was not presented. This relationship was not statistically significant.

Figure 5: Post-stimulus (P3) ERPs and difference topographies. Only trials where T1 was correctly reported are included. (A) through (E) show data classified by T2 accuracy (independent of visibility rating), and topographies show activity on T2-correct trials minus activity on T2-incorrect trials. (F) through (J) show data classified by subjective visibility of T2 (independent of T2 accuracy), and topographies show activity on high visibility trials minus activity on low visibility trials. (A) shows ERPs for T2-correct and T2-incorrect trials, time-locked to the onset of T1. T2 appeared at lag 1 (90msec after T1) or at lag 3 (270msec after T1). (B) shows the T1-P3 difference topography at lag 1 in the period 300-500msec after the onset of T1. (C) shows the T2-P3 difference topography at lag 1 in the period 500-700msec after the onset of T1. (D) shows the T1-P3 difference topography at lag 3 in the

period 300-500msec after the onset of T1. (E) shows the T2-P3 difference topography at lag 3 in the period 700-900msec after the onset of T1. (F) shows ERPs for high and low visibility trials, time-locked to the onset of T1. T2 appeared at lag 1 (90msec after T1) or at lag 3 (270msec after T1). (G) shows the T1-P3 difference topography at lag 1 in the period 300-500msec after the onset of T1. (H) shows the T2-P3 difference topography at lag 1 in the period 500-700msec after the onset of T1. (I) shows the T1-P3 difference topography at lag 3 in the period 300-500msec after the onset of T1. (J) shows the T2-P3 difference topography at lag 3 in the period 700-900msec after the onset of T1. For all topographies, the black dot corresponds to the Pz electrode (electrode 62). All difference topographies are averaged across the relevant time window.

Figure 6: Grand average ERPs at lag 1 and lag 3. Only trials where T1 was correctly reported are included. (A) shows ERPs for lag 1 trials. (B) shows ERPs for lag 3 trials. Mean activity was significantly larger on T2-correct/High trials. Mean activity did not differ between T2-correct/low visibility trials and T2-incorrect/Low visibility trials.

