Available online at www.sciencedirect.com
ScienceDirect
Consciousness and Cognition 15 (2006) 662–672
Consciousness and Cognition
www.elsevier.com/locate/concog

The electrophysiology of introspection ☆

Morten Overgaard [a,*], Mika Koivisto [b,c], Thomas Alrik Sørensen [d], Signe Vangkilde [d], Antti Revonsuo [b,c,e,f]

[a] *Hammel Neurocenter, Aarhus University Hospital, Denmark*
[b] *Center for Cognitive Neuroscience, University of Turku, Finland*
[c] *Department of Philosophy, University of Turku, Finland*
[d] *Department of Psychology, University of Copenhagen, Denmark*
[e] *Department of Psychology, University of Turku, Finland*
[f] *School of Humanities and Informatics, University of Skövde, Sweden*

Received 28 February 2005
Available online 23 August 2006

Abstract

To study whether the distinction between introspective and non-introspective states of mind is an empirical reality or merely a conceptual distinction, we measured event-related potentials (ERPs) elicited in introspective and non-introspective instruction conditions while the observers were trying to detect the presence of a masked stimulus. The ERPs indicated measurable differences related to introspection in both preconscious and conscious processes. Our data support the hypothesis that introspective states empirically differ from non-introspective states.

Keywords: Attention; Consciousness; EEG; ERP; Introspection

1. Introduction

It is sometimes claimed that introspection is the *sine qua non* for empirical consciousness research (Jack & Roepstorff, 2003; Overgaard, 2003), although other experimental methodologies to study consciousness may exist (see e.g. Gallagher & Brøsted Sørensen, submitted). By 'introspection,' one often refers to the direct attending to one's own conscious state (say, how a cup is experienced), whereas 'non-introspection' refers to attending to a stimulus (say, a cup as an external object) (e.g. Gallagher & Brøsted Sørensen, submitted; Lyons, 1986). It is important to underline that introspective as well as non-introspective states may be conscious. Although it has been argued that the difference between introspective and non-introspective

☆ Morten Overgaard was supported by a grant from Carlsberg Foundation. Mika Koivisto and Antti Revonsuo were supported by the Academy of Finland (project numbers 45704, 47238, and 205661).
* Corresponding author. Fax: +45 89 49 44 00.
E-mail address: neumov@sc.aaa.dk (M. Overgaard).

1053-8100/$ - see front matter
doi:10.1016/j.concog.2006.05.002

consciousness is merely conceptual (Dretske, 1995), a number of empirical scientists speak of introspective and non-introspective conscious states as empirically different (Jack & Shallice, 2001; Lane et al., 2000; Marcel, 1993, 2003; Lutz, Lachaux, Martinerie, & Varela, 2002; Schooler 2002). Even though there has not yet been accumulated much empirical research regarding this distinction, the few existing papers agree with the distinction.

Overgaard and Sørensen (2004) as well as Overgaard, Kauffmann, and Ramsøy (2001) have shown that there are significant differences between conditions where participants are asked to report in an introspective way (about their experiences) and conditions where they respond non-introspectively. Participants were instructed to identify a displayed figure by pointing out its shape, colour and location on three different scales. The responses of the participants were treated as being either 'correct,' 'incorrect' or 'near correct.' 'Near correct'-responses matched stimulus partially correctly (e.g. when they pointed at the same colour as the one presented, but in a brighter or darker tone). It was shown that participants in the non-introspective condition had significantly more correct and incorrect responses, whereas the introspective participants most often were 'near correct.' In addition, participants in the introspective condition tended to be more liberal about their reports of, for example, colour, while the participants in the non-introspective condition tended to show a more conservative style and conformed to specific colour categories.

Marcel (1993) showed a dissociation between responses when using eye blinks, hand movements and verbal reports. The dissociation was shown in a blindsight patient as well as in normal participants. When the patient and the participants were instructed to introspect, they gave the most accurate reports when using eye blinks for "yes-reports," less accurate when using hand movements, and the least accurate when using verbal reports. The blindsight patient could even reply 'yes, I am aware of a light' while at the same time—during the same stimulus trial—reporting 'no' with hand gestures. This pattern was not present when they were told to report non-introspectively. Overgaard and Sørensen (2004) expanded on this experiment and showed that a dissociation between the response modes used by Marcel (1993) was only found when instructing participants before showing a stimulus. When the order of the instruction and stimulus was reversed, no dissociation was found.

This result of Overgaard and Sørensen (2004) could be interpreted to indicate that introspection changes the participants' behaviour only when the instruction is given prior to the stimulus so that participants can build up an expectation of how to report. When considering Marcel's (1993) result showing a difference between modalities when an introspective instruction was given, one could hypothesise that introspection has an effect on perception rather than on retrospective memory processes.

A governing theory about consciousness and introspection argues that introspection is a 'displaced perception'—we come to know about our own mental states by attending to external objects and not by attending to the mental states (Dretske, 1995; Searle, 1992). According to this theory, dissociations between introspection and non-introspective consciousness would result from a change in subjective response criteria when given an introspective instruction. If this is the case, it may be revealed by signal detection theory (Green & Swets, 1966) as a systematic increase in bias. Bias, according to signal detection theory, reflects the subject's individual set or attitude to what and when to report. In the analysis, values from visual sensitivity and report bias are separated, thus giving a distinct measure of both.

