Title
Attribution of intentions in autism spectrum disorders: a study of event-related potentials

Running Title
Attribution of intentions ERP in autism

Authors
João Miguel Fernandes[1,2]
Sara Soares[3,4]
Ricardo Lopes[2,3]
Rita Jerónimo[3]
J. Bernardo Barahona-Corrêa[1,5] *

Institutional Affiliations
1. NOVA Medical School | Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisbon, Portugal
2. CADIn – Neurodevelopment & Inclusion, Non-Profit Association, Cascais, Portugal
3. Instituto Universitário de Lisboa (Iscte–IUL), CIS_Iscte, Lisbon, Portugal
4. Unité de Recherche en Neurosciences Cognitives (Unescog), Center for Research in Cognition & Neurosciences (CRCN), Université Libre de Bruxelles (ULB), Brussels, Belgium
5. Champalimaud Research & Clinical Centre, Champalimaud Centre for the Unknown, Lisbon, Portugal;
* corresponding author; please address correspondence to:
bernardo.correa@research.fchampalimaud.org

Authorship
JMF and JBB-C conceived and planned the study. SS, RL and RJ helped in designing and implementing the experimental set-up. JMF and SS were responsible for the recruitment of participants and acquisition of data. JMF carried out data analysis, with all authors contributing to the interpretation of data. JMF was responsible for the initial draft of the manuscript which was subsequently revised by the remaining authors. JBB-C and RJ supervised all stages of the work. All authors gave final approval of the version to be published.

Acknowledgments
The authors would like to thank Damien Vistoli (Université Cote d'Azur, Laboratoire d'Anthropologie et de Psychologie Cognitives et Sociales [LAPCOS], Nice, France) and Eric Brunet-Gouet (Centre Hospitalier de Versailles, Versailles, France, Inserm CESP [DisAp/DevPsy]) for generously providing us with his stimulus set of cartoon strip vignettes of the Comic Strips Task, Jeff Miller (Department of Psychology, University of Otago, Dunedin, New Zealand) for generously providing the statistical software for unequal group analyses, and Sofia Frade (Instituto Universitário de Lisboa [Iscte–IUL], CIS–Iscte, Lisbon, Portugal) for her assistance in ERP data collection. This work has been partially funded by a research grant from the Portuguese Society of Psychiatry and Mental Health (Sociedade Portuguesa de Psiquiatria e Saúde Mental). JBB-C is supported by grant FCT-PTDC/MEC-PSQ/30302/2017-IC&DT-LISBOA-01-0145-FEDER, funded by national funds from FCT/MCTES and co-funded by FEDER, under the Partnership Agreement Lisboa 2020 - Programa Operacional Regional de Lisboa.

Conflict of Interest
The authors declare that they have no conflict of interest related to this work.

Lay Summary
To understand why autistic people have difficulties in inferring others' intentions, we asked participants to judge the congruence of the endings of comic strips depicting either intentional actions (e.g., fetching a chair to reach for something) or situations solely following physical rules (e.g., an apple falling on someone's head), while their electrical brain activity was recorded. Autistic individuals had more difficulties in inferring intentions than neurotypical controls, which may reflect impaired attention and contextual integration of social cues.

Abstract (150–250 words)
Autism spectrum disorder (ASD) is characterized by social cognition deficits, including difficulties inferring the intentions of others. Although deficits in attribution of intentions have been consistently replicated in ASD, their exact nature remains unexplored. Here we registered the electrophysiological correlates of a non-verbal social cognition task to investigate attribution of intentions in autistic adults. Twenty-one male autistic adults and thirty male neurotypical volunteers performed a comic strips task depicting either intentional action (AI) or physical causality with or without human characters, while their electroencephalographic signal was recorded. Compared to neurotypical volunteers, autistic participants were significantly less accurate in correctly identifying congruence in the AI condition, but not in the physical causality conditions. In the AI condition a bilateral posterior positive event-related potential (ERP) occurred 200 to 400 ms post-stimulus (the ERP intention effect) in both groups. This waveform comprised a P200 and a P300 component, with the P200 component being larger for the AI condition in neurotypical volunteers but not in autistic individuals, who also showed a longer latency for this waveform. Group differences in amplitude of the ERP intention effect only became evident when we compared autistic participants to a subgroup of similarly performing neurotypical participants, suggesting that the atypical ERP waveform in ASD is an effect of group, rather than a marker of low task performance. Together, these results suggest that the lower accuracy of the ASD group in the AI task may result from impaired early attentional processing and contextual integration of socially relevant cues.

Keywords
Attribution of Intentions; Autism Spectrum Disorders (ASD); Event Related Potentials (ERP); Social Cognition & Theory of Mind.

Main text

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by persistent, pervasive deficits in social communication and interaction, and restricted, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013). Difficulties in social interaction in ASD have traditionally been attributed to impaired social cognition, a complex set of cognitive processes involving correct perception and integration of social cues (Fernández et al., 2018).

A core component of social cognition is Theory of Mind (ToM), the ability to infer one's own and others' mental states that are relevant for social action (Adolphs, 2009; Baron-Cohen, 2000). The psychological and cognitive processes that support ToM remain a matter of debate (Brewer et al., 2017; Rosenthal et al., 2019). Facial and vocal emotion processing, mental state reasoning, and various executive functions such as attention switching, working memory and cognitive flexibility, among others, are all believed to contribute to ToM processing (Brewer et al., 2017; Green et al., 2020), and probably reflect a multilevel structure that involves both cognitive and affective processes and the interactions between them (Schurz et al., 2021). This translates into multiple instruments having been developed to measure ToM, or specific aspects of ToM, including false belief, mental state attribution, and various formats of attribution of intentions (Brewer et al., 2017; Livingston et al., 2019). Although ToM deficits are widely described in autistic children and adults (Brewer et al., 2017; Pino et al., 2017, 2020) and ToM ability predicts levels of support needed in ASD (Hoogenhout & Malcolm-Smith, 2017), the specificity and universality of ToM deficits in ASD have not been empirically demonstrated, and many seminal studies have not been replicated (Gernsbacher & Yergeau, 2019). Instruments that assess ToM are often unable to discriminate ASD from other disorders with impaired social cognition, such as schizophrenia (Fernandes et al., 2018) or to detect ToM deficits in adults, mostly due to ceiling effects resulting from the fact that they were originally developed for children (Livingston et al., 2019). Furthermore, many ToM tasks are highly sensitive to intellectual and verbal ability, underestimating deficits in autistic individuals with higher IQ scores (Livingston et al., 2019). Finally, most of these instruments are behavioral tasks that simply measure accuracy scores, therefore failing to capture subtle sources of variability in ToM processing across different populations and conditions.

