



Implicit and explicit social mentalizing: dual processes driven by a shared neural network

Frank Van Overwalle* and Marie Vandekerckhove

Department of Psychology, Vrije Universiteit Brussel, Brussels, Belgium

Edited by:

Leonhard Schilbach, University
Hospital Cologne, Germany

Reviewed by:

Chris Frith, University College
London, UK
Jon Freeman, Dartmouth College,
USA

*Correspondence:

Frank Van Overwalle, Department of
Psychology, Vrije Universiteit
Brussel, Pleinlaan 2, B - 1050
Brussels, Belgium
e-mail: frank.vanoverwalle@
vub.ac.be

Recent social neuroscientific evidence indicates that implicit and explicit inferences on the mind of another person (i.e., intentions, attributions or traits), are subserved by a shared mentalizing network. Under both implicit and explicit instructions, ERP studies reveal that early inferences occur at about the same time, and fMRI studies demonstrate an overlap in core mentalizing areas, including the temporo-parietal junction (TPJ) and the medial prefrontal cortex (mPFC). These results suggest a rapid shared implicit intuition followed by a slower explicit verification processes (as revealed by additional brain activation during explicit vs. implicit inferences). These data provide support for a default-adjustment dual-process framework of social mentalizing.

Keywords: mentalizing, implicit inferences, fMRI, ERP, dual-process theories

INTRODUCTION

Tell me what you did today, and I'll tell what you want and who you are. Behaviors are quite often the main road to enter people's mind, to infer their intentions and judge their personality traits. Social inferences that rely on insights about other people's mental content such as intentions, desires, beliefs, traits or other high-level characteristics are termed *mentalizing* inferences. Contrary to the old idea that such complex inferences need a lot of explicit mental elaboration, behavioral research in the 80s (Winter and Uleman, 1984) documented that such person inferences, including trait inferences, are often made implicitly and automatically, without awareness or control about the inference process. The fascination for such implicit and rapid inferences comes from the fact they are relatively correct based on brief verbal or non-verbal information (Letzring et al., 2006), although they may at times differ from explicit inferences. What is the true inference then? The aim of the present paper is to demonstrate that current neuroscientific evidence on mentalizing suggests that implicit and explicit person inferences do not rely on strictly distinct neural processes or substrates.

To explain the distinction between implicit and explicit social judgments, social cognition researchers drew parallels with similar distinctions in cognitive psychology, collectively termed dual-process theories (for reviews see Evans, 2008; Evans and Stanovich, 2013). Dual-process theories argue that there are two different modes of processing, termed *implicit* (also termed unconscious, automatic, spontaneous, experiential, heuristic, intuitive, impulsive, and reflexive) and *explicit* (also termed conscious, controlled, rational, systematic, analytical, and reflective; Schneider and Shiffrin, 1977; Chaiken, 1980; Epstein, 1994; Strack and Deutsch, 2004; Lieberman, 2007). The key feature of dual-process theories is that implicit processes are inaccessible to consciousness and control, while explicit processes are accessible to awareness, introspection and flexible control. It is

also assumed that implicit processes are rapid, while explicit processes are slow (Evans, 2008). Some dual-process theorists in social cognition (and other domains) assume that implicit and explicit processes are subserved by different information processing systems (e.g., associative vs. rule-based; Smith and DeCoster, 2000) or subserved by different brain areas (e.g., Satpute and Lieberman, 2006; Lieberman, 2007; Forbes and Grafman, 2013).

DO DUAL PROCESSES IMPLY DISTINCT NETWORKS?

Behavioral research in social cognition accumulated a growing body of evidence in favor of dual-process models by demonstrating that many person inferences occur not always explicitly, but often implicitly (for a review, Uleman, 1999). Recent findings demonstrate that representing other agents' beliefs, is an implicit capacity acquired early at 7 month of age (Kovacs et al., 2010) and implicitly sustained during adulthood (Schneider et al., 2012a), although it requires some minimal executive resources (Qureshi et al., 2010; Schneider et al., 2012b). However, the idea that implicit and explicit social mentalizing are subserved by distinct and exclusive processing systems or brain networks appears unjustified. True, some neural networks are mainly involved in implicit social processing. Recent neuroscientific research has uncovered subcortical mechanisms located in the amygdala and other limbic structures that elicit primitive affective reactions (e.g., rapid impressions of a face; Todorov et al., 2008a,b; Vandekerckhove and Panksepp, 2011; Forbes et al., 2012), as well as mirror-like neural networks responsible for the understanding of non-verbal movements and actions of humans (Iacoboni, 2009; Van Overwalle and Baetens, 2009).