Whereas the discussion about introspection seems specific and isolated from "consciousness studies," the status of introspection has been a haunting ghost in experimental psychology. Cognitive scientists use introspection to inform their own work from the earliest steps of pilot studies to the interpretations of results where one's own experiences may be an important guide to understand others (Jack & Roepstorff, 2002). The very choice of terminology and categorisation of mental states into "visual perception," "auditory perception," etc. is based, or at least supported, by the scientist's introspective evidence. Yet, cognitive scientists have been under much influence from widespread beliefs that there are specific and very strong methodological objections against the use of introspective evidence that make introspection useless in experimental research (e.g. Nisbett & Wilson, 1977). The debate has however moved considerably on since these beliefs were the dominant views in psychology. Although it is beyond the purpose of this article to review this debate, one may consult White (1988), Ericsson and Simon (1980), Jack and Roepstorff (2002), Marcel (2003) and Overgaard and Sørensen (2004) for the most important developments.

Considering the status of introspection in experimental contexts together with findings indicating that introspective and non-introspective reports differ, we need further information about the cognitive and neural substrates of introspection to understand what happens when participants are asked to report about their conscious states. Since we at present do not know anything about the neural basis of introspection, an exploratory study capable of revealing any potentially interesting differences between introspective and non-introspective states would be important. We found the event-related potentials (ERP) method of particular interest in an exploratory study, as no previous work has looked into the relation between introspection and neural events with neurophysiological means. Whereas exact spatial localization of the neural events related to introspection is not possible using ERPs, this method may be useful in timing the possible differences between introspective and non-introspective cognitive processes and in establishing the relations between introspection and other cognitive events as much work has been done to study the relations between cognition and evoked potentials (see e.g. Rugg & Coles, 1995).

In the present study, we measured ERPs during a visual task while the participants were either introspecting or not. Based on the experiments reviewed above and assuming that introspective and non-introspective states are different, one can hypothesise that it is possible to isolate neural and cognitive events that are present when a subject is introspecting but are not present when he or she is not introspecting. Second, as introspection refers to direct attending to one's own conscious state, the electrophysiological effects related to introspection should be similar to those typically associated with visual attention (e.g. enhancement of posterior negativity around 200-300 ms and the P3 component, see Hillyard & Anllo-Vento, 1998; Luck & Ford, 1998). Moreover, the high temporal resolution of ERPs allows us to time the possible effects of introspection either on early perceptual processes or on later post-perceptual processes or both.

2. Methods

We tested 18 participants. Out of the test group, we had to reject eight participants due to either too high or too low performance for permitting EEG analyses separately for each variable. The cut-off criterion was 40 accepted trials for misses, hits, and correct rejections in both introspective and non-introspective conditions. The remaining 10 participants 4 males, mean age = 24.4 years, range = 19–31 were healthy, right-handed (Oldfield, 1971) and had normal or corrected-to-normal vision.

The stimuli were presented on a white background (16 cd/m^2). Each trial began with a black-white pattern mask ($0.80° \times 0.80°$; 0.27 cd/cm^2) in the center of the computer screen for 1200 ms. This forward mask served also as the fixation point and a warning signal for the beginning of the individual trials. In half of the trials, the forward mask was replaced by a stimulus which was a grey dot for 83 ms (0.76° in diameter; 12 cd/m^2). In the other half of the trials, a blank screen was presented for 83 ms instead of the stimulus. The stimulus/blank screen was followed by a blank screen for 83 or 116 ms until a mirror-imaged version of the mask appeared (backward mask) for 1200 ms. The mask was a black square ($0.80° \times 0.80°$; 0.27 cd/m^2) with white rectangles of various size and orientation inside it. The inter-trial interval from offset of the backward mask to the onset of the forward mask of the following trial was 1600 ms. The session was divided into 6 blocks of 160 trials each, 80 stimulus present trials and 80 stimulus absent (catch) trials. Prior to the experimental blocks, the participants were given a 60-trial demonstration and practice block in which the stimulus-backward mask interval was gradually decreased to the level used in the experimental blocks.

The order of the instructions was counterbalanced across the participants. For half of the participants, the introspective instruction was given for the first 3 blocks and the non-introspective instruction for the remaining 3 blocks and for the other half, the order was reversed. In the introspective instruction we asked the participants to report yes or no according to whether they had a conscious experience of something being presented between the two masks or not. Here the observers were not allowed to guess but to respond according to their experience: "You are to report 'yes' if you had a visual experience of seeing a dot, that is, you saw a dot on the screen, and 'no' if you did not. It is crucial that your report concerns your experience. That means that you are not to perform any guesses." In the non-introspective instruction the participants were instructed to report whether they thought something was presented on the screen between the two masks: "You are to report 'yes' if you believe that there appeared a dot on the screen and 'no' if you do not. It is of no interest to us whether you actually see the dot or not—your task is simply to guess whether the dot was shown and to do so as correctly as possible." The interpretation of the instruction was discussed with the subjects to make sure they understood the difference between introspecting and not introspecting. Participants were asked to respond by pressing "yes" or "no" button depending on the instructions they received at the beginning of each block. Half of the participants responded with their left hand, the other half with their right hand.

EEG was recorded using tin electrodes attached to Electro-Cap electrode system (Electro-Cap International, Inc., USA) with 10/20 system sites F1, F2, F3, F4, F7, F8, Fz, P3, P4, Pz, C3, C4, Cz, T3, T4, T5, T6, O1 and O2. In addition, an electrode placed below the left eye was used for monitoring vertical eye movements and blinks an electrode 1.5 cm to the right of the right eye was used for monitoring horizontal eye movements. Furthermore we used an electrode on the nose as reference and an electrode between Fz and Cz as ground. EEG was amplified (SynAmps) using a band pass of 0.05–100 Hz. The computer sampled the filtered EEG signal at >200 Hz. A 50 Hz notch filter was used to reduce 50 Hz interference. The impedance of the electrodes was kept below 5 kΩ. Baseline was corrected to the activity in the—100–0 ms preceding the stimulus. Trials with artefacts (>70 μV) in any of the electrodes were rejected off-line. ERPs were averaged separately for hits, misses and correct rejections (Green & Swets, 1966).