Using neuroimaging and neurophysiologic methods to study ToM deficits in ASD partly circumvents these limitations and provides more detailed insights into the neural mechanisms that underlie social cognitive deficits. Functional magnetic resonance imaging (fMRI) studies in neurotypical samples have identified a "mentalizing network" comprising the medial prefrontal cortex (mPFC), the temporo-parietal junction (TPJ), and the posterior superior temporal sulcus (Moessnang et al., 2020). Functionally, this network is closely related to the "mirror neuron system" (MNS), first described more than a decade ago as a network of parieto-frontal neurons that become active when the participant observes others' actions and infers the intentions behind those actions (Cattaneo & Rizzolatti, 2009). When inferring others' intentions, autistic adolescents and adults show lower functional activation of the TPJ, right inferior frontal gyrus and left premotor cortex than neurotypical individuals (Kana et al., 2014), as well as decreased functional connectivity between the mirroring and mentalizing systems (Cole et al., 2019;

Kana et al., 2014). Electroencephalographic (EEG) studies have tried to shed further light on the possible role of MNS dysfunction in ToM impairments in ASD. However, studies investigating the relationship between *mu* suppression (a marker of MNS activation) and experimental observation/imitation tasks (indicative of MNS function) have been far from consistent, and clearly more research is needed to understand the mechanisms underlying the relations between ToM and neurophysiologic operations in ASD (Andreou & Skrimpa, 2020).

In neurotypical individuals, event-related potentials (ERPs) have also provided valuable insights into the neural mechanisms of ToM (Gagnon et al., 2016; Leuthold et al., 2012; Sabbagh et al., 2004; Van der Cruyssen et al., 2009; Vistoli et al., 2015). Specifically, studies using mental state inference paradigms have attested the relevance of attentional and contextual integration during early stages of attribution of intentions (Van der Cruyssen et al., 2009; Vistoli et al., 2015). One such study used a non-verbal paradigm consisting of four-image comic strips depicting either intentional or physical causality (Vistoli et al., 2015). An "ERP intention effect" emerged, but only after presentation of the third image of the sequence, where information is contextualized with the cues provided by the first two images. This "ERP intention effect" occurred as a bilateral posterior positive waveform 250 to 650 ms post-stimulus that peaked around 300 ms post-stimulus and is probably related to the P300 component family, a group of ERPs traditionally believed to reflect the evaluation and categorization of stimuli (Polich et al., 1985). In another study where neurotypical participants performed a verbal goal-inference task, Van der Cruyssen et al. (2009) observed a P200 waveform stronger for goal-irrelevant or non-goal-directed behavior compared to goal-consistent behaviors. This was interpreted as evidence of spontaneous goal-inferences occurring in the first 200 ms post-stimulus. Source localization showed that this early automatic processing emerged mainly in the TPJ, whereas more deliberative intentional goal inference occurred later, mainly in the mPFC (Van der Cruyssen et al., 2009). Other studies have shown that this later contextual updating process elicits a P300 waveform that has higher amplitude when the participant observes behaviors that are context-incongruent (Fabiani et al., 1986). To our knowledge, visual ERPs associated with attribution of intentions have not been explored in ASD, although there is evidence of reduced amplitude (but not latency) of the P300 component (Cui et al., 2017).

In this study, we used the non-verbal task developed by Vistoli et al. (2015) to compare the ability of autistic adults and neurotypical volunteers to infer the intentions of others, while registering the electrophysiological correlates of that process. A previous version of this task showed that autistic adolescents and adults make significantly more errors in attribution of intentions than neurotypical individuals (Kana et al., 2014). Our primary hypothesis was thus that autistic participants would commit more errors in this attribution of intentions task than neurotypical individuals, and that this would be reflected by a lower amplitude, but not necessarily atypical latency, of the intention ERP effect in autistic compared to neurotypical participants.

Methods

Participants

Twenty-one adult male autistic participants without intellectual disability were recruited from a specialized center for neurodevelopmental disorders (CADIn – Neurodevelopment & Inclusion, Cascais, Portugal) where they were attending various outpatient psychosocial interventions and activities. All autistic participants had been previously

diagnosed at the center, some in infancy, others in adolescence or early adulthood. Diagnoses were not specifically made for this study. The diagnostic procedure at Cadin involves clinical evaluation by an experienced physician (adult psychiatrist, child & adolescent psychiatrist, paediatric neurologist, or developmental pediatrician) and an experienced psychologist, with the support of one or more validated clinical assessment instruments (Autism Diagnostic Observation Schedule, Second Edition [ADOS-2; Lord et al., 2012] and/or Asperger Syndrome Diagnostic Scale [ASDS; Myles, Bock, & Simpson, 2001] and/or Autism Quotient [AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001]). IQ estimates were available for 19 of the 21 autistic participants (full-scale Wechsler Adult Intelligence Scale-III [WAIS-III; Wechsler, 1999] scores for 17 participants and International Cognitive Ability Resource-16 [ICAR16; Condon & Revelle, 2014] scores for a further two). Thirty adult male neurotypical volunteers, with no past or present history of psychiatric conditions or family history of ASD, were recruited from the participating academic institutions. A known history of brain injury or an active neurological disorder was an exclusion criterion in both groups. All participants had normal or corrected-to-normal vision.

All participants signed a written informed consent. The study was approved by the Ethics Committees of the participating institutions and all study procedures followed the principles of the Declaration of Helsinki.

Task and Procedure

We used an adaptation of the paradigm described by Vistoli et al. (2015) in which comic strips are sequentially presented, depicting situations that result from a character's intentional action or that are strictly a consequence of the laws of physics. Each strip includes four black and white pictures, where the first three build up the situation to its resolution in picture four (see Vistoli et al., 2015 for further details) (Figure 1).

Comic strips were displayed on an LCD placed at a standardized distance, using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA).

The procedure was the same for each comic strip (or trial): the first three pictures were showed for 2000 ms each, separated by 200-ms blank screens. The fourth picture was presented for 5000 ms and was differentiated from the previous three by a dark red border marking it as the strip's ending. After presentation of the fourth picture, participants were asked to respond, as fast and accurately as possible, if that picture presented a logical/congruent ending to the sequence of events from the three preceding images. This written question was presented on a white background for 5000 ms or until the participant gave his answer, by pressing specific yes/no keys on a keyboard adapted to the dominant hand. After the participants' answer or after the answering time ran out, a black fixation-cross on a white background appeared for 1500 ms, separating trials. Following Vistoli et al. (2015), our interest was directed at the third image of the sequence, as to avoid contamination by decision making or incongruity assessment made on the fourth image.

There were three different conditions: 1) attribution of intentions (AI), depicting a character whose action was driven by a specific intention (e.g., fetching a chair to reach for something) that participants must infer; 2) physical causality with character (PCCh), depicting situations that solely followed physical rules, but where a human character was present (e.g., an apple falling on someone's head); and 3) physical causality with objects (PCOb), depicting situations strictly related to physical rules, with no human character

(e.g., wind blowing out a candle). In PCCh and PCOb trials, no processing of intentions is required to decide if endings are logical or not. This setup discriminates the ability to decipher the character's intentions from the ability to understand physical causality (AI vs. PCCh/PCOb), while also controlling for the possibility that the ERP of interest simply reflects the presence of human figures (AI/PCCh vs. PCOb).

