Dynamic causal modeling of the effective connectivity between the cerebrum and cerebellum in social mentalizing across five studies

Frank Van Overwalle¹ · Frederik Van de Steen² · Peter Mariën^{3,4}

Published online: 25 October 2018
© Psychonomic Society, Inc. 2018

Abstract

In this analysis we explored the effective connectivity of the cerebellum with the cerebrum in social mentalizing, across five studies ($n = 91$) involving abstract and complex forms of mentalizing, such as (a) person and group impression formation, based on behavioral descriptions, and (b) constructing personal counterfactual events. Connectivity was analyzed by applying dynamic causal model analysis, which revealed effective connectivity between the mentalizing areas of the cerebellum and cerebrum. The results revealed a significant pattern of bidirectional (closed-loop) connectivity linking the right posterior cerebellum with bilateral temporo-parietal junction (TPJ), associated with behavior understanding. These connections are consistent with known anatomical data on closed loops between the cerebellum and cerebrum, although contralateral closed loops typically dominate. This analysis improves on an earlier psychophysiological interaction analysis of this dataset, which had failed to reveal such evidence of closed loops. Within the cerebrum, there were connections between the bilateral areas of TPJ, as well as connections between bilateral TPJ and the (ventral and dorsal) medial prefrontal cortex. The discussion centers on the function of cerebro-cerebellar connections in generating internal cerebellar “forward” models, potentially serving the automatic understanding, prediction, and error correction of behavioral sequences.

Keywords Dynamic causal modeling · Effective connectivity · Social mentalizing · Cerebellum

Current insights into the nonmotor role of the cerebellum are growing rapidly. During the last decade, evidence from different strands of research has convincingly demonstrated that the cerebellum not only is involved in the modulation of sensorimotor responses but also is crucially implicated in higher-level cognitive and affective functions, and only very recently has it become clear that the cerebellum might also play an important role in social cognition. *Social cognition* refers to the process of inferring the mental states of persons on the

basis of their behaviors (i.e., “mind” reading or mentalizing). Although the focus of neuroscientific research on mentalizing during the last decade has been on the cerebrum (for reviews, see Molenberghs, Johnson, Henry, & Mattingley, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle, 2009; Van Overwalle & Baetens, 2009), a fundamental role of the cerebellum has recently been highlighted in a meta-analysis by Van Overwalle, Baetens, Mariën, and Vandekerckhove (2014). This large-scale meta-analysis, including over 350 functional magnetic resonance imaging (fMRI) studies, identified robust clusters in the cerebellum that showed activity in about one third of the fMRI studies on social cognition, and in almost all studies that involved complex and abstract social inferences (cf. Trope & Liberman, 2010). Abstract mentalizing involves, for instance, judgments about a person’s nonobservable traits, as opposed to those available from visual descriptions (e.g., judging “why” vs. “how” a person is reading a book; Baetens, Ma, Steen, & Van Overwalle, 2014), or imagination of a nonobservable past or future, as opposed to the momentary present (Van Hoeck et al., 2013).

A major theoretical issue is whether the cerebellum in these contexts serves a domain-general function that modulates

✉ Frank Van Overwalle
Frank.VanOverwalle@vub.ac.be

¹ Faculty of Psychology and Educational Sciences, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

² Department of Data Analysis, University of Ghent, Dunantlaan 2, 9000 Ghent, Belgium

³ Faculty of Arts, Department of Clinical and Experimental Neurolinguistics, CLIN, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

⁴ Department of Neurology and Memory Clinic, ZNA Middelheim Hospital, Lindendreef 1, B-2020 Antwerp, Belgium

many processes, including social cognition, in a uniform manner (termed the *universal cerebellar transform*; Guell, Gabrieli, & Schmahmann, 2018). A uniform, domain-general role would suggest that social processes activate the cerebellum in a distributed fashion at many locations. An alternative position is that the cerebellum performs domain-specific implementations of an underlying general function, depending on the input coming from the cerebral cortex. The latter position is supported by a sizable body of evidence documenting that some parts of the cerebellum are specialized in motor processing, whereas others are selectively recruited during cognitive and affective functions (Keren-Happuch, Chen, Ho, & Desmond, 2014; Stoodley, MacMore, Makris, Sherman, & Schmahmann, 2016; Stoodley & Schmahmann, 2009, 2010; Stoodley, Valera, & Schmahmann, 2012). This domain-specific role would suggest that social processes activate a specialized function in the cerebellum and recruit limited areas there. Although initially a domain-general modulatory explanation for social cognition was offered (Van Overwalle et al., 2014), novel evidence has substantially favored the view of a domain-specific process (Van Overwalle, Baetens, Mariën, & Vandekerckhove, 2015).

One of the main arguments put forward in support of a domain-specific view of social cognition was that the cerebellar clusters involved in mentalizing show a strong overlap with the default/mentalizing cerebellar network, as identified by Buckner, Krienen, Castellanos, Diaz, and Yeo (2011; see also Buckner, 2013). Buckner et al. investigated the large-scale organization of circuits between the cerebrum and cerebellum using resting-state connectivity for a total sample of 1,000 participants, and they found a topography of similar network structures in the cerebellum as in the cerebrum (Yeo et al., 2011). This finding was corroborated by a recent meta-analytic connectivity analysis involving 34 studies (578 participants) showing unique cerebro-cerebellar links between the mentalizing networks of the cerebellum and cerebrum (Van Overwalle, D'aes, & Mariën, 2015). Note that the default network in the cerebrum and cerebellum identified during a resting state (Buckner et al., 2011; Raichle et al., 2001) typically shows a strong overlap with the social mentalizing network (Schurz et al., 2014; Van Overwalle, 2009), and even more so when social mentalizing about the present and future self and others is included (Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010). In the cerebellum, the correspondence between the default network and social mentalizing areas is very striking (see Van Overwalle, D'aes, & Mariën, 2015).

Important evidence for cerebro-cerebellar links has additionally been provided via a multistudy analysis exploring functional connectivity within individual participants pooled across five published studies (92 participants; Van Overwalle & Mariën, 2016). This analysis has a major advantage over meta-analytic data, because its unit of analysis was at the level

of individuals rather than entire studies, and because it explored individual connectivity dynamically in time rather than by static coactivation between studies. The multistudy analysis investigated social inferences based on a person's traits, group stereotypes, and a person's own past, which, given their high level of abstractness, revealed strong recruitment of the cerebellum. Using psychophysiological interaction (PPI) analysis (Friston, Harrison, & Penny, 2003; O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012), Van Overwalle and Mariën, (2016) analyzed the connectivity between mentalizing areas in the cerebrum (Schurz et al., 2014; Van Overwalle, 2009) and mentalizing areas in the cerebellum as revealed by Buckner et al. (2011). They found reliable functional connectivity between areas of the mentalizing network in the right posterior cerebellum (i.e., the inferior semilunar lobule, with center at Montreal Neurological Institute (MNI) coordinates 25 – 75 – 40) and in the cerebrum, including bilateral temporo-parietal junction (TPJ) and dorsomedial prefrontal cortex (dmPFC). This clearly supports domain-specific functional connectivity between the cerebellum and cerebrum during various mentalizing tasks. However, because PPI connectivity analysis has some methodological limitations, a number of the specific findings are disputable.