2.1. Behavioural results

We devided the subjects' reports into four categories: Hits (stimulus correctly reported), misses (stimuli not reported), false alarms (reported stimulus when none was present), and correct rejections (no stimulus reported when none was present). In Table 1, the data are displayed ordered by instruction.

We calculated d' and Beta values for each subject to distinguish between sensitivity and individual biases according to signal detection theory (Green & Swets, 1966). We did not find any significant differences between d' values across the two instruction conditions using repeated measures ($F < 1$). For bias, participants varied individually and we did not find any indication that one condition gave rise to higher beta values than the other ($F < 1$).

2.2. Electrophysiological results

Fig. 1 presents the grand average ERPs in the introspective condition and Fig. 2 in the non-introspective condition. The time windows for statistical analysis were determined on the basis of grand average waves. The peak amplitudes and latencies for ERPs were analysed in time windows of 100–160 ms (P1), 130–230 ms (N1), 210–370 ms (P300), 300–420 ms (N350) and 450–750 ms (P500). P1, N1, P300 and N350 were analysed in posterior temporal (T5, T6) and occipital electrodes (O1, O2), in which they were most clearly identified. For these time windows, the peak amplitudes and latencies were analysed with 2 (Instruction: Introspection vs. Non-introspection) × 2 (Response) × 2 (Lobe: temporal vs. occipital) × 2 (Hemisphere) analyses of variance (ANOVAs) with repeated measures. P500 showed the strongest peaks in parietal sites and it was analysed with a 2 (Instruction) × 2 (Response) ANOVA in electrode Pz. First we analysed ERPs to hits and misses (Response: hits vs. misses) in the trials which included the stimulus. For easier comparison of the instruction conditions, Figs. 3–5 overlays the ERPs from the introspective and non-introspective conditions in separate panels for hits, misses, and correct rejections in all electrodes. The statistical analysis reported below showed that the ERPs in the instruction conditions differed most clearly over the left posterior sites.

Table 1
Data organised in hits/misses and false alarms/correct rejections according to experimental conditions

Distribution (%)	Introspection		Non-introspection	
	Mean	*SD*	Mean	*SD*
Hits	13,50	6,71	20,77	8,93
Misses	35,02	7,58	23,90	8,84
False alarms	1,86	3,44	3,73	6,86
Correct rejections	47,80	2,55	45,88	6,79
No answer	1,81	2,93	0,73	1,08

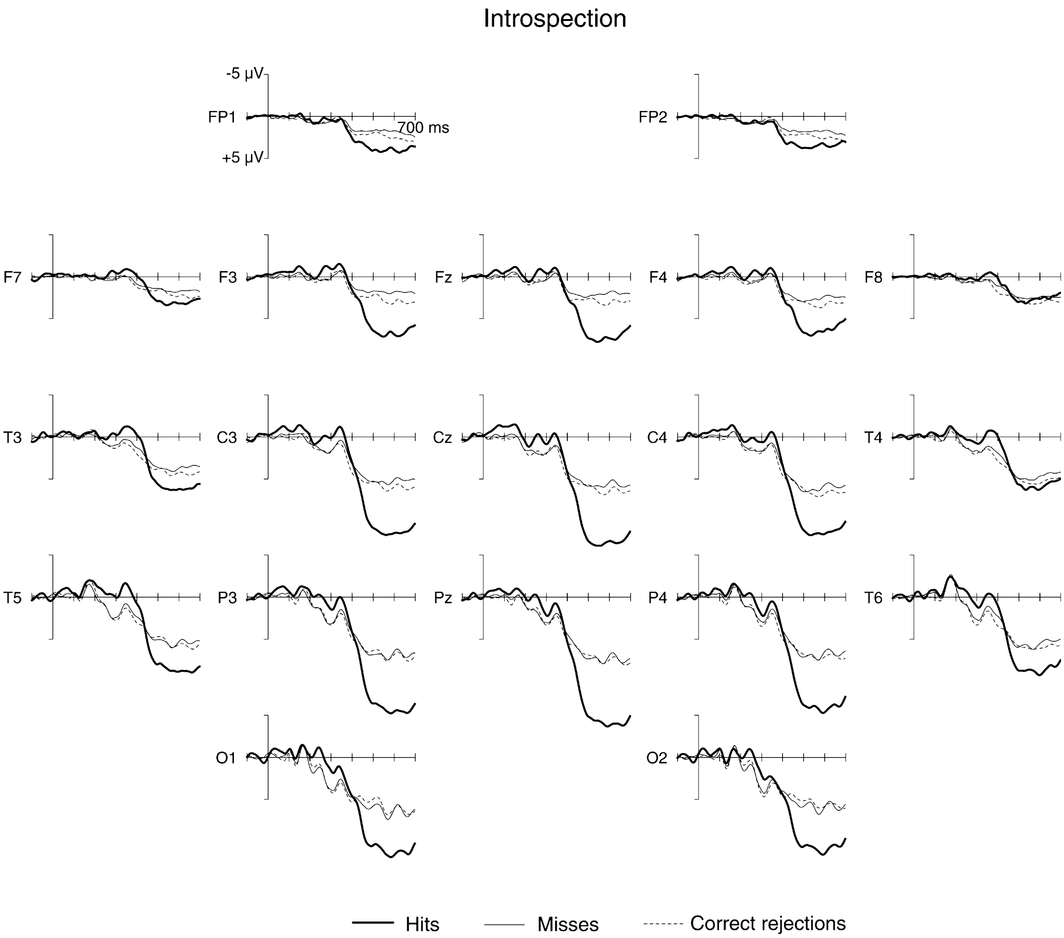

Fig. 1. Event-related potentials (ERPs) to hits, misses and correct rejections in the introspective condition.

