Autism
The International Journal of Research and Practice

Using mobile health technology to assess childhood autism in low-resource community settings in India: an innovation to address the detection gap

Journal:	*Autism*
Manuscript ID	AUT-22-0682.R1
Manuscript Type:	Original Article
Keywords:	Autism, Digital health, LMIC, Global

Abstract:

A diagnosis of autism typically depends on clinical assessments by highly-trained professionals. This high resource demand poses a challenge in low-resource settings. Digital assessment of neurodevelopmental symptoms by non-specialists provides a potential avenue to address this challenge. In this study, we provide the proof of principle for such a digital assessment, with a cross-sectional case-control field study using mixed methods. We developed and tested an app, START, that can assess autism phenotypic domains (social, sensory, motor) through child performance and parent reports. N=131 children (2-7 years old; 48 autistic, 43 intellectually disabled, and 40 non-autistic typically developing) from low-resource settings in India were assessed using START in home settings by non-specialist health workers. The two groups of children with neurodevelopmental disorders manifested lower social preference, higher sensory sensitivity, and lower fine-motor accuracy compared to their typically developing counterparts. Machine-learning analysis combining all START-derived measures demonstrated 78% classification accuracy for the three groups. Qualitative analysis of the interviews with health workers and families of the participants demonstrated high acceptability and feasibility of the app. These results provide feasibility, acceptability, and proof of principle for START, and demonstrate the potential of a scalable, mobile tool for assessing neurodevelopmental conditions in low-resource settings.

Lay abstract

We developed and tested a novel app (START) that can be used by non-specialists to assess features of autism and related conditions in young children, in low-resource settings. The app uses different tasks and parent questionnaires to measure social, sensory, and motor functioning. In this study, non-specialist health workers used this app with 131 children (2-7 years old; 48 autistic, 43 intellectually disabled, and 40 non-autistic typically developing) in home settings. All children were drawn from low-resource suburbs of Delhi, India. Typically, an autism assessment is done by highly-trained professionals. This high resource demand poses a challenge in areas where skilled personnel are scarce and awareness of autism and related conditions is low. The START app is available in several languages and the tasks are designed to look like "games" with simple instructions. Interviews with health workers and families of the participants suggest high acceptability and feasibility of the START app. We observed a consistent pattern of differences between typically and atypically developing children in all three areas of functioning assessed by the app. The two groups of children with neurodevelopmental conditions showed a lower preference for social stimuli, higher sensory sensitivity, and lower accuracy in motor function compared to their non-autistic typically developing counterparts. Parent-report further distinguished autistic from non-autistic children. Machine-learning analysis combining all measures on the app demonstrated that it can accurately (78%) identify children from the three groups (autism, ID, non-autistic TD). This study provides proof of principle for the START mobile app.

Word count: 247

Using mobile health technology to assess childhood autism in low-resource community settings in India: an innovation to address the detection gap

ABSTRACT

A diagnosis of autism typically depends on clinical assessments by highly-trained professionals. This high resource demand poses a challenge in low-resource settings. Digital assessment of neurodevelopmental symptoms by non-specialists provides a potential avenue to address this challenge. In this study, we provide the proof of principle for such a digital assessment, with a cross-sectional case-control field study using mixed methods. We developed and tested an app, START, that can assess autism phenotypic domains (social, sensory, motor) through child performance and parent reports. N=131 children (2-7 years old; 48 autistic, 43 intellectually disabled, and 40 non-autistic typically developing) from low-resource settings in India were assessed using START in home settings by non-specialist health workers. The two groups of children with neurodevelopmental disorders manifested lower social preference, higher sensory sensitivity, and lower fine-motor accuracy compared to their typically developing counterparts. Machine-learning analysis combining all START-derived measures demonstrated 78% classification accuracy for the three groups. Qualitative analysis of the interviews with health workers and families of the participants demonstrated high acceptability and feasibility of the app. These results provide feasibility, acceptability, and proof of principle for START, and demonstrate the potential of a scalable, mobile tool for assessing neurodevelopmental conditions in low-resource settings.