A first limitation is that a PPI analysis only assesses condition-specific changes in the linear relationship between the observed neural signals, and it is therefore not easily amenable to causality interpretations concerning the underlying mechanism (Friston et al., 1997). A second limitation is that PPI is set up as a regression model testing the assumption that an interaction between social reasoning (in comparison with a control condition) and activity in a source area (e.g., the cerebellum) is maximally expressed in a target area (e.g., cerebrum). This regression analysis is set up for each source area independently. This raises two problems. First, a connection might go undetected because it is always “open” and not modulated by the experimental manipulation. Second, a connection between two areas might be “false” because it is actually realized only indirectly, via a third area. PPI is blind to these interpretational difficulties.

To overcome these limitations, in the present connectivity analysis we made use of dynamic causal modeling (DCM; Friston et al., 2003; Stephan et al., 2010), which recently has been augmented with automated tests of all connections in the model for groups of participants (Friston et al., 2016; Friston, Zeidman, & Litvak, 2015). First, a DCM approach allows for specifying “fixed” or “endogenous” connectivity, unmodulated by the experimental manipulation, as well as “modulatory” connectivity regulated by the experimental conditions. Second, DCM allows for estimating and testing all connections at once in a single model, so that “indirect” connections via other areas (which may create “false” connectivity) are better controlled for. Evidently, this is only the case when all relevant (e.g., task-activated) areas are included in the model.

Third, DCM estimates and tests the two directions of connectivity separately (from one area to the other, and vice versa), so that it is possible to make a causal interpretation of neural processing loops.

Given these powerful analytical features, DCM analysis allows for more biologically plausible tests of connectivity and of condition-related changes in connectivity. So far, anatomical studies on animals have indicated that the majority of cerebro-cerebellar connections are characterized by contralateral closed-loop circuits (Kelly & Strick, 2003). In a closed-loop circuit, an area of the cerebrum projects to a contralateral area of the cerebellum and receives input from that same cerebellar area. In addition, some studies have suggested the existence of additional bilateral and ipsilateral loops in rats (Suzuki, Coulon, Sabel-Goedknecht, & Ruigrok, 2012) and humans (Cui et al., 2000; Salmi et al., 2010; Sokolov, Erb, Grodd, & Pavlova, 2014). In support of this anatomical evidence, a human functional connectivity analysis by Krienen and Buckner (2009) revealed that the connections from the cerebral cortex terminate in contralateral areas in the cerebellum 70%–80% of the time, and in ipsilateral areas 20%–30%. Contrary to this evidence, in their PPI analysis mentioned earlier, Van Overwalle and Mariën (2016) failed to find any of such contralateral or ipsilateral closed-loop circuits (see also Fig. 1B below).

By applying a more biologically plausible DCM analysis to the same dataset from Van Overwalle and Mariën (2016), the aim of the present study was to go beyond functional PPI associations and investigate the effective causal connectivity between mentalizing areas of the cerebrum and the cerebellum. Specifically, we expected to find substantial closed-loop connections between the cerebrum and cerebellum that failed to show up in the earlier PPI analysis by Van Overwalle and Mariën. We also expected that these closed loops would mainly be related to mid-section or contralateral cerebral areas, in line with anatomical studies (e.g., Kelly & Strick, 2003). Finally, the DCM analysis would allow us to see whether and which connections between the cerebrum and cerebellum are modulated during mentalizing, as suggested by the earlier PPI analysis. We tested our hypotheses on the neuroimaging data from five studies that were also used in the earlier PPI analysis by Van Overwalle and Mariën.

Method

Selected studies

DCM (Friston et al., 2003; Friston et al., 2016; Friston et al., 2015; Stephan et al., 2010) was applied in five published fMRI studies from our lab (Baetens et al., 2014; Ma, Vandekerckhove, Baetens, et al., 2012a; Ma, Vandekerckhove, Van Hoeck, & Van Overwalle, 2012b; Van

der Cruyssen, Heleven, Ma, Vandekerckhove, & Van Overwalle, 2015; Van Hoeck et al., 2013), which were also used in a PPI analysis by Van Overwalle and Mariën (2016). These studies were selected because they showed activity in the mentalizing network in both the cerebrum and cerebellum and involved higher-level complex social inferences involving person and group traits, as well as a person's own past. All of the studies involved human actions, either depicted visually (e.g., showing a person reading a book) or described verbally by short sentences (e.g., “gives his mother a slap”) or by verbal cues triggering personal memories of the participants. All studies revealed increased activation in core mentalizing areas (including the medial prefrontal cortex [mPFC], bilateral TPJ, and cerebellum) during the critical condition, in comparison with a control condition. The selected studies also used identical experimental and scanning procedures and a software program (SPM) that were used again in the PPI analysis by Van Overwalle and Mariën, so that comparison between the PPI and DCM analyses was straightforward.

Participants

The participants in all studies ($n = 91$) were healthy and right-handed (using the Edinburgh Handedness Inventory; Oldfield, 1971), with no neurological or psychiatric antecedents. One additional participant from the initial pool of Van Overwalle and Mariën (2016) was omitted because of technical problems with the data. The total numbers of participants are given in Table 1. The studies were approved by the Medical Ethics Committees of the University Hospital of Ghent (where all studies were conducted) and the Vrije Universiteit Brussel (of the principal investigator F.V.O.). Written informed consent was obtained from each participant.

Design, stimulus material, and procedure

We will only briefly summarize the essential aspects of the design (primary contrast between the critical experimental vs. control/contrast conditions), materials, and procedure. For more details, we refer readers to Baetens et al. (2014); Ma, Vandekerckhove, Baetens, et al. (2012a); Ma, Vandekerckhove, Van Hoeck, & Van Overwalle (2012b); Van der Cruyssen et al. (2015); and Van Hoeck et al. (2013). The main results of each study are summarized briefly in Van Overwalle and Mariën (2016).

Study 1 (Ma, Vandekerckhove, Baetens, et al., 2012a; $n = 29$): Inconsistent (> consistent) trait inferences based on brief trait-implicating behavioral sentences ($n = 15$ intentional and $n = 14$ spontaneous trait inferences). The participants read 16 sets of three or four behavioral sentences that implied a social trait of a person (e.g., “Jun gives a smile,” implying friendly). In the experimental condition, the last sentence was inconsistent

Fig. 1 (A) Reduced model (with individual studies as covariates) showing significant direct connections (posterior $p > .95$) between the key areas of the right posterior cerebellum, precuneus (Pc), left and right temporo-parietal junction (TPJ), and ventral/dorsal mPFC (v/dmPFC). There are no significant connections from the mPFC to bilateral TPJ, except for a sole link from the vmPFC to right TPJ (-.07), not shown

here. The connectivity estimates correspond to rate constants and are expressed in units of 1/s (Hz). We specified all ROIs receiving external input in both conditions. (B) Results of a PPI analysis on the same dataset by Van Overwalle and Mariën (2016). In both panels, the white area in the cerebellum refers to the mentalizing network according to Buckner et al. (2011)

with the prior sentences, whereas in the contrast condition the last sentence was consistent.

