ORIGINAL ARTICLE



Feasibility and Efficacy of a Virtual Reality Social Prediction Training in Children and Young Adults with Congenital Cerebellar Malformations

Niccolò Butti¹ · Emilia Biffi¹ · Romina Romaniello² · Alessandra Finisguerra³ · Enza Maria Valente^{2,4} · Sandra Strazzer¹ · Renato Borgatti^{2,5} · Cosimo Urgesi^{1,6}

Accepted: 8 April 2024 / Published online: 3 May 2024 © The Author(s) 2024

Abstract

This study tested the feasibility and efficacy of a Virtual Reality (VR) social prediction training (VR-Spirit) specifically designed for patients with congenital cerebellar malformation. The study is a randomised controlled trial in which 28 cerebellar patients aged 7–25 yo were randomly allocated to the VR-Spirit or to a control intervention in VR. The VR-Spirit required participants to compete with different avatars in scenarios that prompted them to form predictions about avatars' intentions. The control intervention consisted of games currently adopted for motor rehabilitation. Social prediction as well as secondary neuropsychological and behavioural outcomes were assessed at the beginning (T0), at the end (T2) and after 2 months (T3). The experimental group showed a significant increase, compared to the control participants, in social prediction assessed through a VR task. Moreover, at least at T3, the VR-Spirit enhanced the use of contextual predictions in a computer-based action prediction task. Importantly, these effects were generalized to secondary neuropsychological outcomes, specifically theory of mind and, only at T2, inhibition. No differences between the interventions were detected on emotional-behavioural problems. Lastly, both interventions showed high feasibility and acceptability. These findings confirm that it is possible to develop condition-specific rehabilitative training on the basis of neurocognitive functions impaired in case of congenital malformation. The VR-Spirit demonstrated to generalize its effects to theory of mind abilities, and it might be thus extended to other neurodevelopmental disorders that present social preception deficits and alterations of predictive processing. **Trial registration**: ISRCTN, ID: ISRCTN22332873. Retrospectively registered on 12 March 2018.

Keywords Social prediction \cdot Cerebellum \cdot Virtual reality \cdot Rehabilitation \cdot Theory of mind \cdot Cerebellar cognitive affective syndrome—CCAS

Niccolò Butti niccolo.butti@lanostrafamiglia.it

- ¹ Scientific Institute, IRCCS E. Medea, Bosisio Parini, Lecco, Italy
- ² IRCCS Mondino Foundation, Pavia, Italy
- ³ Scientific Institute, IRCCS E. Medea, Pasian di Prato, Udine, Italy
- ⁴ Department of Molecular Medicine, University of Pavia, Pavia, Italy
- ⁵ Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy
- ⁶ Laboratory of Cognitive Neuroscience, Department of Languages and Literatures, Communication, Education and Society, University of Udine, Udine, Italy

Introduction

There is extensive evidence that cerebellar alterations result not only in motor impairments, but also in a complex constellation of neuropsychological deficits and abnormalities in behaviour and affect regulation, referred to as the Cerebellar Cognitive Affective Syndrome (CCAS) (Manto & Mariën, 2015; Schmahmann & Sherman, 1998; Tavano et al., 2007). Such a variety of cognitive, social, and behavioural symptoms, characterized by an augmented or diminished response to external and internal stimuli (Schmahmann, 2004), is thought to depend on the prominent role of the cerebellum in predictive processing (Schmahmann, 2019; Sokolov et al., 2017). Through its main functions of providing predictive internal models and error-signalling (Wolpert et al., 1998), the cerebellum would act as a general regulator and monitor of the activity of different brain networks, from sensorimotor processing to language, regardless of the type of information being fed (Peterburs & Desmond, 2016; Siman-Tov et al., 2019). In detail, the cerebellum would extract contextual regularities and provide forward models to cortical networks, with the ultimate outcome to allow overcoming uncertainty and selecting the best matching between external information and internal expectations (Ishikawa et al., 2016; Leggio & Molinari, 2015).

This predictive function might be particularly critical in social contexts, which are inherently dynamic and require continuous adaptation of one's expectations about others' mental states and behaviours (Brown & Brüne, 2012; Stoodley & Tsai, 2021; Van Overwalle et al., 2019a, 2019b). Notably, recent evidence has documented that alterations of cerebellar functions, either in clinical conditions or after non-invasive brain stimulation, affect the processing of social stimuli more than the processing of inanimate objects (Cattaneo et al., 2012; Ma et al., 2021; Oldrati et al., 2021; Van Overwalle et al., 2019). Evidence of structural loops between the lateral posterior cerebellum and the superior temporal sulcus suggests that the cerebellum contributes to understanding biological movements (Sokolov et al., 2010, 2014). The vermis is linked with limbic areas and circuits involved in affect processing and regulation (Adamaszek et al., 2017). The reciprocal connections between the posterior cerebellum and cortical areas engaged by the so-called mentalizing network (Van Overwalle & Mariën, 2016; Van Overwalle et al., 2020) suggest that the cerebellum might apply its predictive computation also on complex social information. Accordingly, an increasing number of studies have investigated social perception in cerebellar patients, reporting deficits at multiple levels of social processing, from understanding biological movement (Abdelgabar et al., 2019) to affect recognition (Adamaszek et al., 2015; Hoche et al., 2016) and theory of mind abilities (Clausi et al., 2019; Sokolovsky et al., 2010). Structural and functional abnormalities of the cerebellum have been documented in clinical populations characterised by social impairments, such as Autism Spectrum Disorder (ASD) (Olivito et al., 2017; Van Overwalle et al., 2020). Early injuries and congenital malformations of the cerebellum can interfere also with the development of the connected cortical nodes of social cognition through mechanisms of developmental diaschisis (Stoodley & Limperopoulos, 2016).

Notably, research on rehabilitative interventions to treat CCAS is scarce (Argyropoulos et al., 2020; Gagliardi et al., 2015; Ito et al., 2010; Ruffieux et al., 2017). To the best of our knowledge, literature does not report interventions specifically designed to target social perception deficits in cerebellar patients. To fill this gap, the Virtual Reality Social Prediction Improvement and Rehabilitation Intensive Training (VR-Spirit) was developed. This intensive training program delivered in virtual reality (VR) aims to enhance social prediction in children, adolescents, and young adults with congenital cerebellar malformations (CCM). Briefly, this training exploits a design based on probabilistic learning of social events in motivating and child-friendly environments. During the training, patients are asked to reach a desired target; importantly, to achieve success in their actions, they must predict the behaviour of other competing virtual avatars. To verify the efficacy and specificity of the VR-Spirit, as well as its feasibility and acceptability, the effects of the VR-Spirit on outcome measures of social prediction ability and neuropsychological functions were compared to the effects induced by a VR motor rehabilitation training, in a randomised controlled trial. The study protocol was previously registered (ISRCTN 22332873) and published (Butti et al., 2020a). A preliminary paper has reported validating evidence on the social effects of the VR-Spirit, highlighting short-term improvements on social prediction (Urgesi et al., 2021). Here, after trial conclusion, we reported the results on the efficacy and specificity of the VR-Spirit on social prediction and its generalisation to secondary neuropsychological outcomes, particularly social perception skills, as well as its feasibility and acceptability. In particular, we hypothesized that the VR-Spirit, compared to the control VR-based motor training, should: (1) enhance social prediction ability resulting in better understanding of other people's intentions and behaviours; (2) improve cognitive performance in specific domains, particularly in social perception; and (3) be feasible and well tolerated in children, adolescents and young adults with CCM.

Methods

Trial Design

The study applied a single-centre, randomised, active controlled trial design, in which patients were allocated to the VR-Spirit or to the control VR-assisted motor rehabilitation intervention. A stratified permuted block randomisation procedure was used for allocation, considering age and cognitive level (i.e., the most recently available Full-Scale Intelligent Quotient, FSIQ), as stratification factors. Participants were assigned to one of six blocks according to the two stratification variables, considering two levels of age $(7.0-12.9, \ge 13.0 \text{ yo})$ and three levels of cognitive functioning (46-60, 61-80, > 80 FSIQ). A sample size of 42 participants was estimated through an a-priori power analysis to detect a large (d=0.8) between-group difference (independent sample t test, two-tailed) with a power of 0.80 and an alpha level set at 0.05 (Butti et al., 2020a). Thus, it was planned to assign up to eight patients per block to achieve and overcome the pre-determined sample size. Within each block, participants were allocated to the VR-Spirit or the active control group on the basis of pre-established permuted sequences, which were generated before enrolling the first participant. Full details of the trial design and randomization procedure have been previously registered and published (Butti et al., 2020a).

