scientific reports

OPEN

# Age-related patterns in the creation of social prediction models: a study using immersive virtual reality

Libera Siciliano[1,2,14]✉, Maria Pyasik[3,14]✉, Gaetano Tieri[4,5], Valentina Ciot[6,7], Claudia Salera[5,8], Nicole Urbini[2,8], Viola Oldrati[9], Andrea Ciricugno[10], Giusy Olivito[2,8], Renato Borgatti[10,11], Marco Iosa[5,8], Zaira Cattaneo[12], Cosimo Urgesi[3,13] & Maria Leggio[2,8]

**Exposure to environmental regularities allows individuals to build internal models, supporting the integration of prior expectations with sensory input in social contexts. However, how social prediction models develop across age and experience remains unclear. To address this, we investigated whether the formation of social predictive models varies across age and experiences in neurotypical individuals (early adolescents, young adults, and adults). Participants completed, in counterbalanced order, two scenarios of an immersive virtual reality (IVR) task that facilitates the creation of top-down predictions during action perception, requiring the participants to implicitly associate the hand movement of four avatars toward one preferred object among three. Self-report questionnaires and standard social cognition tests were used to assess embodiment, presence, usability in the IVR setting, and the direct association between the novel IVR task and social-cognitive tests. Results showed significant performance improvement from the first to the second scenario in all groups, indicating learning and refinement of predictive models. Early adolescents performed significantly worse and showed a lower learning rate than adult groups, who did not differ from each other, suggesting slower formation of social prior expectations in early adolescents. The IVR paradigm was feasible across all age groups and proved valid for assessing social processing. These findings support a developmental trajectory in which the creation of predictive internal models improves with age and experience.**

**Keywords** Immersive virtual reality, Prior expectations, Top-down social predictions, Adolescence, Adulthood

Understanding others' actions and intentions is crucial for social cognition. Efficient social interactions rely not only on understanding others in the moment, but also on the ability to anticipate their upcoming actions or changes in mental and emotional states[1]. These social prediction processes rely on the context of actions or social events, firstly because identical actions or social events can have different meaning in different contexts, and secondly because they are naturally embedded in a certain context and do not occur as a standalone phenomenon[2]. Accordingly, predictive coding accounts of social perception propose that the sensorial information about an action (for example, action kinematics) is compared with the prior expectations about the most likely action or social event[3,4]. In general, prior expectations (priors) are based on the previously observed and experienced statistical regularities inherent to the natural world[5], and the context in which actions or social events occur is

[1]Department of Human Sciences, Link Campus University, Rome, Italy. [2]Ataxia Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy. [3]Scientific Institute, IRCCS E. Medea, Pasian di Prato, Udine, Italy. [4]Virtual Reality Lab, Department of Law and Digital Society, Unitelma Sapienza University, Rome, Italy. [5]SmArt Lab, IRCCS Santa Lucia Foundation, Rome, Italy. [6]Department of Life Sciences, University of Trieste, Trieste, Italy. [7]Laboratory of Cognitive Neuroscience, Department of Languages and Literatures, Communication, Education and Society, University of Udine, Udine, Italy. [8]Department of Psychology, Sapienza University of Rome, Rome, Italy. [9]Scientific Institute, IRCCS E. Medea, Bosisio Parini, Lecco, Italy. [10]IRCCS Mondino Foundation, Pavia, Italy. [11]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy. [12]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy. [13]Department of Human and Social Sciences, Universitas Mercatorum, Rome, Italy. [14]These authors contributed equally: Libera Siciliano and Maria Pyasik. ✉email: l.siciliano@unilink.it; maria.pyasik@lanostrafamiglia.it

one of the sources of these priors[6]. Indeed, priors[7] affect different social prediction processes, such as action prediction or recognition (for review, see[8].

The use of prior expectations, especially in the social domain, undergoes significant maturation during adolescence and is refined across age as it is influenced by the amount of information the brain has collected[9]. Several studies have investigated how predictive abilities in different domains change with ageing. In general, the existing evidence suggests that people of older age rely more on priors and less on sensorial information[10,11]. As a result, older adults tend to benefit more from contextual information, especially in situations of perceptual ambiguity[12]. Conversely, evidence investigating the development of social predictive processes in preadolescence and adolescence is scanty. One study investigated the role of priors in the social domain (emotion discrimination) in adolescents and adults[7]. While adolescents were generally influenced by priors to a similar extent as adults, they showed less bias specifically for the expression of anger, suggesting that social prediction abilities are still developing at that age[7].

The development of social prediction abilities throughout the lifespan may occur in two ways: first, through the maturation of the ability to predict others' actions and social cues based on sensory information (e.g., movement kinematics); and second, through age-related changes in the formation of social priors and contextual expectations that may sustain the disambiguation of sensory information. While previous studies have used computerized tasks to compare the reliance on pre-existing social priors across lifespan stages[7], no study has so far tested whether the formation of new social expectations, independently from pre-existing knowledge and experience, may change across the lifespan. Immersive virtual reality (IVR) offers a unique avenue to compare the formation of social expectations in ecological scenarios. Plenty of literature reports that IVR is a suitable technology for studying complex cognitive functions since it incorporates real-world features and allows for controlled adjustment of contextual stimuli, demonstrating a substantial ecological validity compared to traditional testing tools, given the embodied experience and natural sense of presence conveyed by IVR[13,14]. Accordingly, several studies have demonstrated that IVR interventions implemented in rehabilitation protocols can improve emotional recognition and social skills in children and adolescents, including those with neurological disorders[15,16]. In the vein of predictive processing, an IVR rehabilitative training which targeted social prediction deficits has proven efficacious in enhancing social prediction abilities in children, adolescents and young adults affected by congenital cerebellar malformation[17,18]. Additionally, several studies proved the applicability of IVR for both cognitive testing and rehabilitative intervention in adults and elderly, across a wide variety of cognitive and affective domains[19,20].

The present study aimed at testing the ability to create social predictive models online in three age groups – early adolescents, young adults and adults. To this aim, an IVR task adapted for participants' age groups was implemented. In the task, participants were required to predict the actions (i.e., a choice of one of three possible objects) of four virtual avatars that varied in their appearance; three avatars had a preferred object category, while the fourth avatar was unpredictable. To anticipate the avatars' choices, participants needed to develop predictive models of the object preferred by each avatar based on the contingency between the avatars' hand movement (and afterward the avatars' appearance) and their action outcomes (object choices), without receiving explicit instructions about the cues that indicate the avatars' preferences. After the completion of the IVR tasks, three questionnaires (embodiment questionnaire[21], User Satisfaction Evaluation Questionnaire – USEQ[22]; NASA-Task Load Index - NASA-TLX[23] were administered to validate the sense of embodiment and presence in the virtual scenarios, as well as their usability. Finally, given the novelty of the IVR task, we aimed to test its direct association with widely used social cognition tests. Therefore, participants' social cognition abilities were additionally evaluated with standardized age-appropriate assessments, namely the social perception tests of the NEPSY-II battery[24] for early adolescents, and the Faux Pas Test[25,26] and the Reading the Mind in the Eyes test - RMET[27,28] for both groups of adults (description of social cognition tests are reported in supplementary materials). We hypothesized that the ability to construct predictive internal models of others' actions evolves with age and improves with practice, as assessed by an effective IVR system which facilitates and enhances the creation of forward models during social perception. In addition, we expected to observe different patterns of performance between early adolescents and adults in the IVR task, considering that the prediction processing develops throughout development[29].

