

Trait inferences in goal-directed behavior: ERP timing and localization under spontaneous and intentional processing

Frank Van Overwalle, Sofie Van den Eede, Kris Baetens, and Marie Vandekerckhove
Vrije Universiteit Brussel, Brussels, B-1050, Belgium

This study measured event-related potentials (ERPs) during multiple goal and trait inferences, under spontaneous or intentional instructions. Participants read sentences describing several goal-implying behaviors of a target person from which also a strong trait could be inferred or not. The last word of each sentence determined the consistency with the inference induced during preceding sentences. In comparison with behaviors that implied only a goal, stronger waveforms beginning at ~150 ms were obtained when the behaviors additionally implied a trait. These ERPs showed considerable parallels between spontaneous and intentional inferences. This suggests that traits embedded in a stream of goal-directed behaviors were detected more rapidly and automatically than mere goals, irrespective of the participants' spontaneous or intentional instructions. In line with this, source localization (LORETA) of the ERPs show predominantly activation in the temporo-parietal junction (TPJ) during 150–200 ms, suggesting that goals were detected at that time interval. During 200–300 ms, activation was stronger at the medial prefrontal cortex (mPFC) for multiple goals and traits as opposed to goals only, suggesting that traits were inferred during this time window. A cued recall measure taken after the presentation of the stimulus material support the occurrence of goal and trait inferences and shows significant correlations with the neural components, indicating that these components are valid neural indices of spontaneous and intentional social inferences. The early detection of multiple goal and trait inferences is explained in terms of their greater social relevance, leading to privileged attention allocation and processing in the brain.

Keywords: spontaneous inferences; ERP; LORETA; temporo-parietal junction; medial prefrontal cortex

INTRODUCTION

We look at a person and immediately a certain impression of his character forms itself in us. A glance, a few spoken words are sufficient to tell us a story about a highly complex matter. We know that such impressions form with remarkable rapidity and with great ease

Asch, 1946, p. 258

In social interaction, the perceived intentions and attributes of other people are critical in initiating and maintaining smooth communication and cooperation, as well as avoiding undesirable or dangerous encounters. Sometimes, these social inferences are so essential for our survival that we must make them very rapidly and automatically. Is the other person concealing a present or making a fist? Is he or she friendly or aggressive? Little time to ponder alternatives is available when we are confronted with potentially threatening behaviors. This research is concerned with the processing of goal and trait inferences, the time it requires to make such judgments and which regions in the brain are

involved in this process. Past neuroscientific research has explored how single inferences of other's goals, traits and beliefs are processed in the brain (e.g. Van Duynslaeger *et al.*, 2007; Van der Cruyssen *et al.*, 2008; Van Duynslaeger *et al.*, 2008; for a review see Van Overwalle, 2008). Hence, the critical question addressed in this article is how multiple inferences of this kind interact. Knowing how fast multiple as opposed to single social inferences are made and which brain areas they involve within the same participants, could allow us to gain better insight in the psychological processes and behavioral consequences underlying these inferences.

Given the greater functional relevance of multiple inferences, one might assume that information diagnostic of combined goals and traits is privileged in social processing, leading to faster and stronger brain activation. Such privileged processing might be most relevant under constraints of limited time and resources, when intuitive and quick inferences and decisions are most advantageous. To investigate this, we explored not only multiple inferences when making these intentionally (e.g. when given ample time or incentives to think about them), but also and most critically when making these spontaneously (under little time or incentives, e.g. while doing other routine activities). Many dual-process models in social cognition and social neuroscience distinguish between spontaneous associative

Received 9 July 2008; Accepted 31 December 2008
Advance Access publication 6 March 2009

This research was supported by an OZR Grant of the Free University Brussels to Frank Van Overwalle.
Correspondence should be addressed to Frank Van Overwalle, Department of Psychology, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium. E-mail: frank.vanoverwalle@vub.ac.be.

processes and intentional (controlled) symbolic reasoning (Smith and DeCoster, 2000; Satpute and Lieberman, 2006; Keysers and Gazzola, 2007). Behavioral and neurological research has convincingly demonstrated that goal and trait inferences can be activated spontaneously (e.g. Hassin *et al.*, 2005; Van Duynslaeger *et al.*, 2007; Van der Cruyssen *et al.*, 2008). Making such inferences requires almost no intention or awareness, involves only little mental effort, and is difficult to suppress or modify (for a review, see Uleman *et al.*, 2005).

In line with this functional approach to social inference, recent research has sought more neuroscientific evidence to gain insight in the timing of social brain processes and has turned to event-related potentials (ERP). ERPs are waveforms that reflect electric activity of the brain during responses to specific stimuli, and their timing reveals the processing stage at which social inferences are identified. This research has documented that inferences about another person's goals are made about 200 ms after presenting a critical word implying the goal of an actor (Van der Cruyssen *et al.*, 2008). This timing has been uncovered by comparing behavioral descriptions that are consistent as opposed to irrelevant with previous goal-implicating information about the actor. For instance, after reading about several activities that induce the goal inference 'preparing a party' (e.g. putting up decorations, preparing food, calling friends), our brain very rapidly reacts when irrelevant information is presented (e.g. searching his pet), as shown by early deviations in ERP waveforms. These waveforms show considerable parallels between spontaneous and intentional goal inferences, indicating that goals are inferred automatically at about 200 ms, irrespective of the observer's processing intentions. Research using a similar paradigm has demonstrated that it takes considerably longer—about 400–600 ms post-stimulus—to infer stable personality traits about another person (Van Duynslaeger *et al.*, 2007, 2008). To illustrate, after reading about a person who kisses his daughter and hugs his father which induce the trait inference 'friendly', when that same person subsequently slaps his mother, ERP deviations reveal that our brain detects this trait inconsistency in about half a second.

ERP COMPONENTS

ERPs are manifestations of information processing activities, and different types of ERP components are associated with different functions in this process. Research has revealed two main types which are relevant for social inferences, and which have been uncovered mainly in an 'oddball' paradigm where stimuli are provided that are inconsistent or irrelevant with prevailing stimuli. The first type is the N100–P200 complex which reflects early and automatic feature encoding and classification. The second type is the P300 which occurs later and responds to inconsistencies in comprehension. These characteristics make the N100–P200 and P300 ideally

suited for exploring the neural correlates of automatic (early) and intentional (late) social inferences respectively.

N100–P200 and P300 as indices of inconsistency

The N100–P200 complex consists of an early negative waveform that peaks between 100 and 180 ms after the critical stimulus, followed by a somewhat later positive waveform that peaks at about 200 ms after the critical stimulus. There is evidence that the N100–P200 indexes early sensory stages of feature detection, encoding and classification (Raney, 1993; Luck and Hillyard, 1994; Dunn *et al.*, 1998). For instance, increases in the N100 and P200 peak amplitudes have been related to enhanced attention and resource allocation to processing unexpected or improbable stimuli such as words, auditory and visual information (Raney, 1993; Clark and Hillyard, 1996; Peters *et al.*, 2005). It has also been reported that increased P200 amplitudes are correlated with enhanced retrieval (Smith, 1993; Dunn *et al.*, 1998). These findings have led to the view that the N100–P200 indexes early attention selection and comprehension, often on the basis of physical features of stimuli, in response to new incoming information.

Perhaps more importantly, an enhanced N100 has also been reported for lexical access of low-frequency *vs* high-frequency words (Serenio *et al.*, 1998; Dambacher *et al.*, 2006) and for ambiguous words that are primed with their low-frequency *vs* high-frequency meaning by a preceding sentence context (at 132–192 ms; Serenio *et al.*, 2003). This latter finding indicates that a preceding context may 'direct early, lexical selection of the appropriate meaning' (Serenio *et al.*, 2003, p. 328). This early timing might be important for the present study which involves words that reflect a novel goal or trait meaning induced by a preceding sentence context.

The P300 is a late positive peak that typically initiates around 300 ms (or later) after the critical stimulus and continues till 600 or 1000 ms (and is therefore also termed the *late positive potential*). Research has documented that there is a relation between the P300 and the processing of anomalous, inconsistent or infrequent stimuli presented in a context of otherwise normal or frequent information, as long as this information is relevant for the task. The amplitude of the P300 increases as a function of the amount of discrepancy between the stimulus and the preceding context, and correlates with later recall of the discrepant stimuli, especially when elaborate rehearsal strategies are minimized (Fabiani *et al.*, 1986; Fabiani and Donchin, 1995). These findings have led to the view that the P300 is an index of online updating of working memory after inconsistency detection.

