RESEARCH ARTICLE

The posterior cerebellum supports implicit learning of social belief sequences

Qianying Ma[1] · Min Pu[1] · Elien Heleven[1] · Naem P. Haihambo[1] · Kris Baetens[1] · Chris Baeken[2,3,4] · Natacha Deroost[1] · Frank Van Overwalle[1]

Abstract

Recent studies have documented the involvement of the posterior cerebellar Crus (I & II) in social mentalizing, when sequences play a critical role. We investigated for the first time implicit learning of belief sequences. We created a novel task in which true and false beliefs of other persons were alternated in an adapted serial reaction time (SRT) paradigm (Belief SRT task). Participants observed two protagonists whose beliefs concerning reality were manipulated, depending on their orientation toward the scene (true belief: directly observing the situation) or away from it (false belief: knowing only the prior situation). Unbeknownst to the participants, a fixed sequence related to the two protagonists' belief orientations was repeated throughout the task (Training phase); and to test the acquisition of this fixed sequence, it was occasionally interrupted by random sequences (Test phase). As a nonsocial control, the two protagonists and their orientations were replaced by two different shapes of different colors respectively (Control SRT task). As predicted, the posterior cerebellar Crus I & II were activated during the Belief SRT task and not in the Control SRT task. The Belief SRT task revealed that Crus I was activated during the initial learning of the fixed sequence (Training phase) and when this learned sequence was interrupted by random sequences (Test phase). Moreover, Crus II was activated during occasional reappearance of the learned sequence in the context of sequence violations (Test phase). Our results demonstrate the contribution of the posterior cerebellar Crus during implicit learning and predicting new belief sequences.

Highlights
- The posterior cerebellum Crus (I & II) is engaged in implicit belief sequence learning.
- The temporo-parietal junction also is activated to support implicit belief sequence learning.
- The posterior cerebellum plays a sequence detection role in social domains.

✉ Qianying Ma
  qianying.ma@vub.ac.be

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

[1] Department of Psychology and Center for Neuroscience, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

[2] Faculty of Medicine and Health Sciences, Department of Head and Skin, Ghent Experimental Psychiatry (GHEP) lab, Ghent University, Ghent, Belgium

[3] Department of Psychiatry, University Hospital (UZBrussel), Brussels, Belgium

[4] Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

Keywords Serial reaction time task · False belief · Social sequence learning · Cerebellum · Social cognition

Introduction

Sequence learning is involved in a wide range of human motor, cognitive, and social skills and facilitates efficient and adaptive human behaviors. For example, joint interactions, such as playing football, require continuously monitoring what other people see, know, and believe (Kampis, Fogd, & Kovács, 2017). Moreover, we seem to learn such sequences in an automatic and implicit manner, with little or no awareness.

It has traditionally been believed that the cerebellum plays a major role in training and automatizing sequences of movements (Ito, 2008). A major function of the cerebellum is "sequence detection" (Leggio & Molinari, 2015). The cerebellum builds internal models of repetitive patterns of temporally structured movements and events, or sequences, which include implicit predictions of future consequences based on these temporal representations (Leggio & Molinari, 2015; Pickering & Clark, 2014; Sokolov, Miall, & Ivry, 2017). When movements are novel or predictions based on prior

experiences do not match, these cerebellar internal models send out prediction errors to the cerebral cortex, allowing to minimize prediction errors in the future, thus leading to smooth and automatized movements.

Importantly, accumulating evidence in the past decades has pointed to the important role of the cerebellum in detecting not only motor domain but also nonmotor cognitive (Koziol et al., 2014) and social domain (Guell, Gabrieli, & Schmahmann, 2018; Van Overwalle et al., 2020a). Of interest for the social domain, research has supported the role of the cerebellum in *mentalizing*, that is, understanding others' mental states, which also is referred to as *Theory of Mind* (Premack & Woodruff, 1978; for reviews see Molenberghs, Johnson, Henry, & Mattingley, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle, 2009). This mentalizing capacity allows people to infer others' thoughts, beliefs, and intentions, as if people can read others' minds (Frith & Frith, 2005). During mentalizing, the cerebellum, and especially the posterior cerebellar Crus I and II, are recruited as shown by extensive meta-analyses (Guell et al., 2018; Van Overwalle, Baetens, Mariën, & Vandekerckhove, 2015; Van Overwalle et al., 2020a). These posterior cerebellar areas are located within the default/mentalizing network of the cerebellum (Buckner, Krienen, Castellanos, Diaz, & Yeo, 2011; Van Overwalle et al., 2015), which is part of the larger mentalizing network encompassing the whole brain (Molenberghs, Johnson, Henry, & Mattingley, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle, 2009).

A key ability in social mentalizing is understanding others' *false* beliefs, because this requires distinguishing others' beliefs from one's own view of reality (Kampis et al., 2017; Wimmer & Perner, 1983). One example of a false-belief task, similar to the one used in the present experiment, was developed by Saxe, Schulz, & Jiang (2006). In this false-belief task, a chocolate is moved out of a box and then returned to the same box or moved to another box. Critically, when this happens, a girl is oriented toward or away from the boxes, and the participants were asked to identify "where the girl thinks the chocolate bar is." When the girl was oriented toward the boxes, the participants should realize that the girl could see the final location of the chocolate and therefore held a true belief of reality. Conversely, when the girl was oriented away from the boxes, the participants should realize that the girl could not see any movements, and therefore held a false belief of the chocolate's location, based on her latest true belief of reality (i.e., when she was oriented to the boxes).

Several meta-analyses have shown that the temporoparietal junction (TPJ) is centrally involved in true- and false-belief tasks (Molenberghs, Johnson, Henry, & Mattingley, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle, 2009), suggesting that the TPJ might support the reorientation from one's own perspective to the perspective of others (Cabeza, Ciaramelli, & Moscovitch, 2012; Krall et al., 2015a, b; Özdem, Brass, Van der Cruyssen, & Van Overwalle, 2017).

The critical role of the cerebellum in belief inferences has been convincingly demonstrated in recent research. Patient studies reported that cerebellar patients were less accurate in identifying false beliefs in short narratives than healthy controls (Clausi et al., 2019). Cerebellar patients also were impaired in generating the correct chronological order of false beliefs depicted in randomly ordered cartoons, but they performed close to healthy participants for cartoons depicting non-social mechanical events (e.g., a car accident resulting in material damage; Van Overwalle, De Coninck, et al., 2019a). When healthy participants were asked to generate the chronological order of short narratives or cartoons involving true and false belief, the posterior cerebellar Crus was more strongly activated compared to when the sequences involved nonsocial mechanical events (Heleven, van Dun, & Van Overwalle, 2019). Additionally, it has been demonstrated that the posterior cerebellar Crus is activated in synchrony with the TPJ during belief understanding, via bidirectional functional connections between the cerebellar Crus and the TPJ (Van Overwalle et al., 2019c).

Present study

Previous research most often explicitly asked participants to generate or memorize social sequences, for example, the correct order of social actions implying other persons' beliefs (Heleven, van Dun, & Van Overwalle, 2019) or traits (Pu et al., 2020). However, given that the cerebellum is involved in implicit automatizing processes, we sought to identify whether the posterior cerebellar Crus contributes to implicit as opposed to explicit learning of sequences with social elements, such as inferences of others' beliefs. We used a novel implicit belief learning task, which combines elements of a classic serial reaction time (SRT) task for implicit learning (Nissen & Bullemer, 1987) with social elements from a false belief task (Wimmer & Perner, 1983). Accordingly, we named it a *Belief SRT* task (Ma et al., 2021).

In a classic SRT task, participants have to react to a target's physical characteristic (e.g., location, color, etc.). Unbeknownst to the participants, the appearance of the target always follows a specific sequence. Although participants are not aware that they learning anything, participants typically respond faster when the sequence is repeated. In contrast, they respond slower when the learned sequence is interrupted by a random sequence, followed again by faster responses when the learned sequence is reintroduced (Janacsek et al., 2020; Nissen & Bullemer, 1987).

In the present Belief SRT task, participants were requested to identify the true and false beliefs of two protagonists (see

also Özdem et al., 2019), using a task similar to Saxe et al. (2006) described earlier. Participants saw two protagonists (i.e., Papa Smurf and Smurfette) receiving flowers from one of four little Smurfs positioned at four fixed locations on top of the screen (Fig. 1A). Participants were asked to report how many flowers were given to the protagonist from the protagonist's perspective. When the protagonists were oriented toward the screen, they could see the flowers and thus held true beliefs about the number of flowers. Conversely, when the protagonists were oriented away from the screen (i.e., towards the participants), they could not see any changes and hence potentially held false beliefs about the number of flowers offered. In that case, the correct answer was the same number of flowers the last time when the same protagonist could see the flowers (and held a true belief). Crucially, as in a classical SRT task, unbeknownst to participants, there was a standard sequence of true-false belief orientations linked to the protagonists, which was repeated over the course of the experiment.

Some aspects in the present Belief SRT task are important to clarify. First, the alternations of two protagonists was important, because it forced participants to engage in genuine false belief reasoning and memorizing about each protagonist (i.e., retrieving the last true trial of the same protagonist during false belief trials), rather than simply repeating any answer from the last true trial, which would be the case if there was only one protagonist (Bradford, Jentzsch, & Gomez, 2015, Kaminski, Call, & Tomasello, 2008). Second, participants had to indicate how many flowers the protagonists believed to have received, which required a (motor) response that is independent from the perceived sequence of true-false belief orientations because the number of flowers was random. This effectively rules out potential motor learning confounds, which is especially critical in revealing brain activation related to implicitly learning of perceived belief sequences rather than repeated movement execution (Hardwick, Rottschy, Miall, & Eickhoff, 2013). Third, the purpose of the task was not to investigate implicit belief understanding (Poulin-Dubois et al., 2018; Schneider, Slaughter, & Dux, 2017) but to investigate the implicit learning of a repeated sequence of true-false belief orientations. In fact, participants were explicitly told to infer protagonists' beliefs.

To investigate the preferential role of the posterior cerebellar Crus in social mentalizing, the Belief SRT task was compared against a nonsocial Control SRT task with an identical sequence structure (Fig. 1B). In this Control SRT task, the protagonists were replaced by colored shapes (squares and circles), and the true-false belief orientations of each protagonists were replaced by two color variations of each shape. Thus, the four distinct pictures used in the Belief SRT task were replaced by four distinct pictures of colored shapes in the Control SRT task. A blue square or a green circle indicated that the number of flowers had to be taken from the current trial, an orange square or black circle indicated that the

number of flowers had to be recalled from the previous trial from the same shape. By keeping the structure of sequences identical, observed differences in brain activation can be attributed to process differences related to social mentalizing in the Belief SRT task.

In sum, in the present study, we used a novel Belief SRT task to identify the role of the posterior cerebellar Crus in implicit learning of true-false belief sequences. Based on our review of earlier research on the domain-specific mentalizing role of the posterior cerebellar Crus in synchrony with the cortical TPJ, we hypothesized a critical role of these areas in the Belief SRT task but less so in the Control SRT task. Note that the hypothesis does not imply that social processes are completely dissociated from cognitive processes but rather that the Belief SRT task involves more social processing and, hence, activates more key social brain areas (i.e., Crus and TPJ) than the Control SRT task.

Method

Participants

A total of 40 healthy, right-handed, native Dutch-speaking participants were recruited. To avoid contamination between the Belief and Control SRT tasks, we used a between-participants design, where participants were randomly distributed between the two tasks. Contamination in a within-participant design was suspected, as people are quick to anthropomorphize shapes (e.g., triangles) as engaging in human-like behavior (Heider & Simmel, 1944), with recruitment of social mentalizing areas as a result (Moessnang et al.; 2016; see meta-analysis by Van Overwalle, 2009; Schurz et al., 2014). Critically, another problem with a within-participant design would be transfer of implicit sequence knowledge between the two tasks sharing a similar sequential structure (Geiger, Cleeremans, Bente, & Vogeley, 2018).

All participants had normal or corrected-to-normal vision and color perception. Two participants were excluded because of excessive head movement (outlier scans >5%). Hence, data analysis was based on 18 participants who completed the Belief SRT task (14 females; age 18-28 years, mean age 21.2 ± 2.7) and 20 participants who completed the Control SRT task (15 females; age 18-35 years, mean age 22.7 ± 3.7). All participants gave written, informed consent with the approval of the Medical Ethics Committee at the University Hospital of Ghent. Participants were paid 20 euros and transportation costs in exchange for their participation.