## Materials and methods

### Participants

An a priori power analysis was conducted with the G*power software[30] using the "as in Cohen 1988" option. Healthy participants were recruited for each age group to obtain 90% of power (1-beta) with alpha = 0.05 in our mixed-model design with three age groups (early adolescents, young adults and adults) x two IVR task blocks (Block 1 vs. Block 2). A moderate to large effect size was estimated, f(V) = 0.39, based on the within-subject change of action prediction performance after a social IVR training[18] $r = 0.364$). A minimum sample size of 87 subjects was calculated. Accordingly, we enrolled ≥ 29 participants per group. Specifically, 30 early adolescents (age range: 11–17 years), 31 young adults (age range: 18–35 years) and 33 adults (age range: 36–62 years) were recruited. All participants had no history of neurological disorders or psychiatric illness and had normal or corrected-to-normal visual acuity; all participants were native Italian speakers. The demographics of each group of participants are described in Table 1. The early adolescents and part of the adults and the young adults were recruited at the University of Udine and at secondary schools of Udine, Italy. The young adults and part of the adults were recruited at IRCCS Santa Lucia Foundation, Italy. Thus, participants were recruited from homogeneous high-education context, but no other socioeconomic information was taken. To assess the intellectual level, the adult and pre-adolescent groups with age years ≥ 12 performed the Standard Progressive Matrices (SPM[31], with transformation of raw scores to numerical Intellectual Quotient (IQ) values in relation to age. In contrast, 11-year-old early adolescents performed the Colored Progressive Matrices (CPM[32], with

| Group | *N* | Age (years) | Education (years) | Sex | I.Q. scores |
|---|---|---|---|---|---|
| Early adolescents | 30 | 11.9 (0.76) | 5.9 (0.76) | 16 (F), 14 (M) | 102.9 (16.69) |
| Young adults | 31 | 25.35 (2.69) | 16.74 (1.81) | 19 (F), 12 (M) | 108.57 (10.48) |
| Adults | 33 | 49.52 (7.65) | 15.97 (3.31) | 24 (F), 9 (M) | 119.34 (7.54) |

**Table 1**. Demographic characteristics of the studied groups. Values represent mean and standard deviation (in parentheses) for age, years of education, and IQ scores. N = number of participants per group. Sex indicates the number of female (F) and male (M) participants within each group.

transformation of raw scores to percentiles and then to IQ scores. IQ was used as inclusion criteria to ensure typical intellectual functioning, rather than as a covariate of theoretical interest, given that it was tested with different age-appropriate tests across and within the age groups and, consequently, calculated using different standardization procedures. Therefore, the independence assumption required for covariance analyses was violated[33,34]. Participants, or parents in the case of minors, signed a written informed consent for pseudo-anonymized use of research data. This research study was approved by the Ethics Committees of IRCCS Santa Lucia Foundation and of the University of Udine, according to the principles expressed in the Declaration of Helsinki.

## IVR task

*Stimuli*

The virtual scenarios and objects were developed in Blender 3D (https://www.blender.org/) and implemented in Unity Game Engine (https:/www.unity.com/). The scenarios were designed to represent realistic settings and situations appropriate for the age groups of participants (adolescents and adults). In detail, the scenarios for early adolescents presented a school library and an outdoor playground, while the scenarios for the two groups of adults included a supermarket and a restaurant.

In addition, a practice scenario was created for each age group; its layout, objects and trial structure were identical to those in the experimental scenarios, but the scene did not contain any details to identify the virtual space; the virtual objects were simple 3D geometric figures, and the avatars were represented as grey shapes without distinct features. All scenarios had a similar layout that included a surface (a counter) where the virtual objects were placed, which was located in front of the participant, and a door at the far frontal end of the room/space (Fig. 1A).

Four virtual avatars were created for each set of scenarios (adolescent and adult versions) with MakeHuman (http://www.makehumancommunity.org/) and iClone 3D (www.reallusion.com) software. The avatars' body size and features were age-matched with the groups of participants. In both sets, two avatars were male and two were female, and all avatars differed in certain characteristics such as hair color, hairstyle, and clothing (Fig. 1B). All avatars were animated in Unity Game Engine with the animation of naturalistic walking. The scenarios were presented to the participants via a Meta Quest 2 head-mounted display (HMD) with RGB LCDs with 1832 × 1920 pixels resolution, refresh rate at 120 Hz, the field of view of 110° (diagonal FOV) and 6 degrees of freedom. Participants interacted with the content of the scenarios via Meta Quest 2 hand-held game controllers for both hands.

## Procedure

In all the scenarios, participants observed the virtual scene from the first-person perspective. The only visible part of their virtual body were the hands, which moved synchronously with participants' hand movements performed with the game controllers. In the scene, participants were located above a chair in front of the counter that was reachable with a forward hand movement. At the onset of each trial (activated by the participants via pressing the A key on the right controller), three different objects appeared on the counter in front of the participant. Together with the appearance of the objects, one of four avatars entered the scene through the door located in the far front end of the scene. The avatar approached the counter by walking at a constant speed and grasped one of the objects. In the first four trials, participants were instructed to observe the avatars and the objects without performing any movements. After that, participants' task was to anticipate which object the avatar would grasp and to grasp that object themselves before the avatar (Fig. 1C). Once either the avatar or the participant grasped one of the objects, the avatar and the objects disappeared and the trial was concluded.

The virtual objects in each scenario belonged to one of three categories: in the supermarket, the categories included packaged foods, beverages and household products; in the restaurant, the objects belonged to the categories of foods, beverages and sauces; in the library, the objects were books, pencil cases and stationery objects; in the playground, the objects were stuffed toys, books and flowers; in the practice scenario, the objects were cubes, cylinders and spheres with sizes matching those of the objects used in the other scenarios. In each scenario, three avatars had a preferred category of objects that they grasped in all trials. The fourth avatar picked the objects pseudo-randomly from any category by picking each of the three objects with equal probability (0.33) and therefore an equal number of times per scenario. The avatar-object category mapping was pseudorandomized and differed between the two experimental scenarios that each participant performed; moreover, in addition to the object preferences, avatars' clothes changed between the scenarios. Each of the experimental scenarios included 44 trials and lasted approximately 7 min in total. The two scenarios developed for the age-group were presented in two different blocks, the order of which was counterbalanced across participants.

