The on-line processing of socio-emotional information in prototypical scenarios: inferences from brain potentials

Hartmut Leuthold, Ruth Filik, Kirsty Murphy, and Ian G. Mackenzie

INTRODUCTION

Imagine the following scenario: Your friend, who is a doctor, has just sent you an email telling you that one of his patients intends to put in a formal complaint against him. On reading this, it is very likely that you envisage that your friend must feel upset or angry. If you then learn that he is actually amused by the patient’s actions, you might be surprised by his reaction and try to understand or explain it. What this scenario illustrates is that, although we cannot directly observe the internal states of others, we readily infer other people’s thoughts and feelings, in terms of our ‘mental model’ of their likely behavior. Sometimes new, inconsistent information about the other person may become available, requiring us to update this mental model. The goal of the present study is to investigate some of the brain mechanisms that are associated with such socio-emotional expectations and their on-line evaluation. To this end, different prototypical social scenarios were presented to readers, and event-related brain potentials (ERPs) were analyzed to critical words that specified a character’s socio-emotional reaction to a particular situation.

The capacity of humans to attribute mental states to others is often referred to as mentalizing, mindreading, or theory of mind (ToM). How we infer the mental states of others, their thoughts, beliefs, desires and feelings in order to understand and predict their behavior has been extensively investigated in various areas of psychology (e.g. Heider, 1958; Leslie, 1987; Perner, 1991; Karniol, 2003; Saxe and Wexler, 2005; Frith and Frith, 2003; Siemer and Reisenzein, 2007). Currently, the two major theoretical views about ToM are the theory–theory account, which assumes that we use a naïve theory about the causal interactions of mental states (e.g. Leslie, 1987; Perner, 1991) and simulation theory (e.g. Gordon, 1986; Decety and Grèzes, 2006), which postulates that we simulate others’ minds by putting ourselves (our minds) in their shoes. Whatever the precise mechanisms are, people are likely to take advantage of their memories about social situations similar to the one they are currently encountering. For example, it has been proposed that we use generalized representations in long-term memory (LTM) about how people behave in certain social situations (e.g. in a restaurant), so-called scripts (e.g. Schank and Abelson, 1977), as well as autobiographical information about specific events (e.g. our first job interview) (e.g. Frith and Frith, 2003; Karniol, 2003).

The brain mechanisms contributing to ToM are typically examined by presenting participants with videos or verbal narratives (fictional short stories) that invite inferences about the story characters (Saxe, 2006). This seems justified,
as narratives are assumed to afford cognitive and emotional processes that are similar, if not identical, to those involved during encounters with people in the real world (e.g. Gerrig, 1993; Mar and Oatley, 2008). Employing this approach, functional neuroimaging studies of ToM have repeatedly indicated that the medial prefrontal cortex (mPFC), the right temporal parietal junction (rTPJ) and the posterior superior temporal sulcus (pSTS) are activated in mindreading tasks (for reviews, see Frith and Frith, 2003; Saxe, 2006). The anterior temporal lobe (aTL) has also been implicated in the processing of social information (e.g. Zahn et al., 2007; Simmons et al., 2010; for a review see Olson et al., 2007). For example, Frith (2007) proposed that the aTL is the part of the mentalizing network where general knowledge about social situations (e.g. scripts) is integrated with personal- and context-specific information (but see Ross and Olson, 2010).

Functional neuroimaging studies concerned with text comprehension have shown that a similar brain network is activated while participants read narratives (for reviews, see Ferstl et al., 2008; Mason and Just, 2009). In text comprehension research, it is assumed that information provided by the text is rapidly integrated with information already possessed by the reader (general world knowledge), resulting in the construction and updating of a situation model (Kintsch, 1988). A situation model is held to encode various text dimensions such as time, space, causation, motivation, and, importantly, information about story characters (e.g. Zwaan and Radvansky, 1998; Zwaan, 2004). Ferstl et al. (2005) investigated the brain regions that contribute to the creation and updating of situation models by presenting participants with brief stories and asking them to judge the stories for their temporal or emotional consistency. In-line with behavioral studies measuring sentence reading times (e.g. Gernsbacher et al., 1992, 1998; de Vega et al., 1996), response times after the last word were longer for sentences that contained emotion words that were inconsistent rather than consistent with the emotional state implied by the context (e.g. someone enjoying the perfect end-of-year party feeling sad vs happy, respectively). This result suggests that readers do seem to infer the likely feelings of story characters. Crucially, the right aTL was more strongly activated following an inconsistent than a consistent word, which Ferstl and colleagues took to reflect the higher demands of building a situation model in the case of an inconsistency. In addition, the dorsomedial PFC (dmPFC) was more strongly activated during inconsistent than consistent emotional stories. Ferstl et al. related this dmPFC activity to emotional inference/ToM demands.