Participants

Patients were enrolled at the Child Neuropsychiatry and Neurorehabilitation Unit of the Scientific Institute, IRCCS E. Medea (Bosisio Parini, Italy), to which they were referred for routine medical and rehabilitative checks. Inclusion criteria were: (i) diagnosis of non-progressive congenital cerebellar malformation; (ii) age ranging from 7 to 25 yo; (iii) absence of severe cognitive impairment (FSIQ > 45), as shown by the most recent (<6 months) available FSIQ at the age-corresponding Wechsler scale; iv) absence of severe motor and visual impairments that could interfere with task execution; (v) not being simultaneously involved or having been involved in the previous 6 months in any formal VR intervention study. To broaden the generalizability of the training, we included both participants in developmental age and young adults, so that the rehabilitative training could take advantage of still ongoing neurodevelopmental processes and of the related plasticity (de Lange et al., 2017). Accordingly, regional cerebellar volumes were found to be differently associated with sensorimotor and cognitive task performance in young and older adults (Bernard & Seidler, 2013). Please note that criterion iii was initially FSIQ > 60and it was modified in December 2018 in order to include patients with moderate cognitive delay ($45 < FSIQ \le 55$). This choice was made since preliminary outcomes indicated good feasibility and acceptability of the trial independently by cognitive level. The repetitiveness of the task, the constant presence of a therapist to remember instructions, and the ease of movement in the virtual environment likely helped participants understand and perform the training. However, FSIQ>45 was maintained as an inclusion criterion to ensure patients' adherence to the training and testing procedure. For three participants, the FSIQ was not available due to severe linguistic impairments, so that the IQ was calculated from the perceptual reasoning index at the age-corresponding Wechsler scale.

The first participant was enrolled in February 2018 and the last one in September 2021, with the last follow-up assessment completed in December 2021. The trial was suspended from March to July 2020 due to the Covid-19 pandemic. Although the trial was later resumed and the end date was postponed to January 2022, only five participants were enrolled after July 2020 due to pandemic-related restrictions and safety procedures. These unforeseen events prevented the target sample size of 42 participants from being reached. This target sample size was estimated on the basis of the minimal clinically relevant change expected on the primary outcome measure (for details see Butti et al., 2020a).

An administrative staff member, blind to the trial design, enrolled the 32 participants and organized hospitalisation according to each patient and family's needs and availability. When further screened for eligibility, three patients were excluded from allocation since they did not meet eligibility criteria. In detail, for one patient a diagnosis of progressive ataxia was formulated; for one, the severity of motor impairments prevented task execution; and a third was dismissed due to severe behavioural problems. Moreover, a family declined to participate due to time constraints. A total of 28 patients (21 M; mean age = 13.1 yo) were then assigned to the experimental or control group according to the allocation ratio. Two participants (one per group) did not complete the follow-up assessment due to Covid-19 pandemic, so that data of 26 participants were considered for the follow-up effect (Fig. 1).

A sensitivity analysis conducted with G*Power 3.1 (Faul et al., 2007) showed that the achieved sample size allowed us to detect large effect sizes of at least d=1.1, at T2, and of at least d=1.15, at follow up (T3), with a power of 0.8 and an alpha level set at 0.05.

As our sample did not reach the targeted size, the six blocks of stratified randomisation were unevenly filled (see supplementary Table 1). However, the two groups were balanced in terms of biological sex and stratification variables (i.e., age, FSIQ), confirming successful randomization (Table 1).

Most participants followed a differentiated and/or reduced school programme and were assisted by a special education teacher in accordance with Italian laws (for detailed information of each participant see supplementary tables 2 and 3).

The procedures of this study were approved by the local Ethics Committee and agreed with the principles expressed in the Declaration of Helsinki. Patients and their parents were asked to sign a written informed consent form before starting any procedures.

Interventions

Both the VR-Spirit and control rehabilitation trainings were administered in the Gait Real-time Analysis Interactive Laboratory (GRAIL, Motek, Amsterdam, NL) of the Scientific Institute Medea. The GRAIL system is an integrated platform equipped with a treadmill, an integrated motion-capture system, and a 180° cylindrical projection screen on which the VR scenarios are projected. Both interventions consisted of eight 45-min daily sessions administered over 2 weeks by a GRAIL-patented physical therapist. The therapist placed two reflective markers on the posterior superior iliac spines, which traced patient's movements in the virtual environment.



Fig. 1 CONSORT flow diagram of recruitment

Table 1	Demographic a	nd clinical	information	for the two	groups
---------	---------------	-------------	-------------	-------------	--------

	VR-Spirit	Active control group	Statistical comparison	
				p value
Demographic information				
Biological sex-M:F	11:4	10:3	0.05	0.82
Age—Mean (SD)	13.4 (3.8)	12.8 (5.2)	-0.34	0.74
Clinical information				
IQ—Mean (SD)	72 (26)	67 (21)	-0.56	0.58
Primary cerebellar malformation (N)	Hypoplasia (5), Molar tooth (3), Vermis hypo- plasia (5), Rhomboencephalosynapsis (1), Cerebellar folia dysplasia (1)	Hypoplasia (2), Molar tooth (8), Vermis hypoplasia (3)		
Syndromic/genetic diagnosis (N)	Joubert syndrome (3), ITPR1 gene mutation (1), unknown (11)	Joubert syndrome (8), unknown (5)		

Namely, participants had to move forward or backward to go faster or slower, respectively; they had to shift the pelvis to turn right or left. This movement system limited the impact of gait problems (e.g., ataxia).

The VR-Spirit presented participants with a playground scenario in which they were asked to compete with one of

four avatars in reaching one of three pieces of recreational equipment (a swing, a circular carousel and a rocking carousel). Two avatars were male adolescents and two were female adolescents; each individual avatar was clearly identifiable by hair and t-shirt colours. For each trial of the training, one of the avatars was visibly positioned next to the patient; the avatar then started moving along a 9-m straight-line path and turned into one of the three 3-m long branches directed to one of the recreational pieces. Each avatar's choices of the recreational objects were biased towards pre-established probabilistic associations. In detail, three avatars preferred a specific object in 80% of the trials and moved towards each of the other two objects in 10% of the trials. One avatar moved in a pseudo-random way, according to avatar-object probabilistic associations set at 33%. Each trial lasted 15 s and ended when the avatar reached the chosen object. Participants' velocity was set so that they could pass the avatar only before the crossroad; thereafter the velocity of the participants' movements in the VR-scenario was limited to 2 m/s, being slower than the avatars' movements. This way, participants were forced to anticipate avatars' movements without any perceptual clues about their direction, thus prompting them to form predictive models of the avatarobject associations. Namely, participants had to predict the location where the avatar was walking in order to get to the chosen recreational equipment before him/her. When participants anticipated the avatar in reaching the chosen object, they scored a point in the game and obtained auditory (clapping sound) and visual (activation of the object) reinforcements. The object reached by the avatar was always visible to the participant, in order to provide information on the avatar's preferences also for unsuccessful trials.

Considering 20 trials per avatar, each session consisted of 80 trials in which the pre-established events were presented in a random order. Moreover, four different versions were created according to a counterbalance of the avatar-object associations. Administration of these sessions followed the same order in the 2 weeks but was randomised and counterbalanced between participants. At the end of each session, lasting around 30 min, participants were presented with one of four motor games selected from the GRAIL kit, which were also adopted in the active control training.

Participants assigned to the active control group played a custom-made navigational game, in which they had to conduct a ball out of three mazes, and all the four GRAIL games designed for motor rehabilitation. All games were selected for the absence of social interactions and did not require any form of prediction abilities. The order of administration of these games was kept constant across sessions but counterbalanced between participants (for further details see Butti et al., 2020a).

The training scenario and an example picture of a game from the control motor intervention are reported in Fig. 2.

Assessment and Outcome Measures

All participants underwent assessment of different primary and secondary outcomes at three different time points: before the training (T0), at the end of the training (T2), in a follow-up session 2 months after the end of the training (T3).