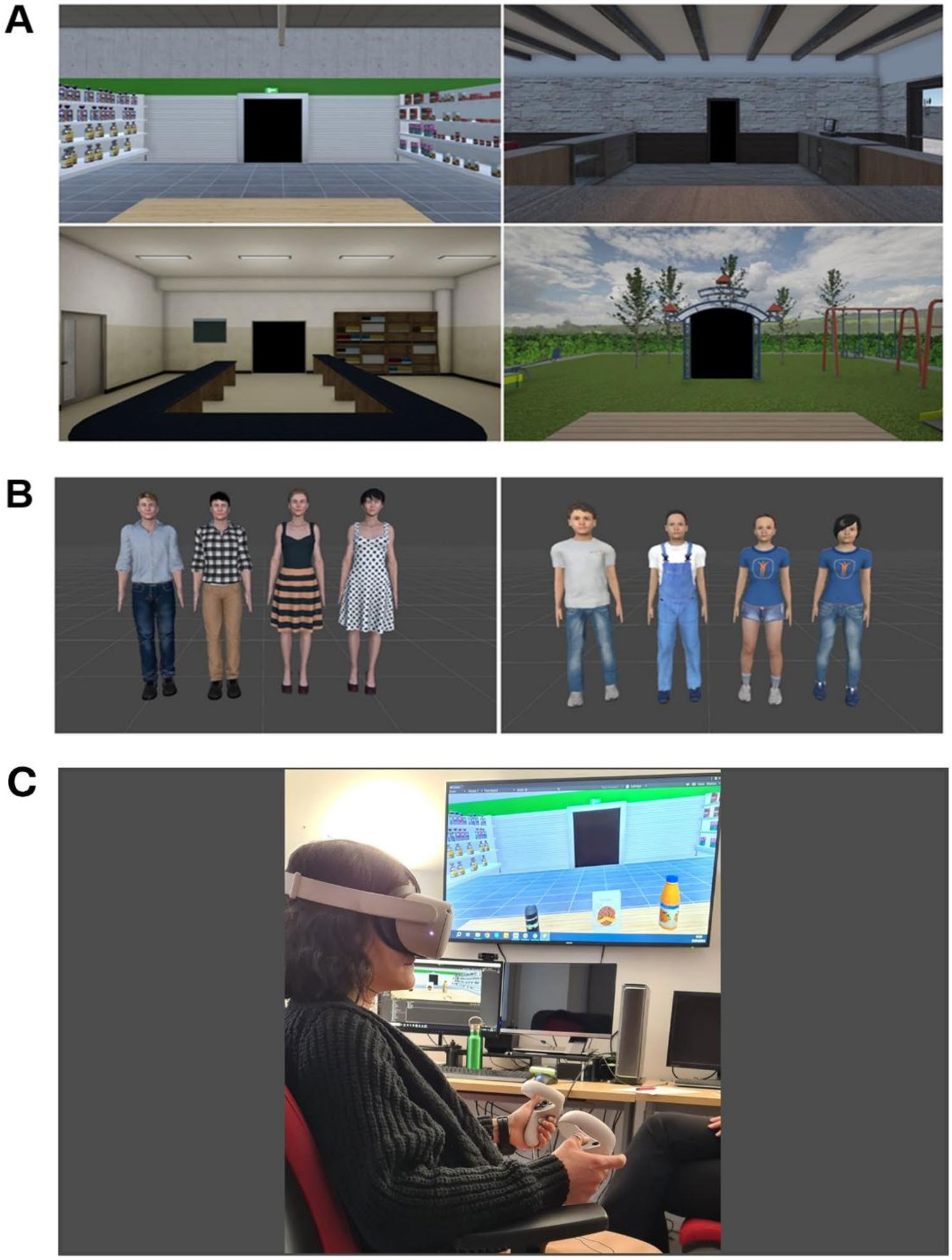


**Fig. 1**. IVR task description and set-up. (**A**) Virtual scenarios for the groups of adults (top left panel: supermarket scenario, top right panel: restaurant scenario) and early adolescents (bottom left panel: library scenario, bottom right panel: playground scenario); (**B**) Virtual avatars for the adults' (left panel) and early adolescents' (right panel) scenarios; (**C**) Experimental setup and procedure, showing a participant wearing the HMD and performing a trial of the supermarket virtual scenario via the hand-held game controllers. Informed consent was obtained to publish the image in an online open access publication.

Before starting the experimental scenarios, participants sat on a chair in a lab with enough surrounding area to safely perform the reaching movements. They were instructed to sit with their back against the back of the chair and their arms placed on the armrests, after which they put on the HMD and took the controllers in each hand. Then, the practice scenario was launched, and the participants could first take a few moments to gain confidence with the virtual environment, performing hand movements, looking around, and describing what they saw. These actions were also performed in the beginning of each experimental scenario and were necessary to create a sense of embodiment and presence in the virtual environment[35]. In the practice scenario, the structure and the aim of the task were explained: participants were instructed how to grasp the virtual objects, how to start the trials, and what the main goal of the task was (predicting the avatars' preferences in picking the objects and grasping that object before the avatar). Crucially, participants were informed that all avatars had a preference for one category of objects. If the participants had grasped the correct object (i.e., the target object of the avatar), a high-pitched sound was delivered via the HMD speakers; otherwise, a low-pitched sound marked erroneous responses. Participants then performed several trials of the practice scenario until they were comfortable with the grasping movement and the procedure. After that, they performed the two experimental scenarios that were presented in a counterbalanced order.

Three questionnaires were administered to validate the sense of embodiment and presence in the virtual scenarios (embodiment questionnaire[21], as well as their usability (USEQ[22], NASA-TLX[23]. Detailed description of statistical analyses on questionnaires and discussed results are reported in the section Additional Results of Supplementary materials.

## Data handling and statistical analysis

*IVR task*

For each IVR scenario, accuracy (proportion of correct responses) and reaction times (RTs, in s) were extracted for each participant. In addition, a learning coefficient was computed by calculating a cumulative accuracy (the sum of correct responses) at each trial, and then fitting a linear model for the cumulative accuracy as a function of trial number; the coefficient of the fitted model was extracted for each participant in each scenario and interpreted as the learning coefficient (learning rate)[36].

Evaluation of the normality of distribution for each analyzed variable was based on skewness and kurtosis; all variables were normally distributed; hence, parametric analyses were performed.

For each performance metric in the IVR tasks (accuracy, RTs, and the learning coefficient), a 2 × 3 mixed-effects ANOVA was performed with Scenario order (block 1, block 2) as a within-subject factor and Age group (early adolescents, younger adults, adults) as a between-subject factor.

A significance threshold of $p = 0.05$ was set for all statistical analyses. Bonferroni correction was applied to all post-hoc paired comparisons whenever they were performed ($p_{corr} = 0.05$). Statistical analyses were performed with R software in R Studio[37].

*IVR task and social cognition tests association*

We assessed the association between the IVR task and widely used social cognitive tests to evaluate whether these tests predict results on the novel tasks. To this aim, participants were administered standardized age-appropriate neuropsychological social tests, namely the social perception tests of the NEPSY-II battery (ToM A, ToM B and Affect Recognition[24] for early adolescents, and the Faux Pas Test[25,26] and the Reading the Mind in the Eyes test (RMET[27,28] for both groups of adults (description of social cognition tests are reported in supplementary materials). Descriptive statistics of social cognition tests and independent-sample t-test (two-tailed) comparing young adults and adults are reported in Table 2. Since the different age-related social cognition tests assessed the same social cognition components, namely the affective component of ToM (ToM B of the NEPSY-II; the affective component of the Faux Pas Test), the cognitive component of ToM (ToM A of the NEPSY-II; the cognitive component of the Faux Pas Test) and emotion recognition (Affect Recognition of the NEPSY-II; the RMET), we transformed participants' scores in z-scores, within each age-group. As concerned with the IVR

| Social cognition tests | Early adolescents | Young adults | Adults | t-value (*p* value) | Adults together |
|---|---|---|---|---|---|
| ToM-A | | | | | |
| Mean | 11.34 | – | – | – | |
| SD | 2.96 | | | | |
| ToM-B | | | | | |
| Mean | 7.71 | – | – | – | |
| SD | 3.60 | | | | |
| AR | | | | | |
| Mean | 7.71 | – | – | – | |
| SD | 4.76 | | | | |
| ToM – Affective Component | | | | | |
| Mean | – | 6.97 | 7.59 | -1.14 | 7.29 |
| SD | | 1.94 | 2.33 | (0.26) | 2.17 |
| ToM – Cognitive Component | | | | | |
| Mean | – | 39.94 | 39.69 | 0.11 | 39.81 |
| SD | | 6.88 | 10.51 | (0.91) | 8.91 |
| RMET | | | | | |
| Mean | – | 27.42 | 25.85 | 1.94 | 26.61 |
| SD | | 3.37 | 3.04 | (0.06) | 3.3 |

**Table 2.** Descriptive statistics of Social Cognition tests and t-test between young adults and adults. The values are reported as mean and standard deviation (SD) of participants scores in the social cognition tests; t-values and p-values of the independent-sample t-test (two-tailed) between adults' groups are reported (results significant at $p < 0.05$).

scores, to simplify analysis for following regression analysis, a Balanced Integration Score (BIS) was calculated for each participant, which reflects overall performance and efficiency since it takes into account both accuracy and RTs[38,39]. First, we converted individual IVR task's accuracy and RTs into z-scores within each age-group, in order to standardize them to a common scale and to balance speed and accuracy. The BIS score (Composite IVR) was calculated as follows: ZAccuracy - ZRTs. This calculation rewards both accuracy and speed, because subtracting a lower RTs z-score actually adds to the finale score[38,39]. Then, we run a Hierarchical Regression Analyses in all participants together, since IVR performance scores were transformed into z-scores and since the different age-related social cognition tests assessed the same social cognition components. The study employed a two-step hierarchical approach. Composite IVR score was entered in the model as the dependent variable. Concerning the first step, the Affective Component of ToM, the Cognitive Component of ToM and Emotion Recognition were entered in the model 1 as independent variables. In the second step, age was also entered in the model 2 as independent variables, to assess whether the addition of age would significantly improve the model. This approach allowed us to control age-related differences while allowing a combined analysis across all participants. However, because different standardized social cognition tests were administered to each age group, z-scoring within groups does not assume measurement invariance across ages. Therefore, the regression results should be interpreted as exploratory associations rather than evidence of convergent validity. Also, this approach might provide more robust and interpretable results compared to separate regression models for each group, given the modest sample size ($N \approx 30$ per group). Semi-partial correlations (Part²) were computed from the regression model to estimate the unique variance explained by each predictor[40]. Due to concerns about normality and sample size, bootstrap resampling (1000 samples) was applied to obtain more robust estimates. A significance threshold of $p = 0.05$ was set for statistical analysis. Adjusted $R^2$ values are reported for each model. In addition, p-values for individual predictors were adjusted using the Benjamini–Hochberg step-up procedure to control for multiple comparisons ($p_{BH}$), given the exploratory nature of this analysis[41]. Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 27.