A crucial limitation of behavioral and fMRI studies is that they do not allow a detailed assessment of the time course of the mechanisms involved in the processing of socio-emotional information. For example, it remains unclear how rapidly the emotional violation is actually detected. Thus, revealing that a certain (emotional) behavior that mismatches the situational context is immediately registered in the brain would provide strong evidence for the view that readers do indeed create a situation model, and infer on-line the story character’s likely thoughts and feelings. In our view, ERPs with their millisecond temporal resolution are ideally suited to address this issue, particularly in the light of previous studies reporting that ERP components can sensitively reveal certain aspects of social information processing (for a review, see Bartholow and Dickter, 2007), as will be outlined in the following.

**ERP components in social cognition**

A number of ERP components seem relevant to the current study, in particular, the N400 and P300. The N400 manifests as a centroparietally distributed, negative-going deflection in the ERP with an onset around 200 ms and a peak at ~400 ms. Words that are unexpected (low cloze probability), or a poor fit with context either at the sentence level or discourse level elicit a larger N400 than those that are expected (high cloze probability) or a good fit (for reviews, see Kutas and Federmeier, 2011; Van Berkum, in press). Such an N400 effect is also triggered by violations of world knowledge (Hagoort et al., 2004; Filik and Leuthold, 2008). Moreover, in a script priming study, Chwilla and Kolk (2005) found a larger N400 to the final target word for script-unrelated (e.g. vacation–trial–dismissal) than for script-related word triplets (e.g. director–bribe–dismissal). Accordingly, this N400 effect appears to sensitively indicate the immediate access of script-related information and its integration into the context.

Currently, the N400 is viewed as relating to the ease of either integrating the meaning of an incoming word into the sentence context and/or retrieving conceptual knowledge from LTM (cf. Kutas and Federmeier, 2011; van Berkum, in press). Interestingly, Van Berkum et al. (2008) found a small but significant N400 effect for stereotype violations concerning age, gender and social class. For example, a larger centroparietal N400 was elicited if a person with an upper class rather than a lower class accent uttered the sentence ‘I have a large tattoo on my back’. In addition, White et al. (2009) reported an N400-like effect for gender stereotype violations in an associative priming paradigm, although it must be noted that this effect was maximal over fronto-central electrodes.

Other researchers, however, have reported the P300 component to be indicative of processes involved in social cognition. For example, Bartholow et al. (2001) asked participants to intentionally infer either positive or negative character traits from a short passage of text (e.g. they might infer that someone is a friendly person). A larger centroparietally distributed ERP positivity, which the authors interpreted as being related to the P300 component (e.g. Donchin and Coles, 1988), was observed after final affective words in the test sentence that described trait-inconsistent (‘...gave his wife a slap’) rather than trait-consistent behavior (‘...gave his mother a kiss’) (see also Van Dyunsleger
On-line socio-emotional processing

et al., 2007, 2008). On the basis of this P300 trait-consistency effect, Bartholow et al. argued that traits had been inferred and activated before the presentation of the critical word. In addition, and in line with the context updating theory (e.g. Donchin and Coles, 1988), they assumed that the P300 effect reflects the revision of one’s person-specific model in working memory.

In summary, a number of ERP components have been reported to reflect subtle differences in the mechanisms underlying social information processing. However, at least to our knowledge, it is currently unclear which brain processes are associated with violations of socio-emotional expectancies.

Objectives of the present study

To address the issues discussed above, we created prototypical scenarios in which an initial sentence was used to establish the social context as in (i) and (ii) below. The target sentence then contained a critical word (in italic in the examples below) that informed the reader of the nature of a character’s socio-emotional response to the situation described in the initial sentence. The described response either matched (i) or mismatched (ii) with what one might typically expect given the situation.

(i) Abbey was a famous, well-respected golf critic, who informed the golf pro that he had no chance of winning the next open.

The golf pro was distraught.

(ii) Abbey was a truthful coach, who informed the golf pro that he had a good chance of winning the next open.

The golf pro was distraught.

Since the materials describe prototypical scenarios, we assume that stored information or scripts representing information about a typical character’s thoughts and feelings, among other things, are probably most critical in this context as opposed to purely text-based information. Thus, we are investigating whether participants rapidly infer the likely feelings of a person based on their general background knowledge (scripts). Because the N400 is known to be larger to unexpected than expected words, to words that are inconsistent than consistent with script-related knowledge and world knowledge, as well as to words that are a poor rather than a good fit to the discourse context (for reviews, see Kutas and Federmeier, 2011; Van Berkum, in press), a violation of a socio-emotional stereotype may be expected to trigger an N400 effect (e.g. Van Berkum et al., 2008; see also White et al., 2009).