2.3. Hits vs. misses

In the P1 time window, the main effect for Lobe was significant for amplitudes ($F(1,9) = 14.81$, $p < .01$) and for latencies ($F(1,9) = 6.88$, $p < .05$), showing that the amplitudes were larger and the latencies were slower over the occipital lobes (1.3 μV, $SD = 1.6$; 132 , $SD = 9.5$) as compared with those over the temporal lobes (0.5 rmuV, $SD = 0.9$; 127 ms, $SD = 9.5$).

The N1 peak amplitudes were more negative in temporal (−2.6 μV, $SD = 1.3$) than in occipital electrodes (−2.1 μV, $SD = 1.6$) ($F(1,9) = 11.79$, $p < .01$), particularly over the right hemisphere (Lobe × Hemisphere: $F(1,9) = 15.97$, $p < .01$). The Response x Hemisphere interaction ($F(1,9) = 5.47$, $p < .05$) suggests that amplitudes to hits and misses differed more over the left (−2.4 μV, $SD = 1.9$ vs. −1.9 μV, $SD = 1.6$) than over the right (−2.5 μV, $SD = 1.6$ vs. −2.3 μV, $SD = 1.6$) hemisphere. However, the amplitude difference between hits and misses was not statistically significant over either of the hemispheres when tested separately. The N1 latencies peaked faster in the introspective condition (175 ms, $SD = 9$) than in the non-introspective condition (182 ms, $SD = 16$) ($F(1,9) = 12.09$, $p < .01$). Furthermore, the Instruction × Hemisphere interaction ($F(1,9) = 6.52$, $p < .05$) showed that the N1 latencies were faster in the introspective condition as compared with the non-introspective condition over the left hemisphere (172 ms, $SD = 13$ vs. 182 ms, $SD = 16$) ($F(1,9) = 17.09$, $p < .01$) but not over the right hemisphere ($F(1,9) = 3.80$, $p = .083$) (178 ms, $SD = 13$ vs. 182 ms, $SD = 19$).

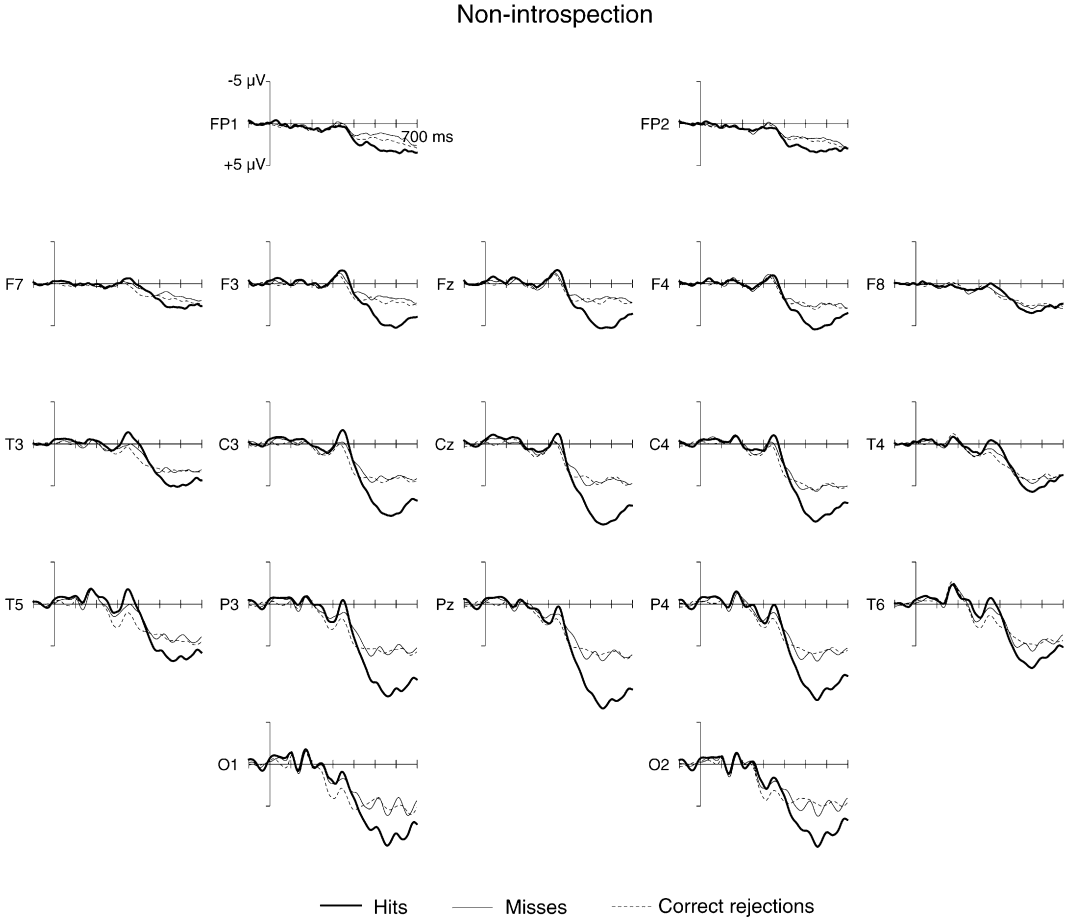

Fig. 2. Event-related potentials (ERPs) to hits, misses and correct rejections in the non-introspective condition.