In contrast, it is acknowledged that higher-level mentalizing brain areas subserve computations that are neither exclusively implicit or explicit (Satpute and Lieberman, 2006; Keyser

and Gazzola, 2007; Forbes and Grafman, 2013). On the contrary, the content of the social inference process seems more crucial than the nature of the process. Each core area in the mentalizing network seems responsible for a distinct computation and appears sensitive to a specific input. The *temporo-parietal junction* (TPJ) seems responsible for judgments on temporary beliefs and intentions, while the *medial prefrontal cortex* (mPFC) seems involved in enduring trait inferences and other stable characteristics (**Figure 1**; see for reviews, Amodio and Frith, 2006; Van Overwalle, 2009; Lombardo et al., 2011; Bzdok et al., 2012; Denny et al., 2012). Each of these areas supports computations of specific social judgments, but none seems predominantly recruited for explicit or implicit inferences.

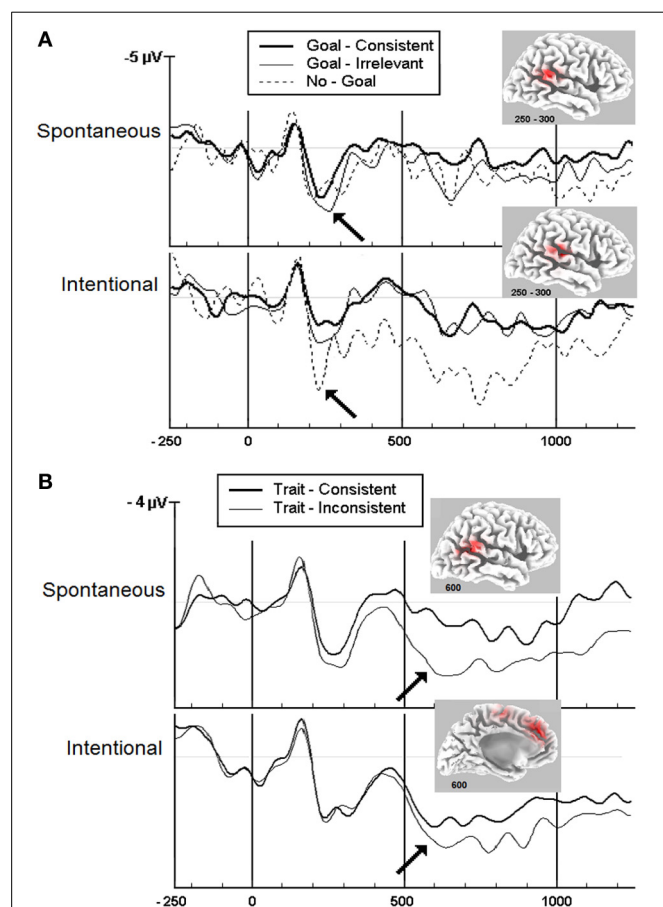


FIGURE 1 | Grand-averaged ERP waveforms showing positive deflections at the central midline (Cz) scalp sites given spontaneous and intentional instructions. The timeline is given in ms. A positive amplitude is shown downward. **(A)** Goal inferences: Goal-consistent words diverge significantly from goal-irrelevant and no-goal words at about 250 ms ($p < 0.05$) as indicated by the arrows. Insets: LORETA source analysis of the TPJ at 250–300 ms. Adapted with permission from Van der Cruyssen et al. (2009). Copyright 2008 by Psychology Press. **(B)** Trait inferences: Trait-consistent words diverge significantly from trait-inconsistent words at about 600 ms ($p < 0.05$) as indicated by the arrows. Insets: LORETA source analysis of the TPJ and mPFC at 600 ms. Adapted with permission from Van Duynslaeger et al. (2007). Copyright 2007 by Oxford Press.