However, the P300 component has also been reported to be of larger amplitude when social expectancies (character traits) are violated than not (e.g. Bartholow et al., 2001; Van Duynslaeger et al., 2007). Thus, it is possible that the present socio-emotional violations will elicit a P300 effect, reflecting the updating of the (person-specific) environmental model in working memory (e.g. Donchin and Coles, 1988), as suggested by Bartholow et al. (2001). Such a P300 effect might also be obtained due to the phasic, locus coeruleus-mediated enhancement of cortical activity after unexpected stimuli (Nieuwenhuis et al., 2005). In order to determine the brain sources contributing to the socio-emotional ERP effect, we applied dipole source analysis of the ERPs. Although the spatial resolution of EEG is much poorer compared to that of fMRI, it should nevertheless be possible to discriminate between relatively distant sources located in the mPFC, TPJ, and aTL.

METHODS

Participants

Twenty-four right-handed native English speakers with no history of neurological dysfunction or reading problems participated (10 males, 14 females, M = 22.6 years). All participants had normal or corrected-to-normal vision and received £12 for participating.

Materials and apparatus

Eighty experimental materials were constructed (for examples, see Supplementary Data). As in (i) and (ii) above, the target word always specified an emotional response. The 80 materials were based on certain well-known careers (e.g. teacher, doctor, etc.) and prototypical situations in order to facilitate character-specific inferences. An off-line test of the materials revealed that our intuitions regarding the likely thoughts and feelings of the target character described in the scenarios were shared by the undergraduate student population of participants (see Supplementary Data). For the ERP study, the materials were arranged in two stimulus files, such that each item appeared either in the match or mismatch condition in a given file, but appeared in both conditions across the two files. Thus, a given file comprised 40 materials in each of the two conditions. Each file also contained 240 filler items, which were of a length similar to the test materials. Of these 240 filler items, 80 materials included referential ambiguities, specifically, they contained pronouns that either did or did not have an explicit referent in the prior discourse. A further 80 materials described well-known fictional characters performing actions that would either be consistent, or inconsistent with our world knowledge. Since these latter 80 items elicited an N400 effect related to world knowledge violation, it was of interest to compare the neural origins of this N400 effect (using dipole source modeling, see below) to the N400 effect elicited by the socio-emotional violations in the current study (for further details of the filler materials containing world knowledge violations, and their accompanying N400 waveforms, please see Supplementary Data). Of the final 80 filler items, 40 of these fillers were simply materials containing no anomalies whatsoever, and 40 items contained descriptions of implausible actions (which did not include socio-emotional information, and did not relate to well-known characters).
Participants were tested in an electrically shielded booth with ambient light kept at a low level. Word stimuli were presented in white 16-point Helvetica font on a black background at the centre of a 21-in. computer monitor at a viewing distance of 65 cm (maintained throughout the experiment by means of a chin rest). Approximately three characters subtended 1° of visual angle.

**Procedure**

Participants were informed about the EEG procedure and the experimental task prior to giving informed consent. Participants were then instructed to read the materials carefully and to avoid making any movements. There were six practice trials, including one comprehension question, to familiarize participants with the procedure. This was followed by 10 experimental blocks, each consisting of 32 trials. The blocks were separated by a break, the duration of which was determined by the participant.

Each trial consisted of the presentation of a context sentence, which was displayed until participants pressed the space bar when they finished reading it, but at least for a minimum duration of 1500 ms. A blank interval of 500 ms followed, after which a fixation cross was presented in the centre of the screen for 1000 ms. Then word-by-word presentation of the second sentence started, during which participants were asked to maintain fixation at the centre of the screen. Each word was displayed centrally for 300 ms, with a 200-ms blank interval between successive word presentations. After the final word was presented, the next context sentence was displayed after a blank interval of 2000 ms. A comprehension question was randomly displayed following 10% of trials. The mean correct response rate was 86.0%, indicating that participants were reading for comprehension.

**EEG recording**

A BIOSEMI Active-Two amplifier system was used for continuous recording of electroencephalographic (EEG) activity from 72 Ag/AgCl electrodes, including dedicated electrodes for recording horizontal and vertical eye movements as well as reference and ground electrodes (for more details, see Filik et al., 2008). EEG and EOG recordings were sampled at 256 Hz. Off-line all EEG channels were recalculated to a linked mastoid reference and trials containing blinks were corrected using the adaptive artifact correction method in Brain Electromagnetic Source Analysis (BESA) software (Ille et al., 2002). The analysis epoch started 200 ms prior to the onset of the critical word and lasted for a total duration of 1,700 ms.

**EEG data analysis**

Trials with nonocular artifacts (drifts, channel blockings, EEG activity exceeding ±60 μV) were discarded. The signal at each electrode site was averaged separately for each of the two experimental conditions, time-locked to target word onset, band-pass filtered (0.01–30 Hz, 6 dB/oct), and aligned to a 100-ms pre-stimulus baseline.