## Results

### IVR tasks

Descriptive statistics of accuracy, RTs and learning coefficient of IVR tasks per group and block are reported in the Supplementary Information (Table S1).

The mean accuracy of the three age groups in the two blocks is presented in Fig. 2A. A 2 × 3 mixed-effects ANOVA showed a significant main effect of Age group [$F(2, 91) = 8.92$, $p < 0.01$, $\eta_p^2 = 0.16$]. Bonferroni corrected post-hoc comparisons revealed lower accuracy in early adolescents compared to both young adults ($p_{corr} <$ 0.001, d = -0.85, mean difference = -0.13, 95% CI for mean difference [-0.21, -0.05]) and adults ($p_{corr} = 0.003$, d = -0.74, mean difference = -0.11, 95% CI for mean difference [-0.19, -0.03]). The two groups of adults did not differ significantly ($p_{corr} = 0.99$, d = -0.11, mean difference = -0.02, 95% CI for mean difference [-0.10, 0.06]). In addition, there was a significant main effect of Scenario order [$F(1, 91) = 18.25$, $p < 0.01$, $\eta_p^2 = 0.17$], with lower accuracy in block 1 compared to block 2. No significant interaction of Age group x Scenario order [$F(2, 91) = 0.85$, $p = 0.43$, $\eta_p^2 = 0.02$] was present.

The pattern of results for the learning coefficient followed that of the results for the accuracy. Specifically, there was a significant main effect of Age group [$F(2, 91) = 8.14$, $p < 0.01$, $\eta_p^2 = 0.15$] which was defined by slower learning in early adolescents compared to both young adults ($p_{corr} < 0.001$, d = -0.83, mean difference = -0.13, 95% CI for mean difference [-0.21, -0.05]) and adults ($p_{corr} = 0.009$, d = -0.61, mean difference = -0.10, 95% CI for mean difference [-0.18, -0.02]); the two groups of adults did not differ significantly ($p_{corr} = 0.99$, d = -0.18, mean difference = -0.03, 95% CI for mean difference [-0.11, 0.05]). A significant main effect of Scenario order [$F(1, 91) = 15.46$, $p < 0.001$, $\eta_p^2 = 0.15$] revealed that learning was slower in block 1 than in block 2. There was no significant interaction of Age group x Scenario order [$F(2, 91) = 0.03$, $p = 0.97$, $\eta_p^2 < 0.001$]; see Fig. 2B.

A 2 × 3 mixed-effects ANOVA for RTs showed a significant interaction of Age group x Scenario order [$F(2, 91) = 3.10$, $p = 0.0498$, $\eta_p^2 = 0.06$], which was driven by a non-significant trend of faster RTs in the young adult's group in the second block compared with the first block ($p_{corr} = 0.11$, d = -0.42, mean difference = 0.59, 95% CI for mean difference [-0.33, 1.51]). There were no significant main effects of Age group [$F(2, 91) = 0.37$, $p = 0.69$, $\eta_p^2 = 0.01$], or Scenario order [$F(1, 91) = 2.94$, $p = 0.09$, $\eta_p^2 = 0.03$] (Fig. 2C).

To summarize, the group of early adolescents had lower performance and learning rate than both groups of adults, which in turn did not differ significantly between each other. In addition, the performance in the second block was significantly better than in the first one in all age groups.

### IVR task and social cognition tests association

The Hierarchical Regression Analysis revealed that model 1 was statistically significant [$F(3, 89) = 3.856$, $p = 0.012$] and explained approximately 11% of the variance in IVR performance ($R^2 = 0.115$, Adjusted $R^2 = 0.085$). Cognitive component of ToM and Emotion Recognition did not reach statistical significance (respectively, $\beta$ = -0.123, t = -0.965, $p = 0.33$, $p_{BH} = 0.33$, Part² = 0.009; $\beta = 0.163$, t = 1.633; $p = 0.11$; $p_{BH} = 0.16$, Part² = 0.027), while the Affective Component of ToM emerged as significant predictor ($\beta = 0.349$, t = 2.729, $p = 0.008$; $p_{BH} = 0.024$, Part² = 0.074) uniquely accounting for approximately 7% of the variance in IVR performance; this indicates that higher scores on this measure was associated with better performance in the IVR task (Fig. 3A).

In model 2, age was added as an additional predictor. The model was significant [$F(4, 88) = 2.869$, $p = 0.028$], and explained an equal proportion of variance ($R^2 = 0.11$; Adjusted $R^2 = 0.075$), whose change was not statistically significant [$\Delta R^2 = 0.000$, $F(1, 88) = 0.036$, $p = 0.85$]. In this model, the Affective Component of ToM remained a significant predictor as in model 1 ($\beta = 0.348$, t = 2.71, $p = 0.008$, $p_{BH} = 0.032$, Part² = 0.074), and it uniquely explained approximately the 7% of the variance, while the Cognitive Component of ToM, Emotion Recognition and age were not significant predictors (respectively, $\beta$ = -0.123, t = -0.959, $p = 0.34$, $p_{BH} = 0.45$, Part² = 0.009;