IMPLICIT AND EXPLICIT MENTALIZING SHARE EARLY TIMING AND CORE BRAIN AREAS

Methods that were commonly used to explore implicit processes in behavioral research, cannot always be directly adapted for neuroimaging research. For instance, subliminal presentation of stories or picture sequences to depict events is impossible, while the use of cognitive load may trigger unwanted neural processes that may interfere with the social inference task (e.g., Qureshi et al., 2010). To illustrate, in contrast to passively watching or explicitly empathizing with another person while reading social information, Morelli and Lieberman (2013) found reduced activity in the TPJ and mPFC when participants were instructed to memorize simultaneously an irrelevant number. Moreover, cognitive load by memorizing further increased activity in a host of areas related to episodic memory and executive control (i.e., precuneus, dorsal anterior and middle cingulate, lateral PFC and inferior parietal cortex). Critically, these increased activations occurred in comparison to either explicit or implicit mentalizing instructions, demonstrating that unlike behavioral research, cognitive load by a secondary task may actually complicate the understanding of neural processes rather than clarifying it. However, the mere presentation of behavioral descriptions that invite most participants to spontaneously infer a social goal, belief or trait under implicit reading instructions, as opposed to explicitly instructions to make these inferences, avoids these methodological difficulties. Nevertheless, passively reading (as well as secondary tasks) cannot eliminate entirely the possibility that some amount of explicit mentalizing might still occur. Using this approach, recent neuroimaging research converges on the view that there is a common underlying mentalizing network that is relatively blind to the implicit or explicit nature of the inference, and that seems more sensitive to the content of the inference.

ERP TIMING: AN IDENTICAL ONSET FOR IMPLICIT AND EXPLICIT INFERENCE

Van Overwalle and colleagues documented that the neural timing (measured by event-related potentials or ERPs) of an early social inference is almost identical under implicit or explicit processing. Under implicit instructions, the participants were told that they had to “read carefully” the material, while participants under explicit instructions had to answer repeatedly the question: “what is the [inference] of this person?” Participants were divided in two separate groups, to avoid contamination of the explicit instructions on implicit judgments. Participants read a number of sentences that all implied the same inference (e.g., “nice” as a trait), and thus provided a strong impression on the target person. At the end of the trial, a critical sentence was presented that was either consistent with the earlier inference (e.g., “gave her sister a *hug*”), inconsistent (e.g., “gave her brother a *slap*”), or neutral (e.g., “gave her mother a *bottle*”). Because timing is crucial in ERP research, the sentences were presented word-by-word in the middle of the screen, and ERP timing started at the beginning of the critical word (see italics in the examples above) that diverged between conditions.

A first ERP study on goal inferences (Van der Cruyssen et al., 2009) reported that goal inferences were made after about 250 ms,

again irrespective of the implicit or explicit instruction (see the significant divergence of wave patterns under consistent vs. inconsistent conditions, **Figure 1A**). A second ERP study on trait inferences (Van Duynslaeger et al., 2007) documented that the onset of trait inferences occurs at about 600 ms irrespective of instructions (**Figure 1B**). Source localization of the ERP waves suggested that the core mentalizing areas (TPJ and mPFC) were most strongly recruited (**Figure 1** insets, using LORETA, Pascual-Marqui et al., 1994; Pascual-Marqui, 1999). Together, these results seem to provide support for a core single-system account with an identical onset at the beginning of mentalizing. Interestingly, the results also reveal that goal inferences are faster than trait inferences, consistent with the proposition by Van Overwalle (2009) that goals involve a quick evaluation of the here-and-now by the TPJ, while traits reflect slower abstractions by the mPFC extracted from behaviors identified in the TPJ (see also Van Overwalle et al., 2011; Ma et al., 2012b).