The analyses of the P300 peak amplitudes revealed a main effect for Lobe ($F(1,9) = 13.40$, $p = .01$), indicating that the amplitudes were the strongest over the occipital lobes. In addition, the Response × Hemisphere interaction was significant ($F(1,9) = 8.07$, $p < .02$), suggesting that the amplitude difference between hits and misses was stronger over the left (2.4 μV, $SD = 2.8$ vs. 3.6 μV, $SD = 2.8$) than the right (3.1 μV, $SD = 2.8$ vs. 3.7 μV, $SD = 2.5$) hemisphere. Further tests of the source of this interaction did not, however, find any statistically significant differences between hits and misses either within each hemisphere or between the hemispheres. Most interestingly, the Instruction × Response interaction was significant ($F(1,9) = 5.61$, $p < .05$). This interaction was due to the fact that the amplitudes to hits were less positive than those to misses in the introspective condition (2.6 μV, $SD = 3.5$ vs. 4.3 μV, $SD = 3.5$) ($F(1,9) = 7.31$, $p < .05$) but not in the non-introspective condition (2.9 μV, $SD = 2.5$ vs. 3.1 μV, $SD = 2.2$) ($F < 1$). The analyses of the P300 peak latencies revealed only a significant main effect for Hemisphere ($F(1,9) = 6.81$, $p < .05$), showing faster latencies over the left (289 ms, $SD = 28$) than the right (296 ms, $SD = 25$) hemisphere.

The N350 peak amplitudes were larger over the temporal than the occipital lobes ($F(1,9) = 22.27$, $p < .001$). The Response × Lobe interaction ($F(1,9) = 7.90$, $p < .05$) suggests that the difference in amplitudes to hits versus misses was larger in the temporal electrodes than in the occipital ones. Further tests showed that hits were associated with larger negativity than misses over the temporal lobes (−1.8 μV, $SD = 3.8$ vs. 0.0 μV, $SD = 2.2$) ($F(1,9) = 5.67$, $p < .05$) but not over the occipital lobes (0.5 μV, $SD = 4.7$ vs. 1.3 μV, $SD = 2.8$) ($F < 1$). The analyses of the N350 peak latencies revealed faster latencies over the left (355 ms, $SD = 16$) than the right (362 ms, $SD = 19$) hemisphere ($F(1,9) = 6.08$, $p < .05$).

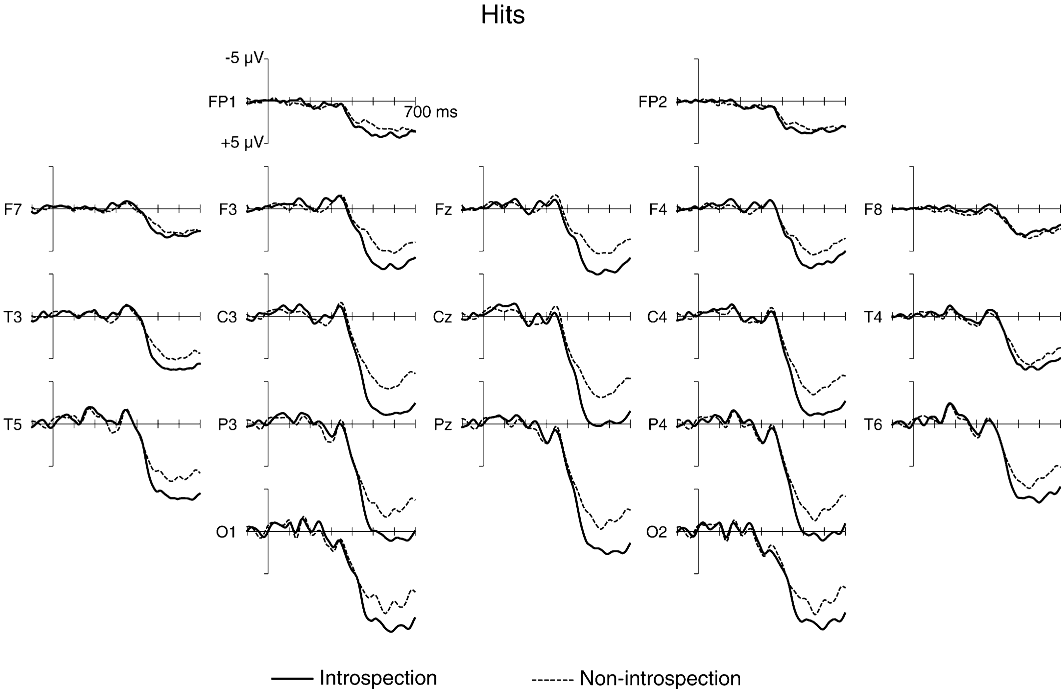

Fig. 3. Event-related potentials (ERPs) in the introspective and non-introspective conditions to hits in all electrodes.