N100–P200 and P300 in social cognition

Several researchers found enhanced ERPs in response to emotional images at very early processing stages. Angry, fearful or happy human faces or armed persons in

comparison to neutral faces or unarmed persons lead to an enhanced N100 (Correll *et al.*, 2006), P200 (Carretié *et al.*, 2001; Eimer and Holmes, 2002; Eimer *et al.*, 2003; Correll *et al.*, 2006; Kubota and Ito, 2007) or P300 (Keil *et al.*, 2002). The N100 is also sensitive to outgroup *vs* ingroup racial differences, while the P200 is additionally sensitive to gender differences, consistent with the idea that outgroup members are often seen as more negative or threatening than ingroup members. Ito and colleagues (Ito and Urland, 2003; Correll *et al.*, 2006; Kubota and Ito, 2007) found that among White participants, Black outgroup faces evoke a larger N100 or P200 than White ingroup faces, and that male outgroup faces evoke a larger P200 than female ingroup faces. Dickter and Bartholow (2007) reported similar race findings among White participants and, importantly, the opposite race pattern among Black participants (i.e. White outgroup faces evoke a larger P200 than Black ingroup faces). This is consistent with the general notion that attention to racial outgroup features increases the P200. The P200 race divergence occurs even when tasks direct attention away from race (Ito and Urland, 2005; Dickter and Bartholow, 2007). Interestingly, the modulation in these early components sometimes predicts behavioral responses, such as faster categorization of angry as compared to happy faces (Kubota and Ito, 2007) or a greater tendency to shoot Black outgroup armed targets as opposed to White ingroup targets in a video game (Correll *et al.*, 2006). These findings of greater attention to more emotional or distinctive social stimuli may reflect an automatic vigilance effect in which attention is quickly and relatively automatically drawn to stimuli with potentially negative implications. None of these studies, however, involves social inferences about the intentions or traits of individual persons.

When considering the more relevant ERP evidence on person inferences, most effects have been documented on the P300, although there is also evidence on earlier components involving the N100 and P200. With respect to goal inferences, the only study to date by Van der Cruyssen *et al.* (2008) described earlier, documented robust P200 peaks reflecting inferences about an actors' goals described in short behavioral sentences. With respect to trait inferences, enhanced N100 peaks during intentional trait inferences following negative behaviors have been reported in one study (Bartholow *et al.*, 2003) but this may reflect automatic attention allocation to negative stimuli that may index a potential threat to the organism. The majority of the evidence points to later processing of trait inferences. To illustrate, Cacioppo and coworkers (Cacioppo *et al.*, 1993; Cacioppo *et al.*, 1994) found that an evaluative inconsistency between a trait word and previously presented trait words (e.g. a negative trait after a sequence of positive traits) elicited a large P300 between ~500 and 1000 ms. As noted earlier, Van Duynslaeger and colleagues (2007, 2008) engaged their participants to form spontaneous or intentional trait impressions (see also Bartholow *et al.*, 2001,

2003) about actors described in short behavioral sentences. The results demonstrated greater P300 activation starting at about 400–600 ms after presenting trait-inconsistent sentences as opposed to trait-consistent sentences, regardless of whether the inference was made spontaneously or intentionally. This implies that traits are generally identified at a late processing stage about 400–600 ms post-stimulus.

SCALP LOCATIONS OF SOCIAL JUDGMENTS

Recent neuroscientific findings involving functional magnetic resonance imaging (fMRI) have established that two regions are predominantly involved in the process of making social inferences. In a recent meta-analysis of social fMRI studies, Van Overwalle (2008; see also Keysers and Perrett, 2004; Keysers and Gazzola, 2007) concluded that the *temporo-parietal junction* (TPJ) implements intuitive and empathic representations about temporary intentions or goals of an actor, perhaps with the aid of an automatic mirror system that recognizes observed intentions of others by their similarity with own behavioral intentions, which are subserved by the *inferior frontal gyrus* (IFG). To illustrate, with respect to goals, when viewing animations of simple objects that move in a human-like fashion, observers have an immediate sense of intentionality (e.g. a triangle 'chases' a square) that activates the TPJ (Martin and Weisberg, 2003; Ohnishi *et al.*, 2004; Schultz *et al.*, 2004). When seeking for or reacting to an appropriate story ending, the TPJ is more strongly activated for stories involving behavioral intentions than mere physical events (Walter *et al.*, 2004; den Ouden *et al.*, 2005; Saxe and Wexler, 2005; Völlm *et al.*, 2006; Blakemore *et al.*, 2007). When reading stories, the TPJ is recruited more when these stories involve beliefs and reasoning by the actors than their physical appearances (Saxe and Powell, 2006). All these tasks do not focus explicitly on the goal underlying the behaviors, and are thus relatively spontaneous.

In contrast, the medial prefrontal cortex (mPFC) is involved in identifying enduring traits of an actor, presumably using more reflective representations and deliberative reasoning. To illustrate, many studies investigating trait inferences found a strong activation of the mPFC (e.g. Mitchell *et al.*, 2005, 2006; Todorov *et al.*, 2007). Generally, this brain area is also involved in controlled processing and action monitoring, and in particular its most posterior part is involved in conflict detection (known as the *anterior cingulate cortex*, but actually encompassing the whole *posterior mPFC*, Van Overwalle, 2008; see also Botvinick *et al.*, 2004).

Rough brain localization of ERPs is now also possible using LORETA source analysis (Pascual-Marqui *et al.*, 1994). Although the spatial resolution of LORETA is poorer than fMRI, it provides the opportunity to slice the activity of the brain into smaller time segments up to milliseconds, unlike fMRI which offers a time resolution in the range of a few seconds. Earlier studies using LORETA have

revealed that goal extraction at about 200 ms mainly involves the TPJ regardless of whether processing is spontaneous or intentional (Van der Cruyssen *et al.*, 2008). In contrast, trait identification at about 600 ms involves the TPJ more strongly under spontaneous inferences while it activates the mPFC more strongly under intentional trait inferences (Van Duynslaeger *et al.*, 2007). Moreover, confirming the functional role of these brain areas, these researchers found a correlation between TPJ and mPFC activity and recall of goals and traits.

To summarize, previous ERP research suggests an early P200 during goal inferences and a late P300 during trait inferences. Moreover, prior fMRI and LORETA imaging research suggests that the strongest brain activity during these social inferences is localized in the TPJ for goal inferences and other spontaneous processes, and in the mPFC for trait and intentional inferences.

PRESENT RESEARCH AND HYPOTHESES

For the present research, we extend the paradigm of Van der Cruyssen *et al.* (2008), and provide short behavioral descriptions that are consistent, inconsistent or irrelevant with the goal implied during preceding behaviors, while some of these descriptions are also consistent, inconsistent or irrelevant with an implied trait about the actor. This latter condition thus implements a novel combination of behaviors that are diagnostic of both goals and traits. We instructed our participants either explicitly to infer the goal or trait of each target person (*intentional social inference*; ISI) or to read the stimulus material carefully, without mentioning anything about a person's goals or traits (*spontaneous social inference*; SSI). If the participants identified the implied goal or trait, then this should show up in the ERP waveforms. Because negative behaviors give rise to more prominent social inferences (Reeder and Brewer, 1979; Cacioppo *et al.*, 1999; Ybarra, 2002) and stronger ERP deflections (Cacioppo *et al.*, 1999; Bartholow *et al.*, 2003; Van Duynslaeger *et al.*, 2007, 2008), we provided only positive goal and trait expectations with negative inconsistencies (e.g. preparing a party is followed by ordering food for oneself only).

When goals and traits interact, it is conceivable that the timing is very similar to previous ERP studies (Van Duynslaeger *et al.*, 2007, 2008; Van der Cruyssen *et al.*, 2008), in which goal inferences reveal an early P200 while trait inferences additionally generate a later P300. However, in line with our functional perspective, we hypothesize that combined goal and trait implications have greater social relevance which invites faster processing, leading to earlier ERPs involving not only a P200, but perhaps even a N100. With respect to localization, however, we believe that there is less flexibility in where and how brain areas are implicated in social computation and therefore suggest the same areas as revealed in earlier research. That is, goal identification mainly involves the TPJ while trait inference or more

generally, intentional social processing, predominantly activates the mPFC (Keysers and Gazzola, 2007; Van Overwalle, 2008). With respect to the impact of instruction, we assume that SSI are made to the same degree as ISI.

As in previous ERP research (Van der Cruyssen *et al.*, 2008; Van Duynslaeger *et al.*, 2007, 2008), we validated the ERP components by a memory measure that proved to be very sensitive in earlier research on spontaneous inferences, cued recall (Winter and Uleman, 1984). After the EEG recording, participants attempt to recall all information presented with the aid of cue words that reflect the implied goal or trait. The idea is that when reading about a behavior that is diagnostic of a goal or trait, the inferred goals or traits are stored in memory together with the sentences from which they are inferred. As a consequence, enhanced recall of consistent behaviors aided by this goal or trait as cue suggests that inferences in terms of goals and traits were made while reading the material. Because earlier ERP research revealed that goals are context dependent and adaptive to new circumstances so that inconsistent behavior may suggest a change of proximal means in the pursuit of the same distal goal (e.g. buying an 'inconsistent' gun for a party suggests a surprise party for Halloween or carnival; see Van der Cruyssen *et al.*, 2008), we expect better memory for consistent and inconsistent behaviors in comparison with irrelevant behaviors. To validate the neural measures, we further predict significant correlations of cued recall with ERP amplitude as well as with LORETA activation, because these neural indices reflect greater brain activity and thus deeper processing, presumably leading to better memory.