Trials with congruent and incongruent endings were presented in pseudo-random order. This ensured that each participant had congruent endings in half of the trials and incongruent endings in the other half, and that half of the participants of each group saw the congruent ending of a comic strip and the other half the incongruent ending of the same comic strip. The same participant never saw the same comic strip twice.
Before the first trial instructions were presented on the LCD screen, briefly describing the goal and structure of the task and explaining how to use the keyboard for giving responses. These instructions were read aloud by the experimenter, who also clarified any questions. Each participant had a 2-minute training session comprising 4 AI, 4 PCCh, and 3 PCOb trials. Training trials were always different from experimental trials. After training, participants were asked if they had understood the task's goal and procedure, and remaining questions were clarified. This was followed by the experimental session with EEG recording, comprising three runs of 24 trials (8 AI, 8 PCCh, and 8 PCOb trials). Each experimental block lasted about 5 minutes and the full recording session took approximately 20 minutes. The three runs of trials were separated by short breaks, whose duration was controlled by participants. Participants remained seated during these breaks and were reminded to reduce blinking and body movements to what was comfortably possible. All these measures were expected to minimize artifacts in the recording.

EEG data acquisition and pre-processing
The EEG signal was recorded continuously during each session, at 1000Hz sampling-rate, connected to a Neuroscan Synamps amplifier (Neurosoft Inc., Sterling, VA, USA, used in 19 autistic participants and 28 neurotypical controls) or an actiCHamp amplifier (Brain Products GmbH, München, Germany, used in 2 autistic participants and 2 neurotypical controls). No online filters were applied. Use of different amplifiers is commonly seen in multi-site studies and is acceptable as long as all EEG related procedures are equal across participants, and the same configurable parameters are used in all amplifiers (Mobascher et al., 2010). Analyses of the restricted data set collected with the Neuroscan Synamps amplifier are presented as Supporting Information.

We used an EasyCap EEG cap (EASYCAP GmbH, Herrsching-Breitbrunn, Germany), with 28 sintered AgCI electrodes mounted according to the international 10–10 system. The reference electrode was placed in AFz and the ground electrode on the right mastoid (Figure 2). Electrode impedance was kept below 10 kΩ.

Data were analyzed with BrainVision Analyzer, version 2.1.2 (Brain Products GmbH, München, Germany). Raw EEG signal was filtered off-line with a 1Hz high-pass filter and a 30Hz low-pass filter. An Independent Component Analysis (ICA) was performed on the whole EEG data, using the FastICA Restricted algorithm from BrainVision Analyzer, which performs an automatic correction based on predefined parameters (Hyvärinen, & Oja, 2000). The main focus of this analysis was correcting ocular movements, especially blinks (detected using the Fp1 channel). Yet, this automatic ICA has a wider range of action, attenuating other sources of noise (e.g., muscle tension). The average proportion of rejected components per subject was 5.2% (range: 3 to7).

Individual epochs were created for each trial from 200 ms before to 700 ms after the third image's onset (Vistoli et al., 2015). Baseline correction was performed using the 200 ms that preceded the third image (-200 to 0), corresponding to a blank screen. Segments with potentials exceeding ±75 μV were rejected. Average ERPs were computed for each participant and condition (AI, PCCh, and PCOb) over a mean of 24 artefact-free trials (range: 21–24) in each condition. Number of accepted trials was similar across conditions in both the ASD group (AI = 23.8 [0.5], PCCh = 23.5 [0.7], PCOb = 23.7 [0.6]; F[2,40] = 2.562, $p = 0.09$), and the neurotypical group (AI = 23.9 [1.2], PCCh = 23.9 [1.4], PCOb = 23.8 [1,2]; F[2,58] = 0.057, $p = 0.945$), and did not differ between groups (F[1,49] = 0.022, $p = 0.883$).

Statistical Analysis

Demographic data and task performance were analyzed using student's t-test or non-parametric tests (Mann-Whitney U and Fischer's Exact Test), as appropriate. Task performance was quantified as the proportion of correctly identified endings (congruent or incongruent) in each condition. To explore correlations between task performance, IQ, and ERP variables, we computed Pearson's or Kendall's Tau coefficients, as appropriate. Normality was assessed with Q-Q plots and Shapiro-Wilk's test.

ERP analysis followed Vistoli et al. (2015). Whereas based on their findings we expected to find a bilateral posterior ERP, we first explored topography effects of the averaged waveforms from each quadrant in the neurotypical control group: left anterior (including Fp1, F7, F3, FC1, C3, and T7 electrodes), right anterior (Fp2, F8, F4, FC2, C4, and T8), left posterior (CP5, CP1, P7, P3, PO7, and O1), and right posterior quadrant (CP6, CP2, P8, P4, PO8, and O2). Statistical analyses using the four groups of electrodes are presented as Supporting Information. As expected, the ERP was exclusively posterior and no differences were found between hemispheres, therefore subsequent analyses focused exclusively on bilateral posterior electrodes (CP5/CP6, CP1/CP2, P7/P8, P3/P4, PO7/PO8, and O1/O2).

ERP components were analysed with repeated measures analysis of variance (rmANOVA), using mean amplitude as the dependent variable and condition as the within-subjects factor for each group. Component amplitude was defined as the mean amplitude within the temporal window of interest (Woodman, 2010). To compare the ERP intention effect across groups we performed a rmANOVA with condition and time as within-subjects factors, and group as the between-subjects factor.

To compare ERP latencies across groups we used the jack-knife approach described by Kiesel et al. (2008). Subsample scores were entered into an ANOVA model, corrected for unequal group sizes using MrFub software (Ulrich & Miller, 2001). We defined latency as the time point when amplitude reached 50% of through-to-peak amplitude. Individual latencies estimates were retrieved from subsample scores following Smulders (2010). We used ANOVA or Kruskal-Wallis H (if non-normal distribution) to test group-differences.

We used Mauchly's tests to assess sphericity (W), and Greenhouse–Geisser (ε) correction where necessary. Observed power (OP) and effect sizes (partial eta-squared [η_p^2]) are also

reported. Post-hoc pairwise comparisons were corrected with Bonferroni procedure. Alpha was set at 0.05.

All analyses were conducted with IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp., Armonk, NY, USA).

Results

Demographic Data

Demographic characteristics of the two groups are summarized in Table 1.
Over 95% of autistic participants were enrolled in high-school or college programs. All IQ scores in autistic participants were > 70. IQ scores were missing in two autistic college students. Mean IQ score in autistic participants was 101.6 ± 17.1.

Twelve autistic participants (against none of the neurotypical participants) were taking the following psychotropic medications: seven were taking antidepressants (escitalopram [4], sertraline, venlafaxine, and mirtazapine), 7 antipsychotics (risperidone [4], quetiapine, aripiprazole, and tiapride), 4 anticonvulsants/ mood-stabilizers (valproate [3], lamotrigine), 2 methylphenidate, 2 low dose benzodiazepines (clobazam and ethyl loflazepate), and 1 clonidine. A separate analysis of autistic participants who were not taking any medication is presented as Supporting Information.

Behavioral Performance

Neurotypical volunteers performed significantly better than autistic participants in the AI condition (Mdn [IQR] accuracy: 89.0 [14] vs. 71.0 [18]; U = 151.5, $p = 0.002$), but not in the PCCh (79.0 [12] vs. 79.0 [21]; U = 302.0, $p = 0.802$) and PCOb (84.0 [10] vs. 86.0 [11]; U = 276.5, $p = 0.455$) conditions (Figure 3). We found no significant correlation between IQ scores and AI task performance (r = 0.340; $p = 0.155$) in the ASD group.