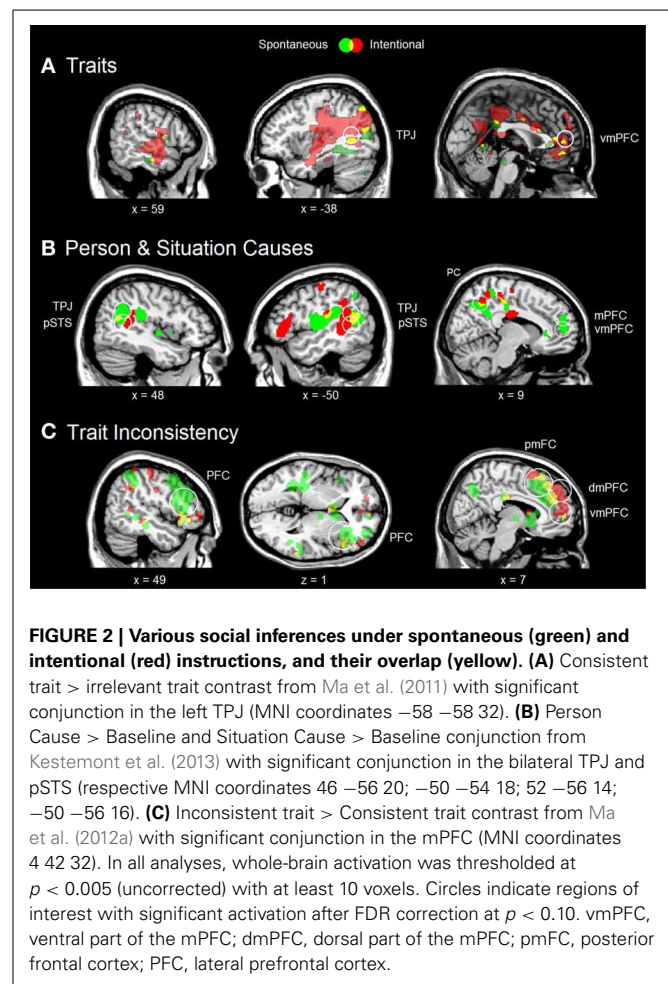
fMRI LOCALIZATION: OVERLAP BETWEEN IMPLICIT AND EXPLICIT INFERENCE

Van Overwalle and colleagues conducted a series of functional imaging studies to explore the overlap between explicit and implicit mentalizing, using a very similar procedure as described before. The results showed significant overlap in mentalizing activity between instructions, even though brains from different participants were compared.

A first fMRI study on implicit and explicit trait inferences (Ma et al., 2011) revealed common activation in the left TPJ, mPFC and bilateral temporal pole (**Figure 2A**). This overlap was statistically significant for the TPJ, indicating a reliable common process across brains from different individuals. Importantly, there were also differences between instructions. Implicit trait inferences significantly recruited only the core mentalizing areas of the TPJ and mPFC, whereas explicit trait inferences additionally recruited other brain areas involved in mentalizing, including the precuneus (autobiographic memory) and posterior part of the superior temporal sulcus (pSTS; biological motion). Analogous findings were reported by Rameson et al. (2010) for implicit and explicit self-descriptions.

In a second fMRI study on person and situation causes (Kestemont et al., 2013), there was an overlap between implicit and explicit instructions in the bilateral TPJ and pSTS, and the precuneus, which was significant for the TPJ and pSTS (**Figure 2B**). Again, there were also differences. Only implicit inferences increased the activation of the mPFC, suggesting a tendency to make dispositional trait attributions to the person irrespective of the situational circumstances that may have constrained or induced the behavior of the agent, known as the *fundamental attribution bias* (Moran et al., unpublished). This biased activation of the mPFC was absent under explicit instructions, consistent with decreased biased processing documented in behavioral research (Gilbert and Malone, 1995) and neuroimaging (Brosch et al., 2013).

In a third fMRI study on trait inconsistencies (Ma et al., 2012a), a significant overlap was found in the dorsal part of the mPFC (**Figure 2C**). Like in the previous trait study, some brain areas were more active only under explicit instructions, including



the left TPJ and pSTS (biological motion) and the precuneus (autobiographic memory). Interestingly, both instructions also revealed a non-significant overlap in the posterior medial frontal cortex (pmFC, including the dorsal part of the anterior cingulate cortex—dACC) and the right PFC. These latter two areas are part of a domain-general conflict monitoring network (Botvinick et al., 2004). Botvinick et al. (2004) proposed that this network detects and resolves conflicts between multiple inputs, and suggested that the pmFC detects inconsistencies, while the lateral PFC resolves these by modifying the attention to the different conflicting inputs.

The findings converge on the notion that under explicit processing, implicit information is enriched by retrieving similar behaviors from the past and imaging more vividly social cues on human action (Ma et al., 2011, 2012a), or by taking in more situational information so that a biased attribution in favor of the person is avoided (Kestemont et al., 2013). However, a limitation of current neuroimaging studies on social mentalizing is that they do not present direct evidence to support the interpretation that additional brain activation during explicit judgments reflect verification or correction after a default intuitive response. There is, however, some neuroimaging evidence on racial attitudes demonstrating this correction process, to avoid stereotypic biases.

