



OPEN Exploring divergent kinematics in autism across social and non-social vitality forms

G. Di Cesare^{1,2,7}, R. Bruschetta^{3,7}, A. Vitale^{2,4,5}, A. Pelosi², E. Leonardi³, F. I. Famà³, M. Mastrogiuseppe³, C. Carrozza³, S. Aiello³, A. Campisi³, R. Minutoli³, P. Chilà³, S. Campisi³, F. Marino³, G. Pioggia³, G. Tartarisco^{3✉}, V. Cuccio^{6,7} & L. Ruta^{3,7}

Vitality Forms (VFs) constitute the dynamic essence of human actions, providing insights into how individuals engage in activities. The ability to perceive and express VFs during interpersonal interactions is pivotal for understanding others' intentions, behaviors, and fostering effective social communication. Despite their ubiquity in all actions, research exploring the role of VFs in neurodivergent conditions related to social and communicative skills, particularly in autism, remains limited. This study aims to investigate the expression of different VFs during the execution of both social and non-social actions in children with an Autism Spectrum Condition (ASC) in comparison to neurotypical children (NT). ASC children and NT children were asked to move a small bottle either towards a target point (non-social context) or moving it towards a receiver (social context) with different VFs specifically neutral, gentle, or rude. Videotaped tasks were subsequently analyzed to study kinematic parameters characterizing VFs. Our results highlighted three main findings: (1) overall, ASC children are able to tune the motor profile of their actions, effectively conveying both gentle and rude VFs; (2) distinct kinematic parameters in the execution of VFs are able to distinguish autistic children from NT children; (3) the social context significantly influences the child's ability to express positive and negative VFs in autism. Taken together, these findings provide new insights to understand how VFs contribute to the complex dynamics of social communication in neurodivergent autistic children, providing a valuable contribution for future interventions and support strategies.

The richness of human interactions is rooted not only in the 'what' and 'why' of our actions but also in the nuanced 'how' with which these actions are performed. Daniel Stern captured the many ways in which an action can be performed with the concept of Vitality Forms (VFs), which convey the affective color of our behavior, ranging from gentle to rude expressions^{1,2}. VFs are an integral part of social communication, allowing individuals to express and interpret behaviors, attitudes, and feelings during interactions by the vital modulation of an action execution. Vitality forms differ from basic emotions in triggering factors, duration, voluntary nature, and brain activity. Basic emotions (joy, anger, sadness, fear, love, disliking, and liking) are short-lasting events triggered by internal (e.g., memory) or external factors (e.g., seeing a lion), typically ending soon after the stimuli cease. These emotions are not voluntary and are not always related to actions; they can be elicited passively (e.g., feeling happy without action). Emotions often induce visceromotor responses^{3,4}. In contrast, vitality forms are voluntary events that modulate behavior (actions or words). For example, based on our attitude toward another person, we can greet them warmly or coldly. Vitality forms are expressed continuously and can be influenced by external factors like social context. Another element differentiating vitality forms from emotions is their neural correlates. Perception and expression of vitality forms activate the dorso-central insula and middle cingulate cortex, while emotions activate different brain areas, such as the amygdala, hippocampus, anterior cingulate cortex, and thalamus.

Previous kinematic studies carried out by Di Cesare et al. showed that, when interacting with others, their gentle and rude VFs modulate our own motor behavior thus stressing their fundamental role in social communication⁵. In this view, a recent kinematic study of Lombardi et al. demonstrated that observing a

¹Cognitive Architecture for Collaborative Technologies Unit, Italian Institute of Technology, Genoa, Italy. ²Department of Medicine and Surgery, University of Parma, Parma, Italy. ³Institute for Biomedical Research and Innovation (CNR-IRIB), National Research Council of Italy, Messina, Italy. ⁴U-VIP: Unit for Visually Impaired People, Istituto Italiano di Tecnologia, Genoa, Italy. ⁵Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Genoa, Italy. ⁶Department of Ancient and Modern Civilizations, University of Messina, Messina, Italy. ⁷These authors contributed equally. ✉email: gennaro.tartarisco@cnr.it

humanoid robot expressing a positive attitude (happy facial expression and gentle action) decreased the velocity, acceleration and maximum height of the agent's movement⁶. In contrast, the observation of the same robot expressing a negative attitude (angry facial expression and rude action) increased the values of these kinematic parameters.

Recent neuroimaging studies have explored the neural mechanisms underpinning the perception and expression of VFs in both action and speech^{7–9}. These findings highlight the involvement of the middle and posterior insula short gyri, suggesting a dedicated mirror mechanism of this insula portion, facilitating the coding and decoding of VFs during action execution and observation^{7–13}. While VFs play a pivotal role in human social communication, their exploration in individuals with neurodivergent conditions such as autism, remains an emerging area of research.

Autistic individuals often struggle with everyday social situations and may find it difficult to understand social meanings in naturalistic interactions. The Enactive Mind (EM) approach¹⁴ views cognition as bodily experiences from an organism's adaptive actions on salient environmental aspects. Reduced attention to social stimuli^{14–16} and altered perception of social salience may partially explain the different acquisition of embodied social cognition in autism. Another framework, that may be relevant to VFs expression and processing, suggests that self/other-awareness and children's emotional relationships with embodied persons are foundational for understanding social meanings in others' actions¹⁷. From an anthropological perspective, it is also important to consider differences in how autistic children show corporeal reflexivity and they perceive and respond to their own bodies in complex social situations. They may exhibit variations in their awareness of themselves as experiencing subjects and physical objects visible to others^{18,19}.

From an anthropological view point, differences in corporeal reflexivity may also be taken into account and autistic children, in naturalistic complex social situations, may show different own body awareness in terms of an experiencing subject and a physical object accessible to the gaze of others.

Both of these neuroscience approaches may be relevant in their contribution to the distinct motor kinematic patterns expressed and understood by autistic individuals.

At a neurological level, previous studies have identified a specific dysfunction in the fronto-parietal mirror mechanism in autism, potentially contributing to the deficits in other's intentional actions decoding, found in autistic individuals²⁰. At a behavioral level, few studies have explored the potential recognition deficits of VFs in autistic children.

Findings suggested that autistic children struggle to identify VFs when observing others' actions^{21,22}, while they do not differ from neurotypical children in discerning the goal of the action²². This impairment of ASC children in VFs recognition, could be due to a different expression of VFs in terms of kinematics.

To date, there is a lack of information on how they perform action VFs in both social and non-social contexts. This study aims to address this gap by investigating two main aspects: (1) examining the spatiotemporal features characterizing the execution of gentle and rude VFs by autistic and neurotypical children in social and non-social scenarios, and (2) exploring, for the first time, whether and how a social context such as the presence of a person receiving the action (receiver) may influence the expression of VFs. Examining VFs in both social and non-social contexts is crucial for understanding how autistic children navigate these expressive nuances, particularly considering the communication and social interaction deficits inherent in the condition.

To this purpose, we carried out a kinematic experiment in which a group of children with an autism spectrum condition (ASC) and a group of neurotypical children (NT) were asked to grasp a bottle and move it (non-social scenario) or moving it towards another person (social scenario) with a gentle, neutral or rude VFs.