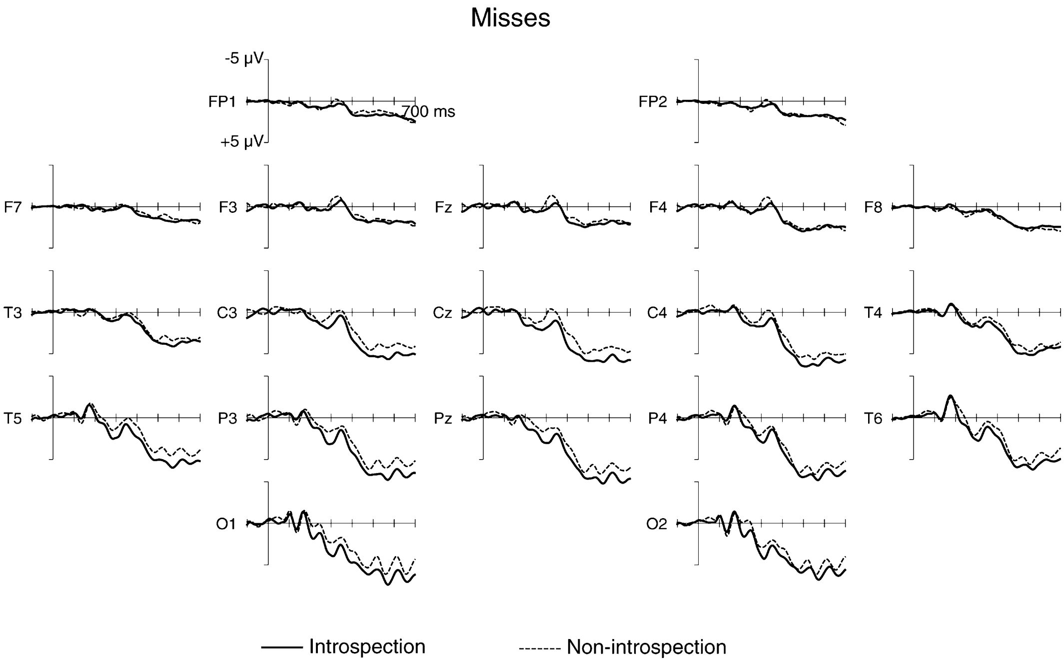

Fig. 4. Event-related potentials (ERPs) in the introspective and non-introspective conditions to misses in all electrodes.

The Instruction × Response ANOVA on the P500 peak amplitudes in Pz found a main effect for Response, $F(1,9) = 22.93$, $p < .01$, showing that the amplitudes to hits were more positive than those to misses. In addition, the Instruction × Response interaction ($F(1,9) = 9.12$, $p < .02$) shows that the difference between hits and

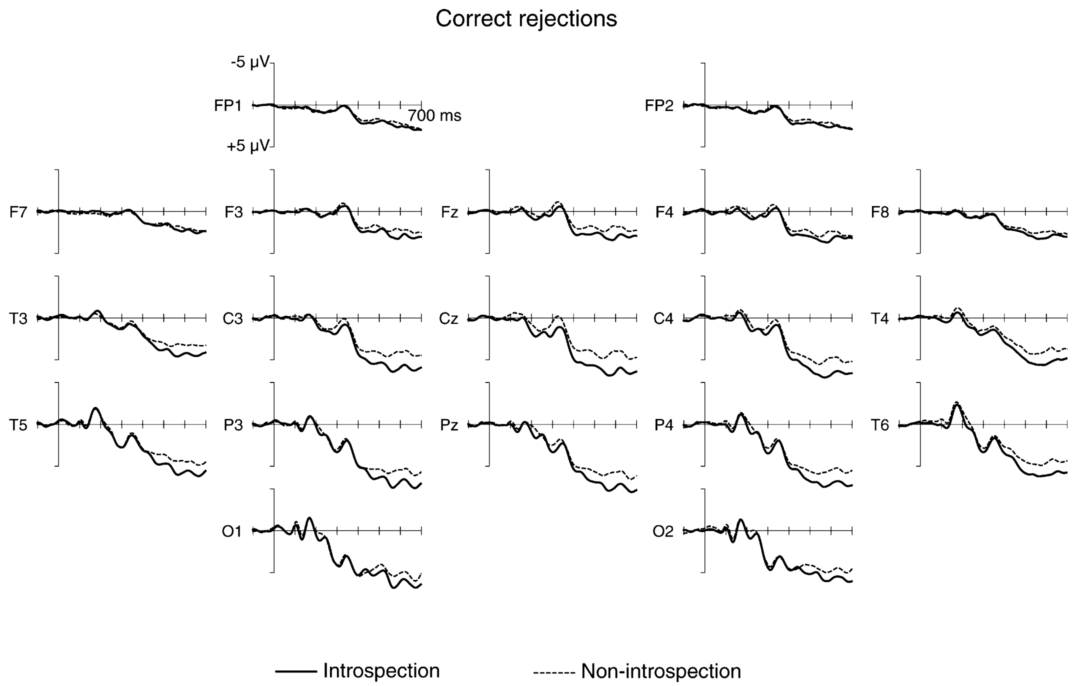

Fig. 5. Event-related potentials (ERPs) in the introspective and non-introspective conditions to correct rejections in all electrodes.

misses is larger under introspection (18.3 μV, $SD = 7.3$ vs. 10.0 μV, $SD = 6.3$) than under non-introspection (14.3 μV, $SD = 5.4$ vs. 8.0 μV, $SD = 2.8$). Further analyses revealed that in spite of the interaction with instruction, the amplitude difference between hits and misses is significant both in the introspection condition ($F(1,9) = 30.35$, $p < .001$) and in the non-introspection condition ($F(1,9) = 13.51$, $p < .01$). Also the analysis of the P500 peak latencies found an Instruction × Response interaction ($F(1,9) = 5.82$, $p < .05$). This interaction was due to slower peak latencies to hits than to misses under introspection (600 ms, $SD = 85$ vs. 526 ms, $SD = 47$) ($F(1,9) = 7.22$, $p < .05$), whereas the latencies did not differ in the non-introspection condition ($F < 1$).

2.4. Misses vs. correct rejections

Implicit processing of the stimulus can be potentially revealed by comparing ERPs to misses (in stimulus-present trials) and correct rejections (in stimulus-absent trials). In both cases, the observers report that there was no stimulus, implying that the stimulus is not represented in visual awareness. Thus, if the comparison shows any differences between the responses to misses and correct rejections, they can be attributed to non-conscious implicit processes. We analysed misses and correct rejections with similar ANOVAs as those involving hits and misses. Here we report only those of the statistically significant effects that include either Instruction or Response as a factor.