Mean ERP amplitudes were measured in average waveforms within time intervals during which specific ERP deflections and experimental effects were most pronounced: 150–200 ms (N1), 200–250 ms (P2), 350–400 ms (N400) and 600–1000 ms, relative to critical word onset. N1 and P2 amplitudes were determined to assess whether brain processes before those eliciting the N400 were influenced by experimental conditions.

ERP amplitude data at midline electrodes were analysed separately from data recorded over lateral electrode sites; the latter were pooled to form 12 regions of interest (ROI). That is, the electrodes were divided along a left-right dimension, an anterior-to-posterior dimension and a dorsal–ventral dimension. The six ROIs over the left hemisphere were defined as follows: left-anterior-ventral (AF7, F7, FT7, F5, FC5), left-anterior-dorsal (AF3, F3, FC3, F1, FC1), left-central–ventral (TP7, T7, C5, CP5), left-central dorsal (C3, CP3, C1, CP1), left-posterior-ventral (PO9’, O9’, P7, PO7, O1) and left-posterior-dorsal (P3, PO3, P1, P5). Six analogous ROIs were defined for homologue electrodes located over the right hemisphere. Statistical analyses were performed by means of Huynh–Feldt corrected repeated measures analyses of variance (ANOVA). ERP amplitudes measured at midline electrodes were subjected to an ANOVA with the variables condition (match, mismatch) and electrode (AFz, Fz, FCz, Cz, CPz, Pz, POz, Oz). ERP data acquired from lateral electrode sites (ROIs) were submitted to an ANOVA with variables condition, hemisphere (left, right), ant-pos (anterior, central, posterior) and vert-wality (ventral, dorsal).

**Dipole source analysis**

We applied spatio-temporal source modeling to examine the brain sources producing the mismatch effect in ERPs during an early time interval (350–400 ms), when N400 effects were maximal for experimental conditions of interest (see below), and a late time interval (600–1000 ms). To estimate the generators of mismatch-related ERP activity and the corresponding source waveforms, equivalent dipole sources were determined using BESA (Version 5.18) with the four-shell spherical head model (cf. Berg and Scherg, 1994; see Supplementary Data). Source locations are reported in approximate Talairach coordinates.

To maximize the signal-to-noise ratio, the dipole-source model was derived for band-pass filtered (0.2–6 Hz) grand average (across participants) difference waveforms that resulted from the following subtraction: $ERP(t)_{Mismatch} - ERP(t)_{Match}$. The source model was iteratively fitted to the difference waveform until a minimum in residual variance (RV) was reached, that is, the percentage of variance in the recorded potential distribution not accounted for by the source model was minimized. As discussed above, for comparison, we derived for the same participants a source model...
for the N400 effect produced by the filler materials in which the critical word indicated a world knowledge violation (see Supplementary Data and Supplementary Figure S1).

**RESULTS**

Figure 1 presents the grand mean ERP waveforms at midline and lateral electrodes for the match and mismatch conditions. The mean ERP difference waveform \( [\text{ERP(mismatch)} - \text{ERP(match)}] \) and corresponding topographic maps across the relevant time windows are displayed in Figure 2. It is apparent from the ERP waveforms that words that violate socio-emotional expectations (mismatch condition) elicit a clear N400 effect starting \( \sim 200 \text{ ms} \) after the onset of the critical word. In addition, a frontal positivity and right-hemispheric posterior negativity for the mismatch compared to the match condition emerged after \( \sim 600 \text{ ms} \).

To assess the time-course of the condition effect, we examined experimental effects on N1 and P2 amplitude. The lateral ROI analysis of mean N1 amplitude (150–200 ms) revealed the typical posterior N1 topography, \( F(2, 46) = 9.82, \epsilon = 1.0, P < 0.001, \eta_p^2 = 0.30, \) however, condition did not influence N1 amplitude, all \( FFs < 2.5, \) \( Ps > 0.09. \) The subsequent P2 (200–250 ms) was maximal over central midline electrodes, \( F(7, 161) = 13.65, \epsilon = 0.22, P < 0.001, \eta_p^2 = 0.37, \) and tended to be of smaller amplitude for the mismatch than the match condition \( (3.49 vs 4.28 \mu\text{V}), F(1, 23) = 3.88, P = 0.06, \eta_p^2 = 0.13. \)