Stimuli material

In the Belief SRT task (Fig. 1A; Ma et al., 2021), four little Smurfs appeared on the top of the screen, which marked the
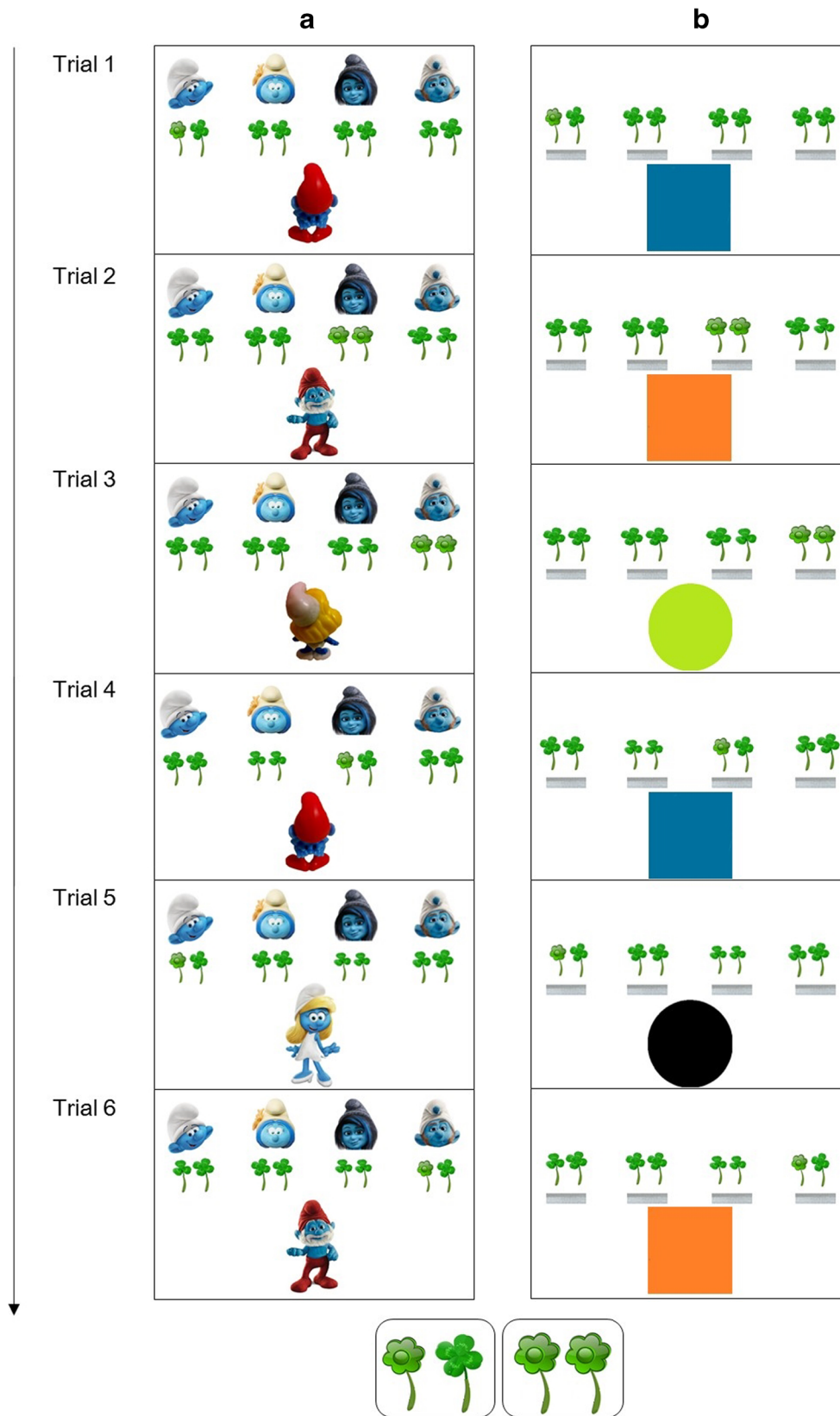
location of the target flowers. The two protagonists, Papa Smurf and Smurfette, were each shown individually at the bottom of the screen with their face directed to or away from the screen. Participants were told that they would "see flowers,

◀ **Fig. 1** Schematic example showing the first 6 trials of the Standard sequence in the Belief SRT task [**A**] and the Control SRT task [**B**]. On each trial, participants had to report the number of flowers as seen by the protagonists (Papa Smurf or Smurfette in the Belief SRT task) or depending on the color variation of the shapes (square or circle in the Control SRT task). Without being informed, belief orientations/color variations followed a Standard sequence. In the Belief SRT task, when the protagonist was oriented to the screen and could see the flowers (true trial), the number of target flowers had to be reported from the current trial; when the protagonist was oriented away from the screen and could not see the flowers (false trial), the number of target flowers had to be reported from the previous true trial from the same protagonist. Similarly, in the Control SRT task, a blue square or a green circle indicated that the number of flowers had to be taken from the current trial (= true trial), whereas an orange square or black circle (= false trial) indicated that the number of flowers had to be taken from the previous true trial from the same shape. The number of flowers was random (1 or 2), making the response unpredictable, and dissociating sequence learning from motor responses. Each trial was self-paced, with all stimuli remaining on screen for 3,000 ms until a response was given, and was followed by a response-stimulus interval of 400 ms before the next trial started. [**Bottom Inset**] The inset shows an enlargement of the target stimulus, consisting of a pair of one or two flowers surrounded by clovers (as a distraction) of approximately the same shape and color. [**Trial 1 - 2**] To illustrate the instructions for the Belief SRT task, in Trial 1, there is one flower that Papa Smurf can see, because he is oriented toward the screen, meaning that the correct response is 1. In Trial 2, there are two flowers. Because Papa Smurf is oriented away from the screen, he cannot see the number of flowers on this trial, hence he still thinks to have received one flower which he last saw on the previous (1st) trial. The correct response is thus again 1. In the Control SRT task, Trial 1, there is one flower. Because the color of the square is blue, the correct response is the observed number of flowers, or 1. In Trial 2, the square is orange, so participants must report the number of flowers from the blue square on the previous (1st) trial. The correct response is 1, again

given by one of four Smurfs at the top of the screen, to Papa Smurf or Smurfette (at the bottom of the screen), and the clovers are of no importance. One of the four little Smurfs will give the flowers while Papa Smurf or Smurfette is watching (facing the screen) or not watching (facing you)." Participants were instructed to indicate as fast and accurately as possible "how many flowers are given (1 or 2) as seen by Papa Smurf or Smurfette." It was further detailed that "throughout the task you have to follow how many flowers Papa Smurf or Smurfette receive. If they are turned with their back to the four Smurfs, you have to indicate how many flowers they (remember that they) received the last time" (translated from Dutch).

In the Control SRT task (Fig. 1B), the following changes were made. On the top of the screen, four sidewalk boards (instead of four little Smurfs) appeared to mark the location of the target flowers. A colored square or circle appeared at the bottom of the screen (instead of Papa Smurf or Smurfette). Participants were told that they "must indicate how many flowers there are (1 or 2) when a blue square or a green circle appears. When the square is orange, repeat the previous number of flowers from the blue square. When the circle is black, repeat the previous number of flowers of the green circle" (translated from Dutch).

The repeated Standard sequence embedded in the two Tasks was identical, and the instructions were structurally similar. While the Standard sequence was fixed, the number of flowers was randomly determined at every trial in order to dissociate sequence learning from motor responses.

Experimental procedure

In both Tasks, the experimental procedure was largely identical. Responses were made with the middle or index finger (i.e., 1 or 2 flowers respectively) of participants' left hand and were collected via a magnet compatible two-button response box. Responses were self-paced and all stimuli remained on screen for 3,000 ms until a response was given. In case of a wrong response, or when no response was given before 3,000 ms, the word "Error" appeared for 750 ms on the screen, and the next trial began. The response-stimulus interval was set at 400 ms (Coomans, Deroost, Zeischka, & Soetens, 2011). After each block, participants received feedback about their average reaction time and error rate, and they were encouraged to make less than 5% errors. Participants got a break of 15 s after every two blocks.

Like a classical SRT task, the current task involved two phases (Fig. 2A). In an initial Training phase, the Standard sequence, consisting of a fixed order of Protagonists (Smurfs/shapes) and Orientations (beliefs/colors; Fig. 2B), was repeatedly presented, and so allowed implicit knowledge to develop (Baetens, Firouzi, Van Overwalle, & Deroost, 2020). This was followed by a Test phase, during which the Standard sequence was occasionally interrupted by blocks of Random sequences (Hardwick et al., 2013).

After a practice phase of 2 blocks of 24 trials (with a different sequence from the main experiment), participants completed the Training phase, consisting of 5 Standard blocks of 32 trials (i.e., twice the 16-trial Standard sequence; Blocks 1-5; Fig. 2A). The Test phase consisted of 25 blocks of 32 trials (Blocks 6-30). Unbeknownst to the participants, Standard blocks, identical as at the Training phase, were interleaved with two types of Random blocks in the Test phase. First, in a Random Orientation block, Orientations (beliefs/colors) were changed into a pseudo-random order while Protagonists (Smurfs/shapes) remained identical as in the Standard blocks. Second, in a Total Random block, Protagonists (Smurfs/shapes) and Orientations (beliefs/colors) were totally randomized with the limitation of at most 2 subsequent trials of the same True or False type, consistent with the Standard blocks. Each Standard block was followed by a Total Random and a Random Orientation block, the latter two blocks presented in two orders, counterbalanced between participants (Fig. 2; Supplementary Table S1 for (pseudo) random sequences). The last block at the end of the whole task was a Standard block. Overall, at the Test phase, there were 9 Standard blocks (288 trials), 8 Random Orientation blocks,

a

Training phase					Test Phase																								
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
					Order 1																								
S	S	S	S	S	S	TR	RO	S	RO	TR	S	TR	RO	S	TR	RO	S	TR	RO	S	RO	TR	S	TR	RO	S	RO	TR	S
					Order 2																								
					S	RO	TR	S	TR	RO	S	RO	TR	S	RO	TR	S	RO	TR	S	TR	RO	S	RO	TR	S	TR	RO	S

b

Protagonist	M	M	Fe	M	Fe	M	Fe	Fe	Fe	Fe	Fe	M	Fe	M	Fe	M
Belief Orientation	T	Fa	T	T	Fa	Fa	T	Fa	Fa	T	Fa	T	T	Fa	Fa	T
Prior true trial		-1			-2	-2		-1	-2		-1			-2	-2	

Fig. 2 Experimental procedure for both Tasks. **A**. Experimental design for the Belief and Control SRT task, with Blocks numbered 1 to 30 on the first data row. In each block, there were 32 trials to which the participants had to respond: Standard (S) blocks with two repetitions of an embedded 16-trial Standard sequence; Random Orientation (RO) blocks with a pseudo-random Orientation sequence; or Total Random (TR) blocks with random sequences of Protagonists and Orientations. The RO and TR blocks were presented in two orders in which the order of the RO and TR blocks were switched (i.e., Order 1 & 2), counterbalanced between participants. **B**. Standard sequence in the Belief SRT task. M = male (Papa Smurf), Fe = female (Smurfette), T = true, Fa = false. In the Control SRT task, the sequence is identical and stimuli were replaced by shapes and colors, as depicted in Fig. 1 (i.e., Male True = Blue Square; Male False = Orange Square, Female True = Green Circle, Female False = Black Circle). See Supplementary Table S1 for the (pseudo) random sequences for Random Blocks

and 8 Total Random blocks (both Random blocks have 256 trials). For each block during the entire experiment, a "Begin of block" and "End of block" message was presented for 4 s and 2 s, respectively.

After scanning, awareness of the sequence was assessed with a funneled questionnaire (Deroost & Coomans, 2018). Participants were asked "Did you notice anything special during the experiment?" Afterwards, they were told that a sequence was imposed on the Protagonists (Smurfs/shapes) and their Orientation (belief/color). They were asked to reproduce as accurately as possible the "order in which papa Smurf and Smurfette appeared (M = Papa Smurf, F = Smurfette)" and whether or not they "could or could not see who gave the flowers (N = could not see; Y = could see)" in the Belief SRT task; and the "order in which squares and circles appeared (BV = blue square, OV = orange square, GC = green circle, ZC = black circle) in the Control SRT task.

Imaging procedure and preprocessing

Images were collected with a Siemens Magnetom Prisma fit 3T scanner system (Siemens Medical Systems, Erlangen, Germany) using a 64-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) running under Windows XP. Participants were placed head first and supine in the scanner bore and were instructed not to move their heads to avoid motion artifacts. Foam cushions were placed within the head coil to minimize head movements. First, high-resolution anatomical images were acquired using a T1-weighted 3D MPRAGE sequence [TR = 2,250 ms, TE = 4.18 ms, TI = 900 ms, FOV = 256 mm, flip angle = 9°, voxel size = 1 × 1 × 1 mm]. Second, a fieldmap was calculated to correct for inhomogeneities in the magnetic field (Cusack & Papadakis, 2002). Third, whole-brain functional images were collected in a single run using a T2*-weighted gradient echo sequence, sensitive to BOLD contrast (TR = 1,000 ms, TE = 31.0 ms, FOV = 210 mm, flip angle = 52°, slice thickness = 2.5 mm, distance factor = 0%, voxel size = 2.5 × 2.5 × 2.5 mm, 56 axial slices, acceleration factor GRAPPA = 4).