The analysis of P1 peak amplitudes showed an Instruction × Lobe × Hemisphere interaction ($F(1,9) = 5.17$, $p < .05$). Further analyses showed that Lobe interacted with Hemisphere only in the non-introspective condition ($F(1,9) = 14.42$, $p < .01$): the difference between left and right hemisphere was stronger in occipital (1.4 μV, $SD = 1.5$ vs. 1.9 μV, $SD = 1.5$) than in temporal (0.7 μV, $SD = 1.3$ vs. 0.6 μV, $SD = 1.0$) electrodes. P1 peak latencies were slightly but statistically significantly shorter in the introspective condition (128 ms, $SD = 9$) than in the non-introspective condition (132 ms; $SD = 6$) ($F(1,9) = 12.38$, $p < .01$).

N1 peak amplitudes did not show any effects involving Instruction or Response. However, also here the peak latencies were shorter in the introspective condition (171 ms, $SD = 6$) than in the non-introspective condition (183 ms, $SD = 16$) ($F(1,9) = 12.38$, $p < .01$). This effect was the strongest over the occipital lobes

(Instruction × Lobe: $F(1,9) = 7.06$, $p = .05$), particularly for misses over the right occipital lobe, Instruction × Response × Lobe × Hemisphere: $F(1,9) = 11.52$, $p = .01$.

P300 peak amplitudes were less positive in response to misses (3.7 μV, $SD = 2.5$) than to correct rejections (4.5 μV, $SD = 2.5$) ($F(1,9) = 5.19$, $p = .05$). The P300 peak latencies were shorter in response to misses (292 ms, $SD = 28$) than to correct rejections (307 ms, $SD = 19$) ($F(1,9) = 6.09$, $p < .05$), particularly in the non-introspective condition, Instruction × Response: $F(1,9) = 6.14$, $p < .05$. The Instruction × Lobe × Hemisphere interaction ($F(1,9) = 13.44$, $p < .01$) is mostly due to the fast latencies in the left temporal electrode under non-introspection (287 ms, $SD = 22$).

The N350 amplitudes were less positive in response to misses (0.7 μV, $SD = 2.5$) than to correct rejections (1.8 μV, $SD = 3.2$) ($F(1,9) = 5.34$, $p < .05$). In addition, the Instruction × Lobe × Hemisphere interaction was significant for N350 amplitudes ($F(1,9) = 6.37$, $p = .05$), but further analyses did not reveal any statistically significant differences between the instruction conditions. The analysis of the P500 peak amplitudes and latencies did not show any effects involving Instruction or Response.

The most interesting part of the electrophysiological data is related to the differences between introspection and non-introspection. To sum up, the N1 latencies peaked faster in the introspective condition. For the P300 amplitudes, the instruction interacted with correctness, so that the amplitudes to hits were more negative than those to misses in the introspective condition, whereas in the N350 time window, hits elicited more negative amplitudes than misses in both instruction conditions. These results regard occipital and posterior temporal recording sites. Furthermore, the P500 peak amplitudes showed that the difference between hits and misses is enhanced in the introspective condition. This latter finding regards electrode Pz.

3. Discussion

To study whether introspective mental states differ from non-introspective mental states, we studied whether the brain's electrophysiological responses differ between introspective and non-introspective instruction conditions, while the participants were trying to detect the appearance of a visual stimulus between forward and backward masks. In the following, we summarize the main findings and discuss the implications of the results for the theories about introspection.

As expected, the introspective instruction elicited different electrophysiological responses as compared with those elicited by the non-introspective instruction. To sum up the most interesting of them, the N1 latencies peaked faster in the introspective condition than in the non-introspective condition. For hits and misses, the difference was significant over the left hemisphere but not over the right. For the P300 amplitudes, the instruction interacted with the correctness of responses in such a way that in occipital and posterior temporal sites the amplitudes to hits were less positive than those to misses in the introspective condition but not in the non-introspective condition, whereas in the N350 range hits were associated with less positivity in both conditions. Furthermore, the P500 peak amplitudes in Pz showed that the difference in positivity between hits and misses was enhanced in the introspective condition.

Introspection is here understood as direct attending to one's own conscious state. The electrophysiological effects of introspection in 210-370 ms (P300) time range fit well to this definition, because attention typically modifies electrophysiological responses by increasing the negativity of amplitudes in this time window in occipito-temporal sites (Hillyard & Anllo-Vento, 1998; Luck & Ford, 1998). Some of the modulations of ERPs by introspection in this time window were related to the difference between hits and misses. Thus, these effects were related to conscious perception, because hit and miss trials were physically identical and they differed only in whether the observers reported having had a conscious experience of the stimulus (hits) or reported not having had a conscious experience of the stimulus (misses). The results are however in contrast with classical notions of introspection in psychology, arguing that introspection is always retrospection (Titchener, 1896). According to this theoretical understanding, electrophysiological effects related to introspection should not be so closely related to attention and perception. Rather, all the electrophysiological effects of introspection should be related to post-perceptual processes. In fact, we observed that the manipulation of instruction modified also (but not only) post-perceptual processes as also mentioned below.