METHOD

Participants

There were 50 participants without prior history of any neurological dysfunction. They consisted of 32 students at the Vrije Universiteit Brussel (VUB) from departments other than psychology, and an additional 18 acquaintances of the (female) experimenter. There were 24 women and 26 men, with an age varying between 18 and 34, and an average of 23.46 ± 4.66 years. The students were recruited via a university-wide electronic mailing system for all students, while the additional participants were recruited on a personal basis. This latter subgroup was equivalent to the university students with respect to level of education, but was somewhat older (around 30 years). In exchange for their participation, they were paid 15 euros. Half of the participants were randomly assigned to the spontaneous instruction while the other half received an intentional instruction. The study was approved by the medical ethics committee of the university, and all participants signed an informed consent.

Some EEG data were incomplete due to technical difficulties during recording or after removal of EEG artifacts (i.e. at least one condition was without valid data), and these

participants were removed from the analysis. Specifically, data from two participants in the spontaneous condition and from four participants in the intentional condition were removed.

Stimulus material

The general design was adapted from Van der Cruyssen *et al.* (2008; see also Van Duynslaeger *et al.*, 2007, 2008). Participants read several events that described the behavior of a fictitious target actor and from which a strong goal and/or trait could be inferred or not. The events involved 10 behaviors that implied a positive goal and trait of the actor, 20 behaviors that implied a goal only, and seven positive events that did not imply a goal or trait. Each event was introduced by showing the name of the actor on the computer screen for 2 s. To avoid any association with a familiar and/or existing name, fictitious 'Star Trek'-like names were used (Van Duynslaeger *et al.*, 2007, 2008). For each event, a series of four behavioral descriptions was presented, each consisting of six words shown in the center of the computer screen. Every word was presented during 300 ms followed by a 350 ms blank (Osterhout *et al.*, 1997), except for the last word of a sentence which was the critical one, and was therefore presented for 350 ms to make sure that it would not be missed.

For the events that implied a goal only, the first three sentences implied a common goal (e.g. preparing a party), whereas the last word of the last sentence determined the degree of consistency with the previously implied goal: goal-consistent (GC), goal-inconsistent (GI) and goal-irrelevant (GIR). For instance, after three introductory sentences describing behaviors implicating a goal (e.g. washing a car), the GC-sentences describe behaviors that are consistent with this goal (e.g. 'Divod rubs with a clammy *sponge*'), GI-sentences are inconsistent (e.g. 'Divod rubs with very heavy *mud*'), and GIR-sentences are irrelevant with respect to the goal (e.g. 'Divod rubs softly in his *hands*'; the illustrations are freely translated from Dutch sentences to provide equivalent examples in English). For goal-implicating events that also implied a trait, the first three sentences additionally imply the same trait of an actor, whereas the last word of the last sentence determines the degree of consistency with the previously implied goal and trait: goal + trait-consistent (GTC), goal + trait-inconsistent (GTI) and goal + trait-irrelevant (GTIR). To illustrate, three introductory sentences describe intentional behaviors that imply a goal (i.e. preparing a party) and a trait (e.g. social), and the GTC-sentences describe behaviors that are consistent with these (e.g. 'Triska looks for enough soft *drinks*'), GTI-sentences are inconsistent (e.g. 'Triska looks for food for *herself*'), while GTIR-sentences describe neutral behaviors (e.g. 'Triska looks for dust on *cupboards*'). As can be seen, the different conditions were carefully matched so that the last sentence always contained the same verb. In addition, no-inference sentences described the actor observing a phenomenon from nature

and did not imply a goal or trait. These sentences were included to serve as a baseline for the analyses because they eliminate any tendency to attribute a goal or trait to the actor. They had the same structure so that three introductory sentences implied a common phenomenon (e.g. snowing), whereas the last sentence determined the consistency with this phenomenon: no-inference consistent (NC), no-inference inconsistent (NI) and no-inference irrelevant (NIR).

The sentences were borrowed from Van der Cruyssen *et al.* (2008; goal-only sentences and no-inference sentences) or newly developed in Dutch. To make sure that these sentences invite the same goal or trait inference, participants ($n = 274$) drawn from a similar population as the main experiment, were requested to write down the goal or trait implied by each event (consisting of the introductory and consistent sentences). We retained for the main experiment the events for which more than 70% of the participants indicated the same goal or trait (or a close synonym). For the novel sentences, a second group of participants ($n = 10$) was requested to rate to what extent the implied goal or trait could be applied on each of the sentences using an 11-point scale (0 = *less applicable* to 10 = *more applicable*) and to rate the valence of the behavior in each sentence using an 11-point scale (0 = *negative* to 10 = *positive*). We retained for the main experiment consistent sentences that were applicable to the goal or trait (mean response ≥ 6) and neutral to positive in valence (mean response ≥ 5), inconsistent sentences that were of little applicability (mean response < 4) and rather negative in valence (mean response < 4), and irrelevant sentences that were of moderate applicability (mean response between 3 and 6) and neutral to positive in valence (mean response ≥ 5). These participants also indicated for each of the goal-only and no-inference sentences the trait it implied. Less than 30% of the participants agreed on a common trait, and so all these sentences were retained for the main experiment.

To make sure that the sentences of the goal + trait and goal-only condition were equivalent, we analyzed the goal and valence ratings (including the ratings obtained by Van der Cruyssen *et al.*, 2008, for the sentences that we borrowed from them). An ANOVA conducted on the mean ratings of the introductory sentences with Inference (goal-only, goal + trait) as factor, demonstrated that they did not differ on the goal or valence ratings, $F(1, 271) < 1$. A similar ANOVA on the mean ratings of the last sentences with Consistency (consistent, inconsistent, irrelevant) as an additional factor revealed that Inference did not interact with Consistency on the goal or valence ratings, $F(2, 84) < 1$. However, there was a strong main effect of Inference, indicating that goal + trait sentences had higher goal ratings overall than goal-only sentences, $F(1, 84) = 13.25$, $P < 0.001$. This effect was not significant for the valence ratings, $F(1, 76) < 1.5$. This indicates that the introductory and experimental sentences were equivalent

across the goal + trait and goal-only conditions with respect to their perceived goal implication and valence, although the experimental goal + trait sentences generally had stronger goal implications than the goal-only experimental sentences.

Procedure

After seating the participant, the electrodes for the EEG and EOG were placed at the correct locations. The participants were informed that they would read stories about several persons and that each story would start with a name followed by several sentences about the actor. Next, they received the crucial instruction to 'try to familiarize yourself with the material of the experiment' (SSI, see also Todorov and Uleman, 2002) or to 'try to make an impression of the person his or her personality trait on the one hand and try to make an impression of the goal that the person wants to reach on the other hand' (ISI). Because participants could not read at their own pace but were shown each word at a fixed pace, they were also told to pay as much attention as possible to each word, because they would get questions about them afterwards. It was also emphasized to move and eye-blink as little as possible to limit artifacts in the EEG (Stern *et al.*, 2001).

During reading the EEG was recorded. Afterwards the electrodes for the EEG and EOG were removed. Next, the participants were given the cued recall task. They had to write as many behavioral sentences as possible with the aid of cue words that consisted of all the implied traits (perfect, gleeful, motivated, sybarite, social, perseverant, generous), implied goals (photographing, winning, making promotion, succeeding, relaxing, preparing a party, paying, washing clothes, losing weight, studying, cleaning a car, mowing the lawn, washing, urinating, borrowing a book, traveling, making a film, solving a murder, recording a song, barbecuing, sleeping, repairing, being elected, marrying) and themes of no-inference events (spring, summer, fall, winter, sunrise, snowing).

Electrophysiological registration and analysis

The EEG was recorded from 19 scalp sites according to the international 10–20 electrode system, using Ag/AgCl electrodes fixed in a stretch head cap (Activeshield) from Advanced Neuro Technology. The montage included three midline sites (Fz, Cz, Pz) and eight sites over each hemisphere (Fp1/Fp2, F3/F4, F7/F8, C3/C4, T3/T4, T5/T6, P3/P4, O1/O2), with the average of all EEG-channels as recording and off-line reference. A ground electrode was located along the midline between the Fz and Cz electrodes. Vertical and horizontal eye movements were recorded bipolarly via electrooculograms (EOGs) using electrodes placed above and below the left eye and 1 cm external to the outer canthus of each eye, respectively. Impedance was kept below 10 k Ω . The EEG was recorded continuously during the whole experiment, with a digitizing rate of 256 Hz.

Stimulus presentation, recording and analysis was done with the hardware (Cognitrace) and software (Eevoke, Eemagine and ASA) from Advanced Neuro Technology. The stimulus material was presented on screen and directly time-locked on the EEG recordings by the Eevoke software. After recording, the raw EEG data were filtered by a 0.03–30 Hz band-pass, horizontal and vertical EOG artifacts were corrected using the SOBI algorithm which is based on independent component analysis (ICA) (Joyce *et al.*, 2004), and remaining artifacts beyond -75 and $75 \mu\text{V}$ at each time unit were removed before analysis. For ERP, the relevant EEG sequences were averaged per participants, channel and condition. Each relevant sequence began 250 ms prior to the presentation of the last critical word in each sentence (prestimulus baseline) and lasted to 1125 ms after the presentation of the critical word. A grand average was then calculated on these individual ERP data across all participants.