Most importantly, ERP amplitude over midline electrodes was significantly influenced by condition in the 350–400 ms interval \( (N400), F(1, 23) = 9.13, P < 0.01, \eta_p^2 = 0.28, \) reflecting the more negative-going waveform for the mismatch than the match condition \( (3.42 vs 4.99 \mu\text{V}). \) The significant Condition \( \times \) Electrode interaction, \( F(7, 161) = 4.12, P < 0.05, \eta_p^2 = 0.15, \) indicated reliable condition effects over central and posterior electrodes \( (\text{FCz, Cz, CPz, Pz, POz and Oz}), \) all \( FFs > 1.60, FPs < 0.05, \) but not over anterior electrodes \( (\text{AFz, Fz}), FFs < 1.54, FPs > 0.22 \) (cf. Figure 2). Analysis of ERP amplitudes over lateral ROIs confirmed the main effect of condition, \( F(1, 23) = 9.95, P < 0.01, \eta_p^2 = 0.30. \) The significant Condition \( \times \) Hemisphere and Condition \( \times \) Verticality interactions, \( FFs(2, 46) > 6.8, FPs < 0.05, \eta_p^2s > 0.23, \) as can be seen in Figure 2, reflected stronger N400 condition effects over the right than the left hemisphere \( (1.6 vs 1.1 \mu\text{V}) \) and over dorsal than ventral ROIs \( (1.7 vs 1.1 \mu\text{V}). \) In separate analyses, however, condition effects were reliable over both left and right hemispheres and over both ventral and dorsal sites, all \( FFs(1, 23) > 6.4, P < 0.05. \)

In the late time interval \( (600–1000 \text{ ms}), \) the midline analysis yielded a significant Condition \( \times \) Electrode interaction, \( F(7, 161) = 6.38, \epsilon = 0.50, P < 0.001, \eta_p^2 = 0.22. \) Decomposing the interaction revealed a larger positivity for the mismatch than match condition over anterior \( (\text{AFz, Fz, FCz}) \) electrodes \( (7.1 vs 6.0 \mu\text{V}), F(1, 23) = 5.75, P < 0.05, \) with the opposite amplitude effect over posterior \( (\text{Pz, POz, Oz}) \) electrodes \( (7.4 vs 7.8 \mu\text{V}) \) producing a trend, \( F(1, 23) = 3.49, P = 0.075 \) (cf. Figure 2). In the lateral ROI analysis, the Condition \( \times \) Hemisphere interaction, \( F(1, 23) = 23.81, P < 0.001, \eta_p^2 = 0.51, \) and the Condition \( \times \) Hemisphere \( \times \) Ant-Pos interaction, \( F(2, 46) = 4.93, \epsilon = 1.0, P < 0.05, \eta_p^2 = 0.18, \) were significant. As indicated by further analyses, the main effect of condition and the Condition \( \times \) Ant-Pos interaction was not significant for left-hemispheric ROIs, \( Fs < 1, \) whereas the significant Condition \( \times \) Ant-Pos interaction for right-hemispheric ROIs, \( F(2, 46) = 6.37, \epsilon = 0.93, P < 0.01, \) was due to a trend for a larger negativity for the mismatch than the match condition over posterior ROIs, \( F(1, 23) = 3.63, P = 0.069 \) (cf. Figure 2), but not over anterior and central ROIs, \( FFs < 2.9, FPs > 0.10. \)

**Dipole analysis**

Figure 3 shows the dipole solution obtained for the grand-average ERP difference waveform. In the light of previous studies which showed N400 generators to be located in the left and right temporal lobes (cf. Van Petten and Luka, 2006), a pair of dipoles with mirror-symmetrical location but independent orientation or amplitude was iteratively fitted to the data in the 350–400 ms time interval. The resulting dipole pair was located in the anterior part of the temporal cortex \( (x = \pm 39 \text{ mm}, y = \mp 5 \text{ mm}, z = \mp 21 \text{ mm}), \) with the model accounting for a large proportion of the data variance \( (93.0\%; \text{RV} = 7.0\%). \) Adding a further dipole or dipole pair minimally reduced RV \( (<1.3\%) \) and instead bears the risk of fitting noise. The temporal lobe dipoles accounted less well for data variance in the 600–1000 ms time interval \( (85.5\%; \text{RV} = 14.5\%). \) Fitting of an additional single dipole, located in the frontomedial cortex \( (x = 5 \text{ mm}, y = 34 \text{ mm}, z = 31 \text{ mm}), \) reduced RV by 8.9% \( (\text{resultant RV} = 5.6\%). \) As can be seen in Figure 3, source activation in bilateral temporal sources peaked between 350 and 500 ms, at about the time when activity in the frontomedial source started.

For the N400 effect elicited by world knowledge violations (from filler materials, discussed above), the lateral dipole pair fitted in the 350–400 ms time interval accounted for \( \sim 88\% \) of data variance \( (\text{RV} = 12.4\%). \) It was located in the middle temporal lobe \( (x = \pm 47 \text{ mm}, y = \mp 30 \text{ mm}, z = \mp 14 \text{ mm}), \) somewhat posterior to the location of the dipole pair obtained for the N400 effect triggered by socio-emotional violations.