ERP Intention Effect in the Neurotypical Control Group

To confirm the validity of the used ERP paradigm , we first looked at the ERP components of the neurotypical group. We found three waveforms: a positive component peaking at around 100 ms, followed by a negative component peaking at around 170 ms, and another positive component extending from 200 to 400 ms post-stimulus (Figure 4-A). Waveforms were remarkably similar in the PCCh and PCOb conditions. In the AI condition, the first two waveforms were similar to those observed in the PCCh and PCOb conditions, unlike the third component (200-400 ms), which was markedly more positive (Figure 4-A) and comprised two distinct peaks at around 220 ms and 310 ms, coincident with the typical time windows of P200 and P300 components, respectively (Breznitz, 2008; Polich, 2007).

RmANOVA with the three conditions (AI, PCCh, and PCOb) and three time intervals (]0 ms to 200 ms],]200 ms to 400 ms], and]400 ms to 600 ms]) as within-subjects factors revealed a main effect of time (F[2, 58] = 15.11; W = 0.64; ε = 0.73; $\eta_p^2 = 0.343$; OP = 0.99; $p < 0.001$), and condition (F[2, 58] = 8.35; W = 0.96; ε = 0.96; $\eta_p^2 = 0.224$; OP = 0.96; $p = 0.001$), and a significant interaction of time * condition (F[4, 116] = 16.20; W = 0.85; ε = 0.92; $\eta_p^2 = 0.358$; OP = 1.00; $p < 0.001$). Post-hoc comparisons for the 0-200 ms interval found a stronger positivity in the PCCh vs. the AI condition ($p = 0.005$). During the 200-400 ms interval the AI condition showed a significantly stronger positivity compared to the other two conditions (both $p < 0.001$), and again in the 400-600 ms interval compared to the PCOb condition only ($p = 0.018$).

ERP Intention Effect in Autistic Participants

The overall morphology of the global averaged ERP in autistic participants was similar to that of the control group (Figure 4-B), although the ERP intention effect in the 200-400 ms interval was visually less evident, with the two peaks found in neurotypical individuals not as clearly present. RmANOVA for the ASD group is reported as Supporting Information (significant differences are highlighted in Figure 4-B). Of note, inclusion of IQ score as a covariate did not influence the main effects of time and condition nor the time * condition interaction, and there were no significant correlations between IQ scores and ERP amplitudes or latencies.

For cross-group comparison of the ERP intention effect we contrasted the AI and PCCh conditions in the two relevant time-intervals (200-400 ms and 400-600 ms). We followed the original rationale of Vistoli et al. (2015) that the PCOb condition, involving no human characters, is irrelevant as a comparator (the analysis including the PCOb condition is presented as Supporting Information). The resulting rmANOVA model with two conditions (AI and PCCh) and two time intervals (]200 ms to 400 ms] and]400 ms to 600 ms]) as within-subjects factors, and group (ASD and neurotypical controls) as the between-subjects factor revealed a main effect of time ($F[1, 49] = 81.49$; $W = 1.0$; $\varepsilon = 1.0$; $\eta_p^2 = 0.624$; $OP = 1.00$; $p < 0.001$) and condition ($F[1, 49] = 33.45$; $W = 1.0$; $\varepsilon = 1.0$; $\eta_p^2 = 0.406$; $OP = 1.00$; $p < 0.001$), and a significant time * condition * group interaction ($F[1, 49] = 7.67$; $\eta_p^2 = 0.135$; $OP = 0.76$; $p = 0.008$). A main effect of group was also observed ($F[1, 49] = 5.25$; $\eta_p^2 = 0.097$; $OP = 0.61$; $p = 0.026$). Post-hoc comparisons showed a significantly less positive waveform in autistic participants versus neurotypical controls in the PCCh condition for both the 200-400 ms interval ($p = 0.018$) and the 400-600 ms interval ($p = 0.003$), but not in the AI condition for the same 200-400 ms ($p = 0.093$) and 400-600 ms ($p = 0.716$) intervals.

P200 and P300 Analyses

As mentioned, the AI waveform in neurotypical participants showed two distinct peaks compatible with P200 and P300 components. For further analysis we considered the mean amplitude in the]180 ms to 250 ms] and]250 ms to 400 ms] intervals, respectively (Breznitz, 2008; Chen et al., 2020; Rischer et al., 2020).

On separate analyses the neurotypical group showed a significant effect of condition ($F[2, 58] = 15.04$; $W = 0.96$; $\varepsilon = 0.96$; $\eta_p^2 = 0.342$; $OP = 1.00$; $p < 0.001$) on P200 amplitude, with a significantly stronger positivity in the AI condition compared to PCCh and PCOb (both $p < 0.001$, respectively) and no difference between the latter. In the ASD group the effect of condition was nonsignificant ($F[2, 40] = 2.97$; $W = 0.94$; $\varepsilon = 0.95$; $\eta_p^2 = 0.129$; $OP = 0.55$; $p = 0.063$). When we compared the two groups in the rmANOVA model (condition [AI, PCCh] * group [AI, neurotypical]), no significant difference was observed ($F[1, 49] = 1.14$; $\eta_p^2 = 0.023$; $OP = 0.18$; $p = 0.291$).

For P300 amplitude we found a significant effect of condition in both the neurotypical ($F[2, 58] = 20.16$; $W = 0.91$; $\varepsilon = 0.91$; $\eta_p^2 = 0.410$; $OP = 1.00$; $p < 0.001$) and the ASD group ($F[2, 40] = 5.95$; $W = 0.70$; $\varepsilon = 0.77$; $\eta_p^2 = 0.229$; $OP = 0.78$; $p = 0.011$). Post-hoc comparisons showed a stronger positivity in the AI compared to PCCh and PCOb conditions in neurotypical volunteers (both $p < 0.001$) and in autistic participants ($p = 0.032$ and $p = 0.014$, respectively). Autistic participants showed smaller P300 amplitude than neurotypical participants ($F[1,49] = 5.38$; $\eta_p^2 = 0.099$; $OP = 0.62$; $p = 0.025$).

P200 latency was significantly longer in the ASD group than in neurotypical volunteers (208.86 ± 19.64 ms vs. 194.82 ± 10.18 ms, respectively; F[1,49] = 11.12; η_p^2 = 0.185; OP = 0.91; p = 0.002), but P300 latency was not (264.36 ± 54.88 ms vs. 283.45 ± 20.18 ms, respectively; H(1) = 0.99; p = 0.320).

ERP Amplitude in Autistic and Neurotypical Participants with Similar Behavioral Performance

As an additional exploratory analysis, we compared ERP amplitude in autistic participants and a subgroup of similarly performing neurotypical (spNT) individuals (n = 10; Mdn [IQR] accuracy in the AI condition = 75.0[8]; U = 98.5, p = 0.787 for ASD vs. spNT) corresponding to the lowest performance tercile of the neurotypical group. Assuming that amplitude of the ERP intention effect is influenced, among other factors, by performance effort and by clinical group affiliation, restricting the comparison to similarly performing participants allows us to control for the former and thus isolate the effect, if any, of diagnostic status.