SPM12 (Wellcome Department of Cognitive Neurology, London, UK) was used to process and analyze the fMRI data. To remove sources of noise and artifact, data was preprocessed. Inhomogeneities in the magnetic field were corrected using the fieldmap (Cusack & Papadakis, 2002). Functional data were corrected for differences in acquisition time between slices for each whole-brain volume, realigned to correct for head movement, and co-registered with each participant's anatomical data. Then, the functional data was transformed into a standard anatomical space (2 mm isotropic voxels) based on the ICBM152 brain template (Montreal Neurological Institute). Normalized data were then spatially smoothed (6-mm full-width at half-maximum, FWHM) using a Gaussian Kernel. Finally, using the Artifact Detection Tool (ART; http://web.mit.edu/swg/art/art.pdf; http://www.nitrc.org/projects/artifact_detect), the preprocessed data

were examined for excessive motion artifacts and for correlations between motion and experimental design, and between global mean signal and experimental design. Outliers were identified in the temporal differences series by assessing between-scan differences (Z-threshold: 3.0 mm, scan to scan movement threshold: 0.5 mm; rotation threshold: 0.02 radians). These outliers were omitted from the analysis by including a single regressor for each outlier. A default high-pass filter was used of 128 s, and serial correlations were accounted for by the default auto-regressive AR (1) model.

Statistical analysis

Statistical analysis of neuroimaging data

The statistical analyses were performed using the general linear model of SPM12 (Wellcome Department of Cognitive Neurology, London, UK). At the first (single participant) level, an event-related design for measuring transient activity across trials was modeled by entering separate regressors for the trials of interest: two regressors for the trials in the Standard blocks at the Training and Test phase, two regressors for the trials in the Total Random blocks and the trials in the Random Orientation blocks at the Test phase, and two additional regressors of no interest for pauses and error trials. This last regressor involved incorrect trials as well as one trial after each incorrect trial, because these latter trials may be affected by error processing on the prior trial.

Contrasts within each Task

As mentioned earlier, a major function of the cerebellum is learning repetitive sequences (Leggio & Molinari, 2015). Prior research on classic SRT tasks revealed different cerebellar areas involved in early and late phases of motor learning (see meta-analysis by Bernard & Seidler, 2013). In line with this, we also distinguish between early and late phases of learning to identify the areas that are preferentially activated in each phase. We also performed a classic test on learning the standard sequence (i.e., sequence-specific learning) by contrasting random versus standard sequences. Consequently, we defined the following contrasts:

1) *General learning:* Cerebellar engagement during the early phase of sequence learning is tested by the contrast: Standard block at Training > Standard block at Test.
2) *Maintenance of learning:* Cerebellar engagement during the late phase of sequence learning in a context of sequence violations (in the Test phase) is tested by the contrast: Standard block at Test > Standard block at Training. Note that this contrast does not test mere late phase of learning, as it also involves reinstating the learned standard sequence after pseudo-random sequences.

3) *Detecting violations:* Cerebellar engagement during violations of the learned Standard sequence (by pseudo-random sequences; in the Test Phase), is tested by two contrasts: Total Random block at Test > Standard block at Test; Random Orientation block at Test > Standard block at Test.

Because we make predictions for these three learning effects in the posterior cerebellar Crus and cortical TPJ for the Belief SRT task, but less so for the Control SRT task, we ran these three contrasts for both tasks.

We conducted a within-participant one-way analysis of variance (ANOVA) and defined t-contrasts between four regressors of interests (i.e., Standard block at Training, Standard block at Test, Total Random block at Test, and Random Orientation block at Test), using a cluster-forming threshold of $p < 0.001$ with minimum cluster extent of 10 voxels, and a cluster-wise significance level of $p < 0.05$, family-wise error (FWE) corrected for multiple comparisons correction.

Contrasts between Tasks

To further test differences between the Belief and Control SRT tasks, we applied the Sandwich Estimator toolbox (SwE; Guillaume et al., 2014a; http://www.nisox.org/Software/SwE/). SwE uses a marginal model to analyze repeated measurements between tasks, taking into account correlations because of repeated measurements, unexplained variations across participants, unbalanced study designs of the variable number of scans, and corrected degrees of freedom.

In this study, we modeled eight covariates in SwE with Task (Belief versus Control) as a between-participant factor orthogonal to the same four regressors of interests as before as a within-participant factor (i.e., Standard block at Training, Standard block at Test, Total Random block at Test, and Random Orientation block at Test). We used the following default SwE options (see http://www.nisox.org/Software/SwE/man): a modified SwE, which assumes that participants in each Belief and Control SRT task share a common covariance matrix, repeated measurements in each within-factor regressor, small-sample adjustment type C2, and degrees of freedom approximation III.

Using the SwE analysis, we first ran simple contrasts between the Belief and Control SRT tasks for each type of block (Standard and Random) and at all Phases (Training and Test). Importantly, to test our hypothesis that implicit learning effects in the cerebellar Crus and TPJ are asymmetric (i.e., present in the Belief SRT task, but absent in the Control SRT task), a series asymmetric interaction effects was defined for each of the learning effects, also known as *spreading interactions*. For example, the spreading contrast for the *general learning* was: Belief Standard block at Training > [Belief Standard block at Test = Control Standard block at Training = Control Standard

block at Test], or expressed in weights: 3 -1 -1 -1 (Table 4). As a comparison, we also ran reverse *spreading interactions* with the predicted asymmetric effect in the Control SRT task (Table 4).

The contrasts between tasks were analyzed using a cluster-forming threshold of $p < 0.005$ with minimum cluster extent of 50 voxels, followed by a voxel level significance of $p < 0.05$, using false-discovery rate (FDR) correction for multiple comparisons (Fleming et al., 2019; Guillaume et al., 2014b).

ROI analyses

To test our specific hypotheses that mainly the posterior cerebellar Crus and TPJ are activated in the Belief SRT task, we defined a number of *a priori* Regions of Interest (ROI) on the basis of similar spherical ROIs in previous fMRI studies. Specifically, the center of these *a priori* ROIs were defined as follows:

1) Meta-analyses and connectivity studies showed significant activations of bilateral Crus II (± 24 -76 -40) and left Crus I (-40 -70 -40) during social reasoning, especially during generating the correct sequence of social events that require the understanding of a person's beliefs (Guell et al., 2018; Van Overwalle, Ma, et al., 2020; Van Overwalle et al., 2020c). Hence, we identified these three coordinates as *a priori* cerebellar ROIs, although it should be noted that the bilateral Crus II is clearly located within the mentalizing network demarcated by Buckner et al. (2011), while the left Crus I is located somewhat more peripherally in the mentalizing network and closer to the executive network (Buckner et al., 2011). Note that for exploratory reasons, we also investigated *a priori* ROI of the right Crus I centered on the same coordinates in the right hemisphere, but this yielded no significant clusters and is not reported.

2) Meta-analyses showed that the cortical bilateral TPJ ± 50 -55 25 was significant activated in understanding people's social beliefs, behavioral intentions, and personality traits (Van Overwalle, 2009; Van Overwalle & Baetens, 2009). Hence, we identified these two coordinates as *a priori* cortical TPJ ROIs.

These *a priori* ROIs were tested with a small volume correction (SVC) using spheres with radius = 10 mm centered around the nearest 2 mm of the coordinates listed above (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Debas et al., 2010), using the same thresholds as for the whole-brain analysis (except that the minimum cluster extent was always set to 10 voxels).

To visualize the differential activation in the posterior cerebellum (Crus I & II) and TPJ during the Training and Test phases, we extracted for each participant the percent signal change using spheres of 10 mm around the peak coordinates of the contrasts between the Belief and Control tasks using the MarsBar toolbox (http://marsbar.sourceforge.net) for all four regressors of interests. These percent signal change data were further analyzed using *t*-test, with threshold $p < 0.05$, two-tailed.

Statistical analysis for behavioral data

We first investigated explicit awareness of the Standard sequence to explore whether learning was implicit. We then tested whether participants learned this Standard sequence. Responses during and immediately after an error were excluded. Mean error rates and reaction times (RTs) were computed for every block.

To assess learning of the Standard sequence, we first used a mixed ANOVA on the RTs of the Standard blocks during Training (Blocks 1-5) as within-participant factors and Task (Belief versus Control) as a between-participant factor. To further test sequence learning, we then ran a mixed ANOVA with Blocks at Test (Blocks 6-29) and Block Type (Standard, Total Random and Random Orientation blocks) as within-participant factors and Task (Belief versus Control) as a between-participant factor. If participants learned the Standard sequence, their RTs should decrease during the Training phase, and more importantly, during the Test phase, their RTs should increase again during the Random blocks compared to the Standard blocks.

For the statistical results, the Greenhouse-Geisser correction is reported when the sphericity assumption was violated. Then, *t*-tests were applied when ANOVA indicated significant differences. The level of significance was set to 0.05, and two-tailed tests were applied.

Results

Explicit awareness

Awareness of the Standard sequence was assessed at the end of the experiment. Participants were first prompted for any awareness (i.e., did you "notice anything particular during the experiment"), and then had to reproduce the whole sequence for the Protagonists (Smurfs / Shapes) and their Orientation (Beliefs / Colors).

In the Belief SRT task, 12 of the 18 participants reported no sequence awareness. The mean length of the longest correct sequence recollection was 1.83/16 for all participants, and 2.83/16 for the aware participants and 1.33/16 for the unaware participants. A *t*-test on the mean length of correct sequences showed no robust difference between aware and unaware participants ($p = 0.09$).

In the Control SRT task, 16 of 20 participants reported no sequence awareness. When they were asked to reproduce the sequence, the mean length of the longest correct recollection was 2.45/16 for all participants, with 3.25 for the aware participants and 2.25 for the unaware participants. A t-test on the mean length of correct sequences showed no difference between aware and unaware participants ($p = 0.20$).

As a consequence of the limited awareness of the Standard sequence, no participant was excluded from further analysis.

Behavioral results

No participants were excluded due to high error rates. The mean error rate was 5.84% over all blocks for the Belief SRT task and 7.41% for the Control SRT task, and did not differ between Tasks at Training or Test (all $ps > 0.20$). The participants learned the Standard sequence, as demonstrated by faster RTs during the Training and Test phase, and slower RTs during the Random blocks compared to the Standard blocks at the Test phase. These results were supported by the following statistical analyses.

We first ran a mixed repeated-measures ANOVA with five Standard blocks at Training as within-participant factor and Task [Belief vs. Control] as between-participant factor. This revealed a main effect of Standard blocks at Training, suggesting faster RTs across five blocks at the Training phase ($F_{(3.1, 111.69)} = 40.54$, $p < 0.001$, $\eta^2 = 0.53$, Fig. 3A left) with no differences between Tasks ($F_{(1, 36)} = 0.98$, $p = 0.33$).

Then, we ran a mixed repeated-measures ANOVA at the Test phase with Blocks at Test and Block Type [Standard, Total Random and Random Orientation blocks] as within-participant factors and Task [Belief versus Control] as a between-participant factor. There was a main effect of

Blocks at Test, suggesting increasingly faster RTs across blocks ($F_{(4.21, 151.41)} = 57.12$, $p < 0.001$, $\eta^2 = 0.61$). A main effect of Block Type confirmed that participants reacted slower in the Total Random and Random Orientation blocks than the Standard blocks ($F_{(2, 72)} = 35.67$, $p < 0.001$, $\eta^2 = 0.50$, Bonferroni post-hoc test: mean difference (MD) $_{\text{Total Random − Standard block}} = 38$ ms, $p < 0.001$; MD $_{\text{Random Orientation − Standard block}} = 49$ ms, $p < 0.001$, Fig. 3A right). Using paired t-test, closer inspection of Block Type (Fig. 3B) revealed slower RTs for the Total Random blocks than the Standard blocks for both Tasks (Belief: $t_{(17)} = 6.09$, $p < 0.001$, Cohen's $d = 1.16$; Control: $t_{(19)} = 2.82$, $p = 0.01$, Cohen's $d = 0.54$), and slower RTs for the Random Orientation blocks than the Standard blocks for both Tasks (Belief: $t_{(17)} = 8.76$, $p < 0.001$, Cohen's $d = 2.11$; Control: $t_{(19)} = 6.12$, $p < 0.001$, Cohen's $d = 1.18$). There was no main effect of Task ($F_{(1, 36)} = 0.14$, $p = 0.71$), indicated no RTs difference between the Belief and Control SRT Tasks.

Neuroimaging results

Contrasts within each Task

We hypothesized significant activations of the posterior cerebellar Crus and TPJ related to implicit learning effects for the Belief SRT task and less so for the Control SRT task. To test our hypotheses, we ran a number of whole-brain contrasts as well as a priori ROI analyses for both tasks. We begin with our specific hypothesis by reporting the results of a priori ROI analysis and then turn to the whole-brain analysis for additional activations.