Keywords

Autism, digital health, LMIC, global

INTRODUCTION

Autism is an early-onset neurodevelopmental condition with a global prevalence of ~1% (Zeidan et al., 2022). It is estimated that India is home to ~5 million families with a child with autism[1] (Arora et al., 2018; Arun & Chavan, 2021; Chauhan et al., 2019; Patra & Kar, 2021; Rudra et al., 2017). Many of these children do not get diagnosed at an appropriate time, or at all, which in turn can reduce their chances to benefit from effective interventions (Divan et al., 2021). Low community awareness about autism leads to reduced help-seeking behaviour (Minhas et al., 2015), and is exacerbated by a number of other challenges to detection. First, there is a paucity of professionals, such as developmental practitioners, psychiatrists, neurologists, and psychologists, to offer diagnostic services to a population of over 1.2 billion (Kumar, 2011). Second, current screening and diagnostic approaches typically involve time-intensive, expensive, and proprietary tools, greatly limiting access (Durkin et al., 2015). While there have been notable efforts to develop locally validated instruments for screening and diagnosis, these too typically need to be administered by specialists (Gulati et al., 2019; Juneja et al., 2014; Mukherjee et al., 2015). Third, social stigma prevents parents' seeking a psychiatric diagnosis for their child (Minhas et al., 2015).

Yet there is emerging evidence from low- and middle-income country settings that non-specialist health-worker delivered, parent-mediated intervention targeting social communication is acceptable and effective in improving outcomes for autistic children (Rahman et al., 2016). In light of such evidence, the detection gap becomes an urgent priority, highlighting the need for proactive screening for autism. The current study aimed at developing a tool that could be used by non-specialists to assess autism risk in low-resource settings, allowing the closing of the detection gap.

Mobile technologies offer a significant advantage in this effort, given their wide penetration and scalability across geographies and socioeconomic strata. Similar efforts have shown promise in high-resource settings (Dawson & Sapiro, 2019a; Egger et al., 2018). In the current study, we develop and provide the proof of principle for an online platform, consisting of a battery of tasks that index

[1] We recognise that the autism community has a diversity of views in using person-first terminology. To reflect this diversity of views, we use 'autistic children' interchangeably with 'children with autism' throughout the manuscript.

various aspects of the autistic phenotype, using a mobile device. In view of the diverse phenotypic domains associated with autism, the mobile platform (app) includes direct assessments of the child on multiple tasks that relate to social behaviour, sensory interest, and motor function. The platform also includes an assessment of parent-reported autistic features through a questionnaire and an observational measure of parent-child interaction. While the broader aim of the project is to develop tools to bridge the detection gap for autism and related neurodevelopmental conditions, the current study constitutes the first step toward this goal by developing this tool and testing its efficacy and feasibility in a field study in children with and without neurodevelopmental disorders. To this end, we have implemented and benchmarked the assessment in the form of a scalable, mobile tool, administered in the community by non-specialists to assess autism-related features in 2-to-7-year-old children in home settings in India.

METHODS

Participants: Three groups of children were recruited: (1) children with a diagnosis of Autism Spectrum Conditions (AS), N=48; (2) children with a diagnosis of Intellectual Disability (ID), N=43; and (3) typically developing (TD) children N=40 (Table 1). The AS and ID groups were recruited through a tertiary clinic and diagnosed by a specialist clinician using DSM-V criteria, while the TD group was recruited from the community. All groups were matched for chronological age. The AS and ID groups were matched on cognitive age using a language-adapted version of the *Developmental Profile-3 (DP3)* (Alpern, 2007). The AS group was contrasted with the other two groups for the severity of autistic symptoms using a locally developed and standardised tool, the *INCLEN Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD)* (Juneja et al., 2014). It is worth noting however, that all children in the AS group also met criteria for intellectual disability. The authors affirm that written informed consent was obtained from a) the primary caregiver of each child participant, and b) each adult participant included for the qualitative data. Research participants/health workers/ primary caregiver (in case of children under 18 years of age), provided written informed consent for publication of the images in Figures 1a, 1b and 2h. All signed consent forms are stored in compliance with local confidentiality laws at the Child Development Group, Sangath, New Delhi, India.

[Insert Table 1 here]

Tools:

The START (Screening Tools for Autism Risk using Technology) task battery was administered on all participants. alongside two standardized tools for assessing autism symptom severity and developmental level. Details of these tools are given below:

The Developmental Profile 3 (DP3) (Alpern, 2007): It is a parental interview scale designed to assess development and functioning across five areas: physical, adaptive, social-emotional, cognitive and communicative. We used the age-equivalent score from the cognitive subscale to estimate development that is not influenced by specific difficulties in social or communicative function.