RESULTS AND DISCUSSION

ERP data

Contrary to our expectations, not the inconsistent descriptions generate the largest peaks, but rather the GTC sentences that confirm the goals and traits implied by the introductory descriptions about a person. These GTC sentences seem to recruit most attention and resources during social processing. Although the direction of the effect is somewhat unexpected, the finding is generally in line with our prediction that a combined inference of goals and traits gets privileged access because of its greater relevance for social interaction and functioning. Figure 1 depicts the waveforms of the midline channels for the crucial goal + trait (GTC) and goal-only (GC) conditions

To statistically analyze the time course of the ERP waveforms, we divided the ERP data of each participant in consecutive time intervals (50–300, 300–450, 450–600 and 600–1000 ms) which are identical to earlier research (Van Duynslaeger *et al.*, 2007; Van der Cruyssen *et al.*, 2008) and identified in each the largest positive and negative peak. These positive and negative peaks were then statistically analyzed for each midline channel separately by means of a Repeated Measures Analysis of Variance (ANOVA) with Consistency (consistent, inconsistent), Inference (goal-only, goal + trait) and Interval (50–300, 300–450, 450–600 and 600–1000 ms) as within-participants factors and Instruction (spontaneous, intentional) as between-participants factor. Reported *P*-levels for this and all other ANOVAs are corrected for violations of the sphericity assumption using the Greenhouse–Geisser correction.

The ANOVA revealed for all midline channels and (positive and negative) peaks a main effect for Interval, $F(3, 123) = 3.07 - 21.41$, $p < 0.05$ (except for negative peaks on Cz), as well as for Inference, $F(1, 41) = 4.50 - 9.05$, $p < 0.05$; and for the Pz channel an effect of Consistency, $F(1, 41) = 5.77 - 15.47$, $P < 0.05$. The positive peaks in

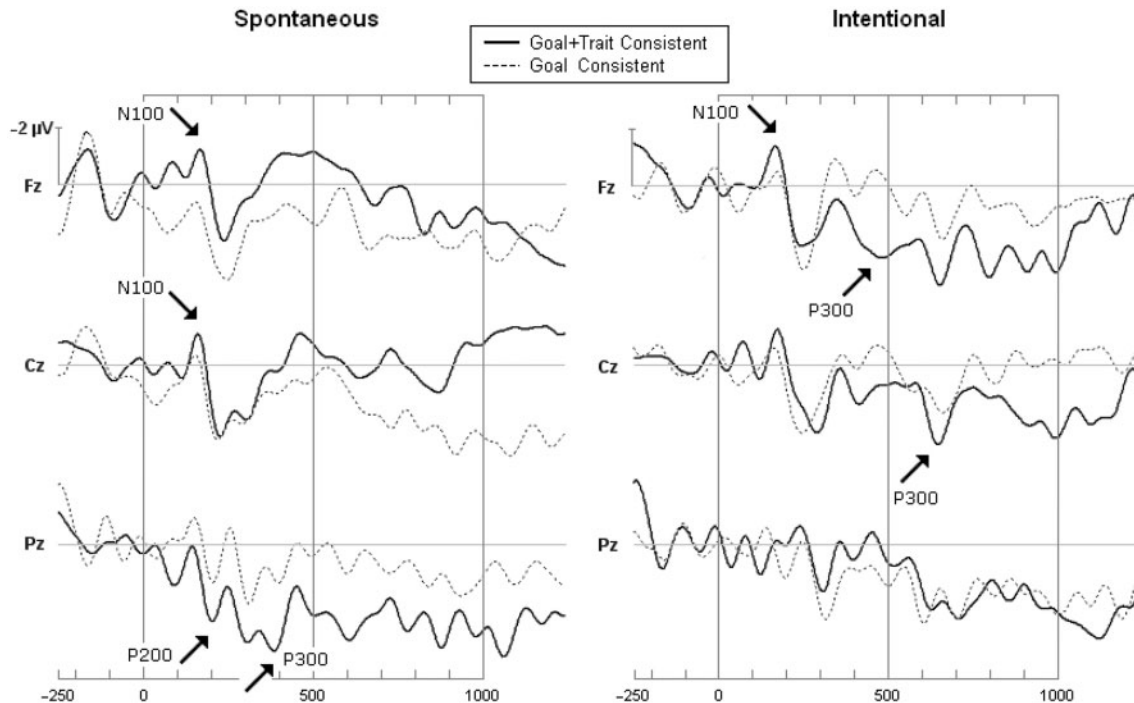


Fig. 1 Effects of inconsistency and type of inference on grand-averaged ERP waveforms showing early negative (N100) and positive (P200) as well as late (P300) positive deflections at the Fz, Cz and Pz scalp sites given spontaneous and intentional instructions. The timeline is given in milliseconds. A positive amplitude is shown downward. The arrows indicate the peak onsets for trait implying words in comparison with no-trait implying words (GTC vs GC), all with significance level $P < 0.05$.

these channels also showed a significant interaction between at least one of these factors and Instruction. In order to test our specific hypothesis concerning the N100, P200 and P300, we conducted simple t -tests comparing the consistent conditions against the inconsistent conditions (i.e. GC vs GI and GTC vs GTI) as well as the consistent conditions amongst one another (GTC vs GC). To control for multiple comparisons, we used a conventional 0.05 α -level for the midline scalp locations in the 50–300 ms interval for the N100 and P200; and in the 300–1000 ms interval for the P300 for which we had *a priori* hypotheses. All significant t -tests were also visually inspected to verify that they reflected peaks in the expected direction. All the other t -tests were corrected for multiple comparisons by a Bonferroni correction, but none of them survived this more stringent criterion. The results for the midline channels are shown in Table 1.

Inconsistencies. Differences between consistent and inconsistent goal + trait conditions are shown in the third part (headed by GTI) of Table 1. There are little significant differences, except for positive ERP deflections under spontaneous instructions at the parietal midline site (Pz) which, as said earlier, go in the direction opposite to expectations, because they are stronger for consistent than inconsistent sentences involving multiple goal and trait inferences (GTC vs GTI). These effects overlap with the effects of multiple inferences reported below, and are

therefore not discussed here. There are no other effects of goal + trait inconsistency. Nor are there effects for goal-only inconsistency as can be seen in the fourth part (headed by GI) of Table 1. The paucity of inconsistency effects might be due to our earlier suggestion that goal inconsistencies might merely suggest a change of means for the same distal intention (i.e. searching for a rifle may imply a Halloween party rather than an ordinary party) and not a real discrepancy (see also Van der Cruyssen *et al.*, 2008).

Multiple vs single inferences. In line with our expectations, multiple inferences of goals and traits elicit stronger ERP deflections than single inferences of a goal only (GTC vs GC). These tests are shown in the second part (headed by GC) of Table 1 and are also depicted in Figure 1. In particular, GTC sentences reveal early negative ERP deflections at frontal midline sites (Fz) for spontaneous and intentional instructions, and also at the central midline site (Cz) for spontaneous instructions. The mean latency of these negative peaks ranges between 130–150 ms for the GTC sentences, indicating an early N100. In addition, these sentences also elicit greater positive peaks at the Pz under spontaneous instructions, with a mean latency at 201 ms indicative of an early P200, and later latencies ranging between 360–819 ms indicative of a late P300. Under intentional instructions, these positive deflections are significant at Fz and Cz sites with a mean latency ranging between 368–788 ms for GTC sentences, indicating only a late P300.

Table 1 Mean Amplitude (in micro volts) of positive and negative ERP peaks as a function of Instruction, Location, Consistency, Inference and Time Interval