Social prediction was set as the primary outcome and assessed through two different paradigms. First, a VR social prediction task was administered at T0 and T2. This task mirrored exactly the design of the VR-Spirit, but adopted a different scenario, namely a sweet stands' environment, and different avatars. This way it was possible to assess improvements of social prediction independently from practice effects. Please note that before starting the first evaluation session, participants were administered a short and effective navigational training to learn how to move within the GRAIL VR environments. In order to avoid repetition of the same events between T0 and T2, two different versions of the task were created by changing the avatar-object associations and their administration was counterbalanced between patients of the same group. A prediction score (VR-prediction score) was computed by assigning one point every time the participant moved according to the avatars' preferred choice (i.e., the 80% avatar-object association). Namely, when the participants reached the object preferred by the avatars before him/her, we considered the trial as a correct prediction. Conversely, a wrong prediction corresponded to the choice of one of the two seldom preferred sweet stands (i.e., the 10% avatar-object association). The trials with the avatar moving in a pseudo-random way (i.e., the 33% avatarobject association) were not considered in the VR-prediction score. This measure ensured ruling out noise due to the use of random strategies, and it was found to be sensitive to changes across time (Urgesi et al., 2021).

Then, social prediction was also assessed at T0, T2 and T3 through a validated, computer-based action prediction task previously adopted in paediatric clinical populations (Amoruso et al., 2019; Butti et al., 2020b). This task consisted of short videos displaying an actor preforming reaching-to-grasp movements towards one of two possible objects (i.e., a glass or an apple) with the intention to drink/eat it or to offer it. The actions were embedded in a scenario containing different contextual cues, namely, coloured tablecloths or coloured plates. Participants were required to recognise which was actor's intention (i.e., to drink/eat vs. to offer). In the familiarization phase, videos stopped when the actor's hand reached the object, so that participants could rely on full kinematics information to understand the action unfolding. Crucially, in this phase, the association between contextual cues and actor's intentions was implicitly biased with pre-established probability of co-occurrence (10%, 40%, 60%, 90%), in order to prompt the formation of contextual predictive models (i.e., priors). During the testing phase, the same videos were presented but occluded well before the hand-object contact, so that participants should rely on the contextual priors to overcome kinematics ambiguity and to understand actor's intention. For each participant, a



Fig. 2 The training scenario and a control motor game. **A** VR-SPIRIT intervention: the playground scenario with a circular carousel (A.1), a rocking carousel (A.2) and a swing (A.3) at the end of the path. **B** One of the games included in the active control intervention: the soccer game

standardised beta coefficient (herein after beta index) was calculated across the trials using a regression analysis with probability and accuracy as the independent and dependent variables, respectively. This index is a proxy of the steepness of the regression line between the four levels of probability, and it thus represents the strength of contextual prior use. Importantly, the beta index was found to be associated with social perception skills in cerebellar patients (Butti et al., 2020b; Urgesi et al., 2021). Accordingly, an increase of moderate effect size in the beta index (0.13 mean change) was considered as a clinically relevant improvement in social prediction, with likely effect on social perception (Butti et al., 2020a). The accuracy in the familiarization phase of the same action prediction task was considered as a secondary outcome measure of implicit learning.



Fig. 3 Timeline of the trial and assessments

Moreover, selected subtests of the Italian version of the NEPSY-II testing battery (Korkman et al., 2007) were administered across the three time points (i.e., T0, T2, T3) to assess different secondary neuropsychological outcomes. The NEPSY-II is one of the most world-wide adopted batteries for neuropsychological testing in developmental age, and it has been used to investigate the neuropsychological functioning of different clinical populations (Ferrari et al., 2022; Narzisi et al., 2013; Rasmussen et al., 2013). In detail, the following subtests were administered: theory of mind and affect recognition in the social perception domain; visual attention and inhibition in the attention and executive functions domain; picture and geometric puzzles for visuospatial processing; memory for designs for visuospatial memory; fingertip tapping for sensorimotor functions. Raw scores of each subtest were transformed into scaled score (mean = 10, SD = 3) according to the normative standardization tables (Urgesi et al., 2011). Average scaled scores (composite score) were calculated for the subtests providing multiple scaled scores for the same outcome, namely: naming, inhibition and switching for inhibition; immediate and delayed recall, for memory for designs; picture and geometric puzzles, for visuospatial processing.

Lastly, parents were asked to fill the Child Behavior Checklist (CBCL) 6–18 before the training (T0) and at the follow-up session (T3). The CBCL is a widely-adopted questionnaire regarding the presence of emotional-behavioural problems in children and adolescents (Achenbach, 2011; Frigerio et al., 2004), providing standardized T-scores (mean = 50, SD = 10) on eight syndromic scales. Two participants were over 18 years and thus they were not administered the questionnaire, while parents of one participant assigned to the experimental group did not fulfil the questionnaire at the follow-up assessment.

A psychologist with expertise in neuropsychological testing, who was not blind to allocation, was responsible for all data collection. A resume of the assessments at the different timepoints is reported in Fig. 3.

Feasibility and Acceptability

Feasibility and acceptability of the VR-Spirit for the target population were assessed. For the former aspect, the number of dropouts, that is the number of participants who declined to complete the training, and the number of sessions completed for each participant were collected. To assess acceptability, self-report and parent-report questionnaires were administered at the end of either training (T2). The questionnaires were adapted from a previous study on cognitive rehabilitation in paediatric patients with neurological disorders (Corti et al., 2018). The self-report questionnaire assessed recommendation level, required effort, pleasantness, comprehensibility of the intervention, difficulties to start and to continue the training. The parent-report version assessed recommendation level and satisfaction, as well as perception of their child's required effort, difficulties to start and to continue the training. Answers were collected on a five-point Likert-like scale, from "not at all" = 1 to "a lot'' = 5. The same Likert scale was used to assess negative (e.g., difficulties to start and continue the training) and positive (e.g., pleasantness, satisfaction) features of the training, thus limiting the impact of a likely tendency to affirmative answers in the self-report questionnaire due to the cognitive impairments.

Data Handling and Statistical Analysis

According to study protocol, all participants who had completed the pre- and post-treatment evaluation sessions were considered in the outcome analysis. Given the limited sample size and observation points, no missing data was imputed (Dziura et al., 2013). Since two patients did not participate in the follow-up session, the analysis on the follow-up effect were performed on 26 participants. Primary and secondary outcomes of the two patient groups were inspected through descriptive statistics, and independent sample t tests were used to verify group differences at baseline (T0). For each outcome measure, the arithmetic differences between T2 and T0 (herein after *training effect delta*) as well as between T3 and T0 (hereinafter *follow-up effect delta*) were calculated (with higher values corresponding to stronger effects). For the primary outcome measure, the delta values of the two groups were compared by means of independent sample ttests, adopting the Bonferroni correction to control for multiple comparisons. For the action prediction task, these delta measures were also inspected individually to check whether and how many patients in each group showed an increase equal to or greater than the pre-determined minimal clinically important change (0.13). Moreover, to further investigate between-group differences after the training, exploratory t tests were conducted on the beta index values at T2 and T3.

As concerns the secondary outcomes, independent sample t tests were executed using the delta values of accuracy in the action prediction familiarization phase as dependent variable, while separate multivariate analysis of variance (MANOVA) models were run for the neuropsychological and the behavioural outcomes, separately for the training effect delta and for the follow-up effect delta, with group as categorical factor. In detail, for the neuropsychological outcomes, the delta measures of theory of mind, affect recognition, visual attention, inhibition (composite score), memory for designs (composite score), visuospatial processing (composite score) and fingertip tapping, were entered as dependent variables. Since three participants were not able to execute the inhibition and fingertip tapping subtests due to linguistic and fine-motor impairments, consequent MANOVAs were performed excluding these subtests. For the behavioural outcomes, we inserted as dependent variables in a MANOVA design, with group as a categorical factor, the follow-up delta score of the T-scores obtained at the syndromic scales of the CBCL 6-18, namely anxiety/depressive symptoms, social withdrawal, somatic complaints, social problems, thought problems, attention problems, rule-breaking behaviour. and aggressive behaviour. For each MANOVA model, univariate results were also calculated to observe differences in specific outcomes. Lastly, each answer at the acceptability questionnaires was compared between groups through Mann–Whitney U tests.