The INCLEN Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD) (Juneja et al., 2014): It is specifically developed for diagnosing autism in 2-9-year-old children in India. It has a high validity against DSM-IV-TR diagnoses and Childhood Autism Rating Scale (Schopler et al., 1980) scores as well as with DSM-V (Vats et al., 2018).

START task battery: It is an Android app presented on a mobile device, that can be administered by non-specialists with minimal training. The app includes a battery of tasks that can be grouped within the following categories: social, motor, sensory, and parent/caregiver report and interaction (see Table 2). This choice of tasks was informed by the developmental differences commonly identified in autistic children. Social and sensory tasks of the battery are included to align with the two domains: social communication and restricted interests, commonly observed to identify autism. Furthermore, the battery includes activities to quantify parental observations and play-based interactions to supplement the information gathered through other tasks. More details of the phenotypic domains and tasks included in the battery are discussed in the sections below.

Social phenotype: Differences in *social* behaviour are a core diagnostic feature of autism. Lab-based experiments designed to measure this aspect of the autistic phenotype have often focussed on presenting social alongside nonsocial stimuli (Dubey et al., 2015; Pierce et al., 2011; Ruta et al., 2017). Such paradigms have revealed that autistic individuals have a reduced preference for social stimuli and make less effort to seek out social over nonsocial stimuli (Hedger et al., 2020). Accordingly, the START task battery includes two measures of social reward responsivity: 1) a passive viewing paradigm similar to the eye-tracking laboratory-based task of Pierce and colleagues (2011), and 2) a choice-based paradigm similar to that of Ruta and colleagues (2017). Reduced looking and responding toward social over nonsocial stimuli have been noted in autistic children in these prior reports. Accordingly, the key metrics of interest from these tasks were those that index the proportion of looking time or button presses toward social compared to nonsocial stimuli.

Sensory phenotype: Atypical *sensory* sensitivity is a commonly reported feature of autism (Ausderau et al., 2014; Ben-Sasson et al., 2007; Posar & Visconti, 2018). It is generally evaluated using parent-report questionnaires or tasks that involve touching/watching objects of special sensory interest (e.g. spinning wheels with illusory contours, pin cushions, musical dome). The START task battery includes an adapted version of one such lab-based task used by (Tavassoli et al., 2016) to measure visual sensory interest. In line with the key metric used in the lab-based version of this task, the dependent variable of interest was the duration for which a child looked at the spinning wheel.

Motor phenotype: Atypical *motor* skills are commonly reported in autism (Anzulewicz et al., 2016; Ghaziuddin & Butler, 1998; Manjiviona & Prior, 1995). Poor spatial coordination and weak adaptation of velocity to reach targets have been suggested to be specific to autism (Forti et al., 2011). Developments in touch sensor technology can help measure spatial coordination and velocity with high precision and ease. The START task battery harnesses this technological development to measure three-dimensional finger movements, providing a fine-grained measure of spatio-temporal performance in fine-motor planning and execution. Three tasks were used to capture variability in motor performance, which includes popping bubbles on a screen, following a butterfly across the screen with a finger, and colouring a pattern with clear outlines. In the bubble popping task, we measured the force with which bubbles were popped as well as the distance of the touch from the centre of the bubble - in line with suggestions from previous research (Anzulewicz et al., 2016; Forti et al., 2011). In the motor following task, we measured the spatiotemporal errors in following a moving target, given the suggested autistic difficulties in motor coordination. Similarly, we measured the number of times that a child crossed over the boundaries of the figure in the colouring task, to provide a proxy for their motor control abilities.

Parent/Caregiver Report and Interactions: Behavioural observations may emerge from parent reports of day-to-day activities of the child, or expert observation of social interaction and play. Brief parent-report tools such as the INCLEN Diagnostic Tool for Autism Spectrum conditions (INDT-ASD) (Juneja et al., 2014), and All India Institute of Medical Sciences (AIIMS)-Modified-INDT-ASD Tool (Gulati et al., 2019) have demonstrated high sensitivity in early screening and diagnosis of autism in an Indian setting. Accordingly, the START app includes a brief questionnaire for primary care givers as well as a provision for video-recording a parent/caregiver-child play session. Dyadic interaction of the child with the caregiver constitutes one of the most ecologically valid metrics of social interaction, and is the primary target of certain types of developmental interventions for autism (Green et al., 2010). In line with previous reports, the key metrics of interest included the number of attempts by the child in initiating interactions, and the number of synchronous responses from the caregiver.