**DISCUSSION**

In the present study, we used ERPs to investigate the on-line processing of socio-emotional information using prototypical scenarios. The key finding was that critical words which mismatched rather than matched with a story character’s likely feelings, as established by the context, elicited a
larger negative-going ERP deflection with a centroparietal maximum between 350 and 400 ms after word onset. We assume that this ERP finding reflects an N400-like effect given its topographic distribution and time-course. Subsequently, violations of socio-emotional expectations triggered a larger sustained positivity over frontal electrodes. Together, these ERP results indicate that readers infer a character’s likely socio-emotional response based on the prototypical information provided by the text.

The present N400 effect accords well with previous observations of a larger N400 to unexpected than expected words, to words that are inconsistent than consistent with script-related knowledge and world knowledge, to words that are a poor rather than a good fit to the discourse context (for reviews, see Kutas and Federmeier, 2011; Van Berkum, in press), as well as similar N400-like effects elicited by stereotype violations concerning age, gender, occupation and social class (Van Berkum et al., 2008; White et al., 2009). Thus, the N400 effect indicates the on-line construction of meaning at the word, sentence and discourse level. Specifically, the early onset of the present N400-like effect, starting at ~200 ms after word onset, provides strong evidence for rapid socio-emotional processing, extending previous N400 research by demonstrating that meaning can include the predictable socio-emotional reactions of others. Thus, we assume that readers immediately infer the likely (emotional) response of a character in a familiar social context, integrating information provided by the text and from

![Fig. 1 Grand average ERP waveforms elicited at midline and lateral electrodes for critical words that matched vs mismatched the socio-emotional context. Positivity is plotted upwards.](image-url)
LTM to establish the situation model. This allows readers to generate on-line a socio-emotional prediction about the character’s response. New incoming information is then checked against this background and when readers encounter information that mismatches rather than matches their prediction, a larger N400 is triggered by the critical word. We assume that this N400 effect reflects the increased integration and semantic memory demands in the case of a violation of a socio-emotional stereotype.

Subsequent to the N400 effect, we observed an enhanced frontal ERP positivity (and right posterior negativity) when socio-emotional predictions were violated compared to
when they were not. Here it is important to note that this effect on the late positivity appears to be different from the P300 effects reported in ERP studies of person perception (e.g. Bartholow et al., 2001; Van Duymslaeger et al., 2007). These previous studies showed that trait-inconsistencies produced an effect on the classical, centroparietal P300 component (e.g. Johnson, 1988), whereas the present effect of socio-emotional violations on the late ERP positivity showed a frontal scalp distribution, presumably reflecting a cognitive mechanism that is different from the one indicated by the classic P300 effect. In relation to this issue, Federmeier et al. (2007) reported a similarly distributed late ERP positivity to unexpected words when the sentence context was highly constraining, whereas Filik et al. (2008) observed such a frontal positivity after referentially ambiguous pronouns. Although the functional significance of the frontal positivity is far from being understood, it might reflect more effortful processing related to either the suppression of an inappropriately predicted word (Federmeier et al., 2007) or the addition of an inferred discourse entity to the discourse model (Filik et al., 2008).

The present finding of an N400 effect rather than a P300 effect also points to differences in the underlying processes in our study compared to the study of Bartholow et al. (2001). One possible reason for these discrepancies might be in terms of the text materials used. That is, the present study used short, well-known scenarios, whereas ERP studies of trait inference processes showing a P300 effect used considerably longer vignettes in order to introduce a specific person to the reader. Thus, in the present study it might have been easier for readers to activate an appropriate situation model and to predict a character’s response, thereby meeting the requirements for eliciting a large N400 in the case of a violation.

Spatio-temporal source modeling, in conjunction with knowledge obtained from functional neuroimaging studies of ToM and of text comprehension, provided further clues as to the significance of the present ERP effects. First, it is worth mentioning that given the spatial accuracy of EEG source localization, with a spatial resolution of 1–2 cm at best (e.g. Krings et al., 1999), the locations reported here should be taken as an approximation. Specifically, the estimate for the z-axis location is likely to be inaccurate, due to the fact that ERPs cannot be recorded from underneath the brain. Second, it is also interesting to note that for the same participants in the same recording session, we found that the N400 effect triggered by violations of world knowledge was produced by a bilateral source located in the middle temporal lobe. Despite the limitations of the present source localization approach (e.g. head model), the bilateral temporal source locations found here for socio-emotional and world knowledge violations correspond relatively well with the N400 sources in more superior regions of the temporal cortex reported in magnetoencephalographic (MEG) studies (e.g. Halgren et al., 2002; Service et al., 2007) and in the anterior medial temporal lobe reported in intracranial recording studies (e.g. McCarthy et al., 1995; Meyer et al., 2005).