All analyses were performed with the Statistica software version 8.0 (Statsoft, Tulsa, OK). Effect sizes were reported with their 95% confidence interval (CI). Significance threshold was set at p = 0.05 for all statistical tests. Effect sizes were estimated and reported as Cohen's *d* for pairwise comparisons, adopting conventional cut-off of 0.2, 0.5, and 0.8 for small, medium, and large effect sizes, respectively, and as partial eta squared (n_p^2) for ANOVA designs, adopting conventional cut-off of 0.14 for small, medium, and large effect sizes, respectively.

Results

Descriptive statistics for the primary and secondary outcomes at the three timepoints are reported in Table 2.

The analyses indicated that the two groups were comparable on all assessed outcomes at baseline (all t < 0.98, all p > 0.33).

Primary Outcomes

The analysis revealed a significantly greater increase of the VR prediction score in the experimental compared to the control group, with large effect size ($t_{26} = -4.85$; p < 0.001; Cohen's d = 1.84, 95% CI 0.91–2.66).

The between-group difference at the beta index for the training effect delta was small to medium and nonsignificant (t_{26} =0.96; p=0.344; Cohen's d=0.37, 95% CI-0.39-1.10), while a medium-to-large effect of greater follow-up delta in the VR-Spirit than in the control group was marginally significant (t_{24} =1.80; p=0.083; Cohen's d=0.71, 95% CI-0.11-1.48). At individual level, only four participants of the control group showed an increase above the clinically significant value (0.13), whereas, for the VR-Spirit group, 7 and 11 patients presented such an improvement at T2 and T3, respectively. Results on the primary outcome measures are reported in Fig. 4.

Exploratory *t* tests confirmed that the between-group difference in the beta index at T2 was medium and non-significant (t_{26} = 1.41; *p* = 0.169; Cohen's *d* = 0.54, 95% CI - 0.23-1.28), but at T3 the beta was greater in the VR-Spirit than in the control group, with a large effect size (t_{24} = 2.25; *p* = 0.034; Cohen's *d* = 0.88, 95% CI 0.05-1.66) (Fig. 5).

Secondary Outcomes

No differences were detected on delta accuracy in the action prediction familiarization phase (all t < |0.62|; all p > 0.54), indicating that the VR-Spirit did not affect action perception in general and the two groups were similarly exposed to the embedded probabilistic context-action associations. Thus, the VR-Spirit effects on the testing phase (our primary outcome measure) were due to an enhanced use of previously learned contextual expectations.

As concerns the neuropsychological outcomes, the two groups were overall comparable on the training effect delta $(F_{7,17} = 1.83; p = 0.145; n_p^2 = 0.43, 95\%$ CI 0.00–0.50), indicating that both groups increased their general performance at the NEPSY-II. However, significant differences were detected for the inhibition $(F_{1,23} = 4.35; p = 0.048; n_p^2 = 0.16, 95\%$ CI 0.00–0.40) and, marginally, for the theory of mind $(F_{1,23} = 4.27; p = 0.05; n_p^2 = 0.16, 95\%$ CI 0.00–0.40)

Table 2 Descriptive statistics (mean and SD) for the primary and secondary outcomes at the three timepoints in the two groups

	VR-Spirit			Active control group		
	TO	T2	T3	TO	T2	T3
Primary outcome measure						
Social prediction						
VR-Prediction score	32.9 (8.4)	51.3 (5.2)		34.4 (10.5)	37.5 (8.1)	
Beta index	-0.03 (0.71)	0.25 (0.51)	0.26 (0.52)	-0.08 (0.53)	-0.05 (0.59)	-0.21 (0.54)
Secondary outcome measure						
Implicit learning						
Action prediction familiarization phase—accuracy	93%	94%	93%	91%	91%	92%
NEPSY-II scaled scores						
Social perception						
Theory of mind	4.7 (3.8)	7.3 (3.8)	7.7 (4.3)	6.2 (4.5)	6.8 (3.9)	5.5 (4.3)
Affect recognition	5.8 (4.1)	7.3 (4.5)	7.6 (3.6)	5.4 (3.1)	6.5 (3.3)	5.5 (3.6)
Visual attention						
Visual attention	4.3 (4.0)	5.7 (5.1)	7.0 (5.6)	4.5 (3.7)	6.3 (4.0)	5.8 (4.8)
Executive functions						
Inhibition	5.6 (3.7)	7.9 (3.4)	8.7 (3.6)	4.8 (2.4)	5.8 (3.3)	6.3 (3.4)
Visuospatial processing						
Picture and geometric puzzles	5.5 (4.2)	7.0 (4.9)	7.7 (5.2)	5.6 (3.2)	7.3 (4.1)	7.3 (4.2)
Visuospatial memory						
Memory for drawings	3.4 (3.7)	6.0 (4.8)	6.4 (5.1)	3.5 (3.9)	6.6 (4.6)	6.5 (4.6)
Sensorimotor functions						
Fingertip tapping	8.2 (4.6)	8.5 (4.7)	9.4 (4.5)	8.5 (3.5)	8.3 (3.5)	8.7 (3.5)
CBCL T-scores						
Emotional-behavioural problems						
Anxious/Depressed	60 (6)		57 (8)	59 (10)		59 (8)
Social withdrawal	58 (9)		61 (11)	58 (7)		60 (7)
Somatic complaints	59 (9)		56 (8)	58 (8)		57 (7)
Social problems	63 (10)		61 (9)	63 (6)		63 (7)
Thought problems	59 (10)		57 (10)	59 (10)		57 (8)
Attention problems	60 (11)		60 (10)	60 (5)		60 (6)
Rule-breaking behaviours	55 (7)		55 (7)	58 (9)		60 (8)
Aggressive behaviours	57 (9)		54 (7)	54 (18)		60 (11)

subtests, both indicating higher increases after the VR-Spirit compared to the active control training. All other comparisons were non-significant (all F < 1.53; all p > 0.22). In the consecutive model excluding inhibition and fingertip tapping, the only significant difference emerged for theory of mind ($F_{1,26}$ =4.80; p=0.037; n^2_p =0.16, 95% CI 0.00–0.39).

For the follow-up delta effect, the analysis indicated a significant group effect ($F_{7,15} = 4.24$; p = 0.009; $n_p^2 = 0.66$, 95% CI 0.08–0.72), which was specifically detected for theory of mind ($F_{1,21} = 14.89$; p < 0.001; $n_p^2 = 0.41$, 95% CI 0.09–0.62), whereas all other comparisons were non-significant (all F < 4.08; all p > 0.056). When excluding inhibition and fingertip tapping, the difference for theory of mind was still the only significant univariate result ($F_{1,24} = 14.83$; p < 0.001; $n_p^2 = 0.38$, 95% CI 0.09–0.59) (Fig. 6).

With regard to the CBCL results, no significant differences emerged (all F < 3.20; all p > 0.088), suggesting that neither the experimental nor the control intervention had an impact on emotional-behavioural problems.

Feasibility and Acceptability

As concerns feasibility, none interrupted participating in either intervention, and only one participant assigned to the VR-Spirit did not complete a session due to a sickness episode not related to the training. Moreover, the two interventions were considered similarly acceptable by both patients and their parents (all Z adjusted <|1.35|, all p > 0.17), with low scores (median for both the groups = 1) in difficulties to start, difficulties to continue the training, and in required



Fig. 4 Boxplot of the training effect delta on the VR-prediction score and training and follow-up effect delta on the beta index in the two groups. The dotted grey line represents the estimated minimal significant improvement on the beta index. Asterisk indicates significant comparison



Fig. 5 Boxplot of the beta index at the three time points in the two groups. Asterisk indicates significant comparison



Fig. 6 Training and follow-up effect Delta on the neuropsychological outcomes in the two groups. Asterisk lines represent significant comparisons



Fig. 7 Patients and parents' ratings of acceptability of the two interventions

effort, while high scores (median for both the groups = 5) were recorded for pleasantness, comprehensibility, recommendation level, and satisfaction (Fig. 7).

Discussion

This randomised, active controlled trial tested the efficacy, feasibility and acceptability of the VR-Spirit, a VR intervention specifically designed to target social prediction abilities in patients with congenital cerebellar malformations, compared to a VR motor rehabilitation training not involving social agents. The experimental group showed a significant increase, compared to the control participants, on the prediction score of a VR social prediction task, with very large effect size. Moreover, at least at the follow-up, the VR-Spirit yielded an improvement in the use of contextual predictions in a computer-based action prediction task, with a large effect size; at individual level, a wide part of the experimental group showed an increase higher than the minimal clinically relevant change. Importantly, these effects were generalized, with a large effect size, to secondary neuropsychological outcomes, specifically theory of mind and, only for the training effect delta, inhibition. No differences between the interventions were detected on parent-reported emotional-behavioural problems. Lastly, both interventions showed high feasibility and acceptability, confirming the potential of VR as a cognitive rehabilitation tool. These results converge in sustaining future applications of VR-Spirit in different contexts and with other neurodevelopmental disorders affecting social perception.