[Insert Table 2 here]

[Insert Figure 1 and Figure 2 here]

Assessment Procedure:

Two high-school graduates with no prior relevant experience were recruited as non-specialist health workers. They underwent a four-day training, with two days in classroom followed by two days of observation and supervised field-training in households. Two psychology postgraduate research assistants were recruited for the project who observed the data collection and ensured adherence to the research procedures. Each health worker was then paired up with a research assistant to visit the participants' houses to collect data, using a Samsung SM P600 tablet. Testing was generally conducted sitting on the floor or bed. Specialist assessment tools (DP-3 and INDT-ASD) were administered by the research assistants.

Research assistants completed a detailed observation schedule noting the environment and circumstances of each data collection, including family involvement and available resources. They interviewed non-specialist health workers both immediately after their training and at the end of data collection, with a focus on challenges faced during data collection and strategies adopted to overcome these. Research assistants interviewed parents of participating children (TD=5, AS=5, ID=5) to explore their experiences with START, including caregivers of children who were able to complete the START assessment tasks and those who were unable to complete them. Separate consent for audio recording was taken prior to these interviews. Further details of the observation and interview schedules are available in the supplementary information sections 1.4 and 1.5.

All the procedures were performed in accordance with relevant guidelines and regulations and approved by the Research Ethics Committee of the University of Reading, UK, as well as the Institutional Review Board for Public Health Foundation of India, and the Indian Council of Medical Research.

Community involvement statement

This project involved an autistic researcher who took part in regular discussions during the analysis of the pilot data collected using the START platform. In addition, the research team organised a dedicated dissemination and discussion event for the autism community stakeholders in India.

Analysis

The project used a mixed-method design. The feasibility and acceptability assessments were done via face-to-face interviews with non-specialist health workers and caregivers. These interviews were then qualitatively analysed using thematic analysis. The efficacy of the task battery in distinguishing children with neurodevelopmental conditions from other groups was done using quantitative methods (using general linear model). The evaluation of the task battery's accuracy in classifying participants into the three groups was done using machine learning methods including XGBoost, logistic regression, and support vector machines.

Feasibility and acceptability

Interviews were conducted in Hindi with non-specialist health workers and caregivers to evaluate the feasibility and acceptability of the START task battery in home settings. Environmental conditions for data capture and the nature and frequency of disruptions during the assessment were recorded from the observation schedule used by the research assistant. All interviews were audio-recorded, transcribed, translated to English, and cross-checked for the accuracy of the translation. In-depth interviews were qualitatively analyzed using thematic analysis (details in Supporting Information, Table S4).

START app data analyses

Pre-set exclusion criteria were applied to the data to ensure quality, resulting in a different number of participants for each task. Detailed information for the analysis of each task and questionnaire measure within the app is provided below.

Preferential Looking Task: Gaze location was identified using a convolutional neural net-based algorithm (Dubey et al., 2022; Krafka et al., 2016). Data were available from 118 of 131 participants (TD = 40, AS = 40, ID = 38). All participants met the inclusion criteria of eye detection for at least 50% of frames and gaze on the tablet for at least 50% of frames. Social preference was computed as a ratio between the number of frames during which a participant was gazing at the social stimulus and the total number of frames in which their gaze was identified to be on either of the two stimuli.

Button Task: Data were available from 116 of 131 participants (TD = 40; AS = 37; ID = 39). Participants who completed fewer than 50% of trials were excluded, resulting in 104 participants (TD=39; AS=27; ID=38) in the final analysis. For each participant, the proportion of social button choice as a fraction of the total number of completed trials was calculated.