First, we verified general learning using a Standard block at the Training > Test contrast. For the Belief SRT task, the a
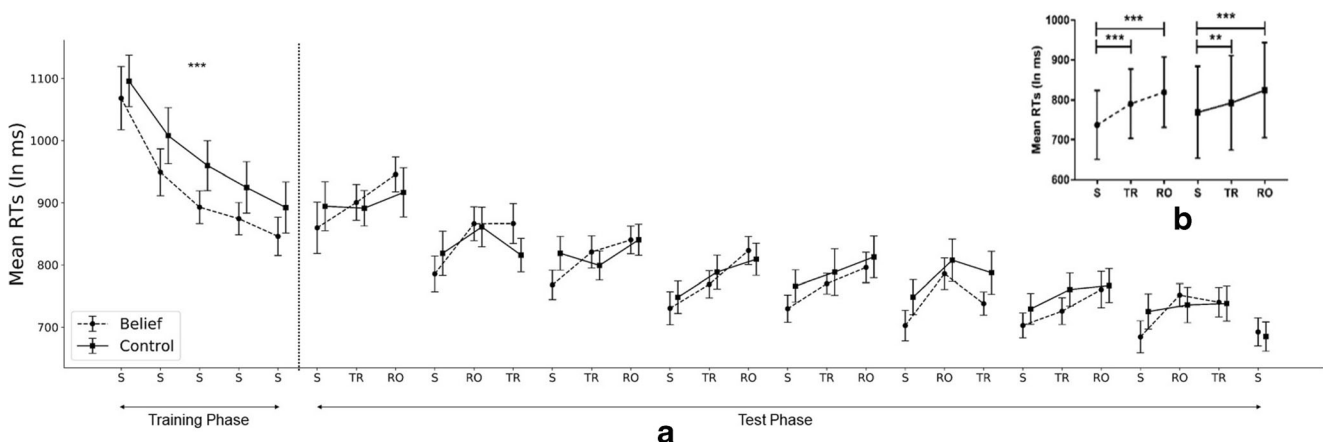
Fig. 3 Behavioral performance demonstrated by mean RTs in the Belief (dashed lines) and Control (full lines) SRT Tasks. S = standard block, TR = total random block, RO = random orientation block. **A.** Mean RTs for each block. Mixed ANOVAs revealed that participants learned the standard sequence demonstrated by faster RTs across standard blocks at the training phase and slower RTs for random blocks compared with the standard blocks at the test phase. **B.** Collapsed RTs for standard and random blocks at the test phase. Paired t-tests revealed slower RTs for the total random (TR) and random orientation (RO) blocks compared with the standard (S) blocks for both belief and control SRT tasks. Error bars are within-tasks standard error of the mean across participants. **$p < 0.01$; ***$p < 0.001$

priori ROI analyses showed activation in Crus I and the TPJ (Figs. 4A and 5A). The whole-brain analysis further revealed activations in the cerebellar lobules VIII, IV/V, and VII and, additionally, cortical activations in the postcentral gyrus, precentral gyrus, insula lobe, and anterior cingulate cortex (Table 1). For the Control SRT task, none of the *a priori* ROI analyses revealed any activation. The whole-brain analysis showed significant cerebellar activation in the lobules VI, VIII, and IX, extending to Lobule I-V and V (Table 1; Fig.
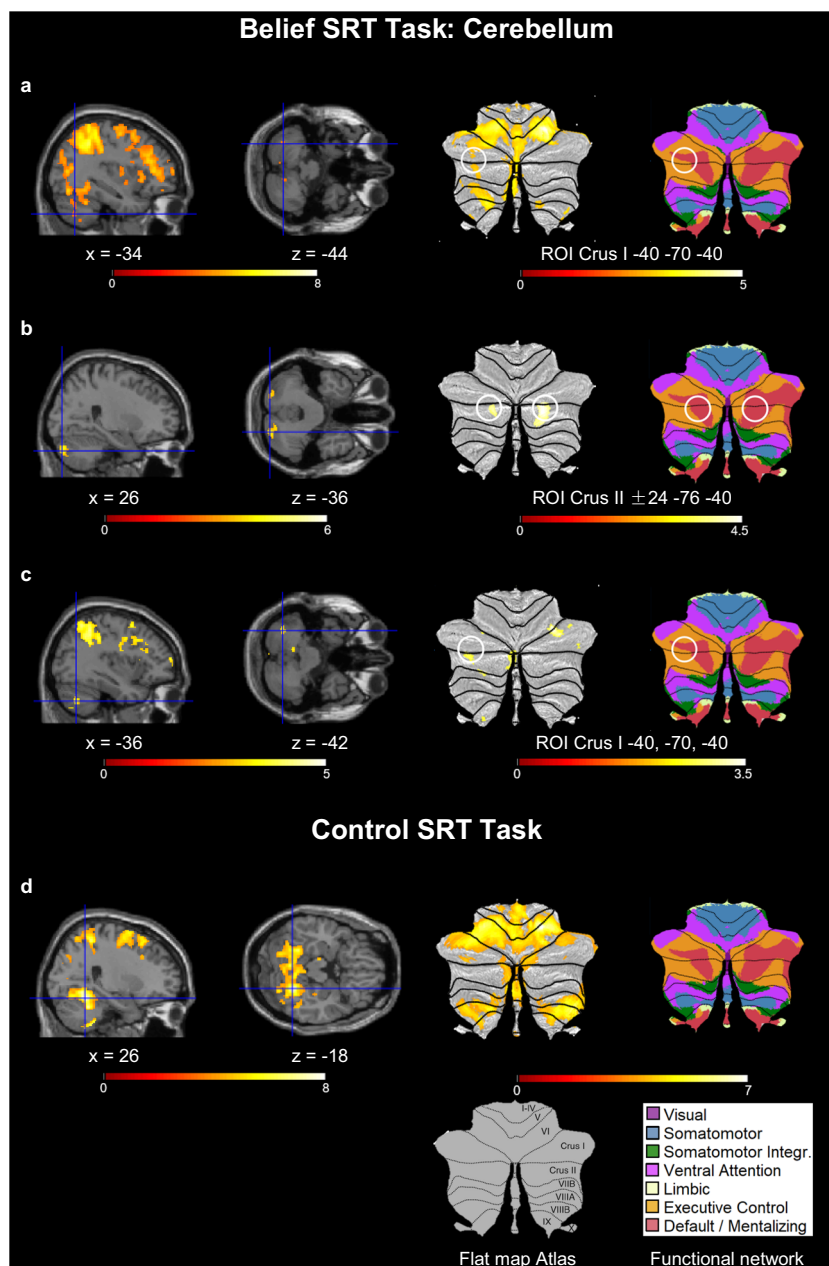
4D) and cortical activation in the temporal gyrus, middle, and superior frontal gyrus.

Second, we verified *maintenance of learning* using a Standard block during Test > Training contrast (Table 2). For the Belief SRT task, the *a priori* ROI and whole-brain analyses showed activation in Crus I & II (Fig. 4B) but not in the TPJ. The whole-brain analysis further showed activation in the temporal pole. For the Control SRT task, none of the *a priori* ROIs were recruited. The whole-brain revealed

Table 1 Whole-brain and ROI activations when learning a new Standard sequence at Training Phase

Contrasts and anatomical label	MNI coordinate			Voxels	max z
	x	*y*	*z*		
Belief SRT task					
Standard block at training > standard block at test					
ROI: L Posterior Cerebellum (Crus I)	-34	-64	-44	16	3.517 *
ROI: R Temporoparietal Junction	46	-62	22	166	4.362***
	46	-50	34		4.083
	52	-46	26		3.459
	50	-54	36		3.424
	44	-50	20		3.276
	54	-58	18		3.153
L Cerebellum (VIII)	-4	-70	-34	352	4.222***
Cerebellar Vermis (IV/V)	0	-62	-14		4.209
Cerebellar Vermis (VII)	0	-66	-26		4.060
R Postcentral Gyrus	46	-36	60	28910	6.908***
R Middle Occipital Gyrus	48	-74	-2		6.638
L Inferior Parietal Lobule	-42	-46	58		6.567
L Precentral Gyrus	-28	-4	62	5148	6.170***
L Middle Frontal Gyrus	-28	14	52		5.623
L Anterior Cingulate Cortex	0	-2	30	326	4.485***
R Insula Lobe	32	20	6	216	4.735*
Control SRT task					
Standard block at training > standard block at test					
L Cerebellum (VIII)	-14	-58	-48	220	4.466***
L Cerebellum (IX)	-16	-46	-54		4.416
L Cerebellum (VIII)	-24	-46	-52		4.122
R Cerebellum (VI)	26	-54	-18	7845	6.141***
L Middle Temporal Gyrus	-50	-72	6	10446	5.761***
R Superior Parietal Lobule	42	-42	56		5.719
L Inferior Parietal Lobule	-36	-46	52		5.638
R Middle Temporal Gyrus	46	-72	0	7845	5.746***
L Middle Frontal Gyrus	-26	-2	62	1464	5.420***
R Superior Frontal Gyrus	24	4	68	850	4.609***
R Middle Frontal Gyrus	30	16	36	281	4.532***

Coordinates refer to the MNI (Montreal Neurological Institute) stereotaxic space. For the Whole-brain and ROI (small volume correction) analysis, we used a cluster-forming threshold of $p < .001$ (uncorrected) with minimum cluster extent of 10 voxels, and a cluster-wise level of $p < .05$, FWE corrected. Only the highest peaks of each cluster are shown, except for the cerebellum, TPJ, and related ROIs, as well as for large clusters > 500 voxels. L = left, R = right. ROIs for the posterior cerebellum are centered around coordinates ±24, -76, -40 (Crus II); and -40, -70, -40 (Crus I) and for the TPJ around coordinates ±50, -55, 25, all with radius 10 mm. * $p < 0.05$, **$p < 0.01$, ***$p < 0.001$ (cluster-wise FWE corrected)
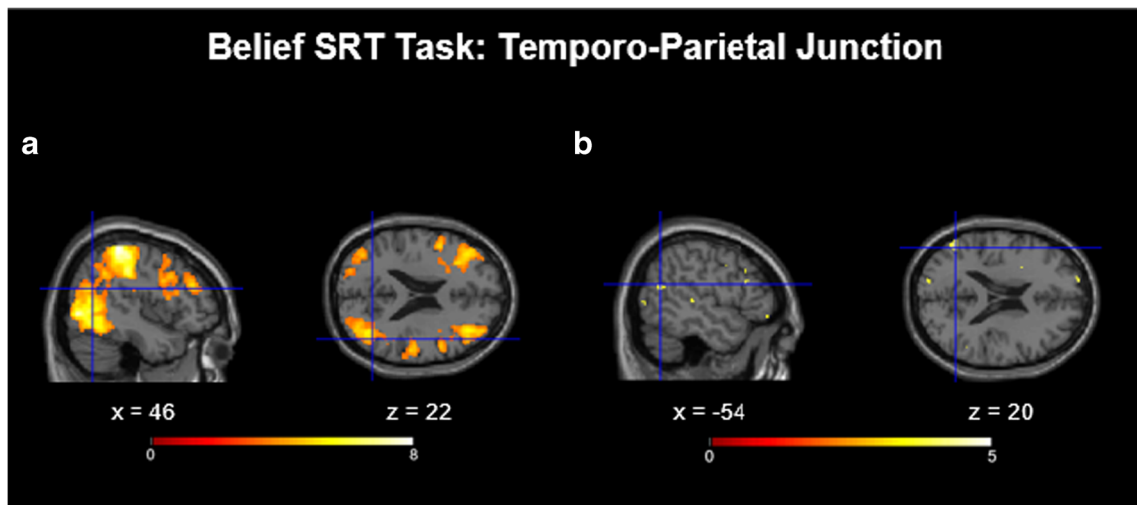
Fig. 4 Activations of three sequence learning effects for the Belief and Control SRT tasks, displayed at an uncorrected threshold of $p < 0.001$, with color bars denoting *t* values. (**Left**) Sagittal and transverse views of whole-brain activation at the peak coordinates of the cerebellar Crus, indicated by crosshairs. (**Right**) Cerebellar activations and *a prior* ROIs drawn on a SUIT flat map (https://www.diedrichsenlab.org/imaging/suit.htm); and on the functional network flatmap from Buckner, Krienen, Castellanos, Diaz, & Yeo, (2011; http://www.diedrichsenlab.org/imaging/AtlasViewer/viewer.html). The three sequence learning effects (**A–C**) show *a priori* ROIs (indicated by white circles) in the posterior cerebellar Crus in or close to the default/mentalizing network (denoted in red on the functional network flatmap). Shown are ROIs that were significant for the Belief SRT task; but note that none were significant for the Control SRT task (**D**). All *a priori* ROIs *shown* are significant at a cluster-forming threshold of $p < 0.001$ (uncorrected), and a cluster-wise level of $p < 0.05$, FWE corrected. Note that not all visible clusters are significant after FWE correction.

cerebral activity in the middle temporal gyrus and medial temporal pole, but there was no activation in the cerebellum.