We hypothesized differences in the kinematics of VFs expressed by autistic children compared to neurotypical children, irrespective of the context (social or non-social); also, in social contexts, the presence of a receiver will impact VFs expression in autistic children.

Methods

Participants

The study was carried out on a group of children with an autism spectrum condition (ASC) ($n=25$) and a group of neurotypical children (NT) ($n=23$) (range age: 7–13 years). Both ASC and NT children were male, right-handed and with an IQ > 70. The two groups were comparable in terms of age and IQ ($p > 0.05$). Exclusion criteria for autistic children included the presence of a neurometabolic or genetic syndrome, epileptic encephalopathies and/or epilepsy, structural malformations of the central nervous system and major movement disorders. Exclusion criteria for NT children included a family history of autism, a personal history of language delay, intellectual disability, or any neurodivergent conditions such as autism, ADHD, motor dyspraxia, dyslexia, autism, anxiety, and so forth. Two ASC children showed outliers kinematic values during the data analysis process and were therefore excluded from the retrospective research. Autistic children were recruited and tested at the National Research Council of Italy, Institute for Biomedical Research and Innovation (CNR-IRIB) in Messina and Catania (Italy), by referring the study to family associations, and via the research center's website. NT children were recruited via mainstream schools in the local territory. The study was approved by the ethics committee of the AOUP of Palermo (protocol number 09/2021), and all the families that voluntarily participated in the study signed a written informed consent and all methods were performed in accordance with the relevant guidelines and regulations.

Neuropsychological assessment

All autistic children had received a clinical diagnosis of an autism spectrum condition by a multidisciplinary team including experienced developmental psychologists and child neuropsychiatrists, supported by gold standard assessments such as the Autism Diagnostic Observation Schedule-2 Edition (ADOS-2)^{23,24}. The Test of Emotion Comprehension (TEC)²⁵, was used to assess the understanding of nine different domains of

emotional comprehension in both the ASC and the NT children. Additionally, the Italian translation²⁶ of the Wechsler Intelligence Scale for Children 4^o edition (WISC-IV)²⁷ was employed to measure intellectual ability of children in the autism spectrum. The WISC-IV allowed us to assess the Verbal Comprehension Index (VCI), Visual Spatial Index (VSI), Fluid Reasoning Index (FRI), Working Memory Index (WMI), and the Processing Speed Index (PSI). In the NT group, the Raven's Standard Progressive Matrices (RPM)²⁸ were used to estimate non-verbal intelligence and logical thinking. None of the children had a clinical diagnosis of developmental coordination disorder (DCD) or motor dyspraxia.

Demographic and clinical characteristics of the sample are reported in Table 1.

To assess the understanding of the meanings associated with “gentle” and “rude,” a specially designed semi-structured survey was administered to each child before commencing the experimental phase. The survey took about 15 min to be completed. The experimenter presented three types of questions. Initially, the child was prompted to provide definitions for the terms “gentle” and “rude” (“*What does the word gentle/rude mean?*”). Subsequently, the child was asked to offer examples of both gentle and rude behaviors (“*Can you please give me an example of being gentle/rude?*”). Following this, the comprehension of the consequences of exhibiting gentle or rude behavior towards others was explored (“*What happens if you are gentle/rude to another person?*”). After this initial survey, all participants received explicit definitions of “gentle” and “rude,” along with their synonyms, aimed at establishing a clear and unambiguous conceptual understanding of the two terms. Subsequently, each child was presented with 12 brief video clips, each lasting a few seconds, depicting actions classified as either gentle or rude (e.g., ‘kicking a ball in a rude way’ or ‘stirring a soup in a gentle way’).

To prevent a possible imitation effect of the vitality forms observed in these videos on the experimental task, the presented actions differed from those performed during the experiment (passing an object). Additionally, some actions were performed with a different effector (foot). After each scene, participants were asked to verbally identify whether the observed action was gentle or rude, and the accuracy of their responses was recorded by the experimenter. Only participants who achieved a minimum score of 75% correct answers (9 out of 12) were considered eligible to proceed to the subsequent experimental phase. Four children were excluded because they fell below the accuracy ratio and their understanding of the concepts of rude and gentle has not been acquired yet.

Experimental paradigm

The setup featured a square table with a small bottle positioned on it, accompanied by two chairs—one for the child and the other opposite the child. A high-resolution HC-X920 digital video camera, angled at 30 degrees on a tripod, recorded the table and the child's hands. Another high-resolution camera on a tripod was placed in front of the child and the table, capturing the scene from the frontal perspective. The experimental design comprised two conditions: social and non-social. In the social scenario, a female experimenter (receiver) occupied the chair opposite the child, while an empty chair was placed in front of the child in the non-social scenario. The child was instructed to sit in the chair facing the table, placing the right hand on the table in a pinching position, aligned with the participant's mid-sagittal plane and 9 cm away from the table edge (Fig. 1, distance a–b). A bottle was positioned 12 cm from the starting position (Fig. 1, distance b–c). Before starting the task, the experimenter explained the instructions. Children were tasked with grasping the bottle and moving it forward with varying VFs (rude, neutral, or gentle), as specified by the experimenter. In the non-social condition, the child moved the bottle forward beyond a marked line on the Table 12 cm far from the bottle (Fig. 1, distance c–d); in the social condition, the child moved the bottle toward a second experimenter seated on the opposite chair (Fig. 1).

To establish a baseline for each participant, children initially performed the bottle movement in both social and non-social conditions for seven repetitions each, using their natural approach without specific instructions on VFs. Subsequently, children executed the same action with either gentle or rude VFs, according to the instruction provided by the experimenter in both social and non-social conditions for seven repetitions each, resulting in a total of 42 actions per child [7 actions × 2 VFs × 2 Contexts + 7 neutral actions (baseline) × 2 Contexts]. To mitigate potential performance bias due to factors like fatigue, boredom, or attention issues, all trials were counterbalanced concerning VFs and conditions.

Video editing

Video footage of all the children was carefully reviewed by a trained researcher. A total of 201 and 143 trials were excluded for ASC and NT respectively, because they did not align with task requirements, such as instances where the child moved the bottle outside the designated area or the camera failed to properly capture the action.

	ASC children (n = 25)	NT children (n = 23)
Age	9.87 ± 2.2	9.63 ± 2.14
IQ total score	98.4 ± 14.4	91.2 ± 8.5
ADOS-2 SA	9.2 ± 4.1	n.a
ADOS-2 RRB	3 ± 2.2	n.a
ADOS-2 total score	12.3 ± 4	n.a
TEC total score	6 ± 1.7	8.3 ± 0.8

Table 1. Demographic and clinical characteristics of the sample. *IQ* intellectual quotient, *SA* social affect, *RRB* restricted repetitive behaviors, *TEC* test of emotion comprehension, *n.a* not applicable.