Wheel Task: Data were available from 125 of 131 participants (TD = 40, AS = 46, ID = 39). Participants who completed fewer than two trials, or whose faces could be detected in only 25% or fewer of the video frames were excluded. This exclusion criterion yielded data from 117 of 125 participants (TD = 37, AS = 41, ID = 39) in the final analysis. Two variables were coded: a) Time spent looking at the wheel, and b) distance of the face from the screen. Time spent looking at the wheel was calculated for every completed trial, summed across trials, and divided by the maximum possible duration of the completed trials. The distance of the face from the screen was calculated using a deep neural network that detected the subject's facial features in each frame (Bishain et al., 2021).

Motor Following Task: Data were available from 120 of 131 participants (TD = 40, AS = 43, ID = 37). Data sets were filtered for completeness by including only participants who finished two or more trials. This criterion yielded 115 participants (TD = 40, AS = 40, ID = 35) for the final analysis. Spatio-temporal difference between the target and the child's motor trajectory was computed as root mean square error (RMSE) to measure accuracy in motor planning and execution. Additionally, we analysed the 'frequency gain' metric for all participants using a Fast Fourier Transformation (FFT), in order to assess the closeness in the source and target motions along the frequency domain (for details see Supporting Information, section 1.1).

Bubble Popping Task: Data were available from 120 of 131 participants (TD = 40, AS = 41, ID = 39). Data were included from all the participants who popped one or more bubbles. The force used while popping the bubbles was recorded using the `getPressure()` parameter recorded by the Android operating system on a Samsung tablet, and averaged across all bubbles popped. The distance between the touch point and the centre of the bubble was calculated to estimate visuomotor targeting accuracy in approaching dynamic stimuli.

Colouring Task: Data were available from 113 of 131 participants (TD = 40, AS = 38, ID = 35). Participants were asked to colour the interior of a target figure. Data sets were included only if participants coloured at least 25% of the pixels on the screen. This criterion yielded 93 participants (TD = 37, AS=29, ID=27) in the analysis. The total number of crossings over the target figure's outlines (movements in and out of the figure) was calculated. Any change in the touch point from inside the figure (pixels identified inside the outline) to the outside or *vice versa* was counted as one crossover.

Parent/Caregiver- Child Interaction: Data were available from 100 of 131 participants (TD = 32; AS = 35, ID = 33). The video recording of the session was coded using the Dyadic Communication Measure for Autism, by three trained independent coders based in India (Green et al., 2010). Two measures were extracted from this data set, one indexing the child's attempts at initiating interactions, and the other indexing synchronous responses from the caregiver. 13% of the videos were coded by all three coders and used to calculate intra-class correlation (ICC) using a 2-way mixed-effects model, based on a single measure, absolute agreement and confidence interval of 95%. A high degree of reliability was found between the coders for scores on parent/caregiver's synchronous interaction as ICC was 0·876 ($p <.0001$, 95% CI [.69, 96]). However, the coders had limited reliability for the scores on child's initiation as ICC was .542 ($p <.0001$, 95% CI [-.04, .85]). Where the videos were coded by more than one coder, we randomly chose codes from any one coder.

START Questionnaire: Data were available from all 131 participants (TD = 40, AS = 48, ID = 43). The items were scored as binary responses. The summed score indicates the number of 'red flag' signs of autism.

For each task, the three groups were contrasted on the dependent variables defined above using analyses of variance (Table 3). The Kruskal-Wallis test was used where the assumption of normality was violated; and Welch and Brown-Forsyth robust tests were run where the assumption of homogeneity of variance was violated. Since the results from these alternative analyses were similar to those obtained with the general linear model, we report in Table 3 results from the standard analysis of variance. Results from the alternative statistical tests are presented in Supporting information Tables S2 and S3. Additionally, we reran the analyses of variance reported in Table 3, including age as a covariate, which had no significant impact on the reported results. Since there were significant group differences in sex and cognitive age, we did not include these variables as covariates in this model.

Machine-learning analysis

This analysis applied a data-driven technique to combine the information from the multiple START metrics in order to optimise discrimination between the three groups (AS, ID, TD). Each dependent variable from the individual tasks constituted a feature vector. These features were then subjected to a set of machine-learning methods including XGBoost, logistic regression, and support vector machines. Each feature vector was first evaluated independently, and then in combination with other feature vectors for its accuracy in classifying individuals into the three groups (see Supporting information section 1.2 for details).