When White egalitarian-motivated people were confronted with Black out-group faces (Cunningham et al., 2004; Forbes et al., 2012), the amygdala was activated only under short (~30 ms) but not under long presentations times (~500 ms). The finding that amygdala activation correlated with a behavioral test of racial attitudes suggests that the initial amygdala response was biased, while later explicit processing reevaluate out-group members in a non-biased manner.

IMPLICIT AND EXPLICIT MENTALIZING AS ITERATIVE REPROCESSING

Taken together, recent neuroscientific data on the neural underpinnings of social mentalizing suggest that implicit and explicit mentalizing share the same early timing and the same core brain areas, but also that explicit inferences may lead to a modulation in some brain areas, reflecting a correction or an enrichment. These data are best explained by more recent *default-interventionist* dual-process theories (Evans, 2011). These dual-process theories argue that implicit processes may provide a quick default solution to an assessment (resulting in an identical timing of early inferences), which may afterwards be either accepted or corrected by explicit reasoning (resulting in longer processing times overall). Cunningham and Zelazo (2007) proposed an interesting implementation of this idea which they termed *iterative reprocessing*. According to these authors, processing occurs on a continuum from relatively implicit to relatively explicit. Increased explicit processing is possible through additional reprocessing cycles, which enables more explicit elaboration of information along a wider and richer range of contexts and constraints, retrieving input from increasingly more brain structures. Inferences are hierarchically arranged, so that lower-level processes continuously provide valenced information, while higher-order processes recruited during subsequent cycles render the inference more explicit. With each iteration cycle, information is passed back and forth between the lower-level and higher-level computations (Cunningham et al., 2007). According to some theories of consciousness, these additional cycles bring the inference process closer to a steady neural attractor state that is accessible to introspection (Dennett, 2001; Timmermans et al., 2012). This iterative cycling theory has been implemented in multilevel bidirectional connectionist computer models of person construal (Freeman and Ambady, 2011) and attitude formation (Ehret et al., unpublished), in which a continuous interaction between lower and higher-level social information dynamically leads to stable person construals (Freeman and Ambady, 2011) and attitudes about them (Ehret et al., unpublished).

It is suggested here that social mentalizing involves similar iterative reprocessing from implicit to explicit (Cunningham and Zelazo, 2007; Freeman and Ambady, 2011). Inferences based on one or a few cycles are relatively implicit and crude intuitions, leaving an early mark in the ERPs and activation in restricted core brain areas. In contrast, inferences based on additional iterations and computations are increasingly rich, balanced and relatively explicit, leaving a broader trace of activation in multiple brain areas. Social inferences are hierarchically arranged, with lower-level brain areas (e.g., amygdala) continuously providing valenced information (Todorov et al., 2008a,b;

Vandekerckhove and Panksepp, 2011) and moderate-level interpretations of behaviors (e.g., TPJ) sending information on the agent's intentions, which feeds higher-level interpretations of the agent in terms of traits (e.g., mPFC; Van Overwalle, 2009; Ma et al., 2012b). These higher-order inferences are recruited during additional processing cycles which make it more likely that situational information is considered and biases are corrected.

How these subsequent cycles touch on novel input and memories, and how this information is integrated and feed back to implicit intuitions at the neural level, are questions for future research. Moreover, it is likely that individual differences in motivation or need for cognition (Epstein et al., 1996; Roets and Van Hiel, 2007, 2011) determine how much cycles are processed until an inference is terminated. Or does this depend more on situational opportunities?

Apart from these basic questions on iterative reprocessing of implicit and explicit mentalizing (Cunningham and Zelazo, 2007; Freeman and Ambady, 2011), there are also novel neuroscientific findings that can be viewed from this perspective. For instance, recent research demonstrated that the ventral mPFC locates a trait code (Ma et al., 2013), that is, a distributed memory that encodes trait information and, when activated, enable access to this stored information (Wood and Grafman, 2003; Northoff and Bermpohl, 2004). The trait code is assumed to have several links to typical behaviors that exemplify the trait, and so facilitate easy understanding and trait attribution on the basis of behavior alone. This suggests that the trait code allows easy access already during early implicit processing (as an intuition). As another example, novel evidence on the role of the dorsal mPFC reveals that this area is strongly involved in generating high construals, that is, abstract categories extracted from lower-level behavior or object information (Baetens et al., 2013). As such, it appears to be the ideal candidate for high-level abstraction in an iterative reprocessing framework. An obvious prediction is that this abstraction function is mainly accessible in explicit as opposed to implicit processing, because only repetitive cycling might allow to reach complex abstractions and awareness about them. Finally, there is also novel evidence on the role of the dorsal mPFC in social working memory (Meyer et al., 2012). Keeping an increasing number of individuals in mind, also increased activity of the dorsal mPFC. Future research can explore how social working memory fits with an iterative reprocessing framework.