Study 2 (Ma, Vandekerckhove, Van Hoeck, & Van Overwalle, 2012b; $n = 13$): Trait ($>$ no trait) inferences based on brief trait-implying behavioral sentences. In the experimental condition, the participants read 20 behavioral sentences that implied a social trait of a person (similar to Study 1); in the contrast condition, they read 20 no-trait sentences describing intransitive behaviors that did not involve any interaction with other objects or persons (e.g.,

“Tolvan moves her right hand”). One participant from this study was removed from the PPI analysis because of technical problems with the data.

Study 3 (Baetens et al., 2014; $n = 18$): Trait inferences ($>$ visual descriptions) based on photos depicting human behavior. In the experimental condition, participants viewed 30 pictures of a person engaged in everyday activities (e.g., a person reading a book), from which they had to infer a trait, whereas in the contrast condition they had to describe the same pictures by its visual features.

Table 1 Total number of participants (n) and number of participants with ROIs identified at reduced thresholds

Study		<i>n</i>	ROIs											
			Cerebellum		Precuneus		ITPJ		rTPJ		dmPFC		vmPFC	
			<.10	<1	<.10	<1	<.10	<1	<.10	<1	<.10	<1	<.10	<1
1	Ma et al. (2012a; explicit)	15	6	3	3	2	3	2	–	3	2	1	2	–
1	Ma et al. (2012a; implicit)	14	3	7	2	–	2	1	2	1	4	–	–	1
2	Ma et al. (2012b)	13	1	2	2	–	–	1	–	1	1	3	–	1
3	Baetens et al.(2014)	18	–	–	–	1	–	–	–	–	–	–	–	–
4	Van der Cruyssen et al. (2015)	18	–	3	1	4	–	2	3	1	1	1	1	3
5	Van Hoeck et al. (2013)	13	–	–	–	–	–	–	–	–	–	–	–	–

Study 4 (Van der Cruyssen et al., 2015; $n = 18$): Categorical group (> person trait) inferences based on brief behavioral sentences. The participants read 40 behavioral sentences that implied a stereotypical characteristic of a social group (e.g., “smiles to the baby” by a kindergarten teacher) versus an equal number of behavioral descriptions that implied a social trait of a person.

Study 5 (Van Hoeck et al., 2013; $n = 13$): Counterfactual imagining (> semantic memory) based on past autobiographic memories induced by brief cue words. In the experimental condition, participants read three cue words that reminded them of 20 negative autobiographic events, from which they had to construct and imagine a more favorable counterfactual past (e.g., how an accident on a holiday might have been avoided). This required mentalizing or thinking about mental states, events, and people’s actions (e.g., if my daughter had been more careful and not daydreaming, she would have noticed the dangerous obstacle while cycling). In the contrast condition, they had to imagine the semantic meaning of each cue word in detail.

In the spontaneous trait condition of Ma, Vandekerckhove, Baetens, et al. (2012a), participants simply read the sentences, and trait inferences were checked by memory measures after scanning. In all other participant groups (including the intentional trait condition of Ma, Vandekerckhove, Baetens, et al. 2012a), participants had to provide a response after reading or viewing the stimuli. Specifically, they had to rate the applicability of a specific trait, such as “friendly,” or social category, such as “kindergarten teacher,” by pressing the appropriate response key, or they had to think their counterfactual thoughts silently and then write them down after scanning. The counterfactual study by Van Hoeck et al. (2013) contained only negative past events; from the study by Baetens et al. (2014), we extracted only positive events; and in the remaining two studies, the sentences varied in valence. The experiments were event-related, and all trials were presented in a random order for each participant.

Imaging procedure

The imaging procedures were identical for all studies. Images were collected with a 3-T Magnetom Trio MRI scanner system (Siemens medical Systems, Erlangen, Germany) using an eight-channel radiofrequency head coil. Stimuli were projected onto a screen at the end of the magnet bore that participants viewed by way of a mirror mounted on the head coil. Stimulus presentation was controlled by E-Prime 2.0 (www.pstnet.com/eprime; Psychology Software Tools) under Windows. Foam cushions were placed within the head coil in order to minimize head movements. A high-resolution T1-weighted structural scan (MP-RAGE) was collected, followed by one functional run (30 axial slices; 4 mm thick; 1 mm skip). The functional scanning used a gradient-

echo echoplanar pulse sequence (TR = 2 s; TE = 33 ms; $3.5 \times 3.5 \times 4.0$ mm in-plane resolution).

Image processing and statistical analysis

The fMRI data were preprocessed and analyzed using SPM5 (Study 1, 2, and 5) or SPM8 (Studies 3 and 4; Wellcome Department of Cognitive Neurology, London, UK). We kept the original data analysis, in order to stick as close as possible to the published results, including the earlier PPI by Van Overwalle and Mariën (2016), given that there were no relevant major changes between these two SPM versions. The data were collected in a single functional run and preprocessed to remove sources of noise and artifacts. The functional data were corrected for differences in acquisition times between the slices for each whole-brain volume, realigned within and across runs to correct for head movement, and coregistered with each participant’s anatomical data. The functional data were then transformed into a standard anatomical space (2-mm isotropic voxels) based on the ICBM 152 brain template (MNI). The normalized data were then spatially smoothed (6-mm full width at half maximum) using a Gaussian kernel. Participants with excessive motion artifacts were excluded from the study and are not included in the counts per study mentioned earlier.

In Studies 3–5, the realigned data were additionally examined for excessive motion artifacts and for correlations between motion or global mean signal and any of the conditions using the Artifact Detection Tool software package (ART; www.nitrc.org/projects/artifact_detect). Outliers were identified in the temporal difference series by assessing between-scan differences (Z threshold, 3.0; scan-to-scan movement threshold, 0.5 mm; rotation threshold, 0.02 radians). These outliers were treated in the analysis by including a single regressor for each outlier, which may have captured and so removed from the estimation some of the noise caused by the motion. The participants included in the studies had less than 10% outliers and showed no correlations between motion and the experimental design or between the global signal and the experimental design.

The timing for the analysis of all conditions was set at the start of the behavioral sentences, memory cues, and/or questions. Statistical analyses were performed using the general linear model of SPM5 or SPM8, in which the event-related design was modeled with one regressor for each condition, using a canonical hemodynamic response function, six movement artifact regressors, and, when using ART, additional outlier vectors. A default high-pass filter of 128 s was used, and serial correlations were accounted for by the default autoregressive AR(1) model.