Most crucially, we found a bilateral dipolar source, which was located in a more anterior part of the temporal lobe for the N400 effect to socio-emotional violations than for the one fitted to the N400 effect reflecting world knowledge violations. This bilateral source in the anterior temporal cortex partially accords with Ferstl et al.’s (2005) finding of higher right-hemispheric aTL activation in the case of emotion violations, which the authors took to reflect the increased demands of building a situation model. Given the different methodologies, procedures and text materials used, however, it is difficult to point to a specific factor that could explain why we failed to find a stronger right temporal lobe source in
our study. However, a recent meta-analysis of functional neuroimaging studies also suggests a role for bilateral temporal lobe regions in text comprehension (Ferstl et al., 2008).

What seems clear is that text comprehension, including the creation of a situation model and the associated inferential processes, very much depends on social conceptual information and other sources of socio-emotional information, for instance, autobiographical memory and emotional, episodic memory. In this respect, the present source localization findings regarding the N400-like effect induced by socio-emotional violations fit well with Frith’s (2007) proposal that the aTL region is responsible for integrating general knowledge about social situations (e.g. scripts) with personal and context-specific information. Certainly, the contribution of this brain region to socio-emotional processing in relation to accessing conceptual knowledge in semantic memory (e.g. Ross and Olson, 2010), to situation model building (e.g. Ferstl et al., 2005) and to the integration of various types of social information (e.g. Frith, 2007) is in need of further investigation.

Source modeling further indicated a single dipolar source located in the frontomedial cortex which partially accounted for the late ERP effect associated with socio-emotional violations. As for the N400-effect, this source location fits well with Ferstl et al.’s (2005) functional neuroimaging study in which they found that the dorsomedial PFC (dmPFC) was more strongly activated when emotional expectations were violated. They proposed that this brain region reflects the higher emotional inference demands in the case of a violation. In addition, the earlier start of activation in bilateral dipoles than the frontomedial dipole is in line with the view that frontomedial brain activity is related to a higher level, integrative function in socio-emotional processing and that the mediofrontal cortex might sensitively reflect ToM demands (for reviews, see Amodio and Frith, 2006; Saxe, 2006).

In conclusion, the present study indicates that vignettes describing familiar social situations elicit clear-cut ERP correlates that are indicative of readers’ rapid socio-emotional processing. Assuming that the process of reading fictional (social) scenarios activates the same or very similar mechanisms to those that are activated when we encounter other people in the real world (e.g. Mar and Oatley, 2008), this methodology provides a promising tool to study the critical situational cues on which socio-emotional processing is based as well as the extent to which person factors such as age, gender and mental health influence these mechanisms.

SUPPLEMENTARY DATA
Supplementary Data are available at SCAN online.

Conflict of Interest
None declared.

REFERENCES
Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. Nature Reviews Neuroscience, 7, 268–77.
Bartholow, B.D., Dickter, C.L. (2007). Social cognitive neuroscience of person perception: a selective review focused on the event-related brain potential. In: Harmon-Jones, E., Winkielman, P., editors. Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior. New York: Guilford Press, pp. 376–400.
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.
Berg, P., Scherg, M. (1994). A fast method for forward computation of multiple-shell spherical head models. Electroencephalography & Clinical Neurophysiology, 90, 58–64.
Chwilla, D.J., Kolb, H.H.J. (2005). Accessing world knowledge: evidence from N400 and reaction time priming. Cognitive Brain Research, 25, 589–606.
Decety, J., Grezes, J. (2006). The power of simulation: imagining one’s own and other’s behavior. Brain Research, 1079, 4–14.
De Vega, M., Leon, I., Diaz, J.M. (1996). The representation of multiple-shell spherical head models. Electroencephalography & Clinical Neurophysiology, 90, 58–64.
Donchin, E., Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? Behavioral and Brain Sciences, 11, 354–6.
Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., Kutas, M. (2007). Multiple effects of sentential constraint on word processing. Brain Research, 1146, 75–84.
Ferstl, E.C., Neumann, J., Bogler, C., von Cramon, D.Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. Human Brain Mapping, 29, 581–93.
Ferstl, E.C., Rinck, M., von Cramon, D.Y. (2005). Emotional and temporal aspects of situation model processing during text comprehension: an event-related fMRI study. Journal of Cognitive Neuroscience, 17, 724–39.
Filik, R., Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: evidence from the N400. Psychophysiology, 45, 554–8.
Filik, R., Sanford, A.I., Leuthold, H. (2008). Processing pronouns without antecedents: evidence from event-related brain potentials. Journal of Cognitive Neuroscience, 20, 1315–26.
Frith, C.D. (2007). The social brain? Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 362, 671–8.
Frith, U., Frith, C.D. (2003). Development and neurophysiology of mentalising. Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 358, 459–73.
Gernsbacher, M.A., Goldsmith, H.H., Robertson, R.W. (1992). Do readers mentally represent characters’ emotional states? Cognition and Emotion, 6, 89–111.
Gernsbacher, M.A., Hallada, B.M., Robertson, R.R.W. (1998). How automatically do readers infer fictional characters’ emotional states? Scientific Studies of Reading, 2, 271–300.
Gerrig, R.J. (1993). Experiencing Narrative Worlds. New Haven, CT: Yale University Press.
Gordon, R.M. (1986). Folk psychology as simulation. Mind & Language, 1, 158–71.
Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M. (2004). Integration of word meaning and world knowledge in language comprehension. Science, 304, 438–41.
Halgren, E., Dhond, R.P., Christensen, N., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. NeuroImage, 17, 1101–1116.
Heider, F. (1958). The Psychology of Interpersonal Relations. Hillsdale, NJ: Erlbaum.
Ille, N., Berg, P., Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. Journal of Clinical Neurophysiology, 19, 113–24.
Johnson, R.Jr. (1988). The amplitude of the P300 component of the event-related potential: review and synthesis. In: Ackles, P.K., Jennings, J.R., Coles, M.G.H., editors. Advances in Psychophysiology, Vol. 3, Greenwich: JAI Press, pp. 69–138.