Third, we verified *detecting violations* in the learned sequence by using a Random block at Test > Standard block at Test contrast (for both Total Random and Random Orientation; Table 3). For the Belief SRT task, the *a priori*

ROI analyses revealed activation in Crus I for the Random Orientation block > Standard block at Test (Fig. 4C), and in the TPJ for the Total Random block > Standard block at Test (Fig. 5B). No other *a priori* ROIs reached significance. The whole-brain analysis revealed activations in the cerebral cortex including the precentral gyrus, postcentral gyrus, superior/

Fig. 5 Sagittal and transverse views of whole-brain activation at the peak coordinates of the TPJ in the Belief SRT task, indicated by crosshairs, displayed at an uncorrected threshold of $p < 0.001$, with color bars denoting t values. The learning effects of general learning (**A**) and detecting violations (**B**) show significant TPJ activation in the Belief SRT task, with a cluster-forming threshold of $p < .001$ (uncorrected), and a cluster-wise level of $p < 0.05$, FWE corrected. Note that not all visible clusters are significant after FWE correction

Table 2   Whole-brain and ROI activations when maintaining the learned Standard sequence at Test Phase

Contrasts and anatomical label	MNI coordinate			Voxels	max z
	x	y	z		
Belief SRT task					
Standard block at test > standard block at training					
ROI: R Posterior Cerebellum (Crus II)	26	-84	-36	120	4.245**
	16	-78	-38		3.321
ROI: L Posterior Cerebellum (Crus II)	-24	-82	-36	27	3.664*
R Posterior Cerebellum (Crus I)	28	-84	-34	169	4.273*
R Posterior Cerebellum (Crus II)	22	-78	-42		3.750
R Posterior Cerebellum (Crus I)	20	-76	-34		3.657
L Temporal Pole	-48	14	-22	535	5.653***
L Superior Temporal Gyrus	-52	-4	-10		4.800
Standard block at test > total random block at test					
--					
Standard block at test > random orientation block at test					
--					
Control SRT task					
Standard block at test > standard block at training					
L Middle Temporal Gyrus	-56	-2	-16	854	5.264***
L Temporal Pole	-54	6	-12		5.256
R Medial Temporal Pole	58	4	-16	142	4.476*
Standard block at test > total random block at test					
--					
Standard block at test > random orientation block at test					
--					

Notes: Coordinates refer to the MNI (Montreal Neurological Institute) stereotaxic space. For the Whole-brain and ROI (small volume correction) analysis, we used a cluster-forming threshold of $p < .001$ (uncorrected) with minimum cluster extent of 10 voxels, and a cluster-wise level of $p < .05$, FWE corrected. Only the highest peaks of each cluster are shown, except for the cerebellum, TPJ, and related ROIs, as well as for large clusters > 500 voxels. L = left, R = right. ROIs for the posterior cerebellum are centered around coordinates ±24, -76, -40 (Crus II); and -40, -70, -40 (Crus I) and for the TPJ around coordinates ±50, -55, 25, all with radius 10 mm. * $p < 0.05$, **$p < 0.01$, ***$p < 0.001$ (cluster-wise FWE corrected)

Table 3 Whole-brain and ROI activations in detecting violations in the learned Standard sequence at Test Phase

Contrasts and anatomical label	MNI coordinate			Voxels	max z
	x	y	z		
Belief SRT task					
Total random block at test > standard block at test					
ROI: L Temporoparietal Junction	-54	-60	20	23	3.604*
L Precentral Gyrus	-28	-8	54	229	4.167***
L Middle Frontal Gyrus	-24	2	60	389	4.077***
L Inferior Frontal Gyrus (pars Opercularis)	-30	8	24	223	4.025***
Random orientation block at test > standard block at test					
ROI: L Posterior Cerebellum (Crus I)	-36	-64	-42	12	3.499*
R Superior Parietal Lobule	20	-58	52	403	4.320***
L Inferior Parietal Lobule	-42	-38	46	990	4.102***
R Superior Frontal Gyrus	28	-2	58	160	4.381*
Control SRT task					
Total random block at test > standard block at test					
--					
Random orientation block at test > standard block at test					
--					

Coordinates refer to the MNI (Montreal Neurological Institute) stereotaxic space. For the Whole-brain and ROI (small volume correction) analysis, we used a cluster-forming threshold of $p < .001$ (uncorrected) with minimum cluster extent of 10 voxels, and a cluster-wise level of $p < .05$, FWE corrected. Only the highest peaks of each cluster are shown, except for the cerebellum, TPJ, and related ROIs, as well as for large clusters > 500 voxels. L = left, R = right. ROIs for the posterior cerebellum are centered around coordinates ±24, -76, -40 (Crus II); and -40, -70, -40 (Crus I) and for the TPJ around coordinates ±50, -55, 25, all with radius 10 mm. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (cluster-wise FWE corrected)

middle frontal gyrus, and superior/inferior frontal gyrus for the two contrasts.

None of the reverse contrasts (e.g., Standard block at Test > Total Random blocks at Test) was significant. Likewise, for the Control SRT task, no significant activation was found for the *a priori* ROI and whole-brain analyses. Overall, the results strongly support the hypothesized activations in the posterior cerebellar Crus and TPJ in the mentalizing network for the Belief SRT task, but not for the Control SRT task.

Contrasts between Tasks: Simple Effects

To further test our hypothesis that implicit belief sequence learning would be preferentially supported by social cerebellar and cerebral areas (i.e., Crus I & II, and TPJ) in the Belief SRT task, and less so in the Control SRT task, we directly compared the Belief and Control SRT tasks using the SwE analysis. We tested a series of simple Belief > Control contrasts for each of Block Type (Standard or Random) at all Phases (Training or Test), followed by the reverse Control > Belief contrasts (Table 4). We again report the *a priori* ROI analysis first and then turn to whole-brain analysis.

The Belief > Control contrast showed several results for our hypothesis. First, for the Standard block at Training, the *a priori* ROI analysis revealed activation of the TPJ (Fig. 6A),

whereas no other *a priori* ROIs reached significance. The whole-brain analysis further showed cortical activations in the lingual gyrus, fusiform gyrus, and superior temporal gyrus. Second, for the Standard block at Test, the *a priori* ROI analysis showed activation of the cerebellar Crus II in the Belief SRT task (Fig. 6B), whereas no other *a priori* ROIs and whole-brain analyses reached significance. Third, for the two Random blocks at Test, the *a priori* ROI analysis also showed activation of the TPJ in the Belief SRT task. Also, the cerebellar Crus II was activated in the Total Random block at Test in the Belief SRT task (Fig. 6C). No other *a priori* ROIs and whole-brain analyses reached significance.

The reverse Control > Belief contrast did not reveal activation in any of the *a priori* ROIs, and additional whole-brain activations were found in the calcarine gyrus, middle occipital gyrus, and inferior parietal lobule in most of the contrasts (Table 4).

To visualize and further explore these differential activations in the posterior cerebellar Crus II and TPJ between the Belief and Control SRT tasks, we extracted percent signal change data using spheres of 10 mm around the peak activations of the strongest simple effects (Table 4). As shown in Fig. 6, these explorative analyses largely confirmed the whole-brain and *a priori* ROI analyses reported above. In general, the percent signal change analyses confirmed the stronger activation in the Belief SRT task than the Control

Table 4 Whole-brain and ROI activations between Belief and Control SRT tasks by SwE analysis

Contrasts and anatomical label	MNI coordinate			Voxels	max z
	x	y	z		
Simple effects					
Standard block at training					
Belief > control					
ROI: R Temporoparietal Junction	52	-46	24	98	4.302**
R Lingual Gyrus	4	-80	-4	323	4.768*
R Fusiform Gyrus	42	-56	-16	221	4.620*
R Superior Temporal Gyrus	50	-44	22	919	4.585*
R Inferior Occipital Gyrus	46	-76	-2		4.132*
Control > belief					
R Calcarine Gyrus	16	-88	2	430	5.390**
Standard block at test					
Belief > control					
ROI: L Posterior Cerebellum (Crus II)	-22	-68	-38	38	3.049*
Control > belief					
L Middle Occipital Gyrus	-12	-96	0	339	4.660**
R Calcarine Gyrus	16	-88	4	619	5.187**
L Inferior Parietal Lobule	-32	-56	44	263	4.360*
Total random block at test					
Belief > control					
ROI: L Posterior Cerebellum (Crus II)	-26	-74	-36	19	2.741*
	-22	-72	-40		2.705*
ROI: R Temporoparietal Junction	50	-46	22	61	3.578*
Control > belief					
L Middle Occipital Gyrus	-12	-96	0	281	4.470*
R Calcarine Gyrus	16	-88	4	448	4.965*
Random orientation block at test					
Belief > control					
ROI: R Temporoparietal Junction	50	-46	22	38	3.693*
Control > belief					
L Middle Occipital Gyrus	-14	-98	2	308	4.633*
R Calcarine Gyrus	16	-90	4	479	4.776*
Spreading interaction of general learning at training					
Belief standard block at training > [belief standard block at test = control standard block at training = control standard block at test]					
ROI: R Temporoparietal Junction	50	-46	26	191	4.694***
	46	-62	24		3.570**
	50	-54	16		3.085 **
L Cerebellum (IV-V)	-10	-54	-20	144	4.041*
L Cerebellum (VI)	-18	-54	-26		3.177*
Cerebellar Vermis (VII)	-2	-64	-26	90	3.063*
L Middle Occipital Gyrus	-46	-76	8	560	4.152*
L Middle Temporal Gyrus	-50	-56	12		3.549*
L Fusiform Gyrus	-22	-76	-6	50	3.461*
L Inferior Occipital Gyrus	-42	-74	-10	112	3.839*
R Precuneus	12	-58	52	2189	4.174*
R Postcentral Gyrus	44	-40	58		4.132*
R Fusiform Gyrus	42	-56	-16	4337	5.115**
R Middle Occipital Gyrus	40	-66	8		4.969**

Table 4 (continued)

Contrasts and anatomical label	MNI coordinate			Voxels	max z
	x	y	z		
R Inferior Occipital Gyrus	46	-74	-2		4.908**
L Postcentral Gyrus	-30	-38	52	125	3.215*
R Middle Frontal Gyrus	48	12	46	240	3.998*
R Superior Frontal Gyrus	24	16	60	217	3.563*
L Inferior Frontal Gyrus (p. Triangularis)	-46	20	10	100	3.551*
L Middle Frontal Gyrus	-36	24	42	82	3.643*
Control standard block at training > [control standard block at test = belief standard block at training = belief standard block at test]					
R Fusiform Gyrus	30	-48	-10	1027	4.699*
R Cerebellum (IV-V)	22	-48	-18		4.018
R Cerebellum (VIII)	8	-72	-38	181	3.444*
L Superior Occipital Gyrus	-14	-90	4	230	3.913*
R Calcarine Gyrus	16	-88	2	787	5.213*
R Cuneus	14	-94	14		4.455*
R Middle Occipital Gyrus	26	-88	18		4.176*
L Precuneus	-2	-64	42	65	3.813*
R Inferior Parietal Lobule	44	-56	-4	61	3.487
L Inferior Parietal Lobule	-34	-56	46	1203	4.221*
L Inferior Temporal Gyrus	-44	-54	-8	99	3.919*
L Fusiform Gyrus	-26	-52	-14	408	4.346*
R Inferior Parietal Lobule	38	-50	40	432	4.050*
R Rolandic Operculum	44	-34	20	175	3.922*
R Postcentral Gyrus	48	-28	42	260	3.590*
R Posterior-Medial Frontal	6	4	72	95	3.621*
L Insula Lobe	-42	8	0	61	3.587*
R Middle Cingulate Cortex	2	16	38	280	3.721*
L Middle Frontal Gyrus	-36	44	32	609	3.809*
Spreading interaction of learning maintenance at test					
Belief standard block at test > [belief standard block at training = control standard block at test = control standard block at training]					
ROI: L Posterior Cerebellum (Crus II)	-22	-74	-38	59	3.590**
	-18	-76	-36		3.169**
Control standard block at test > [control standard block at training = belief standard block at test = belief standard block at training]					
L Middle Occipital Gyrus	-14	-96	0	420	5.010*
R Calcarine Gyrus	14	-90	4	271	4.835*
Spreading interaction of detecting violations at test					
Belief total random & random orientation block at test > [belief standard block at test = control standard block at test = control total random & random orientation block at test]					
ROI: L Posterior Cerebellum (Crus II)	-22	-68	-38	11	3.942*
ROI: R Temporoparietal Junction	50	-46	22	75	4.206**
Control total random & random orientation block at test > [control standard block at test = belief standard block at test = belief total random & random orientation block at test]					
L Superior Occipital Gyrus	-12	-96	2	320	4.510*
R Calcarine Gyrus	16	-90	4	528	5.003*