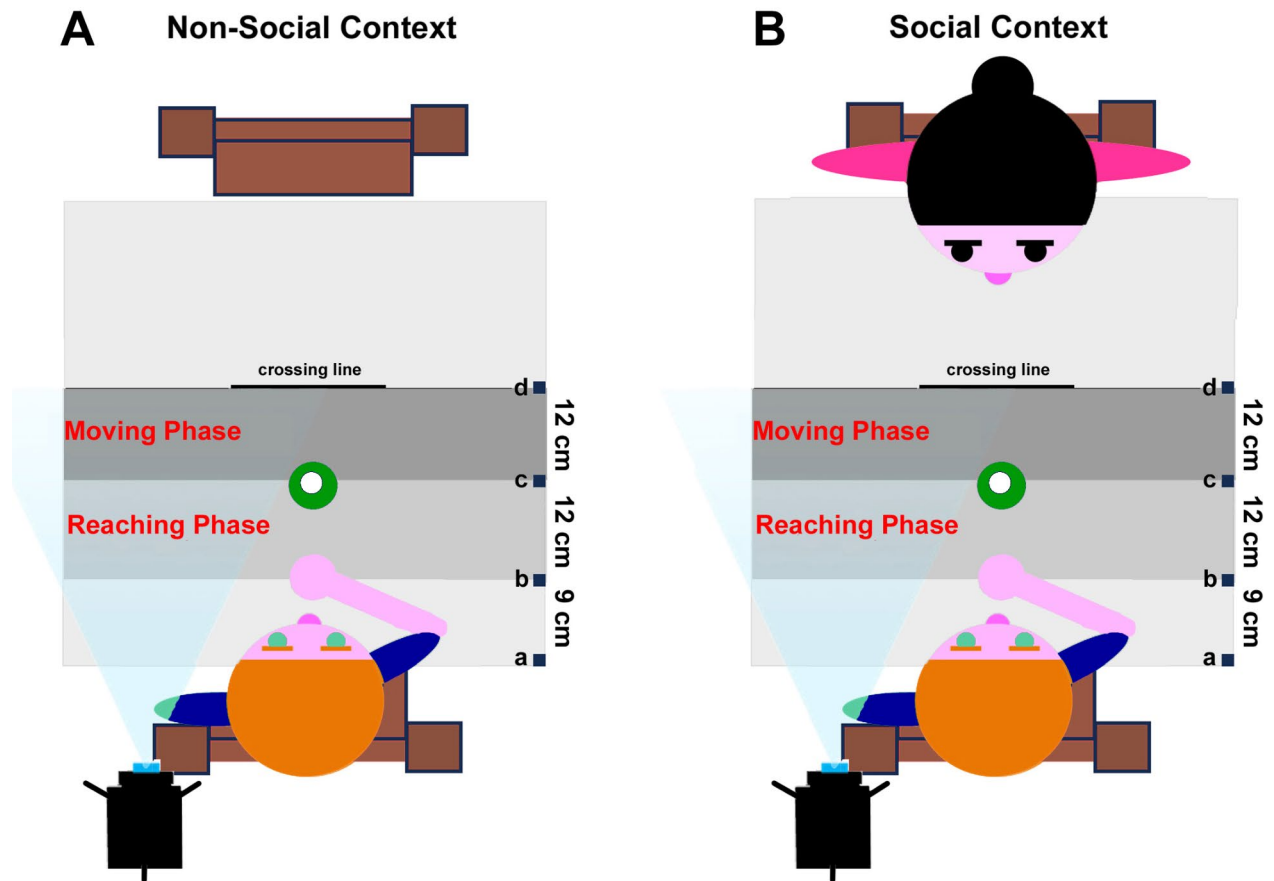


Fig. 1. Experimental Setting adopted during the non-social (A) and social (B) contexts.

Following the cleaning procedures, the remaining number of trials included in the analysis was 849 for ASC and 823 for NT.

Before starting the analysis, each action was divided into two distinct phases using Virtual Dub software:

Phase 1: spanning from the initiation of the reaching movement to the grasping of the bottle (Fig. 2A1,A2, reaching phase);

Phase 2: covering the period from grasping the bottle to completing the action (Fig. 2A3,A4,moving phase).

In the reaching phase, our primary aim was to explore the way children grasped the bottle. In the moving phase, our attention shifted to examining how and to what extent the child manipulated the bottle to move it forward, taking into account the specific VFs and the social and non-social contexts surrounding the action (Fig. 2).

Kinematic parameters extraction

Following the pre-processing stage, each video was analyzed using Deep Learning models for the automatic frame-by-frame tracking of the child's hand movements during the task. The automatic extraction of trajectories was carried out by applying and customizing the MediaPipe Hand Landmarker solution²⁹. MediaPipe Hand Landmarker is a powerful computer vision framework developed by Google's MediaPipe team. It utilizes advanced deep learning techniques, specifically convolutional neural networks (CNNs), to enable real-time hand movements tracking and landmark estimation from video or webcam streams. With MediaPipe Hand Landmarker, developers can leverage pre-trained models and algorithms to identify and locate key reference points on the hand, such as fingertips, knuckles, and the palm. More precisely, this model was able to detect the localization of 21 hand-knuckle coordinates within the detected hand regions. The framework incorporates a multi-stage regression approach, thanks to which the model predicts the landmarks in a sequential manner, progressively refining its estimates at each stage. This process enables to handle challenges like occlusions and improve the accuracy of landmark estimation. Model training was conducted using large annotated datasets: approximately 30K real-world images, as well as several rendered synthetic hand models imposed over various backgrounds. Specifically, human annotators marked the positions of hand landmarks in the training images or frames, creating ground truth labels. Then, the model learned to minimize the difference between its predicted landmarks and the ground truth labels through optimization techniques like backpropagation and gradient descent.