VR-Spirit demonstrated to be effective in boosting social prediction of the avatars' intentions, even when adopting a different scenario. Indeed, participants who completed VR-Spirit learned to predict the motor intentions of another person when choosing which direction to move in a social context. The improvement in social prediction was detectable also, at least partially, in the action prediction task. This task was previously used to investigate implicit predictive processing of social information in healthy and clinical populations (Amoruso et al., 2019; Bianco et al., 2020), and it showed to be sensitive to alterations of cerebellar functionality due to either acquired or congenital disorders (Butti et al., 2020b; Urgesi et al., 2021), as well as to the use of non-invasive brain stimulation on the cerebellum (Oldrati et al., 2021). It is worth noting that the analysis on the delta of the beta index yielded only marginally significant results, probably due to the limited sample size and consequent lack of statistical power, since we failed to reach the desired sample size according to the *a-priori* power analysis, which was conducted on the basis of this outcome (Butti et al., 2020a). Nevertheless, when directly comparing the beta index (and not the delta) at the follow-up assessment, the experimental group showed a greater reliance on contextual priors for action prediction than the control group, and this difference had a large effect size. As mentioned above, the non-significant result at the post-training assessment may depend on the limited sample size. Also, the repetition of the action prediction task in a short time may have yielded a partial improvement even in the control group, blurring betweengroup differences. Data reported in Table 2 and in Fig. 4 and 5 suggest that the VR-Spirit improved the beta index at T2, even though this effect did not reach statistical significance. Furthermore, at individual level, both after training and follow-up, an improvement of the beta index higher than the expected minimal clinically relevant change was detected in a wide part of the sample assigned to the VR-Spirit, but not in those assigned to the control training. Namely, participants who attended the VR-Spirit improved their ability in understanding the motor intentions of another person during the observation of simple every-day life actions.

Even more importantly, the clinical impact of the training was further corroborated by the effects on social perception, specifically on theory of mind. In fact, after the training and even at the follow-up assessment, the experimental group showed a significantly greater improvement in the theory of mind subtest compared to the control group. This means that participants who completed the VR-Spirit showed enhancements in understanding intentions, beliefs, and emotions of another person trough verbal descriptions and pictures representing different social contexts. This result ruled out the confounding effect due to the repeated administration of the same tests, confirming the specificity of the VR-Spirit in targeting the social perception domain.

Extending the preliminary evidence reported in a previous article (Urgesi et al., 2021), these findings suggest that training participants to learn explicit context-behaviour associations may improve social predictions in patients with cerebellar malformations at multiple levels, from the lowerlevel, implicit processing involved in action perception, to the higher-level, conceptual inferences supported by the cerebellum during the attribution of others' mental states (Clausi et al., 2019). For the former, it is important to consider the non-significant difference between groups in the familiarisation phase of the action prediction task. Indeed, both groups were equally exposed to the implicit probabilistic associations, but only the VR-Spirit improved the use of contextual priors during the testing phase, where it was crucial to rely on predictive priors to overcome kinematics ambiguity (Frith & Frith, 2012). In this vein, the experimental training might have enhanced the matching between external, sensory information, in this case kinematics, and the internal, contextual representation formed during the familiarization, resulting into a better and timely prediction of other's motor intentions (Molinari & Masciullo, 2019). At a higher level of social processing, the use of explicit reflective processes to understand contextualized behavioural regularities during the VR task might have trained participants to form conceptual models about others' mental states, leading to better attributions of intentions, beliefs, and emotions in the theory of mind NEPSY-II subtest (Koster-Hale & Saxe, 2013).

The hypothesis that the VR-Spirit might be effective through the enhancement of reflective, metacognitive processes, is corroborated by the specific improvement shown by the experimental group, at least at the after-training assessment, in the inhibition subtest. Indeed, the VR-Spirit required participants to inhibit impulsive responses, which would result in random strategies like merely following the avatar's last choice. Participants were exercised to recognize the avatars, to retrieve information on their preferences, to select the appropriate path, and to switch predictive models between avatars and different sessions. These accounts on how VR-Spirit improved social prediction and theory of mind abilities are in accordance with recent recommendations on rehabilitation of CCAS, which indicate that prompting patients to be aware of their deficits might enhance explicit reflective processes that act as an "external cerebellum", compensating alterations of automatic, implicit cerebellar functions (Argyropoulos et al., 2020; van Dun et al., 2018).

Neither the VR-Spirit nor the active control intervention affected the presence of emotional-behavioural problems. Thus, the effects of the VR-Spirit generalized to theory of mind, but not to emotional-behavioural disorders. On one hand, an increase in neuropsychological skills does not imply the same improvement in the behavioural domain, as they represent different dimensions of psychological functioning. For instance, a child may have adequate social perception skills but appear as withdrawn or anxious, or vice versa. In fact, better abilities to process contextual cues in social contexts may increase behavioural problems, as the emergence and persistence of problem behaviours has communicative functions that are often mediated by social cues (Newcomb & Hagopian, 2018). On the other hand, despite their predominant use as behavioural outcomes, parentreport measures, as the one adopted here, have been mainly developed for healthy populations, and were not designed to detect changes over time, thus showing limited reliability and responsiveness to change in neurodevelopmental disorders (Budimirovic et al., 2017; Hanratty et al., 2015).

In terms of feasibility, only a family declined to participate before group allocation, and only one session of the VR-Spirit was interrupted due to a sickness episode unrelated to the training. Moreover, both participants and their parents reported optimal ratings at the acceptability questionnaire. These results sustain that, although it is based on task repetition, the VR-Spirit is feasible and acceptable for patients with a wide range of cognitive functions, from moderate intellectual disability to (above) average intellectual functioning. Notably, the two interventions were comparable on the acceptability scores, confirming the potential of VR as a cognitive rehabilitation tool in populations with neurodevelopmental disorders (Nossa et al., 2022; Tieri et al., 2018). The high ratings attributed by parents might depend on the lack of specific interventions for children and young adults with congenital cerebellar malformations (Argyropoulos et al., 2020; Romaniello et al., 2022), so that they were highly motivated and enthusiastic to have their son or daughter participate in either VR training. Also, the high ratings in recommendation level and satisfaction provide preliminary

indications on the social validity of the VR-Spirit (Strain et al., 2012; Wolf, 1978). Moreover, the comparable ratings on the acceptability questionnaires, along with the fact that both trainings were delivered in the same VR laboratory, exclude that the differences between groups in the primary and secondary outcomes may be due to a higher placebo effect for the VR-Spirit or to motivational aspects.

Limitations must be acknowledged when discussing results of the study. First, the spread of the Covid-19 pandemic prevented the pre-estimated sample size from being reached, undermining the analysis power for the primary outcome. Nevertheless, the analyses, which were previously defined and published (Butti et al., 2020a), yielded convergent results on the efficacy of the VR-Spirit. Moreover, despite a high variability of intellectual functioning in our sample, representative of populations with cerebellar malformations (Bulgheroni et al., 2016; Summers et al., 2017; Tavano et al., 2007), the limited sample size prevented testing differential effects of the training in relation to cognitive level. Moreover, even though all participants presented with cerebellar malformations, diverse patterns of malformations (e.g., the molar tooth sign associated with Joubert syndrome) might differentially affect the neuropsychological functioning and thus social prediction abilities (Butti et al., 2023; Tavano & Borgatti, 2010), influencing also the effects of the rehabilitative training. Lastly, the neuropsychologist conducting the assessment was not blind to the intervention; however, group allocation was determined on the basis of the order of hospital admission, which was organized by an administrative staff who was blind to the study aims and procedures.