Channel	Goal + trait-consistent (GTC)				Goal-consistent (GC)				Goal + trait-inconsistent (GTI)				Goal-inconsistent (GI)			
	50–300	300–450	450–650	650–1000	50–300	300–450	450–650	650–1000	50–300	300–450	450–650	650–1000	50–300	300–450	450–650	650–1000
Positive peaks																
Spontaneous																
Fz	5.45	2.58	3.83	6.49	5.66	4.09	5.33	6.14	5.13	2.46	4.58	5.00	5.29	2.60	4.29	6.09
Cz	6.35	3.38	3.54	5.13	5.77	3.75	4.69	6.72	6.13	4.11	6.22 ^o	8.11	5.08	2.49	4.00	6.20
Pz	7.68	6.97	6.68	8.04	4.90*	3.99**	3.30**	5.33*	4.80*	3.69*	3.85 ^o	5.38 ^o	6.00	4.56	4.65	5.82
Intentional																
Fz	6.34	6.11	6.82	7.70	5.51	2.56**	4.09*	5.47	7.17	4.39	6.71	8.23	4.42	2.28	3.68	4.81
Cz	7.29	5.11	6.30	6.65	4.67 ^o	2.39 ^o	3.56*	3.73**	6.12	3.14	4.28	5.38	5.22	2.98	4.21	4.58
Pz	5.58	4.63	6.96	7.88	4.95	3.75	5.05	5.88	5.07	3.78	3.83 ^o	6.16	3.72 ^o	2.50	3.78	5.42
Negative peaks																
Spontaneous																
Fz	-4.93	-4.87	-4.55	-5.13	-2.25**	-1.95	-3.16	-3.08	-6.20	-5.84	-5.39	-6.78	-3.67	-4.45	-4.01	-3.99
Cz	-4.37	-4.03	-4.16	-5.46	-2.92*	-2.07	-2.67	-1.72	-4.31	-4.18	-3.36	-3.19	-3.13	-3.80	-3.59	-2.84
Pz	-3.49	-1.28	-1.67	-2.10	-3.84	-2.35	-2.86	-2.55	-5.00	-4.19	-5.34	-4.55	-3.80	-2.19	-2.16	-3.24
Intentional																
Fz	-5.19	-3.07	-2.66	-3.66	-3.53*	-3.09	-2.77	-3.07	-6.18	-4.42	-2.79	-4.90	-5.14	-4.90	-3.71	-5.18
Cz	-4.50	-2.51	-2.97	-5.57	-3.56	-3.08	-2.96	-4.43	-6.24	-6.36	-5.83	-7.56	-4.19	-3.42	-3.55	-4.78
Pz	-4.75	-3.68	-2.73	-2.98	-4.16	-2.33	-1.41	-2.39	-7.22	-5.93	-7.04	-7.84	-5.18	-3.94	-4.04	-3.40

Note: Differences across positive peaks between GTC vs GC, GTC vs GTI, GC vs GI conditions and across negative peaks (for 50–300 ms interval only) between GTC vs GC conditions are indicated at the rightmost conditions and denoted as ^o $P < 0.10$; * $P < 0.05$; ** $p < 0.01$ (uncorrected for multiple comparisons).

Although Figure 1 reveals that these ERPs look somewhat noisy (due to the limited number of trials) so that an interpretation in terms of classic components is hampered, the important point is that significant ERP shifts appear at specific post-stimulus time latencies.

Taken together, the GTC sentences generate an early positive deflection at ~ 200 ms and, even more revealing, an early negative deflection at ~ 150 ms, irrespective of spontaneous or intentional instruction. This indicates that multiple inferences of goals and traits are detected at a very early processing stage, even much earlier than mere goals which are detected at about 200 ms as revealed in previous ERP research by Van der Cruyssen *et al.* (2008). Given that lexical selection begins at about 135 ms (Sereno *et al.*, 2003), the detection at about 150 ms of these social inferences may seem quite fast. However, part of this very early detection is most probably due to the fact that it occurs for consistent sentences, where the implied inference has been primed by the preceding sentences so that identification is speeded up. However, the fact that this speeding up takes place only when goal and trait inferences are combined, suggests that this combination is especially socially relevant and therefore has privileged access to attention and cognitive processing.

Localization of ERP

The localization of electric activity in the cortical brain areas was computed with the LORETA inverse solution method (Pascual-Marqui *et al.*, 1994; version 2003; Esslen *et al.*, 2004), which is a reliable method that estimates the ERP source localization with fairly low errors. However, it

should be pointed out that this solution is not unique, and therefore should be corroborated by a priori hypotheses based on alternative methods such as fMRI (as we do here). In a LORETA analysis, a spatial resolution of 7 mm is used per voxel, resulting in a three-dimensional image consisting of 2394 voxels total. Pascual-Marqui (1999) argued that with as little as 16 electrodes 'localization accuracy is at worst in the order of 14 mm' (p. 85).

We focus our analyses on the consistent inference (GC and GTC) conditions that showed the most robust ERP differences. To eliminate electric activation due to mere sentence reading and comprehension, we subtracted from these consistent conditions the no-inference consistent (NC) condition as a baseline (because even goal-irrelevant behaviors often reveal some sense of goal-directedness, see also Van der Cruyssen *et al.*, 2008). Because the ERP waves in Figure 1 indicate that the timing of the critical ERP deflections start at about 150 ms and 200 ms respectively, we computed LORETA brain maps in a broad 100–600 ms time interval around this critical period. We split this time window up in several interval averages (rather than single time points) to obtain robust results with little noise. We first computed averaged LORETAs for each 50 ms interval up to 300 ms to allow visual inspection of shorter changes, and then computed 100 ms intervals up to 600 ms. For a succinct presentation, successive intervals that did not show substantial LORETA differences were further averaged over larger time periods.

Figure 2 displays the LORETA solutions where red denotes areas of the highest activation relative to other areas. The solution reveals maximal activation for two

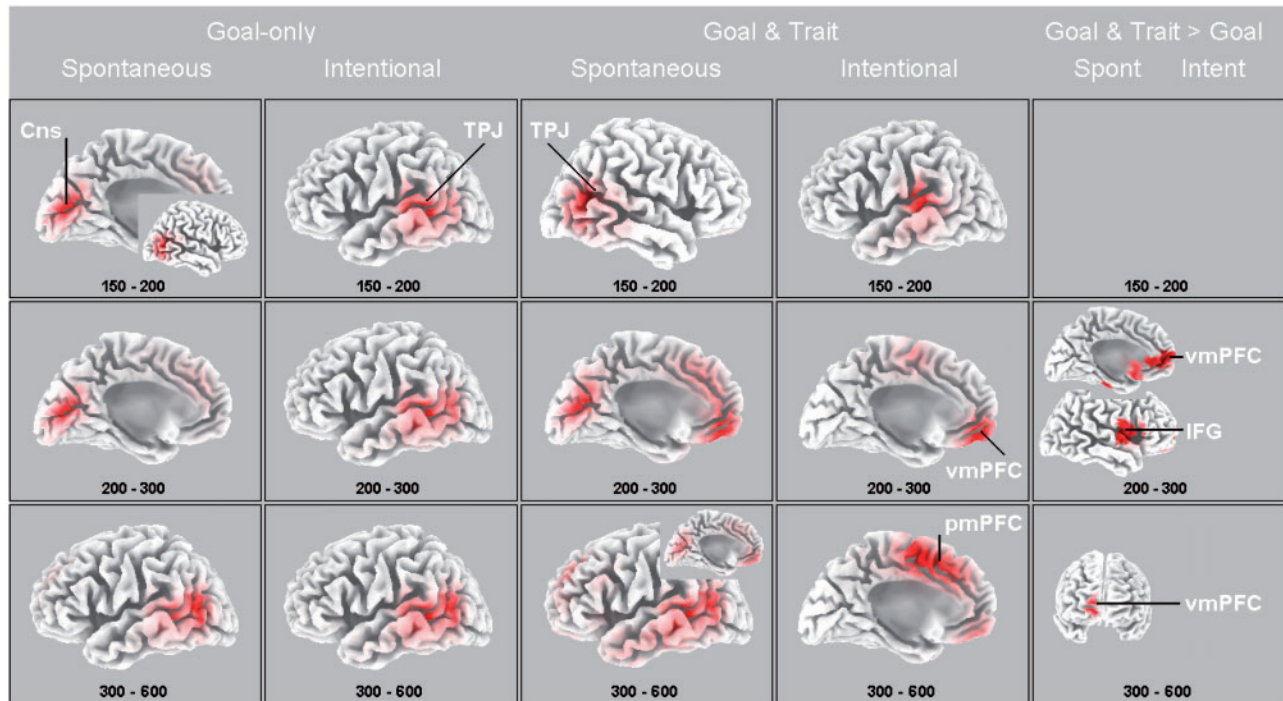


Fig. 2 LORETA source analysis. Columns 1–2 depict the amplitudes of goal-only conditions under spontaneous and intentional instructions, and Columns 3–4 depict the goal + trait conditions under spontaneous and intentional instructions, all averaged during consecutive post-stimulus time segments. The maps are scaled with respect to their minimum and maximum amplitude. The inset in Column 1 reveals that the TPJ is very active during this time interval as well (at least $1.4 \mu\text{V}$ more than in columns 2–4). The last column depicts maps of t -values with red areas indicating significantly greater activation under goal + trait > goal-only conditions under spontaneous (left, $P < 0.05$ uncorrected) and intentional instructions (right, all *ns*). The brain maps display either a medial view of the left hemisphere (oriented from posterior to anterior), a lateral view of the right hemisphere (oriented from posterior to anterior), a lateral view of the left hemisphere (oriented from anterior to posterior), or a frontal view (oriented from right to left). Talairach-coordinates (in x, y, z) of the highest significant differences between goal + trait > goal-only conditions for spontaneous instructions are at 200–300 ms: 4, 50, 8 (mPFC) and 60, 3, 8 (IFG); at 300–600 ms: 11, 66, 1 (mPFC). vmPFC = ventro-medial PFC; pmPFC = posterior medial PFC; for all other abbreviations see text.

predicted brain areas (TPJ and mPFC) and for two other midline areas. One midline area has frequently been revealed in earlier ERP studies (Van Duynslaeger *et al.*, 2007; Van der Cruyssen *et al.*, 2008) and is identified as the *posterior mPFC* (including the *anterior cingulate cortex*, ACC), involved in conflict detection and monitoring (Botvinick *et al.*, 2004; Van Overwalle, 2008). Another midline area, termed the *cuneus* (Cns), has been documented in prior fMRI studies on trait inferences on others or the self (e.g. Fossati *et al.*, 2004; Iacoboni *et al.*, 2004; Johnson *et al.*, 2006; Todorov *et al.*, 2007). The figure displays the LORETA brain maps for goal-only (left panels) and goal + trait (middle panels) conditions. The activation of the TPJ is either shown on the right or left lateral brain surface (depending on where the highest activation was observed), whereas the midline activation is shown on the medial surface. As can be seen, the strongest activation in the TPJ is confined mainly to the inferior part (e.g. Talairach x - y - z coordinates 53, -60 , 15 at 150–200 ms in the spontaneous GTC condition), which is typical for goal inferences (mean coordinates 51, -57 , 25 in the fMRI meta-analysis by Van Overwalle, 2008). The strongest activation of the mPFC is confined mainly to its ventral part (e.g. coordinates -3 , 45, -6 at 200–300 ms in the same condition) typical for inferences on familiar others

(-1 , 49, 4; Van Overwalle, 2008). This latter result may be due to the fact that four sentences were presented about the same actor, so that while reading the last sentence the participants were already somewhat familiar with the actor.