Figures

Figure 1

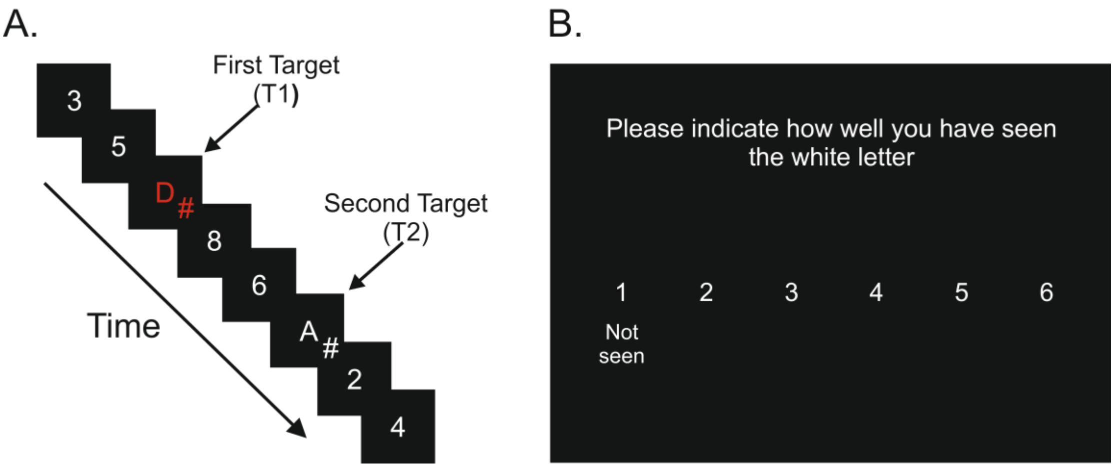


Figure 2

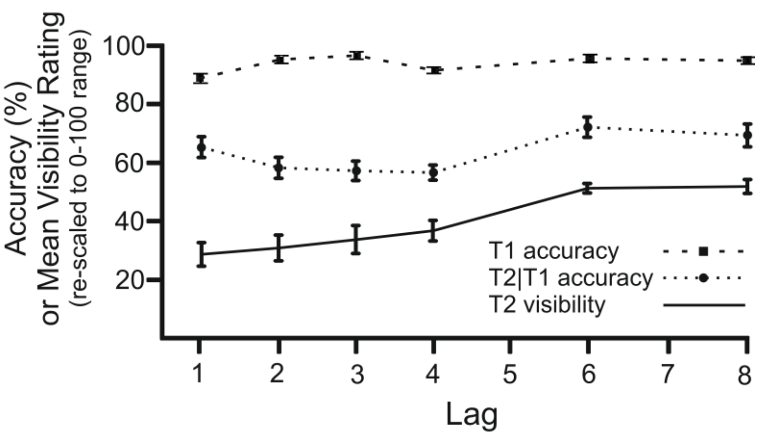


Figure 3

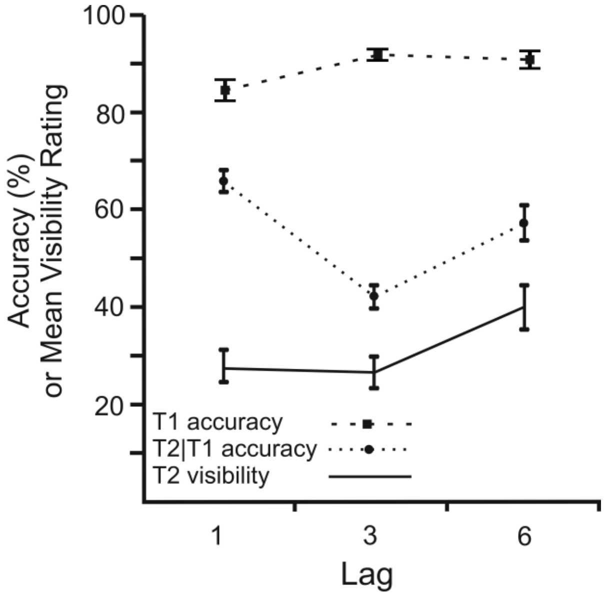

Accuracy (%)
or Mean Visibility Rating
100
80
60
40
20
T1 accuracy
T2|T1 accuracy
T2 visibility
1
3
6
Lag