Karniol, R. (2003). Egocentrism versus protocentrism: the status of self in social prediction. Psychological Review, 110, 564–80.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. Psychological Review, 95, 163–82.

Klings, T., Chiappa, K.H., Coffin, B.N., Cohrs, J.I., Connolly, S., Congrove, G.R. (1999). Accuracy of EEG dipole source localization using implanted sources in the human brain. Clinical Neurophysiology, 110, 106–14.

Kutas, M., Federmeier, K.D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology, 62, 621–47.

Leslie, A.M. (1987). Pretense and representation: the origins of "theory of mind". Psychological Review, 94, 412–26.

Mar, R.A., Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. Perspectives on Psychological Science, 13, 173–92.

Mason, R.A., Just, M.A. (2009). The role of the Theory-of-Mind cortical network in the comprehension of narratives. Language and Linguistics Compass, 3, 157–74.

McCarthy, G., Nobre, A.C., Bentin, S., Spencer, D.D. (1995). Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. Journal of Neuroscience, 15, 1080–9.

Meyer, P., Mecklinger, A., Grunwald, T., Fell, J., Elger, C.E., Friederici, A.D. (2005). Language processing within the human medial temporal lobe. Hippocampus, 15, 451–9.

Nieuwenhuis, S., Aston-Jones, G., Cohen, J.D. (2005). Decision making, the P3, and the locus coerulescere-norepinephrine system. Psychological Bulletin, 131, 510–32.

Olson, L.R., Plotzker, A., Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. Brain, 130, 1718–31.

Perner, J. (1991). Understanding the Representational Mind. Cambridge, MA: MIT Press.

Ross, L.A., Olson, L.R. (2010). Social cognition and the anterior temporal lobe. NeuroImage, 49, 3452–62.

Saxe, R. (2006). Why and how to study Theory of Mind with fMRI. Brain Research, 1079, 57–65.

Saxe, R., Wexler, A. (2005). Making sense of another mind: the role of the right temporoparietal junction. Neuropsychologia, 43, 1391–9.

Schank, R.C., Abelson, R.P. (1977). Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures. Hillsdale, NJ: Erlbaum.

Service, E., Helenius, P., Maury, S., Salmelin, R. (2007). Localization of syntactic and semantic brain responses using magnetoencephalography. Journal of Cognitive Neuroscience, 19, 1193–205.

Siemer, M., Reisenzein, R. (2007). The process of emotion inference. Emotion, 7, 1–20.

Simmons, W.K., Reddish, M., Bellgowan, P.S., Martin, A. (2010). The selectivity and functional connectivity of the anterior temporal lobes. Cerebral Cortex, 20, 813–25.

Van Berkum, J.J.A. (in press). The electrophysiology of discourse and conversation. In: Spivey, M., Joanssone, M., McRae, K., editors. The Cambridge Handbook of Psycholinguistics. Cambridge: Cambridge University Press.

Van Berkum, J.J.A., Brown, C.M., Zwiterlood, P., Kooijman, V., Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 443–67.

Van Berkum, J.J.A., van den Brink, D., Tesink, C.M.J.Y., Kos, M., Hagoort, P. (2008). The neural integration of speaker and message. Journal of Cognitive Neuroscience, 20, 580–91.

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 Petten, C.M., Luka, B.J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain and Language, 97, 279–93.

White, K.R., Crites, S.L.Jr., Taylor, J.H., Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. Social, Cognitive, and Affective Neuroscience, 4, 191–8.

Zahn, R., Moll, J., Krueger, F., Huey, E.D., Garrido, G., Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. Proceedings of the National Academy of Sciences, 104, 6430–5.

Zwaan, R.A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In: Ross, B.H., editor. The Psychology of Learning and Motivation, Vol. 43, New York: Academic Press, pp. 35–62.

Zwaan, R.A., Radvansky, G.A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123, 162–85.