RmANOVA with two conditions (AI and PCCh) and two time intervals (200-400 ms and 400-600 ms) as within-subjects factors and the 2 groups (ASD and spNT) as the between-subjects factor found a main effect for time (F[1, 29] = 39.12; W = 1.0; ε = 1.0; η_p^2 = 0.574; OP = 1.00; p < 0.001) and condition (F[1, 29] = 14.29; W = 1.0; ε = 1.0; η_p^2 = 0.330; OP = 0.96; p = 0.001), as well as a significant time * condition * group interaction (F[1, 29] = 7.52; η_p^2 = 0.206; OP = 0.76; p = 0.010). The test of between-subjects effects was also significant (F[1, 29] = 11.80; η_p^2 = 0.289; OP = 0.91; p = 0.002). Post-hoc comparisons showed a significantly less positive waveform in ASD vs. spNT individuals in the AI condition in the 200-400 ms interval (p = 0.009) and in the PCCh condition for both the 200-400 ms interval (p = 0.004) and the 400-600 ms interval (p = 0.003). The former difference was mostly driven by a smaller P300 (F[1,29] = 11.91; η_p^2 = 0.291; OP = 0.92; p = 0.002), with no differences regarding P200 (F[1, 29] = 3.50; η_p^2 = 0.108; OP = 0.44; p = 0.161) (Figure 5).

Discussion

Our main findings may be summarized as follows: 1) autistic adults performed significantly worse in the AI task than neurotypical volunteers; 2) we replicated an ERP correlate of AI in neurotypical adults, encompassing two distinct components (P200 and P300); 3) this ERP intention effect was equally present in autistic adults, but with subtle differences, namely a longer P200 latency and no effect of condition on P200 amplitude; although P300 amplitude was larger in the AI vs. physical causality conditions in both groups, that amplitude was smaller in autistic participants, and there were no differences in latency; 4) when compared to a subgroup of neurotypical participants with similar accuracy performance in the AI task, autistic participants had a significantly lower mean amplitude of the ERP intention effect.

A lower accuracy in the ability to correctly infer intentions has been previously described in ASD (Cole et al., 2018, 2019; Libero et al., 2014). Concordantly, in our study autistic participants performed significantly worse than neurotypical volunteers in the AI condition, where they had to infer characters' intentions, but not in the two physical causality conditions. This suggests that their lower performance is driven by difficulties in inferring intentions, since disregarding or not understanding task instructions would have resulted in low performance in all three conditions. This is further supported by the

independence between AI task performance and intellectual ability observed in autistic participants.

Although Vistoli et al. (2015) demonstrated that providing explicit instructions has no impact on the ERP intention effect, we do not know whether this also holds for autistic participants. In ASD, ToM deficits have been more consistently observed when using implicit mentalizing tasks (*i.e.*, inferring internal states of others when not specifically told to do so). In contrast, explicit mentalizing tasks that provide instructions concerning which elements should be attended to during the task often fail to detect mentalizing deficits in autistic adults; this seems to be especially the case when these tasks are based on simplistic representations of social interactions that presumably allow higher-functioning ASD individuals to make use of learned strategies (Cole et al., 2018). In our task, though cognisant that they were taking part in a study on social cognition, participants were only instructed to decide whether comic strips endings were congruent or not, without any information or instruction on the nature of conditions (intentional versus physical), nor about the need to explicitly focus on intentions. It is likely that this implicit approach increased the task's sensitivity to group differences in AI performance.

The "ERP intention effect" described here is a bilateral posterior waveform occurring 200 to 400 ms post-stimulus, with greater positive amplitude in the AI condition than in physical causality conditions. In contrast to the single component described by Vistoli et al. (2015) in the 250-650 ms time window and peaking at 300 ms, we identified, in neurotypical controls, two distinguishable waveforms in the 200-400 ms interval that may correspond to P200 and P300 components. Even though differences in experimental setup might explain such differences, distinct P200 and P300 components have been previously described in neurotypical individuals performing a verbal goal-inference task. In that study, Van der Cruyssen et al. (2009) showed that a larger P200 occurred at fronto-central locations in nonintentional compared to intentional situations, and that a larger P300 occurred at centro-parietal midline locations in nonintentional situations and when participants were asked to explicitly make inferences. The authors concluded that the P200 component reflects a rapid, automatic increase in attention to, and processing of, behavioral information, whereas P300 reflects a later, deliberative context-updating process. Different results regarding ERP topography and amplitude modulation likely reflect the use of verbal stimuli and collection of ERPs after a target word in that study, in contrast with the present non-verbal paradigm and collection of ERPs while contextual understanding of the character's intention is still unfolding .

In autistic participants the AI waveform was less pronounced visually, with less evident P200 and P300 components, and, in contrast to neurotypical participants, P200 amplitude did not differ across the three conditions and had a longer latency. This later and more discreet onset of the AI waveform in autistic participants may reflect a reduced motivational response to social stimuli, and thus a lack of differentiation in the attentional processing of social and non-social stimuli. Indeed, P200 amplitude is sensitive to the socially meaningful content of target stimuli (Hajcak et al., 2012; Xu et al., 2017), and is enhanced when uncertainty about the outcome of a given cue is associated to higher selective attention to that cue (Johnen & Harrison, 2020; Xu et al., 2011). Longer P200 latency suggests a slower processing of visual input in autistic participants (Portella et al., 2012) or an over-processing of information needed for accurate stimuli differentiation (Sokhadze et al., 2009). Reduced selective attention to socially meaningful stimuli has been previously described in ASD (Tyndall et al., 2018; Wang et al., 2020), and our

results add to such evidence, by suggesting that less accurate AI in ASD may derive from dysfunction in early stages of selective attention to stimuli that are necessary to aptly anticipate the outcome of the strip. This results in ineffective discrimination between socially relevant and socially irrelevant stimuli and increased processing time (reflected in increased P200 latency) (Wang et al., 2017).

Although P300 amplitude was higher in the AI condition than in the two physical causality conditions in both groups, P300 amplitude (but not latency) was significantly larger in neurotypical than in autistic participants. Benning et al. (2016) reported smaller late positive potential amplitude in response to social stimuli in autistic children and adolescents in comparison with neurotypical controls, and Cox et al. (2015) found an association between autistic traits in typically developing young adults and an attenuated P300 response to the anticipation of social vs. non-social rewards. On the other hand, blunted P300 responses to reward, irrespective of reward type (social vs. monetary), have been observed in children with ASD (Kohls et al., 2011), suggesting the use of alternative mechanisms for processing reward-associated stimuli, with allocation of fewer attentional and working-memory resources. The P300 waveform comprises two subcomponents: P3a, associated with automatic attention to distracter stimuli; and P3b, associated with effortful processing of task-relevant information (Cui et al., 2017; Polich, 2007; Volpe et al., 2007), presumably a working memory-guided target identification mechanism that contributes to goal-directed learning and decision-making (Rac-Lubashevsky & Kessler, 2019). A meta-analysis of P300 amplitude and latency in ASD showed reduced P3b amplitude, but no differences in latency, nor in P3a amplitude or latency, possibly reflecting atypical working memory processing during decision making (Cui et al., 2017). The ERP intention effect described here has also been associated to the retrieval, from working memory, of contextual information provided in the first two images in the strip, which is then integrated with the action depicted in the third image to inform the process of attribution of intentions (Vistoli et al., 2015). Our results with autistic individuals are thus highly concordant with the lower P3b amplitude but normal latency described by Cui et al. (2017) in ASD, and probably reflect atypical or ineffective attention and working memory processing.