Figure 4

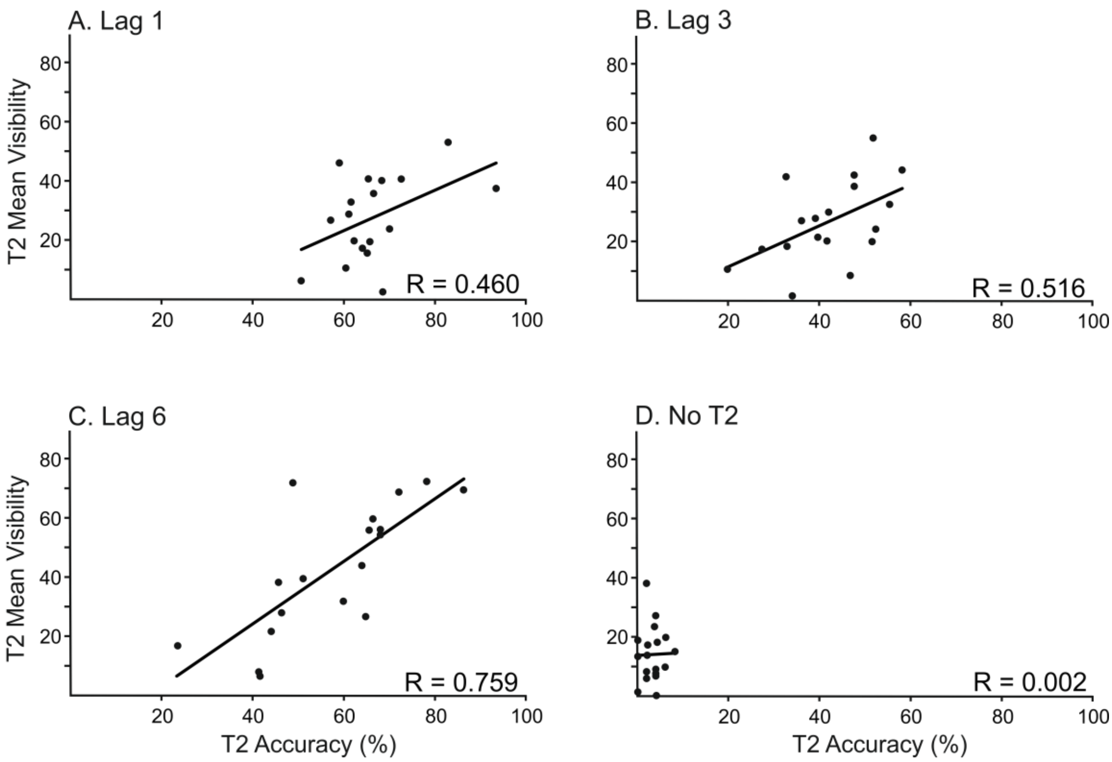

A. Lag 1
B. Lag 3
C. Lag 6
D. No T2
T2 Mean Visibility
T2 Accuracy (%)
R = 0.460
R = 0.516
R = 0.759
R = 0.002
20
40
60
80
100