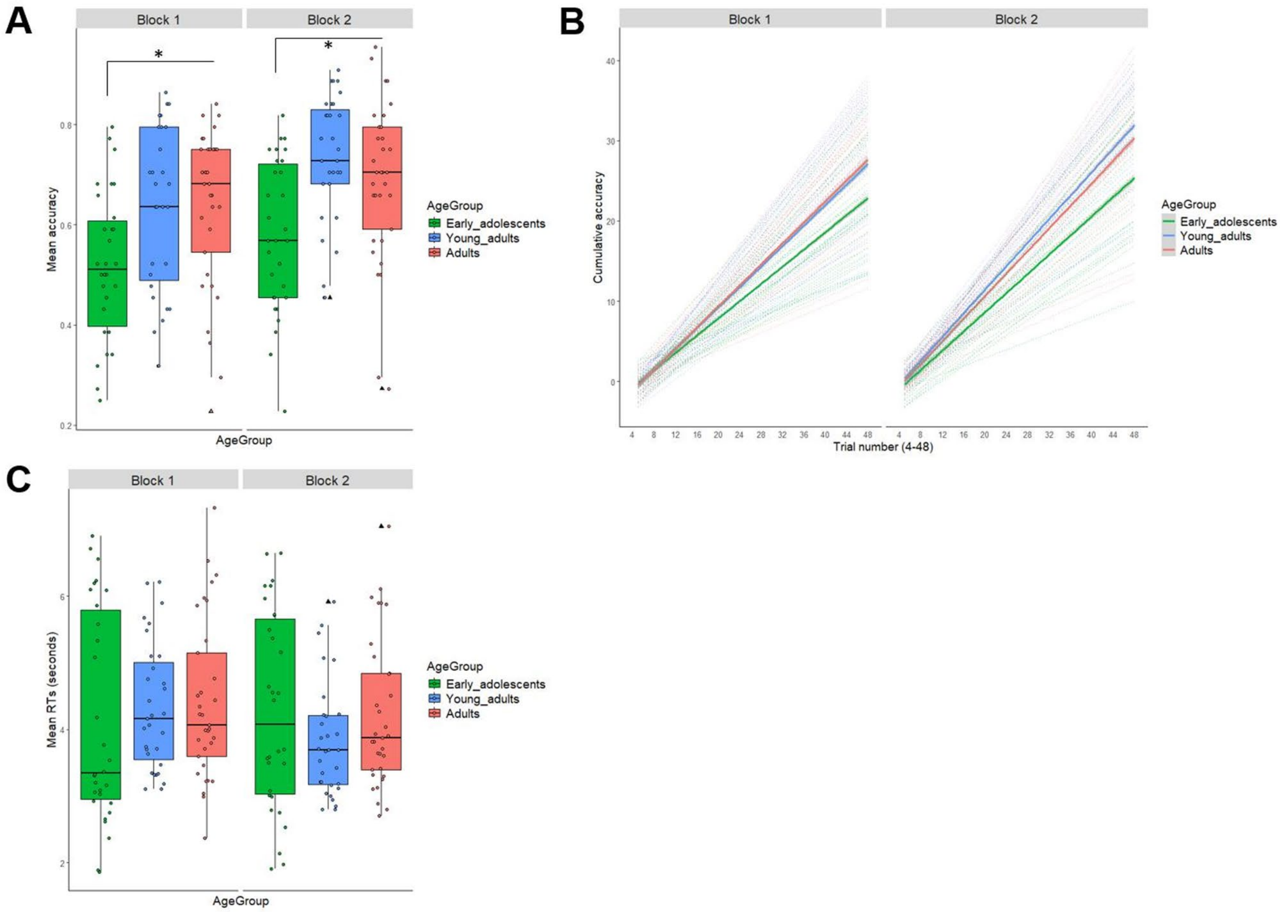


**Fig. 2**. Accuracy and RTs in the IVR tasks presented by age group and task block (IVR scenario order). (**A**) Mean accuracy in each age group and task block; (**B**) Linear fit for cumulative accuracy by trial (learning rate); the solid lines represent the group average and the grey bands around the lines represent the standard error, while the dotted lines represent the single participants' fits; (**C**) Mean RTs in the VR tasks presented by age group and task block (VR scenario order). In the boxplots (panels **A**, **C**), hinges of the box represent the first and the third quartile, with the line in the middle of the box representing the median and the bottom and top whiskers representing the smallest and the largest value no further than 1.5*Interquartile Range (IQR) from the lower and the upper hinges, respectively, the dots represent single participant's values, and the triangles represent outliers. *Significant difference.

$\beta = 0.164$, $t = 1.631$, $p = 0.11$, $p_{BH} = 0.21$, $Part^2 = 0.026$; $\beta = -0.019$, $t = -0.190$, $p = 0.85$, $p_{BH} = 0.85$, $Part^2 = 0.000$); this indicates that the previously observed effect are not accounted for by shared variance with age.

To assess the robustness of these findings, bootstrap confidence intervals (based on 1,000 samples) were examined. In model 1, the Affective Component of ToM remained a significant predictor ($B = 0.485$, $p = 0.009$; $p_{BH} = 0.027$), while Emotion Recognition showed a trend to significance ($B = 0.249$, $p = 0.062$, $p_{BH} = 0.09$). The Cognitive Component of ToM was not statistically significant ($B = -0.174$, $p = 0.34$, $p_{BH} = 0.34$) (Fig. 3B). In model 2, bootstrap estimates showed that the Affective Component of ToM remained a significant predictor ($B = 0.484$, $p = 0.01$, $p_{BH} = 0.04$), Emotion Recognition showed a trend to significance ($B = 0.251$, $p = 0.066$, $p_{BH} = 0.13$) and both the Cognitive Component of ToM and age were not statistically significant (respectively, $B = -0.174$, $p = 0.34$, $p_{BH} = 0.46$; $B = -0.002$, $p = 0.86$, $p_{BH} = 0.86$), confirming that the addition of age did not provide a reliable improvement to the model (Fig. 3B).

Collinearity diagnostics indicated no multicollinearity issues, with all VIF values below 1.643. The Durbin-Watson value (1.766) indicated no autocorrelation in the residuals.

To summarize, Affective Component of ToM, as assessed by conventional social cognition subtests, such as the ToM B subtest of the NEPSY-II for early adolescents and the Affective Component of the Faux Pas Test for adults, revealed significant predictors of IVR performance. When age was added to the model, the predictive effect of the Affective Component remained and age did not significantly contribute to the model, nor did it reduce overall explanatory power.

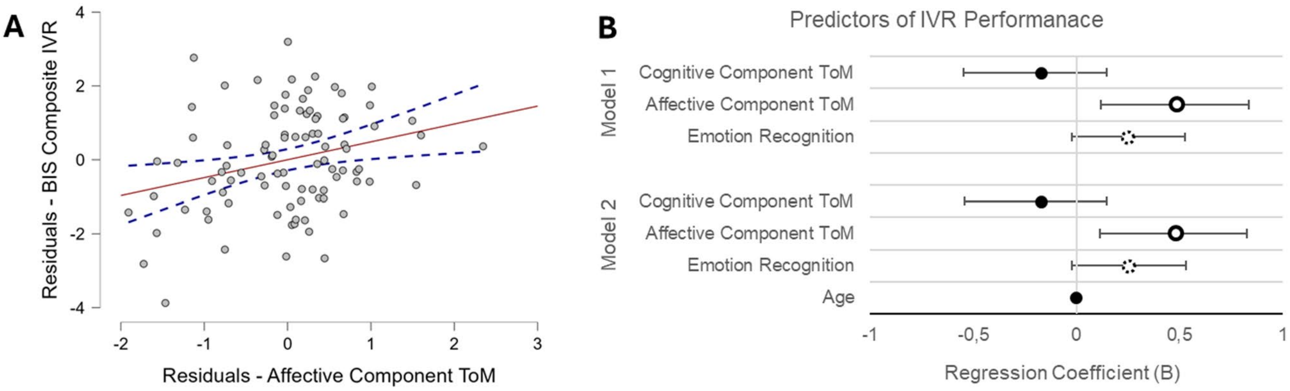


**Fig. 3**. Hierarchical regression analysis plots. (**A**) Partial regression plot showing the relationship of the Affective Component of ToM with the Composite IVR (BIS) after controlling for the other predictors included in model 1 of the first regression analysis. The X-axis represents the partial values of the predictor Affective-ToM and the Y-axis the partial values of the dependent variable Composite IVR (BIS), after removing the effects of the other predictors included in model 1. The slope of the regression line corresponds to the partial regression coefficient (β). Dotted lines indicate the 95% confidence intervals (CIs). (**B**) Unstandardized regression coefficients with 95% bias-corrected bootstrapped confidence intervals (1000 samples). Error bars represent 95% bootstrapped CIs. In the plots, the vertical line at zero indicates no effect. Predictors with CIs that do not cross zero are statistically significant at $p < 0.05$ (half-full dots); Predictors that do cross zero are not statistically significant at ($p > 0.05$; full dots); Predictors represented by dashed dots revealed a trend to significance (B: Model 1, $p = 0.062$; Model 2, $p = 0.066$).

## Discussion

The present study aimed to investigate whether the creation of social prediction models in neurotypical individuals changes across age (early adolescents, young adults, and adults). We specifically aimed to test the hypothesis that the ability to form and implement predictive internal models of others' actions evolves across age and improves with practice, as assessed by an immersive IVR system that facilitates and enhances the creation of forward models during social perception. Additionally, the present study revealed the capacity of IVR task to highlight statistically significant differences in terms of social predictions across age groups and provided exploratory evidence of direct associations with social cognition measures. Also, as reported in detail in supplementary materials, the usability of the IVR task was confirmed.

### Updating prediction models throughout experience

The results of the present study showed that IVR performance of all participants significantly improved from the first to the second block, suggesting learning and refinement of predictive models about others' intentions (Fig. 2A, B). To infer avatars' intentions and to succeed in IVR tasks, participants had to implicitly associate avatar's movement toward the object (and afterwards, the avatar's appearance) with object preference, relying on bottom-up sensorial information. As these associations were proven accurate, participants recognized contextual regularities in the IVR scenario and developed internal models of the avatars' intentions, using top-down predictions adopted as priors to anticipate behavior[42]. Once the predictive model was deemed adequate by practice, the contextual priors induced more accurate expectations and the performance improved, as demonstrated by the progressive increase in learning rate both within and between blocks (Fig. 2B). Consistently, it has been argued that implicitly learnt contextual priors may guide action recognition prevailing actual kinematic information[43–45]. These assumptions are in line with theoretical models of social prediction, which suggest that prior expectations (priors) are based on the formerly observed statistical regularities intrinsic to sensorial input (for example, action kinematics or facial expression) experienced in the natural world[5], and the context of social events or actions is one of the sources of these priors[3,6].