In addition to the effects of instruction on conscious and post-perceptual processing stages, we noted also earlier effects of introspection. The findings showing that introspection had effects on earlier components (P1 and N1 latency) than those in which hits vs. misses (or, we assume, conscious perception vs. unconscious perception) differed, suggest that introspective attention may modify preconscious processes. This finding supports the view that introspection has an effect on early perceptual processes rather than only on post-perceptual processes or memory (Marcel, 1993; Overgaard & Sørensen, 2004). At first, it may seem counter-intuitive that attending to one's own conscious state would influence preconscious processes. However, a large body of electrophysiological findings suggest that attention to supraliminal stimuli enhances early perceptual processes (Hillyard & Anllo-Vento, 1998; Luck & Ford, 1998). Moreover, attention seems to have effects also on processing of masked stimuli that do not reach visual consciousness (Koivisto & Revonsuo, 2004; Koivisto, Revonsuo, & Salminen, 2005). Thus, there is nothing contradictory in assuming that introspection (i.e., attention to one's own conscious state) modifies preconscious processes.

In addition, introspection enhanced the amplitudes of the P500 potential. This potential seems to correspond to the class of classical P3 components which has been associated with updating of working memory and other post-perceptual processes (Donchin & Coles, 1988). Therefore our results suggest that introspection has an effect not only on conscious perception and processing stages preceding it, but also on later post-perceptual processes.

Both in miss and correct rejection trials the observers reported no experience of the stimulus, but the stimulus was present in missed trials. Thus the comparison of misses and correct rejections gives us a handle to analyse whether or not there was any indication of implicit (unconscious) processing of the stimulus. Note that the preconscious effects of introspection in P1 and N1 latency range were not specifically related to processing the stimulus, but they were more general modulatory effects of the instruction. The analyses contrasting the misses and correct rejections found that misses elicited larger negativity than to correct rejections in P300 and N350 time windows, but these differences were not related to the instruction to introspect. The P300 latencies were shorter to misses than correct rejections, particularly in the introspective condition. Attributing these effects to implicit processing should, however, be considered with some caution because hits also showed larger negativity than misses in the same time windows. An alternative explanation would state that the observers were at some elementary level aware of the stimuli in miss trials but that their confidence or subjective conscious experience was weaker as compared to hit trials.

To summarize, the present study shows that brain's electrophysiological responses are differently modified when the participants are asked to introspect as compared with non-introspective instruction. The results suggest that the neural processes associated with introspective and non-introspective states start to differ already in preconscious perceptual stages, and they continue to differ at conscious and late post-perceptual stages. These findings set clear constrains on theories of introspection, suggesting, for example, that the effects of introspection cannot be explained merely by referring to post-perceptual phenomena or to a change in response criteria. Because attention has been shown to modify electrophysiological processes from early perceptual to late post-perceptual processes, the interpretation of introspection as a special kind of attention (attention to one's own conscious state) is consistent with the results showing that introspection has effects on early as well as late processing stages

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.concog.2006.05.002.

References

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 355–372.

Dretske, F. (1995). *Naturalizing the Mind*. MIT Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215–251.

Gallagher, S., & Brøsted Sørensen, J. B. (2006). Experimenting with phenomenology. *Consciousness & Cognition, 15*, 119–134.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley & Sons.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Science United States of America, 95*, 781–787.

Jack A. I., & Roepstorff, A. (2003). Why trust the subject? *Journal of Consciousness Studies*, 10 (9–10), v–xx.

Jack, A., & Roepstorff, A. (2002). Retrospection and cognitive brain mapping: from stimulus-response to script-report. *Trends in Cognitive Science, 6*(8), 333–339.

Jack, A. & Shallice, T. (2001). Introspective physicalism as an approach to the science of consciousness, Cognition special issue edited by S. Dahaene: *The Cognitive Neuroscience of Consciousness*, MIT Press.

Koivisto, M., & Revonsuo, A. (2004). Preconscious analysis of global structure: evidence from masked priming. *Visual Cognition, 11*, 105–127.

Koivisto, M., Revonsuo, A., & Salminen, N. (2005). Independence of visual awareness from attention at early processing stages. *NeuroReport, 16*, 817–821.

Lane, R. (2000). Neural correlates of conscious emotional experience. In R. Lane et al. (Eds.), *Cognitive neuroscience of emotion*. Oxford: Oxford University Press.

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Science United States of America, 95*, 825–830.

Lutz, A., Lachaux, J. P., Martinerie, J., & Varela, F. (2002). Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proceedings of the National Academy of Sciences United States of America, 99*(3), 1586–1691.

Lyons, W. (1986). *The dissapearance of introspection*. MIT Press.

Marcel, A. (1993). Slippage in the unity of consciousness. In G. R. Bock & J. Marsh (Eds.), *Experimental and theoretical studies of consciousness*. John Wiley & Sons.

Marcel, A. (2003). Introspective report: trust, self knowledge and science. *Journal of Consciousness Studies, 9-10*, 167–186.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*, 231–259.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia, 9*, 97–113.

Overgaard, M. (2003). On the theoretical and methodological foundations for a science of consciousness. *Bulletin fra Forum for Antropologisk Psykologi, 13*, 6–31.

Overgaard, M., Kauffmann, O., & Ramsøy, T. Z. (2001). Consciousness and introspection, Toward a science of consciousness: consciousness research abstracts. *Journal of Consciousness Studies*, 182.

Overgaard, M., & Sørensen, T. A. (2004). Introspection distinct from first order experiences. *Journal of Consciousness Studies, 11*(7-8), 77–95.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of Mind—event-related brain potentials and cognition*. Oxford: Oxford University Press.

Schooler, J. (2002). Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences, 6*(8), 339–344.

Searle, J. (1992). *The rediscovery of the mind*. MIT Press.

Titchener, E. B. (1896). *An outline of psychology*. Macmillian.

White, P. (1988). Knowing more about what we can tell. *British Journal of Psychology, 79*, 13–45.