Dynamic causal modeling

To apply DCM, regions of interest (ROIs) representing core mentalizing areas were taken from Van Overwalle and Mariën (2016). The cerebral ROIs were centered around the mean MNI coordinates reported in the meta-analyses by Van Overwalle (2009) and Van Overwalle and Baetens (2009), including the dmPFC (with center at 0 50 35), ventral mPFC (vmPFC: 0 50 5), bilateral TPJ ($\pm 50 - 55 25$), and precuneus/posterior cingulate (Pcun: 0 – 60 40). The cerebellar ROI was identified by Van Overwalle and Mariën and involved the right posterior cerebellar lobe (25 – 75 – 40). As was reported in more detail by Van Overwalle and Mariën, individually tailored cerebral ROIs were created by extracting the time series using the eigenvariate within a sphere with a radius of 8 mm around the nearest local maximum within 15 mm of the corresponding ROI centers listed above, after setting a whole-brain threshold of the contrast at $p < .05$ (uncorrected). All voxels contributing to the ROI were adjusted for the effects of interest (i.e., only for the two conditions that were used in the DCMs; see the Design, Stimulus Material, and Procedure section above). Individually tailored cerebellar ROIs were created in a similar manner, except that the radius was confined to 5 mm around the nearest local maximum within 11 mm of the corresponding ROI center, given the smaller size of the cerebellum (and of its mentalizing network). If the ROI did not contain a peak surviving the threshold, the same procedure was repeated with $p < .10$ and $p < 1.00$ (uncorrected) so that the time series of all ROIs were included for all participants (see Table 1 for the number of ROIs satisfying these criteria). In the latter case (i.e., $p < 1.00$), ROIs were centered around the group-based centers listed above (see also Zhou et al., 2007). The latter procedure of setting increasingly tolerant thresholds differed from the earlier PPI analysis, in which participants were excluded pairwise from the analysis when the threshold ($p < .05$, uncorrected) was not satisfied. The reason was that pairwise exclusion is not possible in a DCM analysis, because time series are needed from the full network, including all ROIs. By setting a more tolerant threshold in DCM, all of the available time series data from all ROIs were included in the analysis, which implies an optimal compromise between maximizing the effect of interest at the individual level while having all participants in the analysis, so that the results are not biased by excluding some participants (Zhou et al., 2007).

To estimate the optimal DCM across all participants and studies, we followed the procedures described in Friston et al. (2016) and Friston, Zeidman, and Litvak (2015) and given in full detail in <https://en.wikibooks.org/wiki/User:Peterz/sandbox>.

First, a full DCM was specified and estimated for each participant using SPM12 (cf. the SPM procedure: `spm_dcm_fit`). A full model allows all connectivity

parameters in all directions to be freely estimated. We specified a bilinear deterministic DCM without centering around the mean (Friston et al., 2003), which included (a) all forward and backward *fixed connections* between the ROIs, (b) all the *modulatory connections* or parameters that reflected condition changes due to the mentalizing condition in particular connections, and (c) *direct input* parameters that reflected the input driving the activity in both conditions in the ROIs. Stated differently, the driving input in Matrix C consists of one vector with all the onsets of both the mentalizing and control conditions combined as one input, and the modularity connections in Matrix B are specified only for the mentalizing condition, so that both matrix inputs are nonredundant to each other (Hillebrandt, Friston, & Blakemore, 2015). The order of the conditions in the DCM for all studies was arranged so that the two conditions of interest were aligned (i.e., the critical experimental and control conditions specified above for each study; see the Design, Stimulus Material, and Procedure section). Note that the connections in DCMs correspond to rate constants and are expressed in units of 1/s (i.e., of hertz).

Second, we constructed a parametric empirical Bayes (PEB) model for the whole group of participants over all parameters (cf. the SPM procedure: `spm_dcm_peb`). This makes it possible to estimate the effective connectivity averaged across all participants (cf. the group average), taking into account the within-participants variability on the connectivity parameters, unlike in a classical test (e.g., a t test), which ignores the estimated uncertainty (variance) about the connection strengths. Moreover, a group-level PEB allows for controlling for differences between sets of studies by treating them as covariates (e.g., differences in the behavioral measures and procedures; Friston et al., 2016). We first controlled for differences between studies by contrasting each study against the first (e.g., Study 1 > 2, Study 1 > 3, etc.). In addition, as listed in Table 2, we controlled for type of judgment—that is, whether judgments involved (a) personal traits or (b) social traits (personal traits and group stereotypes) and (c) were requested explicitly or implied implicitly/spontaneously. We also controlled for (d) the consistency of the behavior—that is, whether events were consistent and shared a common trait versus whether events were inconsistent with respect to a trait or with the past (cf. counterfactual).

Third, we automatically pruned away any connectivity parameter from the group-level PEB that did not contribute to the model evidence, using Bayesian model reduction (cf. the SPM procedure: `spm_dcm_peb_bmc`). This approach has the advantage that any reduced model at the group level can be estimated efficiently without having to reestimate the reduced models at the lower level (single-participant levels), and is therefore recommended (Friston et al., 2015). Specifically, a greedy search iteratively prunes connection parameters from the full model until the model evidence starts to decrease, so that the most relevant nested models from the full PEB model

Table 2 Study contrasts and covariates (i.e., contrasts between studies)

Study		Contrast	Covariates			
			Personal Traits	Social (incl. Personal) Traits	Explicit Traits	Event Consistency
1	Ma et al. (2012a; explicit)	Inconsistent trait > Consistent trait	1	1	1	− 1
1	Ma et al. (2012a; implicit)	Inconsistent trait > Consistent trait	1	1	− 5	− 1
2	Ma et al. (2012b)	Trait > No trait	1	1	1	1
3	Baetens et al.(2014)	Trait > Visual	1	1	1	1
4	Van der Cruyssen et al. (2015)	Group stereotype > Trait	− 2	1	1	1
5	Van Hoeck et al. (2013)	Counterfactual > Semantic	− 2	− 5	1	− 1

are tested (a greedy search is recommended because the model space of all possible nested models is too large to be fully evaluated). Bayesian model averaging of the parameters of the best 256 pruned models is applied and used for group inferences (Zhou et al., 2018, p. 707, their section 2.7.4), and so determines the winning model empirically. We considered connectivity parameters to be significant when their posterior probability $p > .95$ (based on model comparisons with and without each parameter). This Bayesian approach to both the first-level connectivity analysis (DCM) and group-level inference (PEB) on the connectivity parameters eschews the multiple-comparisons problem (Friston et al., 2003, p. 1276, their section 1.3)

Results

Tables 3 and 4 show the fixed and modulatory estimates, respectively, of the connectivity between the cerebral and cerebellar mentalizing ROIs of the reduced model (after pruning), either with or without controlling for the effects of the covariates. The estimates in the off-diagonal cells correspond to rate constants and are expressed in units of 1/s (Hz), or the amount by which activation in the source ROI changes the activation in the target ROI per second. The estimates in the diagonal cells, which are of less interest, express inhibition by self-connections in the format of a log scale. Negative values of the log scale indicate a hertz value between 0 and 1, or weak self-inhibition, whereas positive values indicate a hertz value above 1, or strong self-inhibition.