Figure 5

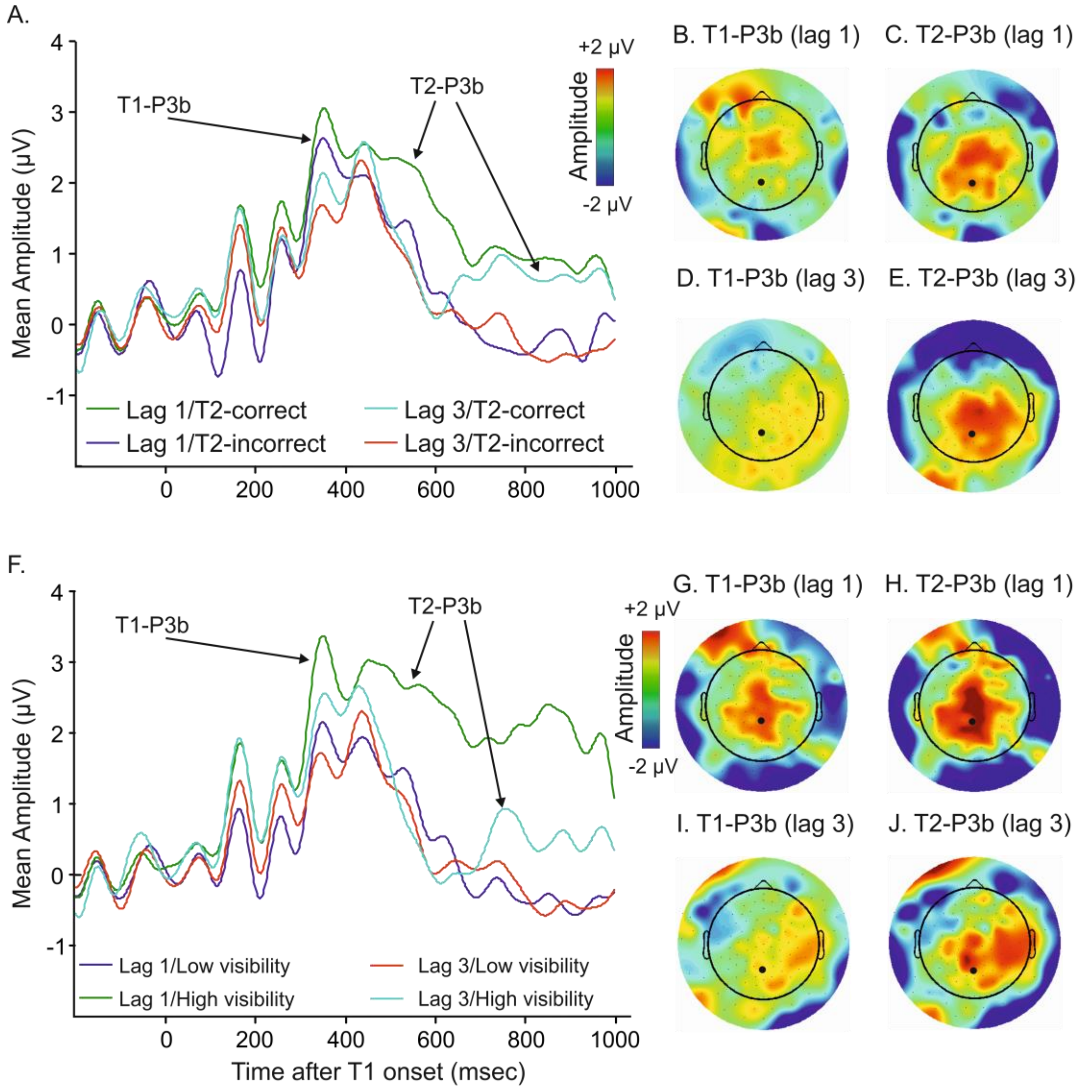
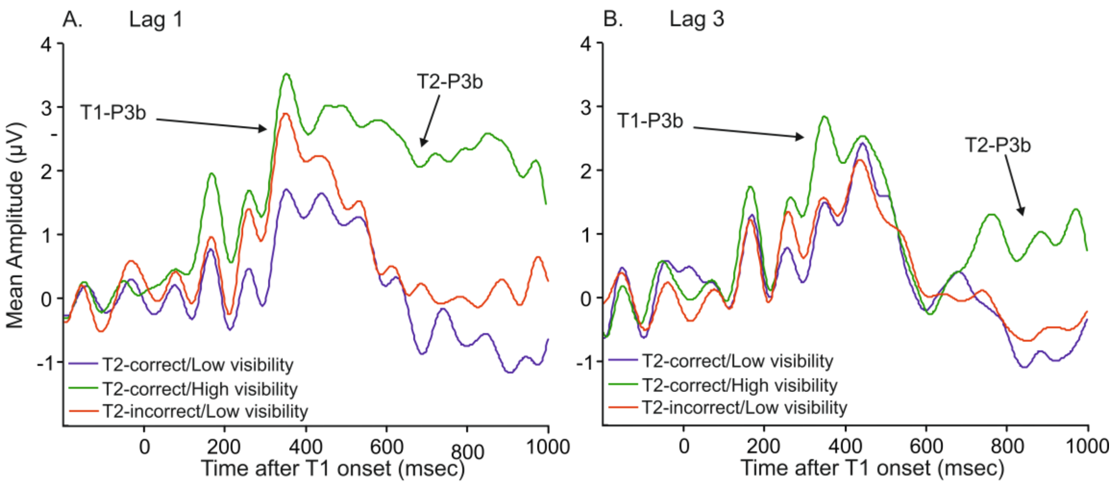
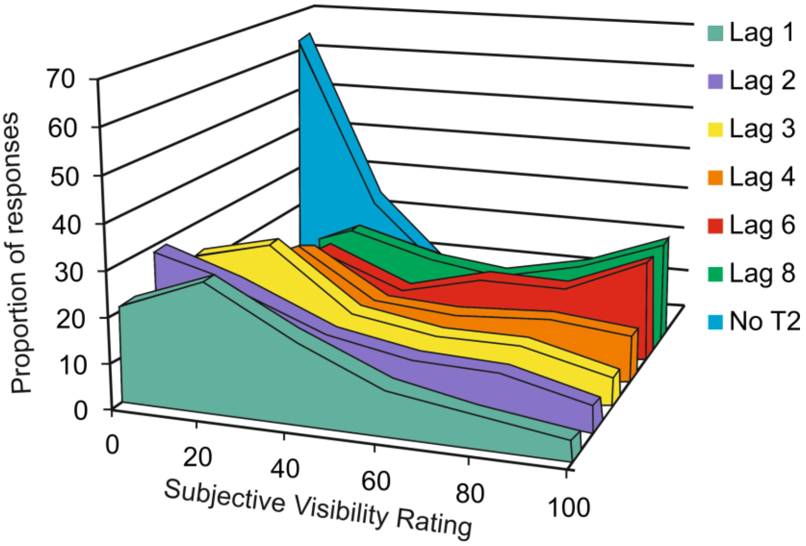