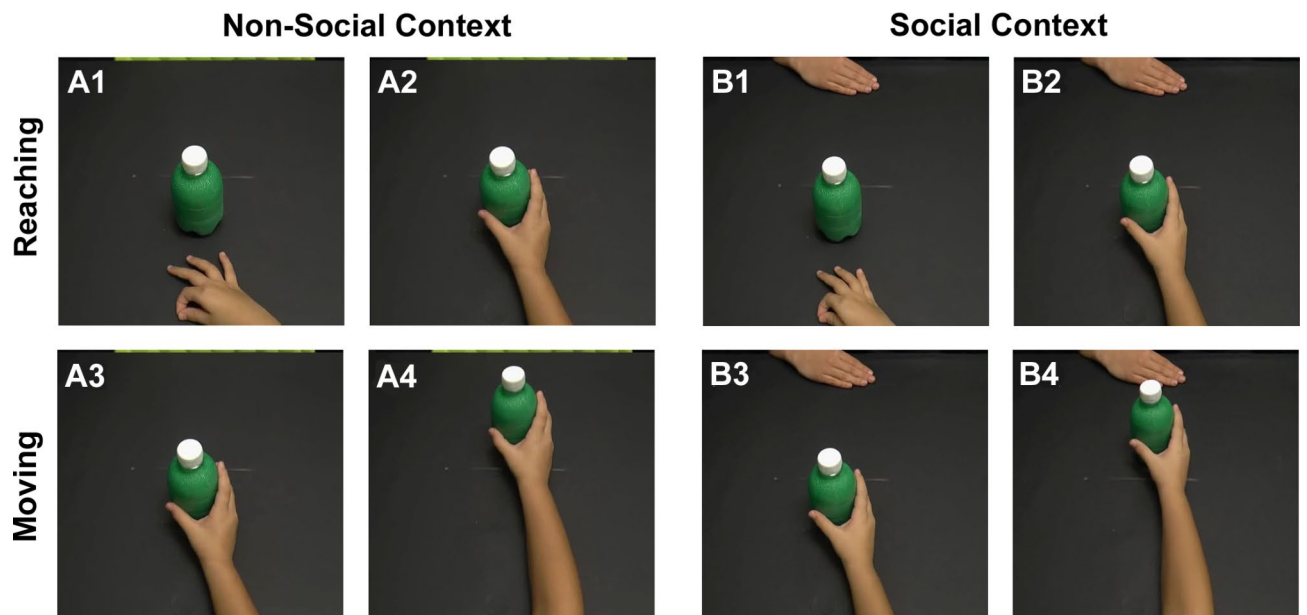


Fig. 2. Phase 1, reaching phase, from the beginning of the movement (A1,B1) until grasping the bottle (A2,B2); Phase 2, moving phase, from grasping the bottle (A3,B3) until the end of the action (A4,B4) during both social and non-social contexts.

The framework is mainly composed of a palm detection model and a hand landmarks detection model. The Palm detection model is responsible for detecting and localizing the palm region in the input image or frame. It is based on a custom-built MobileNetV3 architecture, a lightweight CNN known for its efficiency in real-time applications and designed for efficient palm detection. Once the palm region is detected, the Hand Landmark model is employed to estimate the precise 3D coordinates of the hand landmarks, such as the fingertip points and joints. This model also utilizes a CNN architecture specifically tailored for accurate hand landmark detection. Since running the palm detection model is time consuming, when in video or live stream running mode, Hand Landmarker uses the bounding box defined by the hand landmarks model in one frame to localize the region of hands for subsequent frames. Hand Landmarker only re-triggers the palm detection model if the hand landmarks model no longer identifies the presence of hands or fails to track the hands within the frame. This reduces the number of times Hand Landmarker triggers the palm detection model.

In our task, we modified the network architecture to accurately distinguish between the child's right hand and the experimenter's hand in the social context. Specifically, we customized the network to provide as output the rotation of the rectangular bounding box for each hand (the rectangular region of interest (ROI) that is used to identify and track hand landmarks in an input image or video frame), in order to differentiate between them. This modification allowed for automatic and reliable identification of the child's hand, even during the social context, and facilitated accurate tracking of their movements. The recorded videos have a frame rate of 25 fps, resulting in estimated trajectories that are sampled every 40 ms.

Feature extraction

For the characterization of our specific task, we focused on three reference points (markers) on the hand. Specifically, the wrist was considered as the most stable point for analyzing the kinematic parameters of the whole action (Fig. 3A), while the tips of the thumb and index finger were used to analyze the grasping of the bottle (Fig. 3B).

The features were extracted for each marker using the trajectories in two dimensions, specifically the x and y coordinates of the Cartesian plane, with the origin located at the top left corner of the frame (Fig. 3). Specifically, we extracted the following features.

Velocity and acceleration (both phases) the velocity of the wrist was computed as the Euclidean distance between the location of the reference point in every two subsequent frames divided by their temporal distance. Instead, the acceleration was calculated as the difference between two consecutive velocity samples. For both signals we calculated the x and y components and the total module. Moreover, we resampled them with a frame interval of 5 ms and we applied a moving average filter with a window of three samples to reduce noisy fluctuations (Fig. 3E). Mean and maximum values with the corresponding time were then computed for both the velocity (px/s) and acceleration (px/s²). Furthermore, we also computed the maximum deceleration along with its corresponding time point.

Maximum opening (phase 1) the maximum opening in pixels during the bottle grasping phase was calculated as the maximum Euclidean distance between thumb and index fingertips. The corresponding time was also determined. The feature was normalized by the distance between the wrist and the tip of the middle finger in order to take into account children's dimensions.

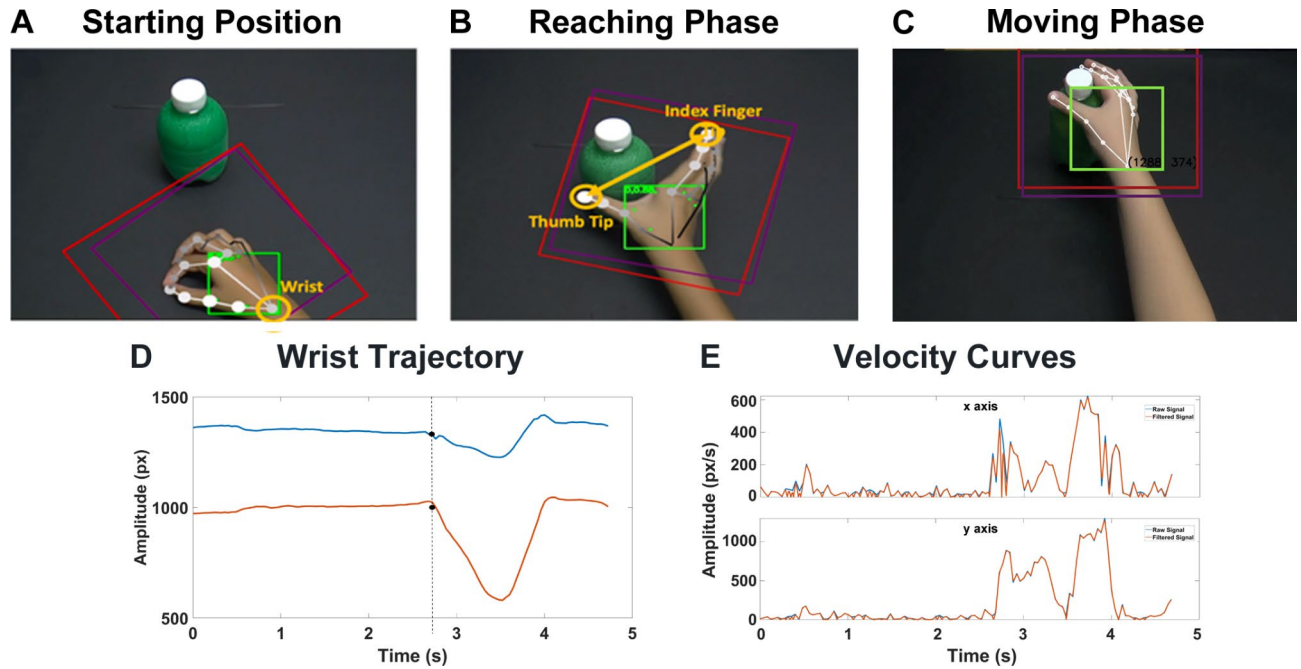


Fig. 3. Hand markers employed for feature extraction: the wrist point was used to get kinematic parameters useful to compute the moving phase of the action (A); thumb and index fingertips used to compute the grasping phase of the action (B). Example of wrist trajectory (x-axis component in blue and y-axis component in red) automatically extracted from one subject (D). The starting point of the action, where the child begins to move, is indicated by the black dots. Example of velocity curve extracted from the x and y components of the wrist trajectory of one subject (E). The original signal is shown in blue, the interpolated signal is indicated in red.