CONCLUSION

Previous neuroscientific research has demonstrated that understanding of another persons' mind involves both implicit and explicit processes located in the mentalizing network (e.g., Keysers and Gazzola, 2007). The old idea still present in some dual-process theories that implicit and explicit thought is inflexibly driven by entirely different underlying systems and brain areas (Smith and DeCoster, 2000; Strack and Deutsch, 2004) is contradicted by recent ERP and fMRI findings on mentalizing. Under implicit and explicit processing instructions, there was a shared early timing (ERP studies) and shared brain activity (fMRI studies) pointing to a single core system of social mentalizing (Van Overwalle, 2009). These data can be explained by

a default-interventionist dual process (Evans, 2008) in terms of iterative reprocessing (Cunningham and Zelazo, 2007; Cunningham et al., 2007; Freeman and Ambady, 2011). According to this framework, there is an implicit default core process that allows observers to make quick social mentalizing

inferences, presumably based on pre-existing learned social knowledge. This implicit core process is subserved by the TPJ and mPFC. Subsequent reprocessing cycles allow to take in more and richer information which enable observers to verify and flexibly correct their original rapid intuition.

REFERENCES

- Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277. doi: 10.1038/nrn1884
- Baetens, K., Ma, N., Steen, J., and Van Overwalle, F. (2013). Involvement of the mentalizing network in social and non-social high construal. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst048. [Epub ahead of print].
- Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546. doi: 10.1016/j.tics.2004.10.003
- Brosch, T., Schiller, D., Mojdehbabksh, R., Uleman, J. S., and Phelps, E. A. (2013). Neural mechanisms underlying the integration of situational information into attribution outcomes. *Soc. Cogn. Affect. Neurosci.* 8, 640–646. doi: 10.1093/scan/nst019
- Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., et al. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct.* 217, 783–796. doi: 10.1007/s00429-012-0380-y
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *J. Pers. Soc. Psychol.* 39, 752–766. doi: 10.1037//0022-3514.39.5.752
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Chris Gatenby, J., Gore, J. C., and Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychol. Sci.* 15, 806–813. doi: 10.1111/j.0956-7976.2004.00760.x
- Cunningham, W. A., and Zelazo, P. D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. *Trends Cogn. Sci.* 11, 97–104. doi: 10.1016/j.tics.2006.12.005
- Cunningham, W. A., Zelazo, P. D., Packer, D. J., and Van Bavel, J. J. (2007). The iterative reprocessing model: a multilevel framework for attitudes and evaluation. *Soc. Cogn.* 25, 736–760. doi: 10.1521/soco.2007.25.5.736
- Dennett, D. (2001). Are we explaining consciousness yet? *Cognition* 79, 221–237. doi: 10.1016/S0010-0277(00)00130-X
- Denny, B. T., Kober, H., Wager, T. D., and Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* 24, 1742–1752. doi: 10.1162/jocn_a_00233
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *Am. Psychol.* 49, 709–724. doi: 10.1037//0003-066X.49.8.709
- Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *J. Pers. Soc. Psychol.* 71, 390–405. doi: 10.1037//0022-3514.71.2.390
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.* 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629
- Evans, J. St. B. T. (2011). Dual-process theories of reasoning: contemporary issues and developmental applications. *Dev. Rev.* 31, 86–102. doi: 10.1016/j.dr.2011.07.007
- Evans, J. St. B. T., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Pers. Psychol. Sci.* 8, 223–241. doi: 10.1177/1745691612460685
- Forbes, C. E., Cox, C. L., Schmader, T., and Ryan, L. (2012). Negative stereotype activation alters interaction between neural correlates of arousal, inhibition and cognitive control. *Soc. Cogn. Affect. Neurosci.* 7, 771–781. doi: 10.1093/scan/nsr052