A.
Mean Amplitude (μV)
T1-P3b
T2-P3b
Lag 1/T2-correct
Lag 1/T2-incorrect
Lag 3/T2-correct
Lag 3/T2-incorrect
+2 μV
Amplitude
-2 μV
B. T1-P3b (lag 1)
C. T2-P3b (lag 1)
D. T1-P3b (lag 3)
E. T2-P3b (lag 3)
F.
Mean Amplitude (μV)
T1-P3b
T2-P3b
Lag 1/Low visibility
Lag 1/High visibility
Lag 3/Low visibility
Lag 3/High visibility
Time after T1 onset (msec)
+2 μV
Amplitude
-2 μV
G. T1-P3b (lag 1)
H. T2-P3b (lag 1)
I. T1-P3b (lag 3)
J. T2-P3b (lag 3)

Figure 6

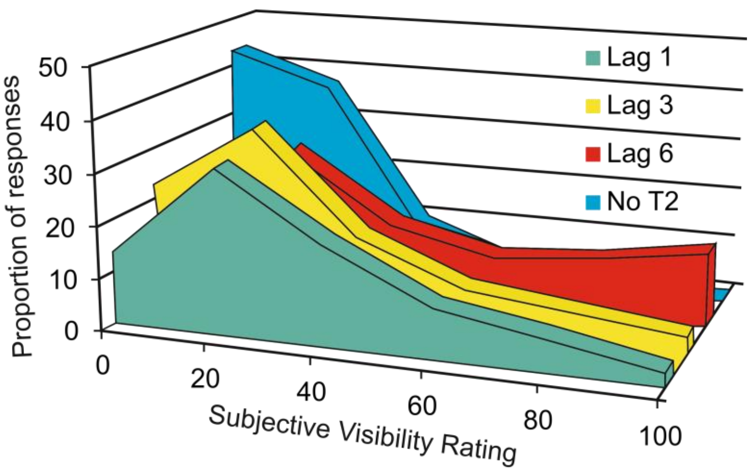

A. Lag 1
B. Lag 3
Mean Amplitude (µV)
T1-P3b
T2-P3b
T2-correct/Low visibility
T2-correct/High visibility
T2-incorrect/Low visibility
Time after T1 onset (msec)
4
3
2
1
0
-1
0
200
400
600
800
1000

Electronic Supplementary Material A

As shown in the figures below, participants used the lower subjective visibility ratings more frequently than they used the higher subjective visibility ratings. For this reason, we created two clusters of subjective visibility, in an attempt to match the number of trials in each condition. To that end, the ‘Low’ subjective visibility cluster consisted of the two lowest subjective visibility ratings, whereas the ‘High’ subjective visibility cluster consisted of the four highest subjective visibility ratings. Two clusters of subjective visibility (Low vs High) were created rather than three clusters (Low vs Middle vs High) or more, to ensure that, for each participant, there were sufficient numbers of accepted trials in each cluster to calculate summary statistics.



Supplementary Figure 1. Frequency of visibility ratings, as a function of lag in Experiment 1.



Supplementary Figure 2. Frequency of visibility ratings, as a function of lag in Experiment 2.