Coordinates refer to the MNI (Montreal Neurological Institute) stereotaxic space. For whole-brain and ROI (small volume correction) analysis in SwE, we used a cluster-forming threshold of $p < .005$ (uncorrected) with minimum cluster extent of 50 voxels (10 voxels for the ROI), followed by a peak level of $p < .05$, FDR corrected. Only the highest peaks of each cluster are shown, except for the cerebellum, TPJ, and related ROIs, as well as for large clusters > 500 voxels. L = left, R = right. ROIs for the posterior cerebellum are centered around coordinates ±24, -76, -40 (Crus II); and -40, -70, -40 (Crus I) and for the TPJ around coordinates ±50, -55, 25, all with radius 10 mm. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (peak FDR corrected)
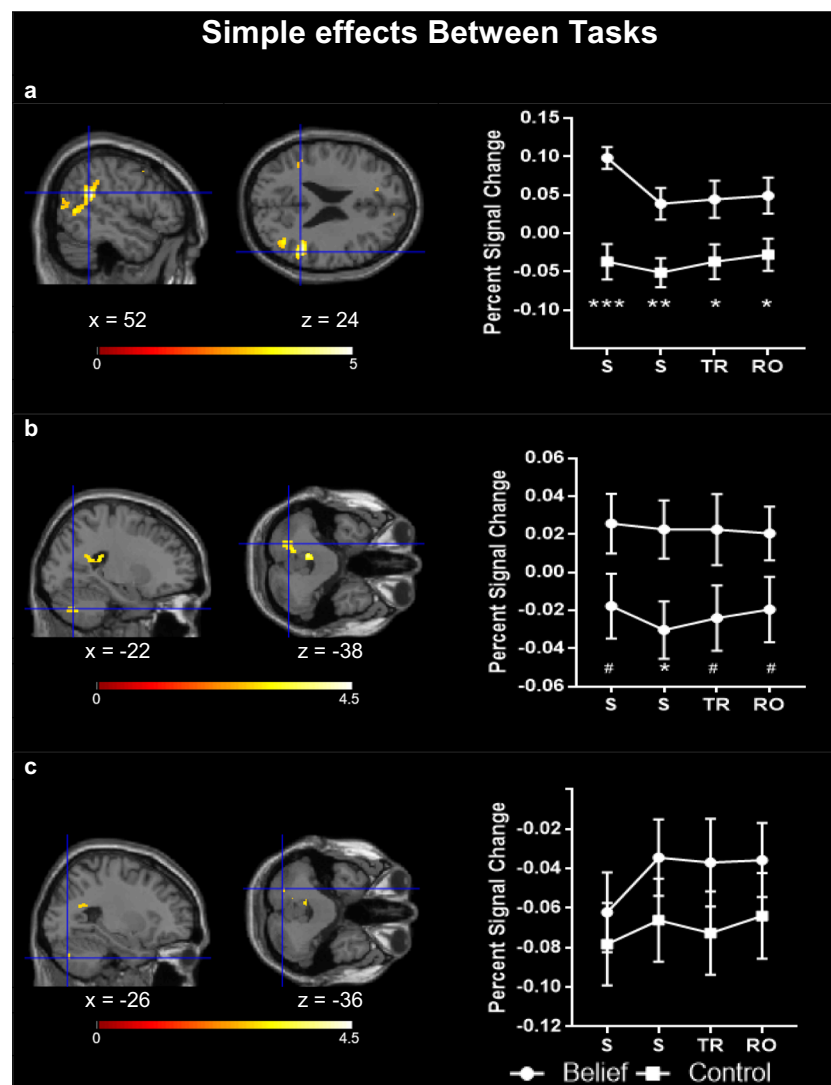
Fig. 6 Simple effects with higher brain activations at the Belief than the Control SRT task. (**Left**) Sagittal and transverse views of brain activation at the peak coordinates of the TPJ and cerebellar Crus II, indicated by crosshairs, displayed at an uncorrected threshold of $p < 0.005$, with color bars denoting SwE z value. (**Right**) Percent signal change at these peak activations in Standard Block at Training (= 1st S), Standard block at Test (= 2nd S), Total Random at Test (= TR) and Random Orientation at Test (= RO), showing stronger activation of the TPJ and Crus II in the Belief than the Control SRT task in all conditions. Error bars denote standard errors of the mean across participants. Symbols at the bottom denote (almost) significant differences between tasks: $^{\#}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

SRT task across all learning phases, and this effect was most prominent for the Standard sequence.

**Asymmetric Contrasts between Tasks: Spreading Interactions**

To test more directly our hypothesis that the cerebellar Crus and the cortical TPJ were significantly activated in the Belief SRT task, but less so in the Control SRT task, we applied a series of asymmetric or *spreading interactions* which assume high activations in the Belief SRT task for a given learning effect, and none in the Control SRT task (Table 4; Figs. 7 and 8).

First, for the spreading interaction of *general learning*, the *a priori* ROI analysis revealed activation in the TPJ (Fig. 8A), and the whole-brain analysis revealed activation in cerebellar lobules

IV-V, VII, and Vermis VII (Fig. 7A). Second, for the spreading interaction of *maintenance of learning*, the *a priori* ROI analysis showed engagement of Crus II (Fig. 7B). Third, for the spreading interaction of *detecting violations,* the *a priori* ROI analysis revealed activation of Crus II and TPJ (Figs. 7C and 8B). No other significant activations for the cerebellum were found. The results from the whole brain analyses are listed in Table 4.

In contrast, the reverse spreading interactions for testing high activations in the Control SRT task, but not in the Belief SRT task (Table 4), revealed no activation in Crus I or II, or in the TPJ for any *a priori* ROI analysis. The whole-brain analysis revealed activation in cerebellar lobules IV-V and VIII for *general learning* only. The results from the whole brain analyses are listed in Table 4.
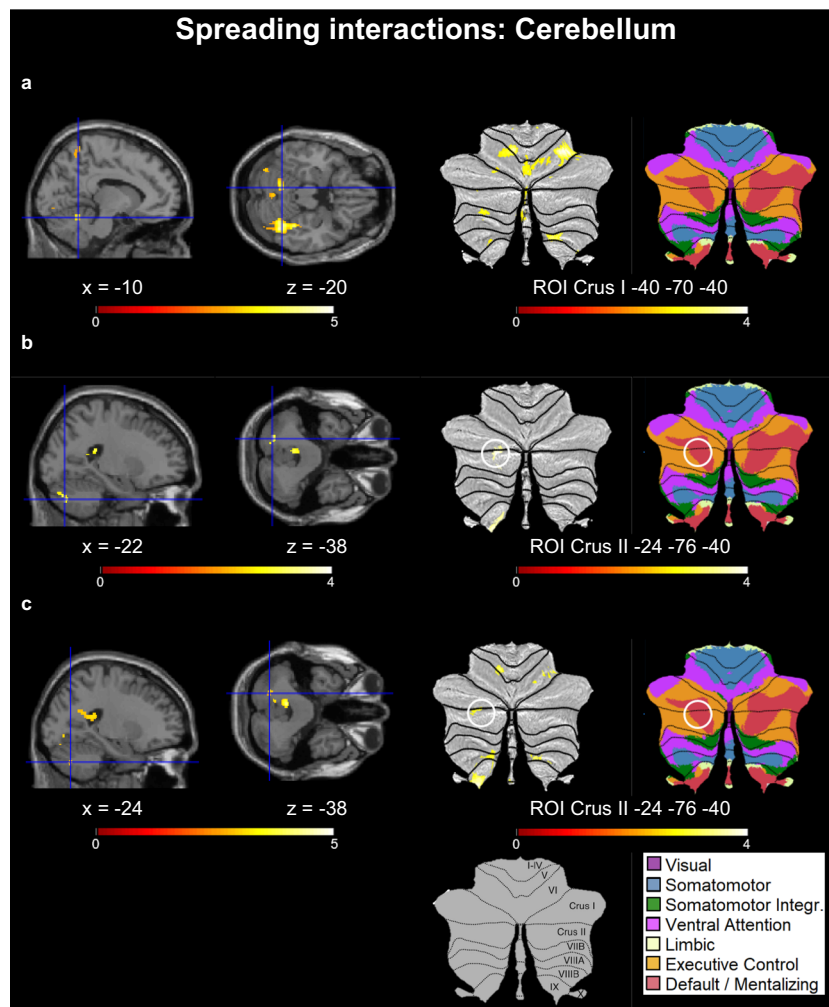
Fig. 7 Spreading interaction with high cerebellar Crus activations at the Belief SRT task, and none at the Control SRT task, displayed at an uncorrected threshold of $p < 0.005$, with color bars denoting SwE z values. (**Left**) Sagittal and transverse views of brain activations at the peak coordinates of the cerebellar Crus, indicated by crosshairs. (**Right**) Cerebellar activations and *a priori* ROIs shown on a SUIT flatmap (https://www.diedrichsenlab.org/imaging/suit.htm) and on the functional network map from Buckner, Krienen, Castellanos, Diaz, & Yeo (2011; http://www.diedrichsenlab.org/imaging/AtlasViewer/viewer.html).

Compared to the Control SRT task, the posterior cerebellar Crus shows stronger *priori* ROI activations (denoted by white circles) in the default/ mentalizing network (denoted in red on the functional network flatmap) in the Belief SRT task for the three learning effects(**A–C**). All *a priori* ROIs shown are significant at a cluster-forming threshold of $p < 0.005$ (uncorrected), and a cluster-wise level of $p < 0.05$, FDR corrected in a SwE analysis. Note that not all visible clusters are significant after FDR correction
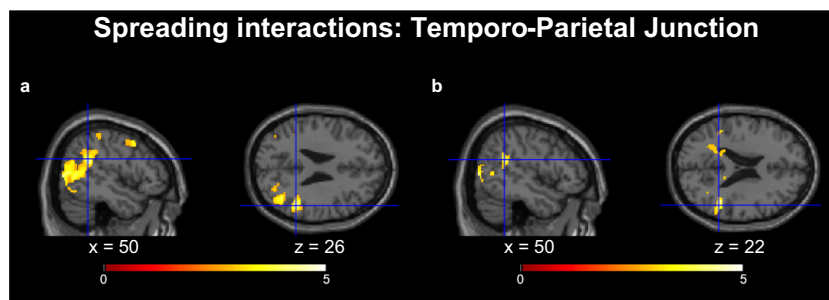
Fig. 8 Spreading interaction showing higher TPJ activation at the Belief SRT task at *general learning* [**A**], and at *detecting violations* [**B**] compared to none at the Control SRT task, displayed at an uncorrected threshold of $p < 0.005$, with color bars denoting SwE z values. Sagittal and transverse views of brain activations at the peak coordinates of the TPJ, indicated by crosshairs. Note that not all visible clusters are significant after FDR correction

Discussion

This study investigated whether the posterior cerebellar Crus is involved in implicit belief sequence learning. By using a Belief SRT task, we tested the hypothesis that this area (as well as the TPJ) is involved in implicitly learning sequences of belief orientations, which require continuous inferences of others' mental states. This research is novel, because the task differs from earlier research that investigated the explicit generation (Heleven, van Dun, & Van Overwalle, 2019) or memorizing of social sequences (Pu et al., 2020).

First, our behavioral results showed that people could implicitly learn belief sequences (i.e., a standard sequence of 16 true-false belief states). This was revealed by faster RTs for the standard sequence and slower RTs for random sequences. Second, and most critically, our neuroimaging results revealed that the posterior cerebellar Crus and the TPJ were preferentially engaged in the Belief SRT task (Table 5). These activations in the posterior cerebellar Crus I/II were located in the default/mentalizing network (Buckner et al., 2011; Van Overwalle et al., 2020a) and differ largely from previous cognitive motor-related implicit SRT tasks, which typically found activation in the anterior cerebellum Lobule V/VI (Bernard & Seidler, 2013). Because the Smurfs were consistently present on the screen during the Belief SRT task and activations were revealed by contrasts within that task, we can safely rule out biological movement or the mere presence of social protagonists as an explanation for the present cerebellar Crus activations (Sokolov et al., 2012).

The results also suggested some division of labor: the posterior cerebellar Crus I and the TPJ were activated during the early phase of implicit learning of the belief sequence and during detection of random violations in this sequence. In contrast, the cerebellar Crus II was activated during maintenance of the sequence at the late learning phase (in a context of sequence violations). This pattern suggests that Crus I may be preferentially involved in detecting novel sequences during early learning or during violations, whereas Crus II may be mainly involved in the formation of an internal model of repeated belief sequences. This is in line with the speculation that the formation of distinct internal models is consolidated during later learning (Bernard & Seidler, 2013). These results are broadly consistent with the different locations and functional profiles of Crus I and II. The ROIs of Crus II are located within the mentalizing network, whereas the ROI of Crus I is located somewhat more peripherally in the mentalizing network and closer to the executive network (Buckner et al., 2011). The meta-analysis by Van Overwalle, Ma & Heleven (2020a) further supports this distinct functional profile, in that the majority of studies recruiting Crus II ROI involved social mentalizing (74%), versus only a minority of Crus I ROI studies (35%). Future research is needed to verify to what extent these differential functions and corresponding areas are robust.