These models are grounded in a general hierarchical model of perception and cognition, such as The Bayesian Predictive Coding (BPC) framework[9,46]. According to this framework, the brain operates as a hierarchical probabilistic inference system that continuously predicts incoming sensory input on the basis of internal models of that input, producing top-down predictions[4,9,46]. When sensory data aligns with these predictions, the system maintains its current model; when discrepancies arise, prediction errors prompt model updating. In this study, the decreasing trend in reaction times and the progressive increase of learning rate are in line with the assumption that the predictive model needs to be revised to reflect each sample of bottom-up evidence until uncertainty is sufficiently minimized and accurate decisions can be made, inducing participants to increasingly rely on predictive priors as internal models are stabilized[47]. According to the predictive model framework, the confirmation of prior expectations enhances top-down gain, not only facilitating faster evidence accumulation but also speeding decision-making[47]. Together with this, the IVR tasks prompted the creation of the social predictions online, since the information about the avatars' object preferences was available only during the task. In this context, despite the general age-related preferences for relying on priors vs. sensorial information, all age groups showed better performance in the second block of the task. This occurred despite the fact that the avatar-object preferences were different in the two blocks of the task, suggesting that the creation of the predictive models based on the context-related cues subsequently consolidates, facilitating the same process even when the cues are different. This outcome aligns with results of a recent study which designed a Virtual Reality Social Prediction Improvement and Rehabilitation Intensive Training (VR-Spirit) to specifically target social prediction difficulties in children, adolescents, and young adults with congenital cerebellar malformations[17]. Specifically, the authors revealed that VR-Spirit administered over two weeks was effective in improving social prediction of the avatars' intentions even when exposed to a different and novel VR scenario[17]. Overall, both these findings suggested that practicing with IVR task successfully boost and consolidate the ability to flexibly switch predictive models among avatars and changing scenarios. On the other hand, while the increased accuracy and learning rate observed in the second scenario points to the training-related refinement of the flexibility of the predictive models, the present study cannot directly indicate the underlying mechanism of this effect. In particular, the present experimental design cannot disentangle between the general refinement of the social prediction models, the increased familiarity with the layout of the IVR task in the second block, or the transfer of specific previously learned associations. The latter seems the least likely because the two scenarios differed not only in the avatar-object associations but also in the avatars' and objects' appearance per se. In a similar vein, since participants familiarized with the IVR scenario in the practice block before the experimental scenarios, the difference between the two blocks is less likely to reflect familiarity with the IVR scenario. Indeed, not only performance accuracy, but also the learning rate increased from block 1 to block 2. However, we cannot definitively rule out a cumulative effect of familiarization across blocks.

## Updating prediction models throughout development

The reliance on top-down information has been shown to increase with age[11]. Nevertheless, thus far no study has specifically investigated top-down prediction processes across ages including early adolescents and using ecological IVR environments.

In line with our hypothesis, we found that early adolescents performed significantly worse and showed a lower learning rate compared to both adult groups, who did not differ significantly from each other (Fig. 2A, B). Although the use of additional relevant information, such as eye-tracking data, would have given direct supporting evidence, the present results might support the view that in early developmental stages, such as childhood and early adolescence, individuals rely more on bottom-up sensory information to interpret the world, which is perceived as chaotic and unpredictable due to limited prior knowledge and expectations about it[11]. Although no neuroimaging methods were employed to directly examine the relationship between IVR task performance and underlying neural mechanisms, some speculative considerations may be advanced to guide future research. Over time, continuous exposure to environmental regularities enables the brain to generate and refine the internal models of the world, facilitating the integration of prior knowledge into perception and cognition and shifting toward greater reliance on top-down predictive mechanisms[29]. Accordingly, older adults have been shown to derive greater benefit from contextual cues in object recognition tasks involving perceptual ambiguity and rely more on top-down processes[12]. Supporting this, as the brain develops, beginning from early adulthood, it generates progressively more sophisticated predictive internal models, thereby reducing the complexity of ongoing computations and streamlining perceptual inference[48].

In this context, peak brain plasticity is known to unfold until late adolescence[49], and the key brain region involved in processing prior expectations (the ventro-medial prefrontal cortex, vmPFC) develops structurally and functionally throughout adolescence[50], which is acknowledged as a critical age for the refinement of social cognition. While prefrontal areas' role in such processes is acknowledged[51], much of recent literature is pointing to a role of the cerebellum as a hub in the predictive brain system[52,53]. Specifically, thanks to the multiple connections the cerebellum exerts with diffuse brain areas, it functions as a broad regulator and integrator across multiple brain networks by generating internal predictive models based on the integration of contextual and sensorial cues[52,54]. As for the social domain, recent advances propose that the cerebellum enables the smoothing of social-cognitive processes by upholding optimal predictions about forthcoming and future social interaction[55]. Accordingly, recent fMRI data on young adults showed a functional double dissociation across brain regions supporting social prediction models creation vs. feedback-based priors updating, with distinctive cerebellar patterns. The study revealed that the generation of a mental model to guide prediction is supported by premotor/parietal cortical systems and sensorimotor-related cerebellar regions, whereas the feedback-based incorporation of prediction error to update the model is supported by prefrontal/mentalizing areas and cerebellar posterior associative regions[56]. As concerned with developmental stages, it has been shown that patients with congenital cerebellar alterations were less accurate in making predictions when the available sensorial information was not sufficient to inform about the incoming events, resulting in an impaired processing of contextual information[18,57]. These results, together with the scarcity of studies that investigated prediction processes in developmental stages, point to the necessity to deepen both behavioral and neural phenomena related to age-related predictive processing, broadening the investigation into cerebellar-cerebral networks.

## Predictive models and social cognition

Overall, although the study's conclusions are speculative and need further verification by integrating appropriate neural measures (e.g., EEG or fMRI data), our findings might be in line with a predictive coding account of social cognition[4,29]. In this vein, the regression analysis results provided exploratory evidence that the IVR task might tap into similar underlying social cognition abilities, namely affective component of ToM, indicating that the IVR task might accurately capture real-life-like social abilities. Importantly, while the semi-partial correlation indicated that the affective ToM uniquely accounted for approximately 7% of the variance in IVR performance, effect sizes of this magnitude are common in individual differences research, particularly when multiple predictors are included in the model[40]. Although this represents a modest effect size, it suggests a specific contribution of affective processes to performance on the IVR task, beyond the influence of other social cognition measures included in the model. This result is in line with a study that assessed the efficacy of VR-Spirit on cerebellar patients' social prediction abilities, which yielded an improvement in social cognition tests outcome (ToM subtest of NEPSY-II) at the end of the two weeks training, as well as an improvement in the use of contextual predictions in a computer-based action prediction task at a 2-months follow-up, proving a generalizing effects of VR-Spirit on social cognition abilities[17]. Generally speaking, formation and use of prediction models built on the consistent exposure to contextualized behavioral regularities during the VR task seems to be related with participants' capacity to develop conceptual representations about others' mental and affective states, as assessed with conventional social cognition tests (ToM B subtest of the NEPSY-II for early adolescents; affective component of the Faux Pas Test for adult groups).

These results agree with previous studies suggesting that VR is a useful tool to measure similar constructs as those measured by clinical or experimental measures[58,59]. Also, as reported in detail in supplementary materials, the use of the IVR task is feasible in all age groups: IVR tool is engaging, since it is perceived as easy to use and satisfying, and IVR scenarios elicit presence and agency, inducing embodiment of the virtual hands. In this vein, as IVR technology keeps improving and advancing, it conveys progressively more accurate and engaging platforms that faithfully emulate real-world complexities and cognitive function. For example, several studies revealed that IVR provides greater cognitive benefits than non-immersive VR[60,61]. Indeed, compared to non-immersive VR paradigms, IVR interventions were more suitable for the development of complex skills in children and adolescents with Autism Spectrum Disorder (ASD)[15], and it might improve theory of mind in neurodevelopmental clinical conditions like ASD and Williams Syndrome[62,63]. Similarly, previous studies showed that VR-Social Cognition Training improved the ability to recognize emotions, to make inference on others' intentions, and to make daily social functioning conversation in both adults and children with ASD[15,64,65].

The use of IVR for cognitive and mental health applications among older adults has been slightly investigated compared to research focused on children, adolescents and young adults. Nevertheless, the current state of the art defines IVR as a promising tool with the potential to test, screen and enhance a wide variety of cognitive and affective functions (spatial cognition, memory, executive functions, activities of daily living) in the aging population [for review, see[19]. In this regard, IVR tools may promote the implementation of next generation testing in social-cognitive domains, as well as novel intervention and rehabilitation techniques, encompassing the whole lifespan. Another advantage of IVR-training tasks could be its use in combination with other restorative rehabilitation interventions, for example coupling IVR-training tasks and Non-Invasive Brain Stimulation (NIBS) to boost the efficacy of intervention programs is a promising tool for rehabilitation[6]. In this vein, future studies should take into account cerebellar stimulation, which is receiving a growing consensus to induce the generation of predictive internal models of others' behaviors and therefore to improve social abilities[6,66].