Electronic Supplementary Material B

Tailoring window placements to the grand average can bias statistical tests of ERPs and thereby inflate the Type I error (i.e. false positive) rate (Kilner, 2013). This practice has similarities to problems highlighted in the fMRI literature of double dipping (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009) and failure to appropriately correct for multiple comparisons (Bennett, Miller, & Wolford, 2009). One way to guard against inflating the Type I error rate is to select window placements from prior literature. This makes window parameter settings a priori justified, and thus, not “fished for” a posteriori. Indeed, if selected in a disciplined fashion, such prior specification should guard against increase in the Type I error rate. Since these placements are tailored to previously collected and not the current data, such a priori selection is likely to inflate the Type II error rate (i.e. failures to reject the null hypothesis when it should be). This is because the window placements may not be optimal for the components (as they present) in the current data. Although this reduces statistical power, an identified significant effect can be relied upon.

Type I errors can also be guarded against by setting window placement parameters in a regular stereotyped fashion, without any tuning of parameters, e.g. by always placing window boundaries at multiples of 100msec and restricting all windows to a fixed size.

We employ both of these approaches to guard against the possibility of inflating the Type I error rate. Specifically, our windows are regular and stereotyped by placing window boundaries at multiples of 100msec, keeping all windows of width 200msec and making them abut one another. Accordingly, we divide a contiguous 600msec region (starting 300 and finishing 900msec post T1 onset) into three 200msec sub-regions – 300-500msec for the T1, 500-700msec for the T2 at lag-1 and 700-900msec for the T2 at lag-3.

Note, there is no “perfect” window placement in the context of a phenomenon such as the attentional blink in which the T1 and T2 responses, at least to some extent, sit on top of each other. However, as just advocated, the fact that the effects sought for obtain for windows placed using prior literature as a precedent, suggests that these placements are appropriate. Specifically, we considered the ERPs presented in Figure 6, Panel A and Figure 7, Panel A of Craston et al. (2009). These provide us with precedent ERP patterns for the T1, see, for example, the T1 P3 (which peaks around 450msec) when the T2 is at lag 8; the T2 at lag 1 and the T2 at lag 3.

So, to confirm, while the care we have taken and restriction thereby imposed on window placement, may reduce statistical power, they should guard against false positives. This in turn means that significant results found should (within the bounds of statistical significance) be true positives, i.e. valid rejections of the null hypothesis.

Electronic Supplementary Material C

Follow-up statistical analyses were conducted to assess the robustness of the interaction between T2 measurement type (accuracy versus visibility) and lag in Experiment 1. These analyses were designed to confirm that the interaction between accuracy and subjective visibility represents a real cognitive effect and not a statistical artifact.

First, we wanted to confirm that ceiling effects on accuracy or floor effects on subjective visibility were not driving the significant interaction between accuracy and subjective visibility. If floor/ceiling effects differentially impact accuracy and visibility measures, then statistically significant (but spurious) interactions may emerge that are not indicative of true cognitive effects. The first follow-up analysis therefore attempted to match the accuracy and subjective visibility curves as close to one another as possible, by minimizing the sum of squares of the residuals of a regression. Specifically, for each participant, accuracy was regressed onto visibility, and that participant's visibility curve was re-scaled using the outputs of the regression model. This rescales the subjective visibility curve, by weighting (multiplying) each lag by the same constant. This can squeeze or stretch the curve vertically, but it cannot change the relationships across lags (horizontally). Since the sum of squares of the residuals is minimized, this cannot create a spurious interaction that did not already exist – as the curves are being made more similar. Thus, this procedure cannot increase the Type 1 error rate but may increase the Type 2 error rate. That is, the rescaling is conservative. If an effect still remains after rescaling, it is real (in the sense that the accuracy and subjective visibility curves do have different shapes across lags) but some statistical power may be lost. Loss of statistical power is a necessary consequence of relating two dependent measures (accuracy and visibility), where the relationship between units in the two measures, and to ceiling and floor effects is unclear.