- Forbes, C. E., and Grafman, J. (2013). Social neuroscience: the second phase. *Front. Hum. Neurosci.* 7:20. doi: 10.3389/fnhum.2013.00020
- Freeman, J. B., and Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychol. Rev.* 118, 247–279. doi: 10.1037/a0022327
- Gilbert, D. T., and Malone, P. S. (1995). The correspondence bias. *Psychol. Bull.* 117, 21–38. doi: 10.1037//0033-2909.117.1.21
- Jacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annu. Rev. Psychol.* 60, 653–670. doi: 10.1146/annurev.psych.60.110707.163604
- Kestemont, J., Vandekerckhove, M., Ma, N., Van Hoek, N., and Van Overwalle, F. (2013). Situation and person attributions under spontaneous and intentional instructions: an fMRI study. *Soc. Cogn. Affect. Neurosci.* 8, 481–493. doi: 10.1093/scan/nss022
- Keysers, C., and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends Cogn. Sci.* 11, 194–196. doi: 10.1093/scan/nss022
- Kovacs, Á. M., Téglás, E., and Endress, A. D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science* 330, 1830–1834. doi: 10.1126/science.1190792
- Letzring, T. D., Wells, S. M., and Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *J. Pers. Soc. Psychol.* 91, 111–123. doi: 10.1037/0022-3514.91.1.111
- Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58, 259–289. doi: 10.1146/annurev.psych.58.110405.085654
- Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., and Baron-Cohen, S. (2011). Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *Neuroimage* 56, 1832–1838. doi: 10.1016/j.neuroimage.2011.02.067
- Ma, N., Baetens, K., Vandekerckhove, M., Kestemont, J., Fias, W., and Van Overwalle, F. (2013). Traits are represented in the medial prefrontal cortex: an fMRI adaptation study. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst098. [Epub ahead of print].
- Ma, N., Vandekerckhove, M., Baetens, K., Van Overwalle, F., Seurinck, R., and Fias, W. (2012a). Inconsistencies in spontaneous and intentional trait inferences. *Soc. Cogn. Affect. Neurosci.* 7, 937–950. doi: 10.1093/scan/nsr064
- Ma, N., Vandekerckhove, M., Van Hoek, N., and Van Overwalle, F. (2012b). Distinct recruitment of temporo-parietal junction and medial prefrontal cortex in behavior understanding and trait identification. *Soc. Neurosci.* 7, 591–605. doi: 10.1080/17470919.2012.686925
- Ma, N., Vandekerckhove, M., Van Overwalle, F., Seurinck, R., and Fias, W. (2011). Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: spontaneous inferences activate only its core areas. *Soc. Neurosci.* 6, 123–138. doi: 10.1080/17470919.2010.485884
- Meyer, M. L., Spunt, R. P., Berkman, E. T., Taylor, S. E., and Lieberman, M. D. (2012). Evidence for social working memory from a parametric functional MRI study. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1883–1888. doi: 10.1073/pnas.1121077109
- Morelli, S. A., and Lieberman, M. D. (2013). The role of automaticity and attention in neural processes underlying empathy for happiness, sadness, and anxiety. *Front. Hum. Neurosci.* 7:160. doi: 10.3389/fnhum.2013.00160
- Northoff, G., and Bermpohl, F. (2004). Cortical midline structures and the self. *Trends Cogn. Sci.* 8, 102–107. doi: 10.1016/j.tics.2004.01.004
- Pascual-Marqui, R. D. (1999). Review of methods for solving the EEG inverse problem. *Int. J. Bioelectromagn.* 1, 75–86.
- Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. J. Psychophysiol.* 18, 49–65. doi: 10.1016/0167-8760(84)90014-X
- Qureshi, A. W., Apperly, I. A., and Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: evidence from a dual-task study of adults. *Cognition* 117, 230–236. doi: 10.1016/j.cognition.2010.08.003
- Rameson, L. T., Satpute, A. B., and Lieberman, M. D. (2010). The neural correlates of implicit and explicit self-relevant processing. *Neuroimage* 50, 701–708. doi: 10.1016/j.neuroimage.2009.12.098
- Roets, A., and Van Hiel, A. (2007). Separating ability from need: Clarifying the dimensional structure of the need for closure scale.