Although findings should be interpreted with caution due to the small size of the spNT subgroup, we found a more marked contrast between autistic and spNT participants in terms of ERP intention effect amplitude. This seemingly counterintuitive finding possibly reflects spNT participants' need to allocate additional processing resources when trying to infer characters' intentions, in line with previous research showing ERP amplitudes to be usually larger when more effort is devoted to a task (Luck, 2014). Consequently, we would expect the intention ERP to be independent from performance even in autistic individuals with high AI accuracy. Having only 4 participants who scored above the neurotypical group's average, we were unable to test this prediction.

Finally, autistic participants showed a negative PCCh waveform during the 400-600 ms interval that significantly differed from both the AI and PCOb conditions. We believe this may reflect a late negative component associated with a re-orienting of attention (Berti, 2008; Berti & Schröger, 2001), elicited by the fact that, in the PCCh condition, the combination of a human character and an "independent" physical action poses additional demands on working memory. This lends further strength to the hypothesis that autistic participants allocate fewer working memory resources, compared to their neurotypical

counterparts, in early stages of the stimulus processing in this task and regardless of the condition.

The present study has several limitations. Intellectual ability was not assessed in neurotypical controls and it is possible that the two groups were unbalanced in that respect; however, groups performed similarly in the non-intentional conditions of the task, and IQ scores did not correlate with AI task performance or significantly influence the rmANOVA model in the ASD group. Despite the fact that we did not control for differences in the use of psychotropic medication, analyses restricted to unmedicated participants produced overlapping results (see Supporting Information). Replication of our findings in larger samples of autistic participants will increase their validity and may help clarify any possible role of medication or underlying comorbidities, which are common in ASD (Doshi-Velez et al., 2014). Methodological limitations of the task, such as the lack of control of low-level visual aspects, like contrast or luminance, and the restricted number of trials without stimuli repetition, have been previously noted (Vistoli et al., 2015). However, the similarity between our results in neurotypical volunteers and those reported by Vistoli et al. (2015) support the validity of the ERP elicited by this task.

In conclusion, our results add to existing evidence that autistic adults have a lower ability to accurately attribute intentions when compared to neurotypical individuals, possibly due to slower early attentional processing and less efficient contextual integration of relevant cues. If additional research confirms our findings, and how they correlate with deficits in specific cognitive processes (e.g., automatic attention to relevant stimuli; working memory), this knowledge can yield important practical implications. We may be able to improve our ability to anticipate and adapt to such processes in social interactions with autistic people and, when required, adapt the social environment in a way that facilitates decisive social cognitive processes such as attribution of intentions. Moreover, if the differences we observed in the ERP intention effect are not replicated in other conditions characterized by ToM impairments, such as schizophrenia (Fernandes et al., 2018), and thus prove to be specific to ASD, they may prove useful as a biomarker to identify homogeneous samples of individuals with ASD. This would bring a much needed improvement to clinical studies and therapeutic trials in a field where clinical heterogeneity contributes decisively to the well-known abundance of negative treatment trials.

References

Adolphs R. (2009). The social brain: neural basis of social knowledge. Annual review of psychology, 60, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

American Psychiatric Association. (2013). Neurodevelopmental Disorders. In *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
https://doi.org/10.1176/appi.books.9780890425596.dsm01

Andreou, M., & Skrimpa, V. (2020). Theory of Mind Deficits and Neurophysiological Operations in Autism Spectrum Disorders: A Review. Brain sciences, 10(6), 393. https://doi.org/10.3390/brainsci10060393

Baron-Cohen, S. (2000). Theory of mind and autism: a review. *International Review of Research in Mental Retardation*, 23, 169-84. https://doi.org/10.1016/S0074-7750(00)80010-5

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders*, 31(1), 5–17. https://doi.org/10.1023/a:1005653411471

Benning, S. D., Kovac, M., Campbell, A., Miller, S., Hanna, E. K., Damiano, C. R., Sabatino-DiCriscio, A., Turner-Brown, L., Sasson, N. J., Aaron, R. V., Kinard, J., & Dichter, G. S. (2016). Late Positive Potential ERP Responses to Social and Nonsocial Stimuli in Youth with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 46(9), 3068–3077. https://doi.org/10.1007/s10803-016-2845-y

Berti S. (2008). Object switching within working memory is reflected in the human event-related brain potential. *Neuroscience Letters*, 434(2), 200–205. https://doi.org/10.1016/j.neulet.2008.01.055

Berti, S., & Schröger, E. (2001). A comparison of auditory and visual distraction effects: behavioral and event-related indices. *Brain Research. Cognitive Brain Research*, 10(3), 265–273. https://doi.org/10.1016/s0926-6410(00)00044-6

Brewer, N., Young, R.L., & Barnett E. (2017). Measuring Theory of Mind in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 47(7), 1927-1941.
https://doi.org/10.1007/s10803-017-3080-x

Breznitz, Z. (2008). The P200 Component. In *Brain Research in Language*. Boston, MA: Springer-Verlag US.

Cattaneo, L., Rizzolatti, G. (2009). The mirror neuron system. *Archives of Neurology*, 66(5), 557-60. https://doi.org/10.1001/archneurol.2009.41

Chen, Z., Jin, J., Daly, I., Zuo, C., Wang, X., & Cichocki, A. (2020). Effects of Visual Attention on Tactile P300 BCI. *Computational intelligence and neuroscience*, 2020, 6549189. https://doi.org/10.1155/2020/6549189

Cole, E.J., Barraclough, N.E., & Andrews, T.J. (2019). Reduced connectivity between mentalizing and mirror systems in autism spectrum condition. *Neuropsychologia*, 122, 88-97. https://doi.org/10.1016/j.neuropsychologia.2018.11.008

Cole, E.J., Slocombe, K.E., & Barraclough, N.E. (2018). Abilities to Explicitly and Implicitly Infer Intentions from Actions in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 48, 1712–1726. https://doi.org/10.1007/s10803-017-3425-5

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, 43, 52–64. https://doi.org/10.1016/j.intell.2014.01.004

Cox, A., Kohls, G., Naples, A. J., Mukerji, C. E., Coffman, M. C., Rutherford, H. J., Mayes, L. C., & McPartland, J. C. (2015). Diminished social reward anticipation in the broad autism phenotype as revealed by event-related brain potentials. *Social Cognitive and Affective Neuroscience*, 10(10), 1357–1364. https://doi.org/10.1093/scan/nsv024

Cui, T., Wang, P.P., Liu, S., & Zhang, X. (2017). P300 amplitude and latency in autism spectrum disorder: a meta-analysis. *European Child & Adolescent Psychiatry*, 26(2), 177-190. https://doi.org/10.1007/s00787-016-0880-z

Doshi-Velez, F., Ge, Y., & Kohane, I. (2014). Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. Pediatrics, 133(1), e54–e63. https://doi.org/10.1542/peds.2013-0819

Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. Psychophysiology, 23(3), 298–308. https://doi.org/10.1111/j.1469-8986.1986.tb00636.x

Fernandes, J.M., Cajão, R., Lopes, R., Jerónimo, R., & Barahona-Corrêa, J.B. (2018). Social Cognition in Schizophrenia and Autism Spectrum Disorders: A Systematic Review and Meta-Analysis of Direct Comparisons. *Frontiers in Psychiatry*, 9, 504. https://doi.org/10.3389/fpsyt.2018.00504