After applying the rescaling measure, participant's original accuracy and scaled subjective visibility data were entered into a Type x Lag ANOVA. Confirming our original findings, the interaction between measurement type and lag was significant ($F(5,85)=7.275$, $p<.001$, $\eta^2=.300$). Post hoc contrasts confirmed that the magnitude of the difference between accuracy and scaled subjective visibility was most evident at lag 1. Specifically, the difference at lag 1 was larger than the difference at lag 4 ($p=.011$), lag 6 ($p=.026$) and lag 8 ($p=.006$). No comparisons between any other pairs of lags were statistically significant (smallest $p=.167$).

The second follow-up analysis attempted to hold T2 accuracy constant. To that end, we included data from lags 1 and 8 only, because mean accuracy was equivalent across these two lags. A repeated measures ANOVA with measurement type (accuracy versus visibility) and lag (lag 1 versus lag 8) was employed. Original subjective visibility ratings (and not scaled subjective visibility) were used. The interaction effect was statistically robust ($F(1,17)=36.836$, $p<.001$, $\eta^2=.684$). Pairwise post hoc contrasts confirmed that there were no accuracy differences between lag 1 and lag 8 ($p=.384$). Despite equivalent accuracy across lags 1 and 8, subjective visibility ratings were significantly reduced at lag 1 ($p<.001$).

The third analysis attempted to demonstrate that the interaction between accuracy and subjective visibility was driven by the early lags (especially lag 1). The ANOVA comparing accuracy and subjective visibility across lags was repeated, but data from the early lags were progressively dropped from the analysis. The ANOVA was first conducted without the lag 1 data, then another ANOVA was run without lag 1 and lag 2 data, and so on. As a reminder, when all lags were included, the interaction between accuracy and visibility was statistically robust ($F(5,85)=13.355$, $p<.001$, $\eta^2=.440$). The interaction between accuracy and subjective visibility remained significant without lag 1 data, but the effect size was reduced ($F(4,68)=5.332$, $p=.001$, $\eta^2=.239$). Without lag 1 and lag 2 data, this interaction was also significant, and effect size was further reduced ($F(3,51)=3.009$, $p=.039$, $\eta^2=.150$). Without lags 1, 2 and 3, there was no statistically robust interaction between accuracy and visibility, suggesting that this interaction was indeed driven by the earlier lags ($F(2,34)=1.709$, $p=.196$, $\eta^2=.091$).

Third, we considered whether an integrated percept of T1 and T2 could account for our findings. We reasoned that on integrated percept trials, information regarding the temporal order of T1 and T2 would be lost, so T1 and T2 would be reported in the correct order on half of the integrated percept trials and in an incorrect (swapped) order on the other half of integrated percept trials. Therefore, all trials where T1 and T2 were reported in a swapped order were considered to involve an integrated percept, and were removed from analysis. Given that there should be an equivalent number of integrated percept trials where T1 and T2 are swapped, we removed an equivalent number of trials where T1 and T2 were correctly identified but not swapped. In order to run a conservative analysis (which weighs against finding a significant result) the additional to-be-removed trials were taken from the lowest subjectivity bins. This process of trial removal was conducted individually for each participant. The remaining trials were analysed using a Type x Lag ANOVA. Confirming our findings, and the interaction between measurement type and lag remained significant for both Experiment 1 ($p<.001$) and Experiment 2 ($p<.001$).

Finally, we directly investigated T2|T1 accuracy at the lowest levels of subjective visibility. To that end, T2|T1 accuracy was calculated only on trials where participants selected a visibility rating of 1 (the lowest possible visibility rating, indicating ‘not seen’). For each lag, T2|T1 accuracy was compared with the degree of accuracy expected due to chance (4.76%), using one sample t-tests. In other words, we investigated whether T2|T1 accuracy was greater than 4.76%, at each lag. This analysis was conducted for lags 1 and 3 in Experiment 2, as that is where the trial counts were sufficiently large to examine a specific subjective visibility (200 trials for each of those lags). As expected, accuracy was significantly greater than chance, despite participants indicating that the subjective viability of the target was nil (lag 1: $p<.001$; lag 3: $p=.003$).