In conclusion, these findings confirm that it is possible to develop condition-specific rehabilitative trainings on the basis of the predictive processing exerted by the cerebellum (Bhanpuri et al., 2014). The VR-Spirit demonstrated to generalize its effects to theory of mind abilities, and it might be thus extended to other neurodevelopmental disorders that present social perception deficits and alterations of predictive processing, such as ASD (Amoruso et al., 2019; Pellicano & Burr, 2012) and Williams Syndrome (Sparaci et al., 2014). By boosting the ability to predict other's motor intentions in a virtual environment, the VR-Spirit may be effective in improving navigation skills, which are often impaired in these neurodevelopmental disorders and in conditions of early brain damage (Broadbent et al., 2014; Nossa et al., 2022; Pavlova et al., 2007). Lastly, the Covid-19 pandemic has urged the medical and scientific community to develop new strategies and tools for remote assessment and rehabilitation. Tele-rehabilitation indeed provides an opportunity to keep fragile individuals safe, as well as to improve outreach and ensure compliance with a rehabilitation plan. This is particularly true for those patients that require long-term care and follow-up, as for cerebellar patients. Although the need of further research on this topic, the application of VR may offer new and promising possibilities for tele-rehabilitation of social skills in populations with intellectual and developmental disabilities (Montoya-Rodríguez et al., 2023). The use of VR for supporting social skills in neurodevelopmental disorders may also shorten acquisition times than traditional interventions (Frolli et al., 2022). Considering VRbased technologies, immersive VR viewers have proven to be able to foster more efficient functional performance and cognitive improvement than non-immersive VR (Chatterjee et al., 2022; Cikajlo & Peterlin Potisk, 2019). Notably, these technologies are affordable and more and more available in everyday life because of the spreading of immersive VR for gaming. Accordingly, the VR-Spirit has been recently adapted to commercial VR-viewers, becoming accessible for a larger number of patients, and it will be tested for telerehabilitation (Malerba et al., 2023).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10803-024-06349-8.

Acknowledgments The authors are grateful to all participants and their parents for taking part into this study. The authors would also thank Daniele Panzeri and Claudio Corbetta for their help in data collection.

Author Contributions Niccolò Butti: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing—Original draft, Visualization. Emilia Biffi: Conceptualization, Methodology, Software, Visualization, Resources, Writing—Review and Editing, Funding acquisition. Romina Romaniello: Conceptualization, Resources, Writing—Review and Editing, Supervision. Alessandra Finisguerra: Resources, Writing—Review and Editing, Visualization. Enza Maria Valente: Resources, Writing—Review and Editing. Sandra Strazzer: Resources, Writing—Review and Editing. Renato Borgatti: Conceptualization, Methodology, Writing—Review and Editing, Resources, Supervision, Project administration. Cosimo Urgesi: Conceptualization, Methodology, Software, Formal analysis, Resources, Writing—Original draft, Supervision, Project administration, Funding acquisition.

Funding This work was supported by grants from the Italian Ministry of Health (Ricerca Finalizzata 2016: GR-2016-02363640 to CU; Ricerca Corrente 2023–2024—Scientific Institute, IRCCS E. Medea) and from the Italian Ministry of University and Research (PRIN 2020; Prot. 20203LT7H3 to RB).

Data Availability The study protocol and summary of basic results are registered and published: https://www.isrctn.com/ISRCTN22332873. Data analysed in this study are available through a public repository at the following link: https://zenodo.org/record/8192485.

Declarations

Competing interests The authors report there are no competing interests to declare.

Ethical Approval The study procedures were in accordance with the Declaration of Helsinki and approved by the ethical committee of the IRCCS Medea (Prot. N.17/18).

Informed Consent Before starting any procedure, participants or their parents/legal guardians were informed about aims and methods of the study and asked to sign informed consent.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Abdelgabar, A. R., Suttrup, J., Broersen, R., Bhandari, R., Picard, S., Keysers, C., De Zeeuw, C. I., & Gazzola, V. (2019). Action perception recruits the cerebellum and is impaired in spinocerebellar ataxia patients. *Brain*, 142(12), 3791–3805. https://doi.org/10. 1093/brain/awz337
- Achenbach, T. M. (2011). Child Behavior Checklist. Encyclopedia of clinical neuropsychology (pp. 546–552). New York: Springer. https://doi.org/10.1007/978-0-387-79948-3_1529
- Adamaszek, M., Kirkby, K. C., Dagata, F., Olbrich, S., Langner, S., Steele, C., Sehm, B., Busse, S., Kessler, C., & Hamm, A. (2015). Neural correlates of impaired emotional face recognition in cerebellar lesions. *Brain Research*, *1613*, 1–12. https://doi.org/10. 1016/j.brainres.2015.01.027
- Adamaszek, M., D'Agata, F., Ferrucci, R., Habas, C., Keulen, S., Kirkby, K. C., Leggio, M., Mariën, P., Molinari, M., Moulton, E., Orsi, L., Van Overwalle, F., Papadelis, C., Priori, A., Sacchetti, B., Schutter, D. J., Styliadis, C., & Verhoeven, J. (2017). Consensus paper: Cerebellum and emotion. *The Cerebellum*, 16(2), 552–576. https://doi.org/10.1007/s12311-016-0815-8
- Amoruso, L., Narzisi, A., Pinzino, M., Finisguerra, A., Billeci, L., Calderoni, S., Fabbro, F., Muratori, F., Volzone, A., & Urgesi, C. (2019). Contextual priors do not modulate action prediction in children with autism. *Proceedings of the Royal Society b: Biological Sciences*, 286(1908), 20191319. https://doi.org/10.1098/ rspb.2019.1319
- Argyropoulos, G. P. D., van Dun, K., Adamaszek, M., Leggio, M., Manto, M., Masciullo, M., Molinari, M., Stoodley, C. J., Van Overwalle, F., Ivry, R. B., & Schmahmann, J. D. (2020). The cerebellar cognitive affective/schmahmann syndrome: A task force paper. *The Cerebellum*, 19(1), 102–125. https://doi.org/10.1007/ s12311-019-01068-8
- Bernard, J. A., & Seidler, R. D. (2013). Relationships between regional cerebellar volume and sensorimotor and cognitive function in young and older adults. *Cerebellum*, 12(5), 721–737. https://doi. org/10.1007/S12311-013-0481-Z/METRICS
- Bhanpuri, N. H., Okamura, A. M., & Bastian, A. J. (2014). Predicting and correcting ataxia using a model of cerebellar function. *Brain*, 137(7), 1931–1944. https://doi.org/10.1093/brain/awu115
- Bianco, V., Finisguerra, A., Betti, S., D'argenio, G., & Urgesi, C. (2020). Autistic traits differently account for context-based predictions of physical and social events. *Brain Sciences*. https://doi. org/10.3390/brainsci10070418
- Broadbent, H. J., Farran, E. K., & Tolmie, A. (2014). Egocentric and allocentric navigation strategies in Williams syndrome and typical

- Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience*, 6(May), 1–19. https://doi.org/10.3389/fnhum.2012.00147
- Budimirovic, D. B., Berry-Kravis, E., Erickson, C. A., Hall, S. S., Hessl, D., Reiss, A. L., King, M. K., Abbeduto, L., & Kaufmann, W. E. (2017). Updated report on tools to measure outcomes of clinical trials in fragile X syndrome. *Journal of Neurodevelopmental Disorders*, 9(1), 1–36. https://doi.org/10.1186/ S11689-017-9193-X
- Bulgheroni, S., D'Arrigo, S., Signorini, S., Briguglio, M., Di Sabato, M. L., Casarano, M., Mancini, F., Romani, M., Alfieri, P., Battini, R., Zoppello, M., Tortorella, G., Bertini, E., Leuzzi, V., Valente, E. M., & Riva, D. (2016). Cognitive, adaptive, and behavioral features in Joubert syndrome. *American Journal of Medical Genetics, Part A*, 170(12), 3115–3124. https://doi.org/10.1002/ ajmg.a.37938
- Butti, N., Biffi, E., Genova, C., Romaniello, R., Redaelli, D. F., Reni, G., Borgatti, R., & Urgesi, C. (2020a). Virtual reality social prediction improvement and rehabilitation intensive training (VR-SPIRIT) for paediatric patients with congenital cerebellar diseases: Study protocol of a randomised controlled trial. *Trials.* https://doi.org/10.1186/s13063-019-4001-4
- Butti, N., Corti, C., Finisguerra, A., Bardoni, A., Borgatti, R., Poggi, G., & Urgesi, C. (2020b). Cerebellar damage affects contextual priors for action prediction in patients with childhood brain tumor. *Cerebellum*. https://doi.org/10.1007/s12311-020-01168-w
- Butti, N., Oldrati, V., Ferrari, E., Romaniello, R., Gagliardi, C., Borgatti, R., & Urgesi, C. (2023). New insights into the neuropsychological profile and intellectual quotient variability in Joubert syndrome compared to other congenital cerebellar malformations. *The Cerebellum*. https://doi.org/10.1007/s12311-023-01580-y
- 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
- Chatterjee, K., Buchanan, A., Cottrell, K., Hughes, S., Day, T. W., & John, N. W. (2022). Immersive virtual reality for the cognitive rehabilitation of stroke survivors. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30, 719–728. https://doi. org/10.1109/TNSRE.2022.3158731
- Cikajlo, I., & Peterlin Potisk, K. (2019). Advantages of using 3D virtual reality based training in persons with Parkinson's disease: A parallel study. *Journal of NeuroEngineering and Rehabilitation*, 16(1), 1–14. https://doi.org/10.1186/S12984-019-0601-1/FIGUR ES/7
- 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*. https://doi.org/10. 3389/fncel.2018.00510
- Corti, C., Poggi, G., Romaniello, R., Strazzer, S., Urgesi, C., Borgatti, R., & Bardoni, A. (2018). Feasibility of a home-based computerized cognitive training for pediatric patients with congenital or acquired brain damage: An explorative study. *PLoS ONE*, 13(6), 1–16. https://doi.org/10.1371/journal.pone.0199001
- de Lange, A. M. G., Bråthen, A. C. S., Rohani, D. A., Grydeland, H., Fjell, A. M., & Walhovd, K. B. (2017). The effects of memory training on behavioral and microstructural plasticity in young and older adults. *Human Brain Mapping*, 38(11), 5666–5680. https:// doi.org/10.1002/HBM.23756
- Dziura, J. D., Post, L. A., Zhao, Q., Fu, Z., & Peduzzi, P. (2013). Strategies for dealing with missing data in clinical trials: From design to analysis. *Yale Journal of Biology and Medicine*, 86(3), 343–358.

- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. https://doi.org/10.3758/bf03193146
- Ferrari, E., Butti, N., Gagliardi, C., Romaniello, R., Borgatti, R., & Urgesi, C. (2022). Cognitive predictors of social processing in congenital atypical development. *Journal of Autism and Devel*opmental Disorders. https://doi.org/10.1007/s10803-022-05630-y
- Frigerio, A., Cattaneo, C., Cataldo, M. G., Schiatti, A., Molteni, M., & Battaglia, M. (2004). Behavioral and emotional problems among Italian children and adolescents aged 4 to 18 years as reported by parents and teachers. *European Journal of Psychological Assessment*. https://doi.org/10.1027/1015-5759.20.2.124
- Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. Annual Review of Psychology. https://doi.org/10.1146/annur ev-psych-120710-100449
- Frolli, A., Savarese, G., Di Carmine, F., Bosco, A., Saviano, E., Rega, A., Carotenuto, M., & Ricci, M. C. (2022). Children on the autism spectrum and the use of virtual reality for supporting social skills. *Children*, 9(2), 181. https://doi.org/10.3390/children9020181
- Gagliardi, C., Brenna, V., Romaniello, R., Arrigoni, F., Tavano, A., Romani, M., Valente, E. M., & Borgatti, R. (2015). Cognitive rehabilitation in a child with Joubert syndrome: Developmental trends and adaptive changes in a single case report. *Research* in Developmental Disabilities, 47, 375–384. https://doi.org/10. 1016/j.ridd.2015.09.013
- Hanratty, J., Livingstone, N., Robalino, S., Terwee, C. B., Glod, M., Oono, I. P., Rodgers, J., Macdonald, G., & McConachie, H. (2015). Systematic review of the measurement properties of tools used to measure behaviour problems in young children with autism. *PLoS ONE*, 10(12), e0144649. https://doi.org/10.1371/ JOURNAL.PONE.0144649
- Hoche, F., Guell, X., Sherman, J. C., Vangel, M. G., & Schmahmann, J. D. (2016). Cerebellar contribution to social cognition. *The Cerebellum.* https://doi.org/10.1007/s12311-015-0746-9
- Ishikawa, T., Tomatsu, S., Izawa, J., & Kakei, S. (2016). The cerebrocerebellum: Could it be loci of forward models? *Neuroscience Research*. https://doi.org/10.1016/j.neures.2015.12.003
- Ito, T., Osawa, A., Maeshima, S., Kikuchi, Y., & Kigawa, H. (2010). Rehabilitation approach for the patient with cerebellar cognitive affective syndrome. *International Journal of Stroke*, 5, 291. https://doi.org/10.1111/j.1747-4949.2010.00480.x
- Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY—Second Edition (NEPSY-II). Harcourt Assessment.
- Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, 79(5), 836–848. https://doi.org/10.1016/j. neuron.2013.08.020
- Leggio, M., & Molinari, M. (2015). Cerebellar sequencing: A trick for predicting the future. *Cerebellum*, 14(1), 35–38. https://doi.org/ 10.1007/s12311-014-0616-x
- Ma, Q., Pu, M., Heleven, E., Haihambo, N. P., Baetens, K., Baeken, C., Deroost, N., & Van Overwalle, F. (2021). The posterior cerebellum supports implicit learning of social belief sequences. *Cognitive*, *Affective*, & *Behavioral Neuroscience*, 21, 1–23. https://doi. org/10.3758/s13415-021-00910-z
- Malerba, G., Bellazzecca, S., Urgesi, C., Butti, N., D'Angelo, M. G., Diella, E., & Biffi, E. (2023). Is social training delivered with a head-mounted display suitable for patients with hereditary ataxia? *Brain Sciences*, 13(7), 1017. https://doi.org/10.3390/brainsci13 071017
- Manto, M., & Mariën, P. (2015). Schmahmann's syndrome—identification of the third cornerstone of clinical ataxiology. *Cerebellum* and Ataxias. https://doi.org/10.1186/s40673-015-0023-1
- Molinari, M., & Masciullo, M. (2019). The Implementation of predictions during sequencing. *Frontiers in Cellular Neuroscience*. https://doi.org/10.3389/fncel.2019.00439

- Montoya-Rodríguez, M. M., de Souza Franco, V., Tomás Llerena, C., Molina Cobos, F. J., Pizzarossa, S., García, A. C., & Martínez-Valderrey, V. (2023). Virtual reality and augmented reality as strategies for teaching social skills to individuals with intellectual disability: A systematic review. *Journal of Intellectual Disabilities*, 27(4), 1062–1084. https://doi.org/10.1177/1744629522 1089147/ASSET/IMAGES/LARGE/10.1177_174462952210891 47-FIG3_JPEG
- Narzisi, A., Muratori, F., Calderoni, S., Fabbro, F., & Urgesi, C. (2013). Neuropsychological profile in high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-012-1736-0
- Newcomb, E. T., & Hagopian, L. P. (2018). Treatment of severe problem behaviour in children with autism spectrum disorder and intellectual disabilities. *International Review of Psychiatry*, 30(1), 96–109. https://doi.org/10.1080/09540261.2018.1435513
- Nossa, R., Gagliardi, C., Panzeri, D., Diella, E., Maghini, C., Genova, C., Turconi, A. C., & Biffi, E. (2022). Could an immersive virtual reality training improve navigation skills in children with cerebral palsy? A pilot controlled study. *Journal of Clinical Medicine*, *11*(20), 6146. https://doi.org/10.3390/jcm11206146
- Oldrati, V., Ferrari, E., Butti, N., Cattaneo, Z., Borgatti, R., Urgesi, C., & Finisguerra, A. (2021). How social is the cerebellum? Exploring the effects of cerebellar transcranial direct current stimulation on the prediction of social and physical events. *Brain Structure* and Function. https://doi.org/10.1007/s00429-020-02198-0
- Olivito, G., Clausi, S., Laghi, F., Tedesco, A. M., Baiocco, R., Mastropasqua, C., Molinari, M., Cercignani, M., Bozzali, M., & Leggio, M. (2017). Resting-state functional connectivity changes between dentate nucleus and cortical social brain regions in autism spectrum disorders. *The Cerebellum*, 16(2), 283–292. https://doi.org/10.1007/s12311-016-0795-8
- Pavlova, M., Sokolov, A., & Krägeloh-Mann, I. (2007). Visual navigation in adolescents with early periventricular lesions: Knowing where, but not getting there. *Cerebral Cortex*, 17(2), 363–369. https://doi.org/10.1093/cercor/bhj153
- Pellicano, E., & Burr, D. (2012). When the world becomes "too real": A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510. https://doi.org/10.1016/j.tics. 2012.08.009
- Peterburs, J., & Desmond, J. E. (2016). The role of the human cerebellum in performance monitoring. *Current Opinion in Neurobiol*ogy, 40, 38–44. https://doi.org/10.1016/j.conb.2016.06.011
- Rasmussen, C., Tamana, S., Baugh, L., Andrew, G., Tough, S., & Zwaigenbaum, L. (2013). Neuropsychological impairments on the NEPSY-II among children with FASD. *Child Neuropsychology*. https://doi.org/10.1080/09297049.2012.658768
- Romaniello, R., Gagliardi, C., Desalvo, P., Provenzi, L., Battini, R., Bertini, E., Bonati, M. T., Briguglio, M., D'Arrigo, S., Dotti, M. T., Giordano, L., Macaluso, C., Moroni, I., Nuovo, S., Santucci, M., Signorini, S., Stanzial, F., Valente, E. M., & Borgatti, R. (2022). Challenges and resources in adult life with Joubert syndrome: Issues from an international classification of functioning (ICF) perspective. *Disability and Rehabilitation*, 44(18), 4966– 4973. https://doi.org/10.1080/09638288.2021.1922516
- Ruffieux, N., Colombo, F., Gentaz, E., Annoni, J. M., Chouiter, L., Roulin Hefti, S., Ruffieux, A., & Bihl, T. (2017). Successful neuropsychological rehabilitation in a patient with cerebellar cognitive affective syndrome. *Applied Neuropsychology: Child*, 6(2), 180–188. https://doi.org/10.1080/21622965.2015.1092087
- Schmahmann, J. D. (2004). Disorders of the cerebellum: Ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 16(3), 367–378. https://doi.org/10.1176/jnp.16.3.367