Fernández, M., Mollinedo-Gajate, I., & Peñagarikano, O. (2018). Neural Circuits for Social Cognition: Implications for Autism. *Neuroscience*, 370, 148-162. https://doi.org/10.1016/j.neuroscience.2017.07.013

Gagnon, J., Aubin, M., Emond, F.C., Derguy, S., Bessette, M., & Jolicoeur, P. (2016). Neural mechanisms underlying attribution of hostile intention in nonaggressive individuals: An ERP study. *International Journal of Psychophysiology*, 110, 153-162. https://doi.org/10.1016/j.ijpsycho.2016.08.007

Gernsbacher, M.A., & Yergeau, M. (2019). Empirical Failures of the Claim That Autistic People Lack a Theory of Mind. *Archives of Scientific Psychology*, 7(1), 102-118. https://doi.org/10.1037/arc0000067

Green, C. C., Brown, N. J., Yap, V., Scheffer, I. E., & Wilson, S. J. (2020). Cognitive processes predicting advanced theory of mind in the broader autism phenotype. *Autism research*, 13(6), 921–934. https://doi.org/10.1002/aur.2209

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In E. Kappenman & S. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 441– 472). New York, NY: Oxford University Press.

Hoogenhout, M., & Malcolm-Smith S. (2017). Theory of mind predicts severity level in autism. *Autism*, 21(2), 242-252. https://doi.org/10.1177/1362361316636758

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. Neural networks : the official journal of the International Neural Network Society, 13(4-5), 411–430. https://doi.org/10.1016/s0893-6080(00)00026-5

Johnen, A.K., & Harrison, N.R. (2020) Level of uncertainty about the affective nature of a pictorial stimulus influences anticipatory neural processes: An event-related potential (ERP) study. *Neuropsychologia*, 146:107525. https://doi.org/10.1016/j.neuropsychologia.2020.107525

Kana, R.K., Libero, L.E., Hu, C.P., Deshpande, H.D., & Colburn, J.S. (2014). Functional brain networks and white matter underlying theory-of-mind in autism. *Social Cognitive and Affective Neuroscience*, 9(1), 98-105. https://doi.org/10.1093/scan/nss106

Kiesel, A., Miller, J., Jolicoeur, P., & Brisson, B. (2008). Measurement of ERP latency differences: a comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, 45(2), 250–274. https://doi.org/10.1111/j.1469-8986.2007.00618.x

Kohls, G., Peltzer, J., Schulte-Rüther, M., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., & Konrad, K. (2011). Atypical brain responses to reward cues in autism as revealed by event-related potentials. *Journal of Autism and Developmental Disorders*, 41(11), 1523–1533. https://doi.org/10.1007/s10803-011-1177-1

Leuthold, H., Filik, R., Murphy, K., & Mackenzie, I.G. (2012). The on-line processing of socio-emotional information in prototypical scenarios: inferences from brain potentials. *Social Cognitive and Affective Neuroscience*, 7(4), 457-66. https://doi.org/10.1093/scan/nsr029

Libero, L.E., Maximo, J.O., Deshpande, H.D., Klinger, L.G., Klinger, M.R., & Kana, R.K. (2014). The role of mirroring and mentalizing networks in mediating action intentions in autism. *Molecular Autism*, 5(1), 50. https://doi.org/10.1186/2040-2392-5-50

Livingston, L.A., Carr, B., & Shah, P. (2019). Recent Advances and New Directions in Measuring Theory of Mind in Autistic Adults. *Journal of Autism and Developmental Disorders*, 49(4), 1738-1744. https://doi.org/10.1007/s10803-018-3823-3

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule (ADOS-2)*. Torrance, CA: Western Psychological Services.

Luck, S.J. (2014). *Overview of Common ERP Components. In An Introduction to the Event-Related Potential Technique* (2nd ed.). Cambridge, MA: MIT Press.

Mobascher, A., Brinkmeyer, J., Warbrick, T., Wels, C., Wagner, M., Gründer, G., Spreckelmeyer, K.N., Wienker, T., Diaz Lacava, A., Dahmen, N., Böttcher, M., Thuerauf, N., Clepce, M., Kiefer, F., De Millas, W., Gallinat, J., & Winterer, G. (2010). The P300 event-related potential and smoking--a population-based case-control study. *International Journal of Psychophysiology*, 77(2), 166-75. https://doi.org/10.1016/j.ijpsycho.2010.06.002

Moessnang, C., Baumeister, S., Tillmann, J., Goyard, D., Charman, T., Ambrosino, S., Baron-Cohen, S., Beckmann, C., Bölte, S., Bours, C., Crawley, D., Dell'Acqua, F., Durston, S., Ecker, C., Frouin, V., Hayward, H., Holt, R., Johnson, M., Jones, E., Lai, M.C., Lombardo, M.V., Mason, L., Oldenhinkel, M., Persico, A., Cáceres, A.S.J., Spooren, W., Loth, E., Murphy, D.G.M., Buitelaar, J.K., Banaschewski, T., Brandeis, D., Tost, H., Meyer-Lindenberg, A., EU-AIMS LEAP group. (2020). Social brain activation during mentalizing in a large autism cohort: the Longitudinal European Autism Project. *Molecular Autism*, 11(1), 17. https://doi.org/10.1186/s13229-020-0317-x

Myles, B., Bock, S., & Simpson, R. (2001). Asperger Syndrome Diagnostic Scale Manual. North Tonawanda, NY: Multi-Health Systems Inc.

Pino, M.C., Masedu, F., Vagnetti, R., Attanasio, M., Di Giovanni, C., Valenti, M., & Mazza, M. (2020). Validity of Social Cognition Measures in the Clinical Services for Autism Spectrum Disorder. *Frontiers in Psycholology*, 11, 4. https://doi.org/10.3389/fpsyg.2020.00004

Pino, M.C., Mazza, M., Mariano, M., Peretti, S., Dimitriou, D., Masedu, F., Valenti, M., & Franco F. (2017). Simple Mindreading Abilities Predict Complex Theory of Mind: Developmental Delay in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 47(9), 2743-2756. https://doi.org/10.1007/s10803-017-3194-1

Polich, J., Howard, L., & Starr, A. (1985). Stimulus frequency and masking as determinants of P300 latency in event-related potentials from auditory stimuli. *Biological Psychology*, 21(4), 309–318. https://doi.org/10.1016/0301-0511(85)90185-1

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology, 118, 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Portella, C., Machado, S., Arias-Carrión, O., Sack, A. T., Silva, J. G., Orsini, M., Leite, M. A., Silva, A. C., Nardi, A. E., Cagy, M., Piedade, R., & Ribeiro, P. (2012). Relationship between early and late stages of information processing: an event-related potential study. *Neurology International*, 4(3), e16. https://doi.org/10.4081/ni.2012.e16

Proverbio, A. M., & Carminati, M. (2019). Electrophysiological markers of poor versus superior math abilities in healthy individuals. *The European Journal of Neuroscience*, 50(2), 1878–1891. https://doi.org/10.1111/ejn.14363

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://www.pstnet.com.