By comparing the Belief against the Control SRT task, our results showed that the posterior cerebellar Crus plays a social domain-specific role. Indeed, the design and sequence structure were identical in both tasks, except that participants in the Belief SRT task had to continuously infer the protagonists' beliefs, whereas this social inference process was absent in the Control SRT task (given the absence of social agents and social behaviors). These findings support our hypothesis that the posterior cerebellar Crus is preferentially involved in social sequences, not only at an explicit level as demonstrated in earlier fMRI studies (Cattaneo et al., 2012; Heleven et al., 2019; Pu et al., 2020), but also at an implicit level. This parallels previous fMRI findings on implicit and explicit social attributions, which showed a great overlap in key mentalizing

Table 5 Summary of ROI results (Tables 1, 2, 3 and 4)

Contrasts	Within task			Between tasks				
	Contrasts			Contrasts		Spreading interactions		
	Belief		Control	Belief > control		Belief		Control
General learning at training Standard block at training >Standard block at Test (Tables 1 & 4)	L Crus I	R TPJ	---	---	R TPJ	---	R TPJ	---
Learning maintenance at test Standard block at test > standard block at training (Tables 2 & 4)	R/L Crus II	---	---	L Crus II	---	L Crus II	---	---
Detecting violations at test Total random block at test > standard block at test (Tables 3 & 4)	---	L TPJ	---	L Crus II	R TPJ	L Crus II	R TPJ	---
Detecting violations at test Random orientation block at test > standard block at test (Tables 3 & 4)	L Crus I	---	---	---	R TPJ			

Tables 1, 2 and 3 show the results from the within-task analysis by SPM and Table 4 shows the results from the between-task analysis by SwE.

L = Left; R = Right.

areas at the cortical level (for a review, see Van Overwalle & Vandekerckhove, 2013). More generally, our results support the "sequence detection" hypothesis put forward by Leggio & Molinaro (2015), and extended by Van Overwalle et al. (2019b) to the social domain by proposing that the posterior cerebellum allows people to predict and automatize social action sequences, and detect disruptions in these sequences (Van Overwalle et al., 2020b).

As a key cortical area for inferring other's mental states, the TPJ was activated during the present Belief SRT task, but not during the Control SRT task. This is in line with previous research that the TPJ serves to compute another's perspective and infer the content of mental states (e.g., beliefs; Saxe et al., 2006; Van Overwalle, 2009). More importantly, the TPJ was simultaneously engaged with the posterior cerebellum when learning a belief sequence or detecting violations, and when the sequence was learned, activation of the TPJ was reduced. This is in line with the assumption of the closed-loops between the cerebellum and TPJ (Kelly & Strick, 2003). Namely, when making belief inferences, the TPJ sends signals to the cerebellum to create internal models of a sequence. Critically, when the cerebellum detects violations of the expected sequence, it sends error signal to the TPJ leading to adjustments in social expectations. Thus, unlike motor sequence learning, which heavily rests on cortical input and exchange with the basal ganglia (Caligiore et al., 2019), in the social domain, the TPJ plays a key role in acquiring and providing information on others' mental states (e.g. beliefs). Further research needs to investigate in more depth the functional and anatomical connections between these mentalizing areas in the cerebellum and cerebral cortex during implicit sequence learning in the social domain.

In addition to the cerebellar Crus and the cortical TPJ, additional activations in the visual, primary motor, and somatosensory cortices were observed in the Belief SRT task. The activations of the dorsal premotor cortex in the present Belief and Control SRT task were close to activations revealed in meta-analysis on sequence learning (when controlling for potential motor confounds; Hardwick et al., 2013). These activations may be related to selecting and updating appropriate sequence knowledge according to visual cues (Hardwick et al., 2013). In addition to this, activation in parietal and frontal cortices in both Belief and Control SRT tasks may be related to increased attention and working memory (Cross, Stadler, Parkinson, Schütz-Bosbach, & Prinz, 2013; Kelley, Serences, Giesbrecht, & Yantis, 2008). However, future research is needed to investigate how consistent these activations are across various social or nonsocial SRT tasks, and which processes underlie these activations.

Like prior SRT research, we applied a behavioral post-test procedure to measure sequence awareness. Participants who reported having noticed "something special" most likely refer to a "feeling of familiarity" rather than explicit sequence knowledge because they showed low performance in a free recall test (Werheid, Zysset, Müller, Reuter, & Von Cramon, 2003). To ensure that awareness was not a mitigating factor, we re-ran all analyses for participants who did not report any awareness (i.e., 12 participants in the Belief SRT task; 16 participants in the Control SRT task). Although the results are somewhat less strong with this reduced number of participants (Supplementary material Tables S2–S4), the activations are similar to the results from all participants (e.g., peak activations in the TPJ and posterior cerebellum were closed to each other, 4-8 mm, in the Belief SRT task). Hence, as is typical in previous studies, we did not exclude participants who noticed "something special" (Dennis & Cabeza, 2011; Gheysen, Van Opstal, Roggeman, Van Waelvelde, & Fias, 2011). To investigate potential differences, future research might recruit more participants in order to distinguish between those with and without any awareness.

Perhaps another interesting issue for future neuroimaging research might be to develop an explicit instruction of the present implicit study (i.e., explicitly informing participants about the existence of true-false belief orientation sequences) and compare the results of the implicit and explicit instructions. By doing so, we could further investigate whether the involvement of the posterior cerebellar Crus is similar or different for implicit and explicit belief-related sequence learning. It also will be of interest for future research to measure individual capacities for mentalizing, and test the relationship between these capacities and implicit learning of belief sequences. Also, future research could consider to investigate the neural time course of the posterior cerebellar Crus and TPJ related to behavioral time course of belief-related sequence learning.

An important concern is to what extent the present Belief SRT task reflects social processes. As noted earlier, the TPJ was preferentially activated in the Belief SRT task in comparison with the Control SRT task, strongly suggesting that social processing was involved, and that participants did infer the protagonists' beliefs rather than only relying on cognitive executive functions. In addition, this TPJ activation is in line with the study by Saxe, Schulz, & Jiang (2006), mentioned earlier, which investigated an analogous false belief design, similar to the present Belief SRT task. To recall, in Saxe's false belief task, a chocolate was moved out of a box and then returned to the same or another box, and participants were asked to identify "where the girl thinks the chocolate bar is" when she was oriented toward or away from the boxes. Conversely, in the control task, the instruction read: "If the girl is facing the boxes at the end of the trial, press the button for the last box. If the girl is looking away from the boxes, press the button for the first box." Both conditions recruited brain regions associated with domain-general attention, response selection and inhibitory control. Importantly, despite the structural equivalence and identical stimuli of the two

conditions, only the belief instruction recruited the right TPJ, close to our TPJ activations (approximately 2-4 mm away, Tables 1 & 3). This finding supports the idea that participants in our study truly made belief attributions in the Belief SRT task. Note that we do not suggest that the Belief SRT task is devoid of other cognitive processes, because belief attribution also requires attention and executive functions (Krall et al., 2015a, 2015b; Van Overwalle, 2011).

However, we cannot entirely exclude the possibility that participants completed the Belief SRT task by applying other, purely perceptual rules. For instance, a rule to switch to the previous trial with the same protagonist when orientation changes. For future research, one could take a similar approach as in the study of Saxe et al. (2006) for a control task, by keeping the stimulus material itself constant in addition to the sequential structure. For example, participants would see protagonists with colored clothes and get the instruction to observe or repeat the number of flowers based on the colors of protagonists' clothes. However, note that Saxe et al. (2006) avoided contamination of cognitive and social processes by first testing the control task, followed by the novel (and suddenly introduced) false belief task. Moreover, note that participants required an extensive 30-minute practice on the control task, but not the false-belief task, before entering the scanner. Perhaps another control task could be developed based on elements of the false-photograph task, which requires participants to infer whether a photo taken by a camera is consistent with current reality or outdated (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2007; Ma et al., 2021; Saxe, 2006). For example, a camera that is orientated toward flowers (representing a "true" or current photo) or away from flowers (representing an "false" or outdated photo).

The current results also provide additional insights for future research. By demonstrating that key mentalizing areas are engaged while people implicitly learn repeated false and true belief orientations, this study provides an important proof-of-principle on people's ability to implicitly learn and use repeated patterns of behavioral cues to predict others' mental states. This supports the idea that people quite often are unintentionally affected by the temporal sequence of others' actions, facial expressions, eye gazes, etc. even without explicit attention to such behavioral cues. This is a crucial ability whereby people come to learn the stable regularities in dynamic social stimuli (Lieberman, 2000) intuitively, which may help them to anticipate behaviors of others and the consequences for themselves, and to recognize deviations that can lead to modifications in future interactions. For future research, various social stimuli could be combined with the SRT task, such as sequential facial expressions or eye gaze directions (Geiger, Cleeremans, Bente, & Vogeley, 2018). These "social" SRT tasks could provide more evidence on implicit learning abilities in dynamic social contexts.

Also, the present Belief SRT task might be useful in clinical studies. Previous research has shown that individuals with autism can implicitly learn motor responses required to identify cognitive patterns in an SRT task (e.g., 12-item sequences of locations, Brown, Aczel, Jiménez, Kaufman, & Grant, 2010; see also Foti, De Crescenzo, Vivanti, Menghini, & Vicari, 2015; Zwart, Vissers, & Maes, 2018). As individuals with autism are characterized by social impairments, it is an open question whether they are also capable of implicit learning of sequential social cues such as true and false belief orientations. Moreover, although research on cerebellar patients has demonstrated that cerebellar injury leads to impairments in social functioning (Wang, Kloth, & Badura, 2014), implicit social learning might still be relatively preserved, suggesting possibilities for improved social treatment of patients. In this perspective, our findings implicate potential neural targets (cerebellar Crus I/II) for improving social cognitive functioning by brain stimulations (Brady et al., 2020). Also, although the present Belief SRT task focused on implicit learning in a restricted way which may still be far away from real social interaction, it might constitute an important diagnostic test on implicit social sequence learning. For example, our novel Belief SRT task could be conducted with patients with autism spectrum disorders or other neurodevelopmental disorders or cerebellar patients characterized by social impairments, and so begin to answer to what extent these patients have impairments in implicit social sequence learning.

Conclusions

Our study demonstrated that the posterior cerebellum, and more in particular the Crus I and II, subserves implicit learning of belief sequences, which supports the cerebellar "sequence detection" hypothesis applied in the social domain. Unlike previous fMRI research on the role of the cerebellum in social sequences which was hereto explicit, this study demonstrated for the first time that sequences of mental states also can develop unintentionally with practice, by recruiting the posterior cerebellum and TPJ in the same areas as explicit mentalizing sequences, without the involvement of overt sequential movements and somatosensory responses. Future investigations could deepen our understanding of how social interaction can be facilitated through learning sequences of social attributions supported by the cerebellum, and therefore deepen our understanding of other functions of the cerebellum in social cognition. Importantly, our novel Belief SRT task can begin to answer questions, such as why and how patients with autism or cerebellar damage experience their social deficits. This task could be further developed and investigated as part of diagnostic tools for assessing implicit social impairments.

References

Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W. L., & Humphreys, G. W. (2007). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal, "reality-unknown" false belief and false photograph tasks. *Cognition*, *103*(2), 300–321. https://doi.org/10.1016/j.cognition.2006.04.012

Baetens, K., Firouzi, M., Van Overwalle, F., & Deroost, N. (2020). Involvement of the cerebellum in the serial reaction time task (SRT) (Response to Janacsek et al.). *NeuroImage*, *220*(February), 117114. https://doi.org/10.1016/j.neuroimage.2020.117114

Bernard, J. A., & Seidler, R. D. (2013). Cerebellar contributions to visuomotor adaptation and motor sequence learning: An ALE meta-analysis. *Frontiers in Human Neuroscience*, *7*(FEB), 27. https://doi.org/10.3389/fnhum.2013.00027

Bradford, E. E. F., Jentzsch, I., & Gomez, J. C. (2015). From self to social cognition: Theory of Mind mechanisms and their relation to Executive Functioning. *Cognition*, *138*, 21–34. https://doi.org/10.1016/j.cognition.2015.02.001

Brady, R. O., Beermann, A., Nye, M., Eack, S. M., Mesholam-Gately, R., Keshavan, M. S., & Lewandowski, K. E. (2020). Cerebellar-Cortical Connectivity Is Linked to Social Cognition Trans-Diagnostically. *Frontiers in Psychiatry*, *11*(November). https://doi.org/10.3389/fpsyt.2020.573002

Brown, J., Aczel, B., Jiménez, L., Kaufman, S. B., & Grant, K. P. (2010). Intact implicit learning in autism spectrum conditions. *Quarterly Journal of Experimental Psychology*, *63*(9), 1789–1812. https://doi.org/10.1080/17470210903536910

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.