Grasping time (phase 1) the duration of grasping time in seconds was measured as the time interval between the beginning of the action and the instant of bottle grasping.

Maximum displacement along x and y during action execution (phase 2) we calculated the maximum extent of movement in pixels along the two axes by measuring the difference between the maximum and minimum ordinate, and the maximum and minimum abscissa covered by the hand.

Coordinates of the hand position at the end of the action (phase 2) we checked whether the model correctly performed the identification and tracking of the child's hand in the final frames of the video. In cases where the hand was not successfully detected in the last part of the video, resulting in missing data, we excluded the final hand coordinates from the features. Specifically, we excluded this variable for all cases where the child's hand was not detected for more than two consecutive final frames.

Data analysis

All statistical analyses and graphical visualization were implemented in R (version 4.2.1)³⁰. To check for differences between the ASC and NT groups on demographic variables (age) and IQ, we conducted univariate ANOVAs. After extraction, all kinematic features were normalized to the baseline condition (neutral actions), in which participants performed the task in a neutral way, without receiving any instruction on a specific VFs by the experimenter. To assess possible similarities or differences between actions performed by ASC and NT children, we investigated kinematic features characterizing their gentle and rude actions. To this aim, for each parameter, data were organized to carry out a General Linear Model (GLM), with VITALITY (gentle and rude), and CONTEXT (social and non-social) and GROUP (ASC, NT) as factors of interest. Furthermore, in order to study whether and how ASC children modulate kinematic parameters during the execution of gentle and rude VFs, we carried out contrasts within ASC group.

Finally, correlation analyses were carried out to assess possible relations between neuropsychological measures and VFs kinematic parameters.

Results

Results described below concern the analysis of the following kinematic parameters recorded in both ASC and NT children: mean velocity (V_m), maximum velocity (V_{max}), time maximum velocity (V_{tmax}), mean acceleration (a_m), maximum acceleration (a_{max}), maximum deceleration (D_{max}), time maximum deceleration (D_{tmax}), maximum opening (O_{max}), time opening max (O_{tmax}), maximum displacement X (DX_{max}), maximum displacement Y (DY_{max}). The main significant results are reported in Table 2; Fig. 4 (for more details see also supplementary information).

	Reaching phase			Moving phase		
		F(1,41)	p		F(1,40)	p
V_m	Vitality	80.92	0.001	Vitality	59.76	0.001
	Group*Vitality*Context	3.81	0.06	Context*Group	7.61	0.01
				Group*Vitality*Context	6.59	0.01
V_{max}	Group	3.35	0.07	Vitality	24.69	0.001
	Vitality	66.84	0.001	Context	4.64	0.03
V_{tmax}	Context*Group	6.13	0.02			
a_m	Group	7.94	0.01	Group	6.21	0.02
	Vitality	86.09	0.001	Vitality	21.89	0.001
				Vitality*Group	4.66	0.03
a_{max}	Group	8.19	0.01	Group	4.66	0.003
	Vitality	20.25	0.01	Vitality	13.23	0.001
				Context	5.76	0.02
				Vitality*Group	3.74	0.06
				Context*Vitality	6.09	0.02
D_{max}	Vitality	21.53	0.01	Group	3.76	0.06
	Context*Group	4.46	0.04	Vitality	37.17	0.001
	Group*Vitality*Context	5.81	0.02	Vitality*Group	4.96	0.03
D_{tmax}	Vitality	5.08	0.01	Group	3.74	0.06
				Context*Vitality	3.58	0.06
DX_{max}				Vitality	14.69	0.001
				Group*Vitality*Context	5.85	0.02
DY_{max}				Vitality	27.67	0.001
				Context*Vitality	5.32	0.03
O_{max}	Group	7.11	0.01			
	Vitality	12.24	0.01			

Table 2. Main and interaction effects: significant F and p value.

Analysis between groups

A significant main effect of GROUP was found during the reaching phase for the following kinematic parameters: a_m ($p < 0.01$), a_{max} ($p < 0.01$) and O_{tmax} ($p < 0.01$). Regarding the moving phase a main effect of GROUP was found for a_m ($p < 0.05$), a_{max} ($p < 0.05$), D_{max} ($p < 0.06$) and D_{tmax} ($p < 0.06$, Fig. 4) (see Fig. 4; Table 2).

Interaction effect analyses

The interaction VITALITY * GROUP was significant only for the moving phase for the following kinematic parameters: D_{max} ($p < 0.03$); a_{max} ($p = 0.06$); a_m ($p < 0.03$) (Fig. 4C).

The interaction CONTEXT * VITALITY was significant only for the moving phase in the following kinematic parameters: DY_{max} ($p < 0.05$), TD_{max} ($p = 0.06$), a_{max} ($p < 0.05$), a_m ($p < 0.05$), V_{max} ($p < 0.05$, Table 2).

The interaction CONTEXT * GROUP was significant for the reaching phase for D_{max} ($p < 0.05$) and the V_{tmax} ($p < 0.05$) parameters. The same interaction revealed also a significant difference of V_m ($p < 0.01$) during the moving phase (Fig. 4D).

Finally, the interaction GROUP * VITALITY * CONTEXT revealed significant results during the reaching phase for D_{max} ($p < 0.05$) and V_m ($p = 0.06$), and, also during the moving phase for DX_{max} ($p < 0.05$) and V_m ($p < 0.01$) parameters (Fig. 4E).

Correlation analyses

Table 3 shows results regarding the correlation analyses carried out between neuropsychological measures and VFs kinematic parameters.

Discussion

The present study investigates the communication of positive and negative Vitality Forms (VFs), such as gentle and rude, by ASC and NT children, in both social and non-social contexts. Our investigation extends to exploring kinematic differences between action VFs performed by ASC children and NT children. To address these objectives, we asked both groups to move a small bottle either to a designated point on the table (non-social context) or moving it towards a receiver (social context) while expressing neutral, gentle, or rude VFs. Video recordings of these actions were analyzed using deep learning models to automatically extract and examine kinematic parameters characterizing the VFs.