We predicted that the TPJ would be relatively more involved in the early detection of goals, whereas the mPFC would be more involved in the late detection of traits. Our results are consistent with these predictions. To begin, consider the critical timing at 150–200 ms. As can be seen, in all conditions during that interval, the TPJ (extending to posterior areas, which is presumably due to the small errors inherent in LORETA) or cuneus show the strongest activation, which suggests that this area is involved in goal detection (see also Van der Cruyssen *et al.*, 2008). After that period at 200–300 ms, the TPJ and cuneus areas remain activated, except that when a trait is implied, the mPFC is now most strongly activated. This suggests that the mPFC area is involved in trait identification. Afterwards at 300–600 ms, the mPFC remains strongly active at the intentional trait implying condition (and also quite active under the spontaneous trait condition, see insert in Figure 2), but not for the other conditions where the TPJ is most active. Concurrent with the mPFC activation in the intentional trait condition, there is also activation in its posterior part

(or dorsal part of the ACC), which may indicate that participants were vigilant for any potential violation or weakening of the implied trait (under the inconsistent and irrelevant conditions respectively) during the last sentences of each paragraph and therefore experienced a conflict or ‘surprise’ when a consistent trait implication was presented.

We tested differences between spontaneous and intentional instructions, and goal-only vs goal + trait inferences by a voxel-by-voxel *t*-test of the LORETA images (Esslen *et al.*, 2004), using a nonparametric randomization test (Nichols and Holmes, 2002) which corrects for multiple comparisons. None of the LORETA solutions differed significantly between instructions. However, there were robust differences with respect to the implied inference, $P < 0.05$ corrected. The significant *t*-values are displayed via LORETA images in the right most panel of Figure 2, with red indicating stronger activation for goal+trait conditions in comparison with goal-only conditions.

These *t*-values generally confirm the brain maps from Figure 2 (although differences reach significance only under spontaneous instructions) and are therefore not discussed further. One result is noteworthy, though. In addition to greater activation of the mPFC, we also found concurrent activation in the inferior frontal gyrus (IFG, pars opercularis) in the goal + trait as compared to the goal-only condition at 200–300 ms. The IFG has been implicated in the mirror system for automatic goal detection in action (Iacoboni *et al.*, 2005; Molnar-Szakacs *et al.*, 2005) and may perhaps have aided in the goal identification of the present stories when a firm trait inference was also implied.

Memory validation

The responses on the cued recall task were scored on the basis of verbatim accuracy of the sentence (without the actor’s name) although synonyms were allowed. The proportion of correct responses is shown in Table 2.

Inconsistencies. The cued recall measure was taken to verify the hypothesis that goal and trait inferences were made on the basis of diagnostic behavioral descriptions, and requires higher cued recall for consistent or inconsistent behaviors as compared to irrelevant behaviors. The recall data were analyzed with an ANOVA with Instruction

(spontaneous, intentional) as between-participants factor and Inference (goal-only, goal + trait) and Consistency (consistent, inconsistent, irrelevant) as within-participants factors. The ANOVA revealed the predicted main effect of Inference, $F(1, 42) = 38.08$, $P < 0.001$, $\eta^2 = 0.476$, Consistency, $F(2, 84) = 16.71$, $P < 0.001$, $\eta^2 = 0.285$, and their interaction, $F(2, 84) = 6.25$, $P < 0.01$, $\eta^2 = 0.130$. There was no main or interaction effect of Instruction, as predicted. Planned comparisons (irrespective of instruction) indicated that goal and trait cues elicited no stronger recall of consistent sentences than inconsistent sentences. However, as predicted, recall of consistent and inconsistent sentences was stronger than irrelevant sentences for goal cues, $F(1, 42) = 34.26$, $P < 0.001$, and trait cues, $F(1, 42) = 6.98$, $P < 0.05$. This suggests that the implied goal and, to a lesser extent the implied trait, was more strongly associated in memory with consistent and inconsistent behaviors regardless of the spontaneous or intentional instruction. As can be seen also in Table 2, the recall for (in)consistent goals was much stronger than (in)consistent traits, $F(1, 42) = 30.70$, $P < .001$. Although this might seem puzzling because participants had additional trait cues for trait-implying sentences, one likely reason is that goal cues provide more specific information on the gist of the story (e.g. marrying) in comparison with abstract trait cues (e.g. social).

Correlations with ERP deflections. Our prediction was that there would be significant correlations between the ERP components and recall. To control for individual main effects, we compared consistent conditions against inconsistent conditions. Thus, for each participant, we computed a difference score for the ERPs of the consistent minus inconsistent conditions (i.e. GC – GI; GTC – GTI) and a similar difference score for cued recall. We then computed Pearson correlations between these memory indices and the most relevant ERP peaks depicted in Figure 1, that is, positive (P200 and P300) peaks at each interval at the Pz and negative (N100) peaks in the 50–300 interval at the Fz. For spontaneous instructions, there are significant positive correlations between cued recall and positive deflections at the Pz in the 50–650 ms interval for goal-only sentences ($r = 0.48 - 0.59$, $P < 0.05$). In contrast, for intentional instructions, there is a positive correlation between cued recall and negative deflections at Fz in the 50–300 ms interval for goal + trait sentences ($r = 0.57$, $P < 0.01$). No other correlations were significant. Taken together, this indicates that the ERP deflections are valid indicators of goal and trait processing as measured by traditional behavioral memory tasks.

Correlations with LORETA. We computed a Pearson correlation across all participants between cued recall and the activation at the LORETA voxels. To protect against multiple comparisons, we explored only the two critical 150–175 ms and 200–225 ms post stimulus intervals as described earlier and analyzed only the hypothesized brain areas (TPJ and mPFC) at a 0.05 level. Under goal + trait

Table 2 Percentage of correct memory at cued recall as a function of instruction, inference and consistency

	Spontaneous			Intentional		
	Consistent	Inconsistent	Irrelevant	Consistent	Inconsistent	Irrelevant
Goal-implying	18 _{ab}	18 _{ab}	5 _c	24 _a	23 _a	7 _{bc}
Goal+ trait-implying	1 _{ab}	3 _{ab}	0 _b	6 _a	4 _{ab}	2 _{ab}

Note: Means in a row sharing the same subscript do not differ significantly from each other according to a Fisher LSD test, $P < 0.05$.

conditions, trait-cued recall is correlated with mPFC activation under intentional instructions during the 200–225 ms interval, $r=0.56$, $P<0.01$ (Talairach coordinates $-3, 38, 8$). This confirms that the mPFC as revealed by the LORETA analysis is the dominant area for intentional trait inferences. No other correlations were significant.

DISCUSSION

Exploration of the brain has sometimes led us to exiting surprises, and the present results are no different in this respect. In particular, we found that, in comparison with mere goal implications, the confirmation of combined goal and trait implications are detected very early and automatically at about 150 ms as revealed by stronger shifts in the ERP waveforms. This suggests that the higher social relevance of these combined inferences allows privileged access to the brain, resulting in their very early detection. This early trait identification from goal-directed behaviors was subserved by higher activation in the mPFC, in line with current research indicating that this area of the brain is strongly involved in attributions of stable characteristics of other people (cf. Van Overwalle, 2008). Although these findings certainly need corroboration by future research, the present study is the first attempt at unraveling the dynamics of multiple social inferences in the brain.

Neural indices of goal and trait inference

As noted earlier, Van der Cruyssen *et al.* (2008) found a positive ERP deflection at about 200 ms after presenting a critical word that was irrelevant with the implied goal, while inconsistencies with the implied trait revealed a positive deflection at about 400–600 ms (Van Duynslaeger *et al.*, 2007, 2008). In contrast to these earlier findings, the present data show increased ERP shifts beginning at about 150 ms after confirmation of the implied trait and goal. Although an interpretation of this effect is admittedly preliminary at this point, we see two interesting aspects in the present data.