Several (combinations of) covariates (i.e., differences between studies; see Table 2) were introduced independently. Specifically, we first introduced differences between individual studies by contrasting each study against the first (e.g., Study 1 > 2, Study 1 > 3, etc.). Second, we also controlled for differences between the studies related to the type of mentalizing judgment: personal traits, social traits (including personal traits), and explicit (vs. implicit) traits. Third, we introduced differences between studies related to the

consistency of the event (e.g., whether events consistently implied the same trait or were inconsistent with respect to a shared trait or the past).

We predicted bidirectional connections (i.e., closed loops) between the (right) social cerebellar ROI and key mentalizing cerebral areas. On the basis of prior anatomical evidence, we also predicted that most closed loops would connect the cerebellum with contralateral cerebral areas. As one might expect, we found the most robust DCM results after controlling for all the differences between studies (Fig. 1A), and we therefore focus here on these results. Consistent with our predictions, we found robust closed-loop connections with the contralateral TPJ as well as with ipsilateral TPJ. In contrast, there were no cerebral connections with the mPFC. In addition, we observed a unidirectional connection from the cerebellum to the precuneus, but no reverse link. It is of interest to note that all of the significant connections coming from the cerebellum were negative, whereas all connections going toward the cerebellum were positive, which might indicate that the cerebellum was sending out error signals that turned the connectivity negative. Finally, we found no modulation of any connection with the cerebellum, except for one significant modulation after controlling only for consistency (Table 4).

The connections within the cerebrum were largely as one might expect (Table 3). There were bidirectional connections between the left and right TPJ and from bilateral TPJ to the ventral and dorsal mPFC, but almost no reverse connections from the mPFC (Fig. 1A), except a reverse link from the vmPFC to the right TPJ (– .07; not shown). We also found unidirectional connections from bilateral TPJ to the precuneus, but not the other way around. Finally, modulation of connectivity in the mentalizing condition was very rare (Table 4). If there was modulation, it mostly involved up- as well as down-regulating connectivity involving the dmPFC, and occasionally the left or right TPJ. In addition, none of the driving inputs to the ROIs reached significance, regardless of which covariation was controlled for.

Table 3 Averaged fixed connections with or without controlling for covariates, in units of 1/s (Hz)

to \ from	Cerebellum	Pcun	lTPJ	rTPJ	dmPFC	vmPFC
Fixed connectivity covarying out:						
Individual studies (Study 1 > 2, Study 1 > 3, Study 1 > 4, Study 1 > 5)						
Cerebellum	− 0.22*	0	0.09*	0.06*	0	0
Pcun	− 0.13*	− 0.23*	0.11*	0.21*	0	0.03
lTPJ	− 0.07*	0	− 0.41*	0.18*	0	0
rTPJ	− 0.08*	0	0.12*	− 0.46*	0	− 0.07*
dmPFC	0	0	0.07*	0.13*	− 0.34*	0
vmPFC	0	0	0.12*	0.09*	0.05*	− 0.22*
Fixed connectivity covarying out:						
Personal > Nonpersonal traits						
Social (incl. Personal) > Nonsocial traits						
Explicit > Implicit traits						
Cerebellum	− 0.23*	− 0	0.09*	0.03	0	0
Pcun	− 0.09*	− 0.23*	0.12*	0.20*	0	0.03
lTPJ	0	− 0.07*	− 0.39*	0.19*	0	0
rTPJ	0	0	0.12*	− 0.43*	0	− 0.07*
dmPFC	0	0	0.06*	0.12*	− 0.32*	0
vmPFC	0	0	0.13*	0.13*	0.03	− 0.22*
Fixed connectivity covarying out:						
Consistent > Inconsistent events						
Cerebellum	− 0.25*	0	0.10*	0	0	0
Pcun	− 0.11*	− 0.23*	0.12*	0.20*	0	0.06*
lTPJ	0	− 0.04	− 0.39*	0.18*	0	0
rTPJ	− 0.06*	0	0.15*	− 0.43*	0	− 0.05*
dmPFC	0	0	0.07*	0.11*	− 0.32*	0.03
vmPFC	− 0.03	0	0.13*	0.13*	0.06*	− 0.20*
Fixed connectivity without covariates						
Cerebellum	− 0.24*	0	0.10*	0	0	0
Pcun	− 0.10*	− 0.21*	0.11*	0.21*	0	0.03
lTPJ	0	− 0.04	− 0.36*	0.20*	0	0
rTPJ	0	0	0.13*	− 0.41*	0	− 0.05
dmPFC	0	0	0.05	0.12*	− 0.29*	0.03
vmPFC	0	0	0.13*	0.13*	0.03	− 0.20*

Cell entries refer to fixed connections, with * denoting posterior probabilities $p > .95$; the covariates are specified in Table 2. PCun = precuneus, l/rTPJ = left/right temporo-parietal junction, d/vmPFC = dorsal/ventral medial prefrontal cortex.

Discussion

In this study we investigated whether mentalizing areas in the cerebellum are effectively connected with mentalizing areas located in the cerebrum. A recent meta-analytic connectivity analysis involving various neuroimaging studies (Van Overwalle, D'ae, & Mariën, 2015) and a PPI connectivity analysis within individual participants pooled across five published studies (Van Overwalle & Mariën, 2016; 92 participants) revealed a consistent and robust pattern of functional connectivity between cerebellum and cerebrum. However, some results in the prior PPI analysis did not

conform to known anatomical evidence, and we reasoned that these divergences might be due to the methodological limitations of a PPI analysis with respect to modulation of connectivity by experimental manipulations and control of indirect connectivity (i.e., via other connections).