The study encompasses two distinct phases for each action: the *reaching phase* (from initiation to grasping the object) and the *moving phase* (from grasping the object to completing the action). Remarkable findings emerged from our study. Firstly, ASC children demonstrated a certain ability to adjust the motor profile of

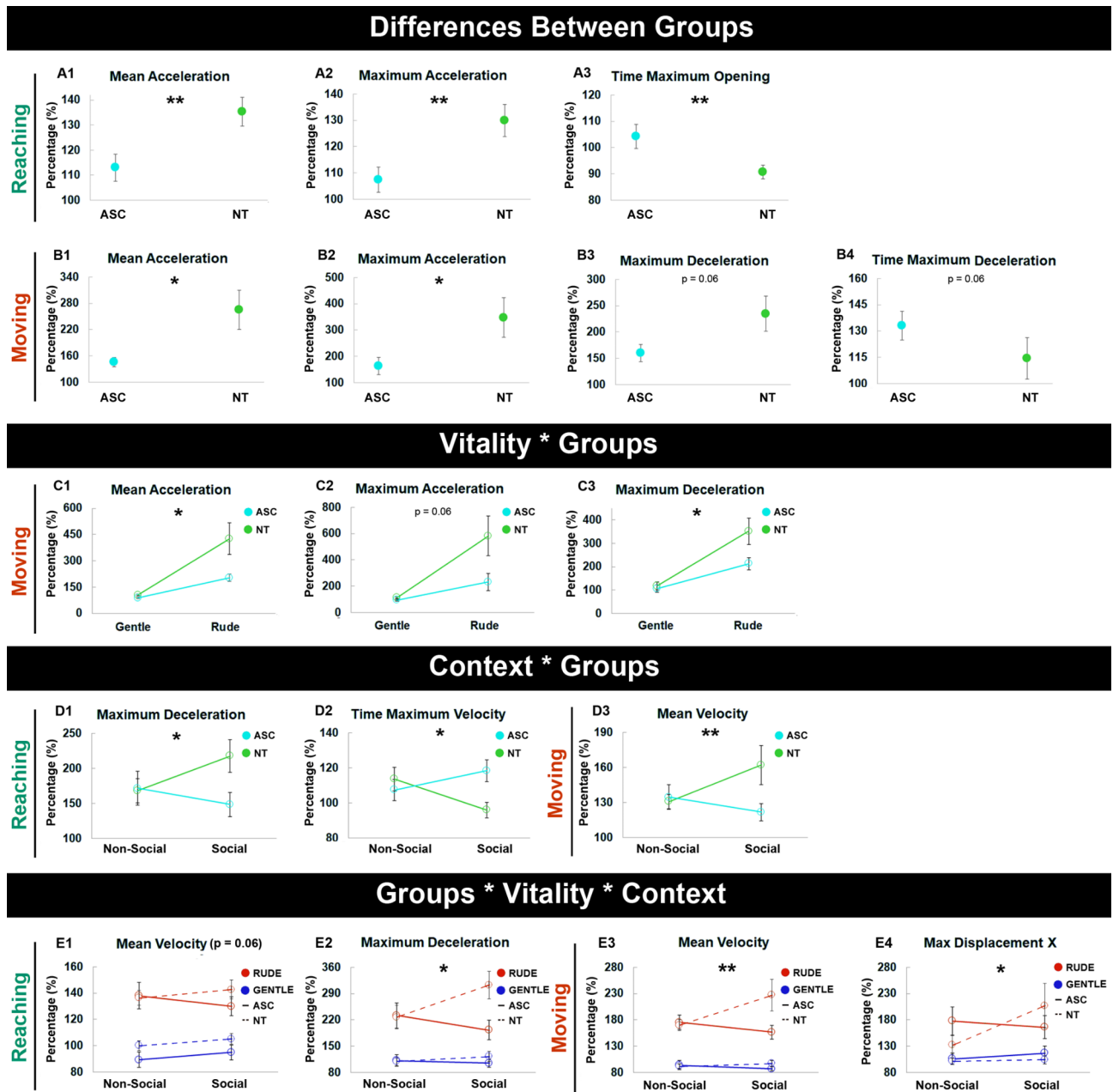


Fig. 4. Overview of the main differences between ASC and NT groups regarding the reaching and moving phases of the vitality forms expression. All the indicated values are normalized to the baseline condition (100% corresponds to the baseline). Vertical bars represent the standard errors. Significance (* $p < 0.05$, ** $p < 0.01$).

their actions to convey gentle and rude VFs, effectively modulating various kinematic parameters such as mean velocity, maximum velocity, mean acceleration, and maximum acceleration. These findings indicate that once ASC children comprehend the meanings of rude and gentle VFs, they exhibit the ability to consequently modify their actions, influencing all pertinent kinematic parameters associated with the specific VFs they intend to express. A noteworthy strength of our study lies in the stringent criteria for participant inclusion, where only children scoring above 75% of accuracy during the semi-structured preliminary interview were admitted to the experiment. This approach serves to mitigate potential confounding effects linked to the lack of VFs expression arising from a misunderstanding of the terms “gentle” or “rude”. Our results are also in line with findings provided by Casartelli and colleagues³¹, who reported that two kinematic parameters such as the peak velocity and peak acceleration changed when ASC children were asked to perform different actions with different VFs.

While ASC children showed the ability to modulate their action VFs, interestingly, our results highlighted that the motor profiles of ASC children diverged from those of TD children. Specifically, ASC children showed reduced mean acceleration, maximum acceleration and maximum deceleration when moving the bottle

ASD children									TD children								
Mean acceleration									Mean acceleration								
	Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.		Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.
ADOS	0.099	- 2.4	- 0.26	- 0.03	0.07	0.10	- 0.28	- 0.09	RAVEN	- 0.13	0.009	0.01	- 0.07	0.04	- 0.22	0.01	- 0.01
TEC	- 0.201	0.28	0.20	0.13	- 0.18	- 0.18	0.23	0.25	TEC	- 0.35	- 0.067	- 0.04	- 0.27	- 0.01	- 0.49*	- 0.04	- 0.16
WISC	0.241	- 0.04	- 0.11	0.17	0.06	0.32	- 0.46	0.19									
AS	0.59**	- 0.23	- 0.33	0.30	0.60**	0.46*	- 0.46*	0.19									
Max acceleration									Max acceleration								
	Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.		Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.
ADOS	0.07	- 0.24	- 0.25	0.04	- 0.07	0.17	- 0.24	- 0.02	RAVEN	- 0.29	0.12	0.10	0.02	- 0.19	- 0.23	0.11	0.10
TEC	- 0.98	0.08	0.07	0.02	0.03	- 0.18	0.07	0.11	TEC	- 0.55**	- 0.11	- 0.10	0.25	0.20	- 0.53**	- 0.10	- 0.12
WISC	0.58**	- 0.13	- 0.17	0.52*	0.54*	0.54*	- 0.18	0.46*									
AS	0.21	- 0.46*	- 0.47*	0.08	0.20	0.19	- 0.48*	0.03									
Time max acceleration									Time max acceleration								
	Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.		Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.
ADOS	0.09	0.14	0.12	0.14	- 0.38	0.15	0.14	0.13	RAVEN	0.16	0.13	0.12	0.33	0.15	0.27	0.09	0.30
TEC	- 0.21	- 0.17	- 0.19	- 0.16	- 0.18	- 0.19	- 0.18	- 0.15	TEC	0.20	0.15	0.18	0.004	0.19	0.17	0.17	- 0.11
WISC	0.17	0.20	0.20	0.19	- 0.04	0.18	0.20	0.19									
AS	- 0.21	- 0.17	- 0.17	- 0.18	- 0.30	- 0.17	- 0.15	- 0.18									
Max deceleration									Max deceleration								
	Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.		Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.
ADOS	0.06	- 0.14	0.05	- 0.24	0.15	0.06	- 0.02	- 0.27	RAVEN	- 0.20	- 0.22	- 0.16	- 0.31	- 0.28	- 0.001	- 0.11	- 0.35
TEC	- 0.33	0.08	- 0.03	- 0.16	- 0.31	0.24	0.14	- 0.03	TEC	- 0.15	- 0.14	- 0.03	- 0.37	0.31	0.11	0.04	- 0.48
WISC	- 0.05	0.07	- 0.15	0.29	- 0.31	0.26	- 0.02	0.18									
AS	0.63**	0.18	0.34	0.45*	0.60**	0.46*	0.09	0.24									
Time max deceleration									Time max deceleration								
	Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.		Gentle	Rude	Social	Non social	Gentle social	Gentle non soc.	Rude social	Rude non soc.
ADOS	- 0.08	0.13	0.08	0.10	- 0.18	0.14	0.16	0.01	RAVEN	0.26	0.18	0.19	0.27	0.18	0.34	0.20	0.14
TEC	- 0.09	- 0.13	- 0.19	0.01	- 0.18	0.10	- 0.15	- 0.04	TEC	- 0.10	- 0.05	- 0.06	- 0.09	- 0.12	- 0.06	- 0.01	- 0.12
WISC	0.33	0.19	0.31	0.18	0.34	0.16	0.22	0.09									
AS	- 0.61**	- 0.27	- 0.36	- 0.65**	- 0.48*	- 0.54*	- 0.22	- 0.37									