First, the early negative-going ERP shift beginning at ~ 150 ms suggests that a combination of goals and traits increases attention and processing because of its high functional relevance. This is consistent with prior research on attentional selection showing ERP shifts as a function of attention to salient social stimuli that are of inherent interest. To illustrate, seeing salient and potentially threatening Black outgroup faces increases the ERP amplitude as early as 120 ms post-stimulus in comparison with White ingroup faces (Ito and Urland, 2003, 2005). Enhanced ERPs around 150–200 ms have been reported for emotionally salient stimuli of high evolutionary significance, such as erotic images and mutilations (Schupp *et al.*, 2007). Likewise, emotional human faces or armed persons lead to modulation of ERP amplitude after about 200 ms (Carretié *et al.*, 2001; Eimer and Holmes, 2002; Eimer *et al.*, 2003; Schupp *et al.*, 2004; Correll *et al.*, 2006; Kubota and Ito, 2007). Although these

studies present visual stimuli unlike the present verbal material, they demonstrate that stimuli with inherent high significance often generate early ERP shifts, reflecting facilitated perceptual and attentional processing. Perhaps the selective attention towards consistent goal + trait inferences in the present study also explains why the late positive deflection (P300) was enhanced under consistent rather than inconsistent conditions, contrary to what one would expect in ERP research.

Second, the early speed at which the trait consistency is detected is somewhat perplexing. Such early ERP differences (Clark and Hillyard, 1996; Peters *et al.*, 2005) were not expected on the basis of earlier goal and trait inference research involving short behavioral descriptions (Bartholow *et al.*, 2001, 2003; Van der Cruyssen *et al.*, 2008; Van Duynslaeger *et al.*, 2007, 2008). However, this finding is consistent with research on lexical access which shows that low vs high-frequency words lead to enhanced N100 deflections (Dambacher *et al.*, 2006; Sereno *et al.*, 1998) and, more importantly, that a preceding sentence context primes the selection of appropriate low vs high-frequency word meaning as early as 132 ms post-stimulus (Sereno *et al.*, 2003). Applied to the present research, assuming that the implied traits were identified during the preceding sentences, this might have strongly reduced the time necessary to process the appropriate meaning of a critical word in the last sentence, leading to early trait detection and confirmation.

Although the preceding reasoning demonstrates that an early influence on semantic selection is possible, it remains unclear why this occurs specifically for goal-directed behaviors that also imply traits? Perhaps, the reason is that this type of information represents the prototypical case of social processing in the brain as set out in the meta-analysis by Van Overwalle (2008). His hypothesis is that temporarily goal implications are automatically identified in the TPJ, and are then propagated to the mPFC for trait identification. By embedding traits into goal-directed sentences, we have apparently created extremely diagnostic behaviors that fit nicely with this prototypical stream of brain processing. Intuitively, this makes sense. Social traits like friendliness or respect for others do not come without motivation or effort, and therefore require some amount of intention by the actor. Thus, actions that lead one to think that someone invests a great deal of effort to please others are a better precursor for inferring a kindness trait than a simple and perhaps accidental 'hello'. Likewise, executing a murder in several preplanned behavioral stages is more indicative of a cruelty trait than killing someone in the spur of the moment. As these examples tell us, trait inferences derived from premeditated actions have more long-term and functionally significance than trait implications based on one-shot behaviors. A related explanation is that such goal-directed behaviors can be understood as repeating the same trait implications, reinforcing trait attributions by the

consistency principle as proposed in classic covariation theory (cf. Kelley, 1967) and implemented in connectionist approaches to person perception (e.g. Van Overwalle and Labiouse, 2004).

Can this early trait detection be due to some lower-level features of trait-implicating sentences that have little to do with traits in and of themselves? This seems unlikely. The fact that the early ERP deflections were strongest for goal + trait implicating sentences compared to goal-only sentences which all imply positive goals that are neutral to positive in affective content, rules out negative affect as an alternative explanation for our early trait identification (e.g. Eimer and Holmes, 2002; Keil *et al.*, 2002; Eimer *et al.*, 2003; Correll *et al.*, 2006; Kubota and Ito, 2007). The analysis of the pilot data (as reported in the method section) documented that these sentences were equivalent with respect to their perceived valence. Although the goal + trait sentences were generally seen as more goal-directed, this effect appeared across all consistency conditions and thus cannot explain why only consistency of goal + trait led to early processing. Another concern is the reduced amount of critical stimuli for each consistency condition (one-third out of 10 GTC sentences and 20 GC sentences). Although this small number certainly compromises the signal-noise ratio of our ERPs, this was not problematic to obtain significant ERP shifts that were, moreover, replicated across spontaneous and intentional conditions. To conclude, although we cannot exclude lower level perceptual features or other methodological artifacts as an explanation for the high speed of trait identification, there is no evidence for it in the present data.

Brain localization of social inferences

Although localization of ERP signals using LORETA is less precise and limited to the cortex in comparison with fMRI, the present analysis confirms that by focusing on shorter time intervals of these ERP signals, it is possible to detect a progression in time of reliable activation differences between goal and trait inferences, or between spontaneous and intentional instructions that hereto have remained relatively undetected in fMRI research (e.g. Mitchell *et al.*, 2006) because blood de-oxygenation on which fMRI rests, requires at least several seconds.

Our LORETA analysis in Figure 2 indicates that at the early 150–200 ms interval, the most active brain area is the temporo-parietal junction (TPJ) while all differences between trait-implicating and trait-irrelevant material or between spontaneous and intentional mind sets are negligible. That the TPJ dominates brain activation at that stage is consistent with the idea that this area is involved in automatic goal processing at about 200 ms post-stimulus. As noted earlier, the activation of the TPJ is consistent with ERP evidence by Van der Cruyssen *et al.* (2008) as well as other fMRI evidence that seems to suggest that people immediately recognize observed intentions of others

(cf. Van Overwalle, 2008). Behavioral studies have also demonstrated that people can infer goals automatically without awareness or intent (Hassin *et al.*, 2005) and even tend to adopt automatically the perceived goals of others in their own goal pursuit (Aarts *et al.*, 2004).

After this crucial goal-processing stage, at about 200–300 ms, the mPFC is more active under trait processing, irrespective of instruction. Thus although the ERPs indicate that trait identification begins at about 150 ms, it takes about 50 ms more before brain areas responsible for its identification become most active. At 300–600 ms, the mPFC remains most strongly active when traits are inferred intentionally (and also very active when traits are inferred spontaneously), while the TPJ is most active under goal processing or spontaneous instructions. These results are in close agreement with earlier LORETA results documenting more mPFC activation during intentional processing of trait diagnostic information and stronger TPJ activation during spontaneous trait processing (Van Duynslaeger *et al.*, 2007), although in that latter study the differences between spontaneous and intentional trait inferences in the mPFC appeared much later at about 600 ms post-stimulus.

Overall, the results are consistent with earlier fMRI research documenting that the TPJ is recruited during goal identification when observing human-like goal-directed movement (Martin and Weisberg, 2003; Ohnishi *et al.*, 2004; Schultz *et al.*, 2004) or goal-directed behavior (Walter *et al.* 2004; den Ouden *et al.*, 2005; Saxe and Wexler, 2005; Saxe and Powell, 2006; Völlm *et al.*, 2006; Blakemore *et al.*, 2007), whereas the mPFC is implicated in trait processing (Mitchell *et al.*, 2005, 2006; Todorov *et al.*, 2007). The activation of the TPJ during early goal identification followed by the later recruitment of the mPFC provides support for recent insights suggesting that the TPJ involves an early pre-reflective, intuitive and empathic level of representation that does not require conscious reflections, while the mPFC is an area that allows slower explicit and reflective representations that are partly informed by these automatic processes (Keysers and Gazzola, 2007) and subserve trait inferences (Van Overwalle, 2008).

CONCLUSION

The present results attest to the crucial and flexible role of social processing in the brain. Although earlier research showed that goals are identified at an early 200 ms post-stimulus and traits at a later 400–600 ms, the present study suggests that multiple inferences combining both types of social inferences have even more privileged access to brain processing, leading to their detection after only 150 ms. Given that this rapid access involves the confirmation of traits and goals inferred earlier, these results speak to the extremely rapid and automatic nature of the neural computations required for the confirmation of multiple social inferences; they do not tell us how fast these goals and traits were identified in the first place

(during preceding sentences). At present, we assume that the 200 ms and 400–600 ms timings documented in earlier research reflects this type of identification process, whereas the present 150 ms timing of the present study reflects a mere confirmation process after priming. Further research is needed to confirm this early timing, and to exclude lower-level perceptual features that are perhaps correlated with trait-implicating behaviors. Nevertheless, the same brain areas are involved for the same type of inferences and under the same type of processing instructions as in past ERP research. The TPJ is involved during goal confirmation regardless of instructions, while the mPFC is active during trait confirmation, the latter especially under intentional instruction (whereas the TPJ is more involved under spontaneous trait confirmation). The consistent recruitment of specific brain areas for social inference, but at a different timing for goals, traits or their combination, is a unique contribution of recent ERP research.