To overcome these limitations, we employed DCM to explore effective connectivity across the original data from Van Overwalle and Mariën (2016). As we explained in the introduction, DCM has many advantages over PPI. First, and perhaps most importantly, DCM allows for specifying the directionality of the connections so that we can verify the existence of closed loops. Second, it also allows for

Table 4 Averaged modulatory connections with or without controlling for covariates, in units of 1/s (Hz)

to \ from	Cerebellum	Pcun	lTPJ	rTPJ	dmPFC	vmPFC
Modulatory connectivity covarying out:						
Individual studies (Study 1 > 2, Study 1 > 3, Study 1 > 4, Study 1 > 5)						
Cerebellum	− 0.01	− 0.01	− 0.03	0.01	0.04	0.01
Pcun	− 0.02	− 0.02	− 0.02	0.01	0.71*	0
lTPJ	− 0.02	0.01	0	− 0.02	0.58*	− 0.01
rTPJ	0	− 0.01	− 0.02	− 0.58*	0.02	0
dmPFC	0	0.02	− 0.03	0	− 0.58*	0
vmPFC	− 0.02	0	− 0.03	0	0	− 0.59*
Modulatory connectivity covarying out:						
Personal > Nonpersonal traits						
Social (incl. Personal) > Nonsocial traits						
Explicit > Implicit traits						
Cerebellum	− 0.01	− 0.01	− 0.03	0.01	0.03	0
Pcun	− 0.02	− 0.01	0	0	0.52*	− 0.01
lTPJ	− 0.01	0.01	− 0.01	− 0.45*	0.47*	− 0.01
rTPJ	0	− 0.01	0	− 0.02	0.02	− 0.01
dmPFC	0	0.02	− 0.02	0	− 0.02	0
vmPFC	− 0.02	− 0	− 0.02	0	0.01	− 0.61*
Modulatory connectivity covarying out:						
Consistent > Inconsistent events of social traits						
Cerebellum	− 0.01	− 0.01	− 0.02	0	0.47*	0
Pcun	− 0.02	− 0.02	− 0.01	0.01	0.65*	− 0.01
lTPJ	− 0.02	0.01	0	− 0.02	0.50*	− 0.01
rTPJ	0	0	− 0.01	− 0.02	0.01	− 0.01
dmPFC	0	0.02	− 0.01	0	− 0.57*	− 0.01
vmPFC	− 0.02	0	− 0.01	0	0	− 0.25
Modulatory connectivity without covariates						
Cerebellum	− 0.01	− 0.01	− 0.02	0	0.02	0
Pcun	− 0.01	− 0.01	− 0.01	0.01	0.02	0
lTPJ	− 0.01	0.01	0	− 0.02	0.02	− 0.01
rTPJ	0	0	0	− 0.01	0.01	− 0.01
dmPFC	0	0.02	− 0.01	0	− 0.02	− 0.01
vmPFC	− 0.02	0	− 0.01	0	0.01	− 0.02

Cell entries refer to connections, with * denoting posterior probabilities $p > .95$; modulation is shown for the experimental conditions; the covariates are specified in Table 2. PCun = precuneus, l/rTPJ = left/right temporo-parietal junction, d/vmPFC = dorsal/ventral medial prefrontal cortex.

distinguishing between (a) connections that are fixed and independent of the experimental manipulation and (b) modulations of the experimental conditions on these connections. Third, DCM identifies only direct connections, by controlling for indirect effects via other connections specified in the model. Finally, the recently developed empirical Bayesian analysis makes it possible to take into account contrasts between studies that are due to differences in the measures and procedures. In this respect, DCM allows a much more accurate insight into the flow of neural propagation of activity between the cerebrum and cerebellum.

After pruning away connections that were not contributing to the model evidence, the analysis offered a dynamic causal model that provided the best fit with the data (see Fig. 1A). This final model identified bidirectional closed loops between the right posterior cerebellar lobe and bilateral TPJ, a key area of the mentalizing network. This confirms and extends the coactivations between the cerebellum and cerebrum during social reasoning from earlier meta-analytic work (Van Overwalle, D'aes, & Mariën, 2015), but it improves on the earlier PPI analysis, in which no closed loops were identified (Van Overwalle & Mariën, 2016). Interestingly, all connections originating from the cerebellum were negative, which

might indicate that the cerebellum emits a kind of error signal that is used by the cerebral cortex to adjust its activation in the opposite direction.

The closed loop with bilateral TPJ is in line with the role of the TPJ in understanding the socially relevant behavioral information provided in our studies (Schurz et al., 2014; Van Overwalle, 2009). Overall, this connectivity pattern indicates that social processes in the mentalizing network of the cerebrum—involving mainly the TPJ—trigger further activity in a domain-specific mentalizing network in the right posterior cerebellar lobe via closed loops with the left and right TPJ. This provides support for domain-specific connectivity between the mentalizing areas of the cerebrum and cerebellum, because activity in mentalizing areas of the cerebellum seem to be strongly associated with specific input and output in social mentalizing areas of the cerebrum. This speaks against a uniform, domain-general account that would predict no specificity in the activity or connectivity of social processing.

The connections are consistent with recent anatomical research showing predominantly closed loops from the cerebellum terminating in contralateral cerebral areas, in both animals (Kelly & Strick, 2003; Suzuki et al., 2012) and humans, although a minority of connections terminate in ipsilateral cerebral areas (Krienen & Buckner, 2009; Salmi et al., 2010; Sokolov et al., 2014). These ipsilateral TPJ connectivity findings seem to run counter to the anatomical evidence (Kelly & Strick, 2003) and suggest that the cerebro-cerebellar connections of the posterior (social) cerebellum might involve a great deal of ipsilateral connections, as was suggested by earlier research (20%–30%; Krienen & Buckner, 2009). Indeed, research with humans has reported ipsilateral connections mainly in posterior mentalizing cerebellar areas, including Crus 1 during the perception of human movements (Sokolov et al., 2014) or human musical sounds (Salmi et al., 2010). Future research might explore whether the presence of significant ipsilateral TPJ connectivity with the posterior (social) cerebellum is accidental or robust, and what this might imply for the functionality of the cerebellum in social cognition. Structural neuroimaging data, such as diffusion tensor imaging or tracing studies, will be needed in order to make stronger claims about the anatomical tracts along which social mentalizing information might propagate.

Within the cerebrum, the model identified unidirectional connections from bilateral TPJ to the vmPFC and dmPFC, and little reverse connectivity. These findings suggest that input flows mainly from the TPJ to the mPFC, in line with data showing earlier relevant neural activity in the TPJ than in the mPFC, as revealed by both electroencephalographic findings (Van der Cruyssen, Van Duynslaeger, Cortoos, & Van Overwalle, 2009; Van Duynslaeger, Sterken, Van Overwalle, & Verstraeten, 2008; Van Duynslaeger, Van Overwalle, & Verstraeten, 2007) and fMRI studies (Ma, Vandekerckhove, Van Hoeck, & Van Overwalle, et al., 2012). There were also

bidirectional connections between the bilateral TPJ areas. Finally, we observed additional connections to the precuneus, but none originating from this area.