Table 3. Results of correlations analysis carried out in ASC and NT children between neuropsychological tests scores and VFs kinematic parameters (* $p < 0.05$, ** $p < 0.01$). The numbers in bold, are those marked with an asterisk (*/**) and signify statistical significance.

forward. Additionally, significant variations were observed in the time taken to achieve the maximum hand opening during the reaching moment and the point of maximum deceleration. Moreover, disparities were evident in mean acceleration, maximum velocity, and time of maximum opening during the reaching phase. Taken together, these findings suggest challenges in motor control of actions according to VFs for children with ASC, particularly when required to express rude VFs. The observed kinematic differences between ASC and NT children are in line with previous research by other authors^{32,33}, who have shown that ASC individuals tend to take longer to perform actions compared to control groups. Furthermore, these studies have shown that ASC individuals have difficulties in planning sequential motor actions, which are essential for action chains such as reaching, grasping and passing, central to the paradigm of our current study. According to these authors, the challenges faced by ASC individuals in formulating motor plans may arise from an impairment in processing multiple pieces of information, particularly in complex tasks. This cognitive demand may affect both motor efficiency and comprehension.

Another pivotal aspect of our study focused on investigating how ASC children modulate their action VFs in different contexts, such as social and non-social, highlighting potential distinctions from NT children. Particularly, ASC children exhibited reduced mean velocity during the moving phase compared to NT children, suggesting that the presence of a social receiver might have negatively influenced ASC children, impacting the execution of their action VFs. In our opinion, the inclusion of the social context held significant importance

since, as Stern argues, VFs play a crucial role in conveying our internal states to others. Therefore, the affective expression of actions is closely tied to the concept of social interaction, dependent on the presence of another individual distinct from ourselves, with whom we share intentions, interests, and emotions and these aspects are particularly important in a neurodivergent condition involving social cognition such as autism. Furthermore, in terms of interaction effects, we observed that the modulation of gentle and rude VFs tends to decrease in the social context. Essentially, when ASC children were tasked with moving a bottle to another person, the kinematic differences that distinguish rude and gentle VFs showed a significant decrease. In contrast, NT children exhibited an increase in certain kinematic parameters, such as maximum acceleration and mean velocity, when executing gentle and rude VFs during the transition from the non-social to social context. It seems that social interaction diminishes the ability of ASC children to markedly differentiate between the “gentle” and “rude” VFs for the specified parameters. Conversely, children in the NT group demonstrated better adaptability and a more consistent response to requests during the social context, maintaining a clear distinction between gentle and rude actions.

When the influence of autism profile on kinematic expression of VFs was explored, we found that higher social communication difficulties, as indicated by elevated SA scores at ADOS-2, is associated with a lower modulation of VFs, especially in social context and for rude VF. This phenomenon may be attributed to a specific challenge in modulating the vitality of actions when confronted with a social context. Additionally, it may be influenced by elements of social anxiety creating a state of action inhibition, particularly evident in the expression of rude VF. The intersection between social communication difficulties and the modulation of VFs suggests that social communication challenges in ASC children extend beyond the domain of communication and social interactions, into the domain of motor behavior, influencing the nuanced expression of vitality in actions.

Last but not least, our study contributes to the limited body of research investigating how ASC children modulate VFs during genuine social interactions using an innovative, non-invasive approach. Automatic deep learning video tracking allows a comprehensive and continuous record of motion without imposing physical constraints and capturing a more authentic representation of the child’s motor behavior. The use of artificial intelligence algorithms for kinematic motion tracking offers distinct advantages over traditional optical motion capture systems, especially when dealing with children who may resist wearing markers or sensors and who may experience sensory processing atypicalities common in autism. Understanding how social communication skills impact the motor behavior of ASC children holds practical clinical implications. Approaches that encompass a holistic understanding of the interconnected nature of social communication and motor domains stand to offer more effective support for individuals on the autism spectrum.

Limitations

While the study offers valuable insights into the expression of Vitality Forms (VFs) in children with Autism Spectrum Condition (ASC) and neurotypical (NT) children, it is essential to acknowledge certain limitations. Firstly, the study’s sample size, comprising ASC and NT children, may not fully capture the diversity within each group. A larger and more diverse sample would enhance the robustness of the findings and improve the generalizability of the study’s outcomes to a broader population of autistic children and neurotypical development. Also the expression and recognition of VFs may vary between sexes. Restricting the study to only males affect the extent to which your results can be generalized in relation to sex and it is an aspect that should certainly be investigated in future research.

Additionally, the tasks involved a specific scenario, such as moving a bottle either towards a target point (non-social context) or moving it toward a receiver (social context). While this design choice aimed to capture various aspects of VFs expression, the situational specificity may limit the study’s ability to generalize findings to a wider range of everyday social interactions. Also, the cross-sectional nature of the study provides a snapshot of VFs expression in children at a specific point in time. A longitudinal approach could offer a more dynamic understanding of how VFs expression evolves over time, providing insights into developmental trajectories in both autistic and neurotypical populations. Acknowledging these limitations is crucial for contextualizing the study’s findings and guiding future research in this area. Despite these constraints, the study contributes valuable insights into the interplay between VFs expression and social communication in ASC children, opening avenues for further exploration and the development of targeted support strategies.