Rac-Lubashevsky, R., & Kessler, Y. (2019). Revisiting the relationship between the P3b and working memory updating. *Biological Psychology*, 148, 107769. https://doi.org/10.1016/j.biopsycho.2019.107769

Rischer, K. M., Savallampi, M., Akwaththage, A., Salinas Thunell, N., Lindersson, C., & MacGregor, O. (2020). In context: emotional intent and temporal immediacy of contextual descriptions modulate affective ERP components to facial expressions. *Social cognitive and affective neuroscience*, 15(5), 551–560. https://doi.org/10.1093/scan/nsaa071

Rosenthal, I. A., Hutcherson, C. A., Adolphs, R., & Stanley, D. A. (2019). Deconstructing Theory-of-Mind Impairment in High-Functioning Adults with Autism. *Current biology*, 29(3), 513–519.e6. https://doi.org/10.1016/j.cub.2018.12.039

Sabbagh, M.A., Moulson, M.C., & Harkness, K.L. (2004). Neural correlates of mental state decoding in human adults: an event-related potential study. *Journal of Cognitive Neuroscience*, 16(3), 415-26. https://doi.org/10.1162/089892904322926755

Schneider, B. A., Avivi-Reich, M., & Mozuraitis, M. (2015). A cautionary note on the use of the Analysis of Covariance (ANCOVA) in classification designs with and without within-participant factors. *Frontiers in psychology*, 6, 474. https://doi.org/10.3389/fpsyg.2015.00474

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. Psychological bulletin, 147(3), 293–327. https://doi.org/10.1037/bul0000303

Smulders F. T. (2010). Simplifying jackknifing of ERPs and getting more out of it: retrieving estimates of participants' latencies. *Psychophysiology*, 47(2), 387–392. https://doi.org/10.1111/j.1469-8986.2009.00934.x

Tyndall, I., Ragless, L., & O'Hora, D. (2018) Effects of perceptual load and socially meaningful stimuli on crossmodal selective attention in Autism Spectrum Disorder and neurotypical samples. *Consciousness and Cognition*, 60, 25-36. https://doi.org/10.1016/j.concog.2018.02.006

Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, 38(5), 816–827. https://doi.org/10.1111/1469-8986.3850816

Van der Cruyssen, L., Van Duynslaeger, M., Cortoos, A., & Van Overwalle, F. (2009). ERP time course and brain areas of spontaneous and intentional goal inferences. *Social Neuroscience*, 4(2), 165-84. https://doi.org/10.1080/17470910802253836

Vistoli, D., Passerieux, C., El Zein, M., Clumeck, C., Braun, S., & Brunet-Gouet, E. (2015). Characterizing an ERP correlate of intentions understanding using a sequential comic strips paradigm. *Social Neuroscience*, 10(4), 391-407. https://doi.org/10.1080/17470919.2014.1003272

Volpe, U., Mucci, A., Bucci, P., Merlotti, E., Galderisi, S., & Maj, M. (2007). The cortical generators of P3a and P3b: a LORETA study. *Brain Research Bulletin*, 73(4-6), 220-30. https://doi.org/10.1016/j.brainresbull.2007.03.003

Wang, Q., Chang, J., & Chawarska, K. (2020). Atypical Value-Driven Selective Attention in Young Children With Autism Spectrum Disorder. *JAMA Network Open*, 3(5), e204928. https://doi.org/10.1001/jamanetworkopen.2020.4928

Wang, S., Yang, C., Liu, Y., Shao, Z., & Jackson, T. (2017). Early and late stage processing abnormalities in autism spectrum disorders: An ERP study. *PloS one*, 12(5), e0178542. https://doi.org/10.1371/journal.pone.0178542

Wechsler, D. (1997). WAIS-III Administration and scoring manual. San Antonio, TX: The Psychological Association.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, perception & psychophysics*, 72(8), 2031–2046. https://doi.org/10.3758/APP.72.8.2031

Xu, K., Li, S., Ren, D., Xia, R., Xue, H., Zhou, A., & Xu, Y. (2017). Importance Modulates the Temporal Features of Self-Referential Processing: An Event-Related Potential Study. *Frontiers in Human Neuroscience*, 11, 470. https://doi.org/10.3389/fnhum.2017.00470

Xu, Q., Shen, Q., Chen, P., Ma, Q., Sun, D., & Pan, Y. (2011). How an uncertain cue modulates subsequent monetary outcome evaluation: an ERP study. *Neuroscience letters*, 505(2), 200–204. https://doi.org/10.1016/j.neulet.2011.10.024

Tables

Table 1. Demographic Data of Participants.

	ASD (n = 21)	Control (n = 30)	Test statistic	*P*-value
Mean age in years; SD (range)	26.3; 5.2 (18-37)	28.2; 5.2 (18-46)	t = 1.087	0.282
Education level, n (%) - Middle School - High School - College†	 1 (4.8) 8 (38.1) 12 (57.1)	 0 (0.0) 5 (16.7) 25 (83.3)	Fisher's Exact = 4.659	0.07

Note. ASD, autism spectrum disorder; SD, standard deviation. † includes attendance or completion.

Figure legends

Figure 1. Examples of comic strips presented in attribution of intentions (AI), physical causality with characters (PCCh), and physical causality with objects (PCOb) conditions. Each comic strip consists of a short story of three sequential pictures followed by either a congruent, or an incongruent ending. Reproduced with permission by Eric Brunet-Gouet.

Figure 2. Schematic representation (montage) of the 28 electrodes' localization in the 10-10 system (view from top). Gnd, ground electrode; Ref, reference electrode.

Figure 3. Median accuracy on the three conditions of the Comic Strips task in participants with autism spectrum disorders (ASD) and in neurotypical (NT) volunteers. AI, attribution of intentions condition; PCCh, physical causality with characters condition; PCOb, physical causality with objects conditions. Error bars represent interquartile range. * indicates statistically significant differences ($P < 0.05$).

Figure 4. ERP waveforms elicited by the third picture in the (A) neurotypical control group (NT) and (B) in the autism spectrum disorder (ASD) group. Waveforms were plotted from the averaged mean amplitudes of participants in the NT and ASD groups respectively, left and right posterior electrode groups merged. AI, attribution of intentions condition; PCCh, physical causality with characters condition; PCOb, physical causality with objects condition; ms, milliseconds; μV, microvolts. Shaded areas represent standard-error of the mean. The thick grey lines over the horizontal axis represent the time intervals during which significant within-group differences were observed: NT group: * $P = 0.005$ AI vs. PCCh; ** $P < 0.001$ AI vs. PCCh and AI vs. PCOb; *** $P = 0.018$ AI vs. PCOb; ASD group: **** $P < 0.05$ AI vs. PCCh and AI vs. PCOb; ***** $P < 0.05$ PCCh vs. AI and PCCh vs. PCOb; Between-group differences for the PCCh condition were significant only in the 200-400 ms interval ($p = 0.018$) and the 400-600 ms interval ($p = 0.003$).

Figure 5. ERP waveform elicited by the third picture in the attribution of intentions (AI) condition, in the autism spectrum disorders (ASD) group (n = 21) and a subgroup (n = 10) of similarly performing neurotypical participants (CTR). Waveforms were plotted from the averaged mean amplitudes of all participants in each group/subgroup, left and right posterior electrodes groups merged. ms, milliseconds; μV, microvolts. Shaded areas represent standard-error of the mean.