## Limitations

Although promising, the results of the present study must be interpreted within the context of some limitations. First, participants' age could be considered only as a grouping factor, rather than as a continuous predictor, primarily due to the low age variability in the early adolescents' group (11–13 years old), thus limiting the representativeness of this developmental stage. While the age of the present sample was determined by the available possibilities for participant recruitment, it is important to acknowledge that early adolescence and adolescence are characterized by rapid cognitive, physical, physiological, and neuroplastic changes[49], and therefore, considering a broader age range among early adolescents and adolescents might have allowed to capture continuous development of the social prediction abilities. Another limitation concerns the lack of information about the participants' socioeconomic status, which is a known factor that influences cognition, especially during development[67–69]. The socioeconomic status of the two groups of adults might be considered rather homogeneous: the two groups of adults were recruited at a university and in a research center, thus suggesting a common social environment from which they were selected. Secondly, as indexed by the total years of completed education reported in Table 1, most participants in the two groups of adults had university-level education (bachelor's or master's degrees), and the three participants who did not hold a university degree (1 adult and 2 younger adults) completed high school education. On the other hand, less was known about the socioeconomic status of the early adolescents' group. All of them were recruited into general-education secondary schools within one city in a relatively rich area of the country (Udine), but we did not collect any information about their families' income, parents' education, or other related data, which might limit the interpretation of our results. Also, future studies should provide IQ estimates based on the same test and normative scales, to reliably control its influence on task performance and learning speed. Likewise, another limitation to consider concerns the lacking acquisitions of other likely relevant data, such as broader cognitive measures to control for domain-general processes (e.g., processing speed and working memory), as well as eye-tracking data, which could support the interpretation of the main results (differences in the performance and the learning rate in the IVR tasks) as related to the different weighting of predictions and sensorial information. Additionally, although the present study relied on the implicit learning of the statistical regularities of avatar-object associations rather than on explicit learning, future studies should consider including a manipulation check questionnaire to assess whether participants explicitly learned the avatar-object category association rules, thereby providing further insight into participants' strategies. Lastly, samples size should be increased to allow for more robust statistical analysis to appropriately assess convergent validity of the novel IVR task with established social cognition measures.

## Conclusion

The results of the present study suggest that social prediction abilities improve from early adolescence to adult age, as individuals learn to recognize statistical regularities of events in the surrounding environment and incorporate them in accurate experience-based internal models of the world. Therefore, our study supports a developmental model of social prediction in which the implementation of predictive internal models improves with age and experience, consistent with predictive coding accounts[9]. Indeed, our findings support the presence of age differences in social prediction abilities. In detail, early adolescents, compared to younger adults and adults, might show stronger reliance on sensorial information and weaker reliance on prior expectations during the creation of the predictive models based on contextual cues, as revealed by IVR results. Although speculative and needing further proof, these findings shed light on how predictive coding mechanisms might contribute to social cognitive development and how internal models are shaped through both learning and development. Importantly, in all age groups the IVR task was effective in boosting social prediction of the avatars' intentions, and these learning effects transferred across different scenarios. Noteworthy, our results provide exploratory evidence of association between the novel IVR task and conventional social cognition tests, pointing to a generalized effect of IVR employment. Also, the use of the IVR task is feasible in all age groups.

## Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 5 September 2025; Accepted: 11 May 2026
Published online: 15 May 2026

## References

1. Brown, E. C. & Brüne, M. The role of prediction in social neuroscience. *Front. Hum. Neurosci.* **6**, 1–19 (2012).
2. Todd, J., Frost, J. D., Yeark, M. & Paton, B. Context is everything: How context shapes modulations of responses to unattended sound. *Hear Res* **399**, (2021).
3. Kilner, J. M. More than one pathway to action understanding. *Trends Cogn. Sci.* **15**, 352–357 (2011).
4. Rapaport, H. & Sowman, P. F. Examining predictive coding accounts of typical and autistic neurocognitive development. *Neuroscience and Biobehavioral Reviews* **167** (2024).
5. Huang, Y. & Rao, R. P. N. Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science* **2**, 580–593 (2011).
6. Oldrati, V. et al. How social is the cerebellum? Exploring the effects of cerebellar transcranial direct current stimulation on the prediction of social and physical events. *Brain Struct. Funct.* **226**, 671–684 (2021).
7. Barbalat, G., Bazargani, N. & Blakemore, S. J. The influence of prior expectations on emotional face perception in adolescence. *Cereb. Cortex*. **23**, 1542–1551 (2013).
8. Amoruso, L. & Finisguerra, A. Low or High-Level Motor Coding? The Role of Stimulus Complexity. *Frontiers in Human Neuroscience* **13**, 332 (2019)
9. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
10. Chan, J. S. et al. Predictive Coding Over the Lifespan: Increased Reliance on Perceptual Priors in Older Adults—A Magnetoencephalography and Dynamic Causal Modeling Study. *Front Aging Neurosci* **13**, (2021).
11. Trapp, S., Guitart-Masip, M. & Schröger, E. A link between age, affect, and predictions? *European Journal of Ageing* **19**, 945–952 (2022).
12. Lai, L. Y., Frömer, R., Festa, E. K. & Heindel, W. C. Age-related changes in the neural dynamics of bottom-up and top-down processing during visual object recognition: an electrophysiological investigation. *Neurobiol. Aging*. **94**, 38–49 (2020).
13. Mäkinen, H., Haavisto, E., Havola, S. & Koivisto, J. M. User experiences of virtual reality technologies for healthcare in learning: an integrative review. *Behav. Inform. Technol.* **41**, 1–17 (2020).
14. Tieri, G., Morone, G., Paolucci, S. & Iosa, M. Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. *Expert Review of Medical Devices* **15**, 107–117 (2018).
15. Yang, X. et al. Effectiveness of Virtual Reality Technology Interventions in Improving the Social Skills of Children and Adolescents With Autism: Systematic Review. *J Med. Internet Res* **27**, (2025).
16. Zhang, F., Zhang, Y., Li, G. & Luo, H. Using Virtual Reality Interventions to Promote Social and Emotional Learning for Children and Adolescents: A Systematic Review and Meta-Analysis. *Children* **11**, 41 (2024).
17. Butti, N. et al. Feasibility and Efficacy of a Virtual Reality Social Prediction Training in Children and Young Adults with Congenital Cerebellar Malformations. *J. Autism Dev. Disord*. **55**, 2463–2479 (2025).
18. Urgesi, C. et al. 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 (2021).
19. Skurla, M. D. et al. Virtual reality and mental health in older adults: a systematic review. *International Psychogeriatrics* **34** 143–155 (022)
20. Xu, J. et al. Effectiveness of Virtual Reality-Based Well-Being Interventions for Stress Reduction in Young Adults: Systematic Review. *JMIR Mental Health* **11**, 29:e52186. (2024).
21. Tieri, G., Tidoni, E., Pavone, E. F. & Aglioti, S. M. Mere observation of body discontinuity affects perceived ownership and vicarious agency over a virtual hand. *Exp. Brain Res.* **233**, 1247–1259 (2015).
22. Gil-Gómez, J. A. et al. USEQ: A short questionnaire for satisfaction evaluation of virtual rehabilitation systems. *Sensors (Switzerland)* **17**, (2017).
23. Hart, S. G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proc Hum Factors Ergon Soc Annu Meet. **50**, 904–8. (2006)
24. Urgesi, C., Campanella, F. & Fabbro, F. *NEPSY-II, Contributo Alla Taratura Italiana.*. (Giunti O.S., Milano, Italy, (2011).
25. Liverta Sempio, O., Marchetti, A. & Lecciso, F. *Faux Pas: traduzione italiana*. Milan: Theory of Mind Research Unit; Department of Psychology; Catholic University of the Sacred Heart (2005).
26. Stone, V. E., Baron-Cohen, S., & Knight, R. T. Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci.* **10**, 640–656 (1998).
27. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. The Reading the Mind in the Eyes Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *J. Child Psychol. Psychiatry*. **42**, 241–251 (2001).
28. Serafin, M. & Surian, L. Il Test degli Occhi: uno strumento per valutare la 'teoria della mente'. *Giornale italiano di psicologia*. **31**, 839–862 (2004).
29. Bar, M. The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* **11**, 280–289 (2007).

30. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res. Methods* 175–191 (2007).
31. John, R. J. Raven Progressive Matrices. In Handbook of Nonverbal Assessment. (ed McCallum, R. S.). Springer, Boston, MA, (2003).
32. Raven, J. C. *Progressive Matrices. Set A, Ab, B. Board and Book Form* (Lewis, 1947).
33. Miller, G. A. & Chapman, J. P. Misunderstanding analysis of covariance. *J. Abnorm. Psychol.* **110**, 40–48 (2001).
34. Suckling, J. Correlated covariates in ANCOVA cannot adjust for pre-existing differences between groups. *Schizophrenia Research* **126**, 310–311 (2011).
35. Maselli, A. & Slater, M. The building blocks of the full body ownership illusion. *Front Hum. Neurosci* **7**, (2013).
36. Butti, N. et al. News insights into social cognition in Williams syndrome from a comprehensive assessment and a virtual reality task. *Sci Rep* **14**, (2024).
37. R Core Team. R: A language and environment for statistical computing. Vienna, Austria. (2020). https://www.r-project.org/
38. Liesefeld, H. R. & Janczyk, M. Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behav. Res. Methods*. **51**, 40–60 (2019).
39. Liesefeld, H. R., Fu, X. & Zimmer, H. D. Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *J. Exp. Psychol. Learn. Mem. Cogn.* **41**, 1140–1151 (2015).
40. Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Routledge, 2003).
41. Benjamini, Y. Discovering the False Discovery Rate. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **72**, 405-416 (2010).
42. Seriès, P. & Seitz, A. R. Learning what to expect (in visual perception). *Front Hum Neurosci* **7**, 668 (2013).
43. Amoruso, L., Finisguerra, A. & Urgesi, C. Tracking the time course of top-down contextual effects on motor responses during action comprehension. *J. Neurosci.* **36**, 11590–11600 (2016).
44. Amoruso, L., Finisguerra, A. & Urgesi, C. Autistic traits predict poor integration between top-down contextual expectations and movement kinematics during action observation. *Sci Rep* **8**, (2018).
45. Amoruso, L., Finisguerra, A. & Urgesi, C. Contextualizing action observation in the predictive brain: Causal contributions of prefrontal and middle temporal areas. *Neuroimage* **177**, 68–78 (2018).
46. Brown, E. C. & Brüne, M. The role of prediction in social neuroscience. *Front. Hum. Neurosci.* **7**, 668 (2012).
47. Dunovan, K. E., Tremel, J. J. & Wheeler, M. E. Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia* **61**, 210–221 (2014).
48. Moran, R. J., Symmonds, M., Dolan, R. J. & Friston, K. J. The Brain Ages Optimally to Model Its Environment: Evidence from Sensory Learning over the Adult Lifespan. *PLoS Comput. Biol* **10**, (2014).
49. Ismail, F. et al. Cerebral plasticity: Windows of opportunity in the developing brain. *Eur. J. Pediatr. Neurol.* **21**, 23–48 (2017).
50. Nelson, E. E. & Guyer, A. E. The development of the ventral prefrontal cortex and social flexibility. *Developmental Cognitive Neuroscience* **1,** 233–245 (2011).
51. Wang, B. A., Drammis, S., Hummos, A., Halassa, M. M. & Pleger, B. Modulation of prefrontal couplings by prior belief-related responses in ventromedial prefrontal cortex. *Front Neurosci* **17**, (2023).
52. Leggio, M. & Molinari, M. Cerebellar Sequencing: a Trick for Predicting the Future. *Cerebellum* **14**, 35–38 (2015).
53. Vandervert, L. et al. The Evolution of the Optimization of Cognitive and Social Functions in the Cerebellum and Thereby the Rise of Homo sapiens Through Cumulative Culture. *Cerebellum* **23**, 1981–1992 (2024).
54. Ishikawa, T., Tomatsu, S., Izawa, J. & Kakei, S. The cerebro-cerebellum: Could it be loci of forward models? *Neuroscience Research* **104**, 72–79 (2016).
55. Van Overwalle, F. et al. Consensus Paper: Cerebellum and Social Cognition. *Cerebellum* **19**, 833–868 (2020).
56. Wang, Y., Davidow, J. Y., Lane, R. D. & Satpute, A. B. Disentangling Prediction and Feedback in Social Brain Networks: A Predictive Processing Approach. Preprint at bioRxiv (2025).
57. Friston, K. Prediction, perception and agency. *International Journal of Psychophysiology* **83**, 248–252 (2012).
58. Henry, M., Joyal, C. C. & Nolin, P. Development and initial assessment of a new paradigm for assessing cognitive and motor inhibition: The bimodal virtual-reality Stroop. *J. Neurosci. Methods*. **210**, 125–131 (2012).
59. Corriveau Lecavalier, N., Ouellet, É., Boller, B. & Belleville, S. Use of immersive virtual reality to assess episodic memory: A validation study in older adults. *Neuropsychol. Rehabil*. **30**, 462–480 (2020).
60. Chatterjee, K. et al. Immersive Virtual Reality for the Cognitive Rehabilitation of Stroke Survivors. *IEEE Trans. Neural Syst. Rehabil. Eng.* **30**, 719–728 (2022).
61. Cikajlo, I. & Peterlin Potisk, K. Advantages of using 3D virtual reality based training in persons with Parkinson's disease: A parallel study. *J Neuroeng. Rehabil* **16**, (2019).
62. Amoruso, L. et al. Contextual priors do not modulate action prediction in children with autism. *Proceedings of the Royal Society B: Biological Sciences* 286, (2019).
63. Sparaci, L., Stefanini, S., D'Elia, L., Vicari, S. & Rizzolatti, G. What and Why understanding in autism spectrum disorders and williams syndrome: Similarities and differences. *Autism Res.* **7**, 421–432 (2014).
64. Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T. & Chapman, S. B. Virtual reality social cognition training for young adults with high-functioning autism. *J. Autism Dev. Disord*. **43**, 34–44 (2013).
65. Didehbani, N., Allen, T., Kandalaft, M., Krawczyk, D. & Chapman, S. Virtual Reality Social Cognition Training for children with high functioning autism. *Comput. Hum. Behav.* **62**, 703–711 (2016).
66. Oldrati, V. et al. Excitatory cerebellar transcranial direct current stimulation boosts the leverage of prior knowledge for predicting actions. *Soc Cogn. Affect. Neurosci* **19**, (2023).
67. Bradley, R. H. & Corwyn, R. F. Socioeconomic status and child development. *Annu. Rev. Psychol.* **53**, 371–399 (2002).
68. Na'amnih, W. et al. Associations of early-life exposures and socioeconomic status with cognitive development at preadolescence. *Pediatr. Neonatol.* **64**, 192–200 (2023).
69. González, L. et al. Socioeconomic position, family context, and child cognitive development. *Eur. J. Pediatr.* **183**, 2571–2585 (2024).



## Acknowledgements
The authors are grateful to the participants of the study. The authors thank the Scuola Media Statale di Pasian di Prato (UD) and the Comitato di Udine della Croce Rossa Italiana for their contribution to participants recruitment.


## Author contributions
Siciliano L. and Pyasik M.: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing—Original draft, Visualization; Tieri G.: Conceptualization, Methodology and Software (designed and realized the applications in virtual reality), Investigation, Data curation, Writing—Review and Editing; Ciot

V., Salera C., Urbini N., Oldrati V., Ciricugno A and Olivito G.: Formal analysis, Investigation, Data curation, Writing—Review and Editing; Borgatti R. and Cattaneo Z.: Conceptualization, Methodology, Resources, Writing—Review and Editing, Funding acquisition; Iosa M.: Conceptualization, Methodology, Software, Resources, Writing—Review and Editing; Urgesi C.: Conceptualization, Methodology, Software, Resources, Supervision, Writing—Review and Editing, Funding acquisition; Leggio M.: Conceptualization, Methodology, Software, Resources, Supervision, Writing—Review and Editing, Project administration, Funding acquisition.

## Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-53205-z.

**Correspondence** and requests for materials should be addressed to L.S. or M.P.

**Reprints and permissions information** is available at www.nature.com/reprints.

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