Conclusion

In conclusion, our study highlights the multifaceted relationship between vitality forms (VFs), social communication and motor behavior in ASC children compared to NT children. In particular, ASC children demonstrated an appropriate ability to modulate the motor profile of their actions, effectively expressing both gentle and rude VFs. However, distinct kinematic parameters emerged as discriminators between autistic and neurotypical children during VFs execution. Importantly, the social context exerted a significant influence, shaping the expression of positive and negative VFs in autism. These findings provide new insights into the complex dynamics of social communication in neurodivergent autistic children and extend our understanding of how VFs operate in the context of autism. The study highlights the importance of considering motor behaviour alongside social communication skills, paving the way for targeted interventions. Future strategies aimed at supporting autistic children in social interactions may benefit from recognising the role of VFs and their kinematic nuances. Overall, our research represents a valuable step towards improving interventions and promoting a more comprehensive understanding of the social experiences of individuals with autism.

Data availability

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

Received: 28 February 2024; Accepted: 24 September 2024

Published online: 15 October 2024

References

1. Stern, D. *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology* (New York Basic Books, 1985).
2. Stern, D. N. *Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy, and Development* (Oxford University Press, 2010).
3. James, W. What is an emotion?. *Mind* **34**, 188–205 (1884).
4. Malezieux, M., Klein, A. S. & Gogolla, N. Neural circuits for emotion. *Annu. Rev. Neurosci.* **46**, 211–231. <https://doi.org/10.1146/annurev-neuro-111020-103314> (2023).
5. Di Cesare, G., De Stefani, E., Gentilucci, M. & De Marco, D. Vitality forms expressed by others modulate our own motor response: A kinematic study. *Front. Hum. Neurosci.* **11**, 565. <https://doi.org/10.3389/fnhum.2017.00565> (2017).
6. Lombardi, G., Sciutti, A., Rea, F., Vannucci, F. & Di Cesare, G. Humanoid facial expressions as a tool to study human behaviour. *Sci. Rep.* **14**(1), 133. <https://doi.org/10.1038/s41598-023-45825-6> (2024).
7. Di Cesare, G., Di Dio, C., Marchi, M. & Rizzolatti, G. Expressing our internal states and understanding those of others. *Proc. Natl. Acad. Sci.* **112**(33), 10331–10335 (2015).
8. Di Cesare, G., De Stefani, E., Gentilucci, M. & De Marco, D. Vitality forms expressed by others modulate our own motor response: A kinematic study. *Front. Hum. Neurosci.* **11**, 565 (2017).
9. Di Cesare, G. et al. The middle cingulate cortex and dorso-central insula: A mirror circuit encoding observation and execution of vitality forms. *Proc. Natl. Acad. Sci.* **118**(44), e2111358118 (2021).
10. Rizzolatti, G. & Craighero, L. The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).
11. Iacoboni, M. & Dapretto, M. The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* **7**(12), 942–951 (2006).
12. Rizzolatti, G. & Fabbri-Destro, M. The mirror system and its role in social cognition. *Curr. Opin. Neurobiol.* **18**(2), 179–184 (2008).
13. Keysers, C. & Fadiga, L. The mirror neuron system: New frontiers. *Soc. Neurosci.* **3**(3–4), 193–198 (2008).
14. Ami, K., Warren, J., Robert, S. & Fred, V. The enactive mind, or from actions to cognition: Lessons from autism. *Philos. Trans. R. Soc. Lond. B* **358**, 345–360. <https://doi.org/10.1098/rstb.2002.1202> (2003).
15. Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G. & Jones, W. Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* **459**, 257–261 (2009).
16. Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J. & Brown, E. Children with autism fail to orient to naturally occurring social stimuli. *J. Autism Dev. Disord.* **28**(6), 479–485. <https://doi.org/10.1023/a:1026043926488> (1998).
17. Hobson, R. P. Explaining autism: Ten reasons to focus on the developing self. *Autism* **14**(5), 391–407. <https://doi.org/10.1177/1362361310364142> (2010).
18. Ochs, E. & Solomon, O. Autistic sociality. *Ethos* **38**(1), 69–92 (2010).
19. Ochs, E. Corporeal reflexivity and autism. *Integr. Psychol. Behav.* **49**, 275–287. <https://doi.org/10.1007/s12124-015-9306-6> (2015).
20. Oberman, L. M. et al. EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cogn. Brain Res.* **24**(2), 190–198 (2005).
21. Di Cesare, G. et al. Differences in action style recognition in children with autism spectrum disorders. *Front. Psychol.* **8**, 1456 (2017).
22. Rochat, M. J. et al. Impaired vitality form recognition in autism. *Neuropsychologia* **51**(10), 1918–1924 (2013).
23. Gotham, K., Pickles, A. & Lord, C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J. Autism Dev. Disord.* **39**(5), 693–705 (2009).
24. Lord, C. et al. Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *J. Autism Dev. Disord.* **19**(2), 185–212 (1989).
25. Pons, F. & Harris, P. Longitudinal change and longitudinal stability of individual differences in children's emotion understanding. *Cogn. Emot.* **19**(8), 1158–1174 (2005).
26. Orsini, A., Pezzuti, L. & Picone, L. *Wechsler Intelligence Scale for Children IV Edizione Italiana* (Giunti, 2012).
27. Wechsler, D. *Wechsler Intelligence Scale for Children; Manual* (1949).
28. Raven, J. Raven progressive matrices. In *Handbook of Nonverbal Assessment* 223–237 (Springer, 2003).
29. Google. *MediaPipe [Computer software]*. Retrieved from <https://mediapipe.dev/> (2023).
30. RStudio Team. *RStudio: Integrated Development for R*. RStudio, PBC. <http://www.rstudio.com/> (2020).
31. Casartelli, L. et al. Vitality form expression in autism. *Sci. Rep.* **10**(1), 17182 (2020).
32. Forti, S. et al. Motor planning and control in autism. A kinematic analysis of preschool children. *Res. Autism Spectr. Disord.* **5**(2), 834–842 (2011).
33. Cattaneo, L. et al. Impairment of actions chains in autism and its possible role in intention understanding. *Proc. Natl. Acad. Sci.* **104**(45), 17825–17830 (2007).

Acknowledgements

This work has been supported by the INTER PARES project, EU POC METRO 2014/2020 (Azione I.3.1.—Codice Progetto ME I.3.1.b.); EARLY START project (Deliberazione n. 1508/2021); READS project (MIMIT, Prog. n. F/180026/01-04/X43). Giuseppe Di Cesare is founded by National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.2—Call for tender No. 247 of 19/08/2022 and No. 367 of 07/10/2022 of Italian Ministry of University and Research funded by the European Union—NextGenerationEU (Project code SOE_0000049). Roberta Bruschetta is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma; Simona Campisi is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVIII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma.

Author contributions

Conceptualization: G.D.C., G.T. and L.R.; Methodology: L.R., G.T., V.C., M.M; Formal analysis: A.P., R.B., G.T., S.C., A.V. and G.D.C.; Investigation: E.L., A.V., F.I.F., V.C., R.M., P.C., A.C., S.A., C.C.; Writing - original draft

preparation: G.D.C., A.V., V.C., S.C. and R.B.; Writing - review and editing: L.R., G.T.; Supervision: L.R., F.M.; Funding acquisition: L.R., G.P.

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/s41598-024-74232-8>.

Correspondence and requests for materials should be addressed to G.T.

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.

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

© The Author(s) 2024