- Schmahmann, J. D. (2019). The cerebellum and cognition. In Neuroscience Letters. https://doi.org/10.1016/j.neulet.2018.07.005
- Schmahmann, J. D., & Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain*. https://doi.org/10.1093/brain/ 121.4.561
- Siman-Tov, T., Granot, R. Y., Shany, O., Singer, N., Hendler, T., & Gordon, C. R. (2019). Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. *Neuroscience and Biobehavioral Reviews*, 105(August), 262–275. https://doi.org/10.1016/j.neubiorev.2019.08.012
- Sokolov, A. A., Gharabaghi, A., Tatagiba, M. S., & Pavlova, M. (2010). Cerebellar engagement in an action observation network. *Cerebral Cortex*, 20(2), 486–491. https://doi.org/10.1093/cercor/bhp117
- Sokolov, A. A., Erb, M., Grodd, W., & Pavlova, M. A. (2014). Structural loop between the cerebellum and the superior temporal sulcus: Evidence from diffusion tensor imaging. *Cerebral Cortex*, 24(3), 626–632. https://doi.org/10.1093/cercor/bhs346
- 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
- Sokolovsky, N., Cook, A., Hunt, H., Giunti, P., & Cipolotti, L. (2010). A preliminary characterisation of cognition and social cognition in spinocerebellar ataxia types 2, 1, and 7. *Behavioural Neurology*, 23(1–2), 17–29. https://doi.org/10.3233/BEN-2010-0270
- Sparaci, L., Stefanini, S., D'Elia, L., Vicari, S., & Rizzolatti, G. (2014). What and why understanding in autism spectrum disorders and Williams syndrome: Similarities and differences. *Autism Research*, 7(4), 421–432. https://doi.org/10.1002/AUR.1370/ SUPPINFO
- Stoodley, C. J., & Limperopoulos, C. (2016). Structure–function relationships in the developing cerebellum: Evidence from early-life cerebellar injury and neurodevelopmental disorders. *Seminars in Fetal and Neonatal Medicine*, 21(5), 356–364. https://doi.org/10. 1016/j.siny.2016.04.010
- Stoodley, C. J., & Tsai, P. T. (2021). Adaptive prediction for social contexts: The cerebellar contribution to typical and atypical social behaviors. *Annual Review of Neuroscience*, 44(1), 475–493. https://doi.org/10.1146/ANNUREV-NEURO-100120-092143
- Strain, P. S., Barton, E. E., & Dunlap, G. (2012). Lessons learned about the utility of social validity. *Education and Treatment of Children*, 35(2), 183–200. https://doi.org/10.1353/ETC.2012.0007
- Summers, A. C., Snow, J., Wiggs, E., Liu, A. G., Toro, C., Poretti, A., Zein, W. M., Brooks, B. P., Parisi, M. A., Inati, S., Doherty, D., Vemulapalli, M., Mullikin, J. C., Vilboux, T., Gahl, W. A., & Gunay-Aygun, M. (2017). Neuropsychological phenotypes of 76 individuals with Joubert syndrome evaluated at a single center. *American Journal of Medical Genetics Part A*, 173(7), 1796–1812. https://doi.org/10.1002/ajmg.a.38272
- Tavano, A., & Borgatti, R. (2010). Evidence for a link among cognition, language and emotion in cerebellar malformations. *Cortex*, 46(7), 907–918. https://doi.org/10.1016/j.cortex.2009.07.017
- Tavano, A., Grasso, R., Gagliardi, C., Triulzi, F., Bresolin, N., Fabbro, F., & Borgatti, R. (2007). Disorders of cognitive and affective development in cerebellar malformations. *Brain*. https://doi.org/ 10.1093/brain/awm201
- Tieri, G., Morone, G., Paolucci, S., & Iosa, M. (2018). Virtual reality in cognitive and motor rehabilitation: Facts, fiction and fallacies. *In Expert Review of Medical Devices*. https://doi.org/10.1080/17434 440.2018.1425613
- Urgesi, C., Campanella, F., & Fabbro, F. (2011). NEPSY-II, Contributo alla Taratura Italiana (2nd ed.). Giunti OS.
- Urgesi, C., Butti, N., Finisguerra, A., Biffi, E., Valente, E. M., Romaniello, R., & Borgatti, R. (2021). Social prediction in pediatric patients with congenital, non-progressive malformations of the

cerebellum: From deficits in predicting movements to rehabilitation in virtual reality. *Cortex*, *144*, 82–98. https://doi.org/10. 1016/j.cortex.2021.08.008

- van Dun, K., Overwalle, F. V., Manto, M., & Marien, P. (2018). Cognitive impact of cerebellar damage: Is there a future for cognitive rehabilitation? CNS & Neurological Disorders—Drug Targets. https://doi.org/10.2174/1871527317666180110125043
- Van Overwalle, F., & Mariën, P. (2016). Functional connectivity between the cerebrum and cerebellum in social cognition: A multi-study analysis. *NeuroImage*, 124, 248–255. https://doi.org/ 10.1016/j.neuroimage.2015.09.001
- Van Overwalle, F., De Coninck, S., Heleven, E., Perrotta, G., Taib, N. O. B., 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., Manto, M., Leggio, M., & Delgado-García, J. M. (2019b). The sequencing process generated by the cerebellum crucially contributes to social interactions. *Medical Hypotheses*. https://doi.org/10.1016/j.mehy.2019.05.014
- Van Overwalle, F., Manto, M., Cattaneo, Z., Clausi, S., Ferrari, C., Gabrieli, J. D. E., Guell, X., Heleven, E., Lupo, M., Ma, Q., Michelutti, M., Olivito, G., Pu, M., Rice, L. C., Schmahmann,

J. D., Siciliano, L., Sokolov, A. A., Stoodley, C. J., van Dun, K., & Leggio, M. (2020a). Consensus paper: Cerebellum and social cognition. *The Cerebellum*, *19*(6), 833–868. https://doi.org/10. 1007/s12311-020-01155-1

- Van Overwalle, F., Van de Steen, F., van Dun, K., & Heleven, E. (2020b). Connectivity between the cerebrum and cerebellum during social and non-social sequencing using dynamic causal modelling. *NeuroImage*. https://doi.org/10.1016/j.neuroimage. 2019.116326
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart 1. *Journal of Applied Behavior Analysis*, 11(2), 203–214. https://doi. org/10.1901/jaba.1978.11-203
- Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9), 338–347. https://doi.org/10.1016/S1364-6613(98)01221-2

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.