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: An integrative theoretical account. *Trends in Cognitive Sciences*, *16*(6), 338–352. https://doi.org/10.1016/j.tics.2012.04.008

Caligiore, D., Arbib, M. A., Miall, R. C., & Baldassarre, G. (2019). The super-learning hypothesis: Integrating learning processes across cortex, cerebellum and basal ganglia. *Neuroscience and Biobehavioral Reviews*, *100*, 19–34. https://doi.org/10.1016/j.neubiorev.2019.02.008

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral Cortex (New York, N.Y. : 1991)*, *15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Cattaneo, L., Fasanelli, M., Andreatta, O., Bonifati, D. M., Barchiesi, G., & Caruana, F. (2012). Your actions in my cerebellum: Subclinical deficits in action observation in patients with unilateral chronic cerebellar stroke. *Cerebellum*, *11*(1), 264–271. https://doi.org/10.1007/s12311-011-0307-9

Clausi, S., Olivito, G., Lupo, M., Siciliano, L., Bozzali, M., & Leggio, M. (2019). The cerebellar predictions for social interactions: Theory of mind abilities in patients with degenerative cerebellar atrophy. *Frontiers in Cellular Neuroscience*, *12*(January), 510. https://doi.org/10.3389/fncel.2018.00510

Coomans, D., Deroost, N., Zeischka, P., & Soetens, E. (2011). On the automaticity of pure perceptual sequence learning. *Consciousness and Cognition*, *20*(4), 1460–1472. https://doi.org/10.1016/j.concog.2011.06.009

Cross, E. S., Stadler, W., Parkinson, J., Schütz-Bosbach, S., & Prinz, W. (2013). The influence of visual training on predicting complex action sequences. *Human Brain Mapping*, *34*(2), 467–486. https://doi.org/10.1002/hbm.21450

Cusack, R., & Papadakis, N. (2002). New robust 3-D phase unwrapping algorithms: Application to magnetic field mapping and undistorting echoplanar images. *NeuroImage*, *16*(3), 754–764. https://doi.org/10.1006/nimg.2002.1092

Debas, K., Carrier, J., Orban, P., Barakat, M., Lungu, O., Vandewalle, G., … Doyon, J. (2010). Brain plasticity related to the consolidation of motor sequence learning and motor adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(41), 17839–17844. https://doi.org/10.1073/pnas.1013176107

Dennis, N. A., & Cabeza, R. (2011). Age-related dedifferentiation of learning systems: An fMRI study of implicit and explicit learning. *Neurobiology of Aging*, *32*(12), 2318.e17-2318.e30. https://doi.org/10.1016/j.neurobiolaging.2010.04.004

Deroost, N., & Coomans, D. (2018). Is sequence awareness mandatory for perceptual sequence learning: An assessment using a pure perceptual sequence learning design. *Acta Psychologica*, *183*(January), 58–65. https://doi.org/10.1016/j.actpsy.2018.01.002

Fleming, D., Belleville, S., Peretz, I., West, G., & Zendel, B. R. (2019). The effects of short-term musical training on the neural processing of speech-in-noise in older adults. *Brain and Cognition*, *136*(August), 103592. https://doi.org/10.1016/j.bandc.2019.103592

Foti, F., De Crescenzo, F., Vivanti, G., Menghini, D., & Vicari, S. (2015). Implicit learning in individuals with autism spectrum disorders: A meta-analysis. *Psychological Medicine*, *45*(5), 897–910. https://doi.org/10.1017/S0033291714001950

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology : CB*, *15*(17), 644–645. https://doi.org/10.1016/j.cub.2005.08.041

Geiger, A., Cleeremans, A., Bente, G., & Vogeley, K. (2018). Social Cues Alter Implicit Motor Learning in a Serial Reaction Time Task. *Frontiers in Human Neuroscience*, *12*(May), 1–12. https://doi.org/10.3389/fnhum.2018.00197

Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., & Fias, W. (2011). The Neural Basis of Implicit Perceptual Sequence Learning. *Frontiers in Human Neuroscience*, *5*(November), 1–12. https://doi.org/10.3389/fnhum.2011.00137

Guell, X., Gabrieli, J. D. E., & Schmahmann, J. D. (2018). Triple representation of language, working memory, social and emotion processing in the cerebellum: convergent evidence from task and seed-based resting-state fMRI analyses in a single large cohort. *NeuroImage*, *172*(January), 437–449. https://doi.org/10.1016/j.neuroimage.2018.01.082

Guillaume, B., Hua, X., Thompson, P. M., Waldorp, L., & Nichols, T. E. (2014a). Fast and accurate modelling of longitudinal and repeated measures neuroimaging data. *NeuroImage*, *94*, 287–302. https://doi.org/10.1016/j.neuroimage.2014.03.029

Guillaume, B., Hua, X., Thompson, P. M., Waldorp, L., Nichols, T. E., & Initiative, N. (2014b). *The SwE Toolbox : a Toolbox for the Analysis of Longitudinal and Repeated Measures Neuroimaging Data*. 4.

Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage*, *67*, 283–297. https://doi.org/10.1016/j.neuroimage.2012.11.020

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, *57*(2), 243. https://doi.org/10.2307/1416950

Heleven, E., van Dun, K., & Van Overwalle, F. (2019). The posterior Cerebellum is involved in constructing Social Action Sequences: An fMRI Study. *Scientific Reports*, *9*(1), 11110. https://doi.org/10.1038/s41598-019-46962-7

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4), 304–313. https://doi.org/10.1038/nrm2332

Janacsek, K., Shattuck, K. F., Tagarelli, K. M., Lum, J. A. G., Turkeltaub, P. E., & Ullman, M. T. (2020). Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *NeuroImage*, 207(May 2019), 116387. https://doi.org/10.1016/j.neuroimage.2019.116387

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Kampis, D., Fogd, D., & Kovács, Á. M. (2017). Nonverbal components of Theory of Mind in typical and atypical development. *Infant Behavior and Development*, 48, 54–62. https://doi.org/10.1016/j.infbeh.2016.11.001

Kelley, T. A., Serences, J. T., Giesbrecht, B., & Yantis, S. (2008). Cortical Mechanisms for Shifting and Holding Visuospatial Attention. *Cerebral Cortex*, 18(1), 114–125. https://doi.org/10.1093/cercor/bhm036

Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, 23(23), 8432–8444. https://doi.org/10.1523/jneurosci.23-23-08432.2003

Koziol, L. F., Budding, D., Andreasen, N., D'Arrigo, S., Bulgheroni, S., Imamizu, H., … Yamazaki, T. (2014). Consensus paper: The cerebellum's role in movement and cognition. *Cerebellum*, 13(1), 151–177. https://doi.org/10.1007/s12311-013-0511-x

Krall, S. C., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P. T., Eickhoff, S. B., … Konrad, K. (2015a). The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Structure and Function*, 220(2), 587–604. https://doi.org/10.1007/s00429-014-0803-z

Krall, S., Volz, L. J., Oberwelland, E., Grefkes, C., Fink, G. R., & Konrad, K. (2015b). The right temporoparietal junction in attention and social interaction: A transcranial magnetic stimulation study. *Human Brain Mapping*, 00. https://doi.org/10.1002/hbm.23068

Leggio, M., & Molinari, M. (2015). Cerebellar Sequencing: a Trick for Predicting the Future. *The Cerebellum*, 14(1), 35–38. https://doi.org/10.1007/s12311-014-0616-x

Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin*, 126(1), 109–136. https://doi.org/10.1037/0033-2909.126.1.109

Ma, Q., Heleven, E., Funghi, G., Pu, M., Baetens, K., Deroost, N., & Van Overwalle, F. (2021). Implicit Learning of True and False Belief Sequences. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.643594

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(April), 276–291. https://doi.org/10.1016/j.neubiorev.2016.03.020

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 32(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

Özdem, C., Brass, M., Schippers, A., Van der Cruyssen, L., & Van Overwalle, F. (2019). The neural representation of mental beliefs held by two agents. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-019-00714-2

Özdem, C., Brass, M., Van der Cruyssen, L., & Van Overwalle, F. (2017). The overlap between false belief and spatial reorientation in the tempero-parietal junction: The role of input modality and task. *Social Neuroscience*, 12(2), 207–217. https://doi.org/10.1080/17470919.2016.1143027

Pickering, M. J., & Clark, A. (2014). Getting ahead: Forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, 18(9), 451–456. https://doi.org/10.1016/j.tics.2014.05.006

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., … Ruffman, T. (2018). Do infants understand false beliefs? We don't know yet – A commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognitive Development*, 48(September), 302–315. https://doi.org/10.1016/j.cogdev.2018.09.005

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Pu, M., Heleven, E., Delplanque, J., Gibert, N., Ma, Q., Funghi, G., & Van Overwalle, F. (2020). The posterior cerebellum supports the explicit sequence learning linked to trait attribution. *Cognitive, Affective and Behavioral Neuroscience*, 20(4), 798–815. https://doi.org/10.3758/s13415-020-00803-7

Saxe, R. (2006). Why and how to study Theory of Mind with fMRI. *Brain Research*, 1079(1), 57–65. https://doi.org/10.1016/j.brainres.2006.01.001

Saxe, R. R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: dissociating theory of mind and executive control in the brain. *Social Neuroscience*, 1(3–4), 284–298. https://doi.org/10.1080/17470910601000446

Schneider, D., Slaughter, V. P., & Dux, P. E. (2017). Current evidence for automatic Theory of Mind processing in adults. *Cognition*, 162, 27–31. https://doi.org/10.1016/j.cognition.2017.01.018

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., Gharabaghi, A., Grodd, W., Tatagiba, M. S., & Pavlova, M. A. (2012). Biological motion processing: The left cerebellum communicates with the right superior temporal sulcus. *NeuroImage*, 59(3), 2824–2830. https://doi.org/10.1016/j.neuroimage.2011.08.039

Sokolov, A. A., Miall, R. C., & Ivry, R. B. (2017). The Cerebellum: Adaptive Prediction for Movement and Cognition. *Trends in Cognitive Sciences*, 21(5), 313–332. https://doi.org/10.1016/j.tics.2017.02.005

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, 30(3), 829–858. https://doi.org/10.1002/hbm.20547

Van Overwalle, F. (2011). A dissociation between social mentalizing and general reasoning. *NeuroImage*, 54(2), 1589–1599. https://doi.org/10.1016/j.neuroimage.2010.09.043

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, 48(3), 564–584. https://doi.org/10.1016/j.neuroimage.2009.06.009

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(4), 337–344. https://doi.org/10.1080/17470919.2015.1005666

Van Overwalle, F., De Coninck, S., Heleven, E., Perrotta, G., Taib, N. O. Ben, Manto, M., & Mariën, P. (2019a). The role of the cerebellum in reconstructing social action sequences: A pilot study. *Social Cognitive and Affective Neuroscience*, 14(5), 549–558. https://doi.org/10.1093/scan/nsz032

Van Overwalle, F., Ma, Q., & Heleven, E. (2020a). The posterior crus II cerebellum is specialized for social mentalizing and emotional self-experiences: a meta-analysis. *Social Cognitive and Affective Neuroscience*, 15(9), 905–928. https://doi.org/10.1093/scan/nsaa124

Van Overwalle, F., Manto, M., Cattaneo, Z., Clausi, S., Ferrari, C., Gabrieli, J. D. E. E., … Leggio, M. (2020b). Consensus Paper: Cerebellum and Social Cognition. *The Cerebellum*, 19(6), 833–868. https://doi.org/10.1007/s12311-020-01155-1

Van Overwalle, F., Manto, M., Leggio, M., & Delgado-García, J. M. (2019b). The sequencing process generated by the cerebellum crucially contributes to social interactions. *Medical Hypotheses*, *128*, 33–42. https://doi.org/10.1016/j.mehy.2019.05.014

Van Overwalle, F., Van de Steen, F., & Mariën, P. (2019c). Dynamic causal modeling of the effective connectivity between the cerebrum and cerebellum in social mentalizing across five studies. *Cognitive, Affective and Behavioral Neuroscience*, *19*(1), 211–223. https://doi.org/10.3758/s13415-018-00659-y

Van Overwalle, F., Van de Steen, F., van Dun, K., & Heleven, E. (2020c). Connectivity between the cerebrum and cerebellum during social and non-social sequencing using dynamic causal modelling. *NeuroImage*, *206*(July), 116326. https://doi.org/10.1016/j.neuroimage.2019.116326

Van Overwalle, F., & Vandekerckhove, M. (2013). Implicit and explicit social mentalizing: Dual processes driven by a shared neural network. *Frontiers in Human Neuroscience*, *7*(SEP), 1–6. https://doi.org/10.3389/fnhum.2013.00560

Wang, S. S. H. S.-H., Kloth, A. D., & Badura, A. (2014). The Cerebellum, Sensitive Periods, and Autism. *Neuron*, *83*(3), 518–532. https://doi.org/10.1016/j.neuron.2014.07.016

Werheid, K., Zysset, S., Müller, A., Reuter, M., & Von Cramon, D. Y. (2003). Rule learning in a serial reaction time task: An fMRI study on patients with early Parkinson's disease. *Cognitive Brain Research*, *16*(2), 273–284. https://doi.org/10.1016/S0926-6410(02)00283-5

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Zwart, F. S., Vissers, C. T. W. M., & Maes, J. H. R. (2018). The Association Between Sequence Learning on the Serial Reaction Time Task and Social Impairments in Autism. *Journal of Autism and Developmental Disorders*, *48*(8), 2692–2700. https://doi.org/10.1007/s10803-018-3529-6