The identification of close-loop circuits between the cerebellum and cerebrum, potentially with reverse error signals originating from the cerebellum, is in line with the cerebellar theory of Ito (2008) and further strengthens our ideas on the domain specificity during social cognition of an underlying general function of the cerebellum. We reasoned that an evolutionary older function of the cerebellum is forward control of action, implemented by constructing internal models of motor and sensory processes that facilitate automatic planning and understanding of action sequences. Extending on this earlier function, an evolutionarily recent adaptation is forward control of purely mental processes during cognitive operations and social reasoning, in which event sequences play an important role (Ito, 2008; Pisotta & Molinari, 2014). By building internal cerebellar models of social events, humans can anticipate action sequences during social interaction in an automatic and intuitive way. This role of internal sequence prediction is perhaps most prominent in mental reconstructions, devoid of direct observations, of autobiographic hypothetical counterfactual events, and of nonobservable traits and stereotypes. When these internal anticipations fail on the basis of novel evidence, it is assumed that the cerebellum fine-tunes the existing anticipations and makes corrective adjustments while the action is unfolding, and so creates an improved internal model (Ito, 2008; Pisotta & Molinari, 2014). Taken together, our connectivity data provide strong evidence for a circuit in which trait inferences, group stereotypes, and autobiographic events, extracted in the cerebrum, are further propagated to an internal model in the cerebellum, which—to foster fluent and automatic social interaction—acts as a forward controller that sends error signals to improve predictions concerning the implicated behavioral sequences.

This study also has several limitations. First, the results identified mainly fixed connections between the cerebrum and cerebellum, with very little experimental modulation, contrary to Van Overwalle and Mariën (2016; see Fig. 1B), in which a PPI connection implies, by definition, modulation of connectivity by the mentalizing as opposed to the control condition. The lack of modulation seems to speak against the view that the present cerebro-cerebellar connections are specifically related to social mentalizing. Perhaps this is because most conditions (including the control conditions) contained substantial social elements (i.e., persons), and only differed in the mentalizing judgment requested. Moreover, the finding of substantial fixed connectivity may be interpreted differently. It may suggest that the observed increase of activation during mentalizing (as opposed to control conditions) in specific mentalizing areas occurs at the same pace and with the same amount of synchrony in both the cerebellum and the cerebrum; this provides perhaps even stronger evidence for the

close connectivity between these two brain parts. Thus, one might argue that the lack of modulation strengthens our domain-specific hypothesis that the observed connections are always specifically attuned to social mentalizing at particular cerebral locations, whether the current input involves stronger or weaker social elements. If true, this interpretation suggests that the same connections should also be observed during the resting state. Anyway, DCM is a more biologically inspired model of connectivity, which could explain the discrepancies from the PPI analysis.

Second, combining several datasets to reveal robust connectivity may come at a price, since not all tasks, statistical contrasts, and treatments of motion artifacts were completely similar. On the other hand, the experimental conditions were similar, in that they all required increasing abstraction away from the here and now, in comparison with the control conditions, which typically elicits stronger cerebellar activation (Van Overwalle et al., 2014). To recall, our studies involved trait-implicating as opposed to non-trait-implicating events (Study 2–3); traits from groups rather than individuals (Study 4); events inconsistent with an implied trait, as opposed to consistent events (Study 1); and counterfactual events that implied a different past, inconsistent with the true past (Study 5). Finding common patterns of connectivity within this variety of tasks and within a model of this complexity should increase our confidence in the robustness and stability of the present results. Moreover, the covariate analysis demonstrated that after controlling for systematic differences across the studies, the connections became stronger, illustrating the robustness of the DCM approach and the present results. In fact, the identification of closed loops depended very much on our covariation analysis, which controlled for all differences between studies. Another way of testing robustness might be to see how the model behaves when some ROIs are eliminated (Ushakov et al., 2016). However, this procedure is not neutral, and it is less convincing because the results depend very much on whether the eliminated ROIs play a peripheral or central role in the network, in which case we might expect, respectively, little or drastic changes in the model estimates. Nevertheless, it is still possible that the present pattern of connectivity is unique to the present tasks and participants. Moreover, although our covariation analysis ruled out some potential sources of variation, still, many other things apart from the study design and task could have affected the present results. This issue can only be resolved by future research.

Conclusion

The present dynamic modeling analysis supports the recently acknowledged critical role of the cerebellum in social mentalizing. It demonstrates that there are domain-specific mentalizing closed loops between the cerebrum and the

cerebellum. Future research will be needed in order to explore questions concerning the basic social cognition processes that take place in the cerebellum. We suggest that the cerebellum is a forward controller that matches internal cerebellar models of event sequences against external sources and contexts in the cerebrum.

Author note This research was funded by a Strategic Research Program (SPR15) awarded by the Vrije Universiteit Brussel, Belgium. Our dear and warm colleague, P.M., died at the end of 2017.

References

- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional–anatomic fractionation of the brain’s default network. *Neuron*, 65, 550–562. <https://doi.org/10.1016/j.neuron.2010.02.005>
- Baetens, K., Ma, N., Steen, J., & Van Overwalle, F. (2014). Involvement of the mentalizing network in social and non-social high construal. *Social Cognitive and Affective Neuroscience*, 9, 817–824. <https://doi.org/10.1093/scan/nst048>
- Buckner, R. L. (2013). The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. *Neuron*, 80, 807–815. <https://doi.org/10.1016/j.neuron.2013.10.044>
- Buckner, R., Krienen, F., Castellanos, A., Diaz, J. C., & Yeo, B. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106, 2322–2345. <https://doi.org/10.1152/jn.00339.2011>
- Cui, S. Z., Li, E. Z., Zang, Y. F., Weng, X. C., Ivry, R., & Wang, J. J. (2000). Both sides of human cerebellum involved in preparation and execution of sequential movements. *NeuroReport*, 11, 3849–3853. <https://doi.org/10.1097/00001756-200011270-00049>
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6, 218–229. <https://doi.org/10.1006/nimg.1997.0291>
- Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modeling. *NeuroImage*, 19, 1273–1302. [https://doi.org/10.1016/S1053-8119\(03\)00202-7](https://doi.org/10.1016/S1053-8119(03)00202-7)
- Friston, K. J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., Van Wijk, B. C. M., . . . Zeidman, P. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. *NeuroImage*, 128, 413–431. <https://doi.org/10.1016/j.neuroimage.2015.11.015>
- Friston, K., Zeidman, P., & Litvak, V. (2015). Empirical Bayes for DCM: A group inversion scheme. *Frontiers in Systems Neuroscience*, 9, 164. <https://doi.org/10.3389/fnsys.2015.00164>
- Guell, X., Gabrieli, J. D. E., & Schmahmann, J. D. (2018). Embodied cognition and the cerebellum: Perspectives from the dysmetria of thought and the universal cerebellar transform theories. *Cortex*, 100, 140–148. <https://doi.org/10.1016/j.cortex.2017.07.005>
- Hillebrandt, H., Friston, K. J., & Blakemore, S.-J. (2015). Effective connectivity during animacy perception—Dynamic causal modelling of Human Connectome Project data. *Scientific Reports*, 4, 6240. <https://doi.org/10.1038/srep06240>
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9, 304–313. <https://doi.org/10.1038/nrn2332>
- Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of*