REFERENCES

- Aarts, H., Gollwitzer, P.M., Hassin, R.R. (2004). Goal contagion: Perceiving is for pursuing. *Journal of Personality and Social Psychology*, *87*, 23–37.
- Asch, S.E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–90.
- Bartholow, B.D., Fabiani, M., Gratton, G., Bettencourt, B.A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science*, *12*, 197–204.
- Bartholow, B.D., Pearson, M.A., Gratton, G., Fabiani, M. (2003). Effects of alcohol on person perception: A social cognitive neuroscience approach. *Journal of Personality and Social Psychology*, *85*, 627–38.
- Blakemore, J.S., den Ouden, H., Choudhury, S., Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive Affective Neuroscience*, *2*, 130–9.
- Botvinick, M.M., Cohen, J.D., Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: Un update. *Trends in Cognitive Sciences*, *8*, 539–46.
- Cacioppo, J.T., Crites, S.L. Jr., Gardner, W.L., Berntson, G.G. (1994). Bioelectrical echoes form evaluative categorizations: I. A late positive brain potential that varies as a function of trait negativity and extremity. *Journal of Personality and Social Psychology*, *67*, 115–25.
- Cacioppo, J.T., Crites, S.L., Berntson, G.G., Coles, M.G.H. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychological Science*, *4*, 108–12.
- Cacioppo, J.T., Gardner, W.L., Berntson, G.G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, *76*, 839–55.
- Carretié, L., Mercado, F., Tapia, M., Hinojosa, J.A. (2001). Emotion, attention, and the ‘negativity bias’, studied through event-related brain potentials. *International Journal of Psychophysiology*, *41*, 75–85.
- Clark, V.P., Hillyard, S.A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, *8*, 387–402.
- Correll, J., Urland, G.R., Ito, T.A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology*, *42*, 120–8.
- Dambacher, M., Kliegl, R., Hofmann, M., Jacobs, A.M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, *1084*, 89–103.
- den Ouden, H.E.M., Frith, U., Frith, C., Blakemore, S.-J. (2005). Thinking about intentions. *NeuroImage*, *28*, 787–96.
- Dickter, C.L., Bartholow, B.D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Social, Cognitive and Affective Neuroscience*, *2*, 189–98.
- Dunn, B.R., Dunn, D.A., Languis, M., Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain and Cognition*, *36*, 355–76.
- Eimer, M., Holmes, A. (2002). An ERP study on the time course of emotional face processing. *NeuroReport: For Rapid Communication of Neuroscience Research*, *13*, 427–31.
- Eimer, M., Holmes, A., McGlone, F.P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cognitive, Affective, and Behavioral Neuroscience*, *3*, 97–110.
- Esslen, M., Pascual-Marqui, R.D., Hell, D., Kochi, K., Lehmann, D. (2004). Brain areas and time course of emotional processing. *NeuroImage*, *21*, 1189–203.
- Fabiani, M., Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 224–40.
- Fabiani, M., Karis, D., Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology*, *23*, 298–308.
- Fossati, P., Hevenor, S.J., Lepage, M., et al. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage*, *22*, 1596–604.
- Hassin, R.R., Aarts, H., Ferguson, M.J. (2005). Automatic goal inferences. *Journal of Experimental Social Psychology*, *41*, 129–40.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., Rizzolatti, G. (2005). Grasping the intentions of others with one’s own mirror neuron system. *PLoS Biology*, *3*, e79.
- Iacoboni, M., Lieberman, M.D., Knowlton, B.J., et al. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage*, *21*, 1167–173.
- Ito, T.A., Urland, G.R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*, 616–26.
- Ito, T.A., Urland, G.R. (2005). The influence of processing objectives on the perception of faces: an ERP study of race and gender perception. *Cognitive, Affective, and Behavioral Neuroscience*, *5*, 21–36.
- Johnson, M.K., Raye, C.L., Mitchell, K.J., et al. (2006). Dissociating medial frontal and posterior cingulate activity during self-reflection. *Social Cognitive and Affective Neuroscience*, *1*, 56–64.
- Joyce, C.A., Gorodnitsky, I.F., Kutas, M. (2004). Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology*, *41*, 313–25.
- Keil, A., Bradley, M.M., Hauk, O., Rockstroh, B., Elbert, T., Lang, P.J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, *39*, 641–9.
- Kelley, H.H. (1967). Attribution in social psychology. *Nebraska Symposium on Motivation*, *15*, 192–238.
- Keysers, C., Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences*, *11*, 194–6.
- Keysers, C., Perrett, D.I. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences*, *8*, 501–7.
- Kubota, J.T., Ito, T.A. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology*, *43*, 738–52.
- Luck, S.J., Hillyard, S.A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*, 291–308.
- Martin, A., Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cognitive Neuropsychology*, *20*, 575–87.
- Mitchell, J.P., Banaji, M.R., Macrae, C.N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*, 1306–15.
- Mitchell, J.P., Cloutier, J., Banaji, M.R., Macrae, C.N. (2006). Medial prefrontal dissociations during processing of trait diagnostic and

- nondiagnostic person information. *Social, Cognitive and Affective Neuroscience*, 1, 49–55.
- Molnar-Szakacs, I., Iacoboni, M., Koski, L., Mazziotta, J.C. (2005). Functional segregation within pars opercularis of the inferior frontal gyrus: Evidence from fMRI studies of imitation and action observation. *Cerebral Cortex*, 15, 986–94.
- Nichols, T.E., Holmes, A.P. (2002). Nonparametric permutations tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15, 1–25.
- Ohnishi, T., Moriguchi, Y., Matsuda, H., et al. (2004). The neural network for the mirror system and mentalizing in normally developed children: An fMRI study. *NeuroReport*, 15, 1483–7.
- Osterhout, L., Bersick, M., McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition*, 25, 273–85.
- Pascual-Marqui, R.D. (1999). Review of methods for solving the EEG inverse problem. *International Journal of Bioelectromagnetism*, 1, 75–86.
- Pascual-Marqui, R.D., Michel, C. M., Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18, 49–65.
- Peters, J., Suchan, B., Zhang, Y., Daum, I. (2005). Visuo-verbal interactions in working memory: Evidence from event-related potentials. *Cognitive Brain Research*, 25, 406–15.
- Raney, G.E. (1993). Monitoring changes in cognitive load during reading: An event-related brain potential and reaction time analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 51–69.
- Reeder, G.D., Brewer, M.B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, 86, 61–79.
- Satpute, A.B., Lieberman, M.D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Research*, 1079, 86–97.
- Saxe, R., Powell, L.J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, 17, 692–9.
- Saxe, R., Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, 43, 1391–9.
- Schultz, J., Imamizu, H., Kawato, M., Frith, C.D. (2004). Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *Journal of Cognitive Neuroscience*, 16, 1695–705.
- Schupp, H.T., Öhman, A., Junghöfer, M., Weike, A.I., Stockburger, J., Hamm, A.O. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion*, 4, 189–200.
- Schupp, H.T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A.I., Hamm, A.O. (2007). Selective visual attention to emotion. *The Journal of Neuroscience*, 27, 1082–9.
- Sereno, S.C., Brewer, C.C., O'Donnell, P.J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science*, 14, 328–33.
- Sereno, S.C., Rayner, K., Posner, M.I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *NeuroReport*, 9, 2195–200.
- Smith, E.R., DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108–31.
- Smith, M.E. (1993). Neurophysiological manifestation of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience*, 5, 1–13.
- Stern, R.M., Ray, W.R., Quigley, K.S. (2001). *Psychophysiological recording*, 2nd edn. New York: Oxford University Press.
- Todorov, A., Uleman, J.S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, 83, 1051–64.
- Todorov, A., Gobbini, M.I., Evans, K.K., Haxby, J.V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, 45, 163–73.
- Uleman, J.S., Blader, S.L., Todorov, A. (2005). Implicit impressions. In: Hassin, R.R., Uleman, J.S., Bargh, J.A., editors. *The New Unconscious*. New York: Oxford University Press, pp. 362–92.
- Van der Cruyssen, L., Van Duynslaeger, M., Cortoo, A., Van Overwalle, F. (2008). ERP time course and localization of spontaneous and intentional goal inferences. *Social Neuroscience*.
- Van Duynslaeger, M., Sterken, C., Van Overwalle, F., Verstraeten, E. (2008). EEG components of spontaneous trait inferences. *Social Neuroscience*, 3, 164–77.
- Van Duynslaeger, M., Van Overwalle, F., Verstraeten, E. (2007). Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, 2, 174–88.
- Van Overwalle, L. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review*, 8, 28–61.
- Van Overwalle, F. (2008). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, in press.
- Völlm, B.A., Taylor, A.N.W., Richardson, P., et al. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, 29, 90–8.
- Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., Bara, B.G. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience*, 16, 1854–63.
- Winter, L., Uleman, J.S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, 47, 237–52.
- Ybarra, O. (2002). Naïve causal understanding of valenced behaviors and its implications for social information processing. *Psychological Review*, 128, 421–41.