- Pers. Soc. Psychol. Bull.* 33, 266–280. doi: 10.1177/0146167206294744
- Roets, A., and Van Hiel, A. (2011). Allport's prejudiced personality today: need for closure as the motivated cognitive basis of prejudice. *Curr. Dir. Psychol. Sci.* 20, 349–354. doi: 10.1177/0963721411424894
- Satpute, A. B., and Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Res.* 1079, 86–97. doi: 10.1016/j.brainres.2006.01.005
- Schneider, D., Bayliss, A. P., Becker, S. I., and Dux, P. E. (2012a). Eye movements reveal sustained implicit processing of others' mental states. *J. Exp. Psychol. Gen.* 141, 433–438. doi: 10.1037/a0025458
- Schneider, D., Lam, R., Bayliss, A. P., and Dux, P. E. (2012b). Cognitive load disrupts implicit theory-of-mind processing. *Psychol. Sci.* 23, 842–847. doi: 10.1177/0956797612439070
- Schneider, W., and Shiffrin, R. (1977). Controlled and automatic human information processing: i. detection, search, and attention. *Psychol. Rev.* 84, 1–66. Available online at: <http://psycnet.apa.org/journals/rev/84/1/1/>
- Smith, E. R., and DeCoster, J. (2000). Dual-process models in social and cognitive psychology: conceptual integration and links to underlying memory systems. *Pers. Soc. Psychol. Rev.* 4, 108–131. doi: 10.1207/S15327957PSPR0402_01
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803_1
- Timmermans, B., Schilbach, L., Pasquali, A., and Cleeremans, A. (2012). Higher order thoughts in action: consciousness as an unconscious re-description process. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1412–1423. doi: 10.1098/rstb.2011.0421
- Todorov, A., Baron, S. G., and Oosterhof, N. N. (2008a). Evaluating face trustworthiness: a model based approach. *Soc. Cogn. Affect. Neurosci.* 3, 119–127. doi: 10.1093/scan/nsn009
- Todorov, A., Said, C. P., Engell, A. D., and Oosterhof, N. N. (2008b). Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* 12, 455–460. doi: 10.1016/j.tics.2008.10.001
- Uleman, J. S. (1999). "Spontaneous versus intentional inferences in impression formation," in *Dual-Process Theories in Social Psychology*, eds Y. Chaiken and S. Trope (New York, NY: The Guilford Press), 141–160.
- Vandekerckhove, M., and Panksepp, J. (2011). A neurocognitive theory of higher mental emergence: from anoetic affective experiences to noetic knowledge and auto-noetic awareness. *Neurosci. Biobehav. Rev.* 35, 2017–2025. doi: 10.1016/j.neubiorev.2011.04.001
- Van der Cruyssen, L., Van Duynslaeger, M., Cortoos, A., and Van Overwalle, F. (2009). ERP time course and brain areas of spontaneous and intentional goal inferences. *Soc. Neurosci.* 4, 165–184. doi: 10.1080/17470910802253836
- Van Duynslaeger, M., Van Overwalle, F., and Verstraeten, E. (2007). Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Soc. Cogn. Affect. Neurosci.* 2, 174–188. doi: 10.1093/scan/nsm016
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858. doi: 10.1002/hbm.20547
- Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584. doi: 10.1016/j.neuroimage.2009.06.009
- Van Overwalle, F., Van Duynslaeger, M., Coomans, D., and Timmermans, B. (2011). Spontaneous goal inferences are often inferred faster than spontaneous trait inferences. *J. Exp. Soc. Psychol.* 48, 13–18. doi: 10.1016/j.jesp.2011.06.016
- Winter, L., and Uleman, J. S. (1984). When are social judgments made? evidence for the spontaneousness of trait inferences. *J. Pers. Soc. Psychol.* 47, 237–252. doi: 10.1037//0022-3514.47.2.237
- Wood, J. N., and Grafman, J. (2003). Human prefrontal cortex: processing and representational perspectives. *Nat. Rev. Neurosci.* 4, 139–147. doi: 10.1038/nrn1033

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 June 2013; accepted: 22 August 2013; published online: 13 September 2013.

Citation: Van Overwalle F and Vandekerckhove M (2013) Implicit and explicit social mentalizing: dual processes driven by a shared neural network. *Front. Hum. Neurosci.* 7:560. doi: 10.3389/fnhum.2013.00560

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Van Overwalle and Vandekerckhove. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.