- Neuroscience, 23, 8432–8444. <https://doi.org/10.1523/JNEUROSCI.23-23-08432.2003>
- Keren-Happuch, E., Chen, S.-H. A., Ho, M.-H. R., & Desmond, J. E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Human Brain Mapping*, 35, 593–615. <https://doi.org/10.1002/hbm.22194>
- Krienen, F. M., & Buckner, R. L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral Cortex*, 19, 2485–2497. <https://doi.org/10.1093/cercor/bhp135>
- Ma, N., Vandekerckhove, M., Baetens, K., Van Overwalle, F., Seurinck, R., & Fias, W. (2012a). Inconsistencies in spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, 7, 937–950. <https://doi.org/10.1093/scan/nsr064>
- Ma, N., Vandekerckhove, M., Van Hoeck, N., & Van Overwalle, F. (2012b). Distinct recruitment of temporo-parietal junction and medial prefrontal cortex in behavior understanding and trait identification. *Social Neuroscience*, 7, 591–605. <https://doi.org/10.1080/17470919.2012.686925>
- Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, 65, 276–291. <https://doi.org/10.1016/j.neubiorev.2016.03.020>
- O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: Psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, 7, 604–609. <https://doi.org/10.1093/scan/nss055>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Pisotta, I., & Molinari, M. (2014). Cerebellar contribution to feedforward control of locomotion. *Frontiers in Human Neuroscience*, 8, 475:1–5. <https://doi.org/10.3389/fnhum.2014.00475>
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98, 676–682. <https://doi.org/10.1073/pnas.98.2.676>
- Salmi, J., Pallesen, K. J., Neuvonen, T., Brattico, E., Korvenoja, A., Salonen, O., & Carlson, S. (2010). Cognitive and motor loops of the human cerebro-cerebellar system. *Journal of Cognitive Neuroscience*, 22, 2663–2676. <https://doi.org/10.1162/jocn.2009.21382>
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, 42, 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>
- Sokolov, A. A., Erb, M., Grodd, W., & Pavlova, M. A. (2014). Structural loop between the cerebellum and the superior temporal sulcus: Evidence from diffusion tensor imaging. *Cerebral Cortex*, 24, 626–632. <https://doi.org/10.1093/cercor/bhs346>
- Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage*, 49, 3099–3109. <https://doi.org/10.1016/j.neuroimage.2009.11.015>
- Stoodley, C. J., MacMore, J. P., Makris, N., Sherman, J. C., & Schmahmann, J. D. (2016). Location of lesion determines motor vs. cognitive consequences in patients with cerebellar stroke. *NeuroImage: Clinical*, 12, 765–775. <https://doi.org/10.1016/j.nicl.2016.10.013>
- Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *NeuroImage*, 44, 489–501. <https://doi.org/10.1016/j.neuroimage.2008.08.039>
- Stoodley, C. J., & Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex*, 46, 831–844. <https://doi.org/10.1016/j.cortex.2009.11.008>
- Stoodley, C. J., Valera, E. M., & Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: An fMRI study. *NeuroImage*, 59, 1560–1570. <https://doi.org/10.1016/j.neuroimage.2011.08.065>
- Suzuki, L., Coulon, P., Sabel-Goedknecht, E. H., & Ruigrok, T. J. H. (2012). Organization of cerebral projections to identified cerebellar zones in the posterior cerebellum of the rat. *Journal of Neuroscience*, 32, 10854–10869. <https://doi.org/10.1523/JNEUROSCI.0857-12.2012>
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117, 440–463. <https://doi.org/10.1037/a0018963>
- Ushakov, V., Sharaev, M. G., Kartashov, S. I., Zavyalova, V. V., Verkhlyutov, V. M., & Velichkovsky, B. M. (2016). Dynamic causal modeling of hippocampal links within the human default mode network: Lateralization and computational stability of effective connections. *Frontiers in Human Neuroscience*, 10, 528:1–14. <https://doi.org/10.3389/fnhum.2016.00528>
- Van der Cruyssen, L., Heleven, E., Ma, N., Vandekerckhove, M., & Van Overwalle, F. (2015). Distinct neural correlates of social categories and personality traits. *NeuroImage*, 104, 336–346. <https://doi.org/10.1016/j.neuroimage.2014.09.022>
- Van der Cruyssen, L., Van Duynslaeger, M., Cortoos, A., & Van Overwalle, F. (2009). ERP time course and brain areas of spontaneous and intentional goal inferences. *Social Neuroscience*, 4, 165–184. <https://doi.org/10.1080/17470910802253836>
- Van Duynslaeger, M., Sterken, C., Van Overwalle, F., & Verstraeten, E. (2008). EEG components of spontaneous trait inferences. *Social Neuroscience*, 3, 164–177. <https://doi.org/10.1080/17470910801907226>
- 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, 174–188. <https://doi.org/10.1093/scan/nsm016>
- Van Hoeck, N., Ma, N., Ampe, L., Baetens, K., Vandekerckhove, M., & Van Overwalle, F. (2013). Counterfactual thinking: An fMRI study on changing the past for a better future. *Social Cognitive and Affective Neuroscience*, 8, 556–564. <https://doi.org/10.1093/scan/nss031>
- Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, 30, 829–858. <https://doi.org/10.1002/hbm.20547>
- Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, 48, 564–584. <https://doi.org/10.1016/j.neuroimage.2009.06.009>
- Van Overwalle, F., Baetens, K., Mariën, P., & Vandekerckhove, M. (2014). Social cognition and the cerebellum: A meta-analysis of over 350 fMRI studies. *NeuroImage*, 86, 554–572. <https://doi.org/10.1016/j.neuroimage.2013.09.033>
- Van Overwalle, F., Baetens, K., Mariën, P., & Vandekerckhove, M. (2015). Cerebellar areas dedicated to social cognition? A comparison of meta-analytic and connectivity results. *Social Neuroscience*, 10, 337–344. <https://doi.org/10.1080/17470919.2015.1005666>
- Van Overwalle, F., D'aeys, T., & Mariën, P. (2015). Social cognition and the cerebellum: A meta-analytic connectivity analysis. *Human Brain Mapping*, 36, 5137–5154. <https://doi.org/10.1002/hbm.23002>
- Van Overwalle, F., & Mariën, P. (2016). Functional connectivity between the cerebrum and cerebellum in social cognition: A multi-study analysis. *NeuroImage*, 124, 248–255. <https://doi.org/10.1016/j.neuroimage.2015.09.001>
- Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner, R. L. (2011). The organization of the

- human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106, 1125–1165. <https://doi.org/10.1152/jn.00338.2011>
- Zhou, Y., Liang, M., Tian, L., Wang, K., Hao, Y., Liu, H., . . . Jiang, T. (2007). Functional disintegration in paranoid schizophrenia using resting-state fMRI. *Schizophrenia Research*, 97, 194–205. <https://doi.org/10.1016/j.schres.2007.05.029>
- Zhou, Y., Zeidman, P., Wu, S., Razi, A., Chen, C., Yang, L., . . . Friston, K. J. (2018). Altered intrinsic and extrinsic connectivity in schizophrenia. *NeuroImage: Clinical*, 17, 704–716. <https://doi.org/10.1016/j.nicl.2017.12.006>