RESULTS

Results are presented below in three sections: a) feasibility and acceptability, b) group comparisons, and c) group classification accuracy using machine-learning analysis.

a) **Feasibility and Acceptability**

High completion rates (>70%) were obtained for all task measures collected (Supporting Information, Figure S2). The two main drivers behind missing data were a) children's unwillingness to play a game, seen more often in atypical children compared to typically developing ones, and b) app malfunctions for specific tasks. While none of the children who did not complete a task had any documented visual, motor, or auditory impairments, visual inspection of the data suggests that those who did not complete were more likely to be younger and of lower cognitive age than those who completed the tasks. Triangulation of data from the observation schedule and in-depth interviews highlighted the challenges in assessments such as limitations of space, variations in lighting, background noise, and interruptions. Health workers identified the importance of the involvement of the family in meeting these challenges, and that of written standard operating protocols for guiding assessments. App-based assessment seems to have high acceptability for children, who actively played the "games" on the tablet and enjoyed its child-friendly design elements. Parents also found START to be acceptable but questioned the credibility of an app-based assessment of child development (see Supporting Information Table S4 for the list of themes).

b) **Group comparisons**

We examined group differences in social, sensory, motor functions, parent/caregiver report and dyadic interaction. For each of these domains, the three groups were contrasted on the stated dependent variables (Table 3). In the social domain, an effect of group membership is seen on the preferential-looking task, as AS and ID children looked at the social stimuli less than the TD group did. However, no such group difference was seen in the button task. In the sensory domain, children with AS and ID looked at the spinning wheel longer than their TD counterparts did. In the motor domain, both AS and ID groups were distinguished from TD by force in the bubble-popping task and by visuomotor accuracy across all the motor tasks. Finally, an effect of group membership was found in measures of parent/ caregiver-report and interaction. Parents of autistic children endorsed higher numbers of items from the START questionnaire than parents of either ID or TD children. Inspection of Table 3 suggests a consistent pattern of difference between the two groups with neurodevelopmental conditions and the TD groups.

[Insert Table 3 here]

c) **Machine-learning analysis**

The classification accuracies, the sample proportions for each group and other details as determined in the machine-learning analysis are provided in Supporting Information Table S1. Based on these results, the Motor Following task (RMSE in following the butterfly trajectory) was the most promising independent task with 60% overall classification accuracy into three groups (TD, ID, AS), superior to a random chance classification accuracy of approximately 33%. This discrimination accuracy is at par with that reported by the questionnaire measure (Figure 3).

[insert figure 3 here]

While the classification accuracy for individual START metrics is relatively weak, combining the metrics yields a significant improvement, resulting in an overall classification accuracy of 78% (Table 4). The combination of metrics yielding the best classification consisted of the following: RMS error in the visuomotor following, boundary crossings in colouring, and force in bubble-popping; time watched and variation in distance from the display in the wheel task; both gaze and choice measures of social preference; and video-coded and questionnaire measures of autistic behaviour.

[Insert Table 4 here]

DISCUSSION

We tested a battery of tasks, questionnaires, and observational measures administered by a non-specialist on a mobile platform (app) in three groups of children with and without neurodevelopmental conditions. This app was found to be both feasible for delivery by non-specialists in home settings and acceptable to all users including community health workers, parents, and children. We find strong evidence for group differences on the majority of the measures between children with and without neurodevelopmental conditions.

Task measures

The task measures focused on social, sensory, and motor functioning. Specifically in the social domain, greater attention to social over non-social rewards was noted in non-autistic typically developing children. This pattern of results is consistent with reports on similar paradigms applied in laboratory settings, using standard infra-red eye trackers (Dubey et al., 2022; Hedger et al., 2020). In contrast to the preferential looking task, the button task did not show a difference between the three groups. This absence of a group difference could be driven by differences in the administration of the task between the current and the original report on this paradigm (Ruta et al., 2017).

Strong group differences were noted in task measures of motor function. The non-autistic typically developing group performed more accurately than both the autistic and ID groups in the motor following task, as indexed by lower spatial errors (RMSE). Convergent findings indicating poorer visuomotor control in autistic children compared to the non-autistic typically developing group were demonstrated as greater numbers of boundary crossings in the colouring task, and lower accuracy in reaching a dynamic target in the bubble-popping task. Additionally, the autistic group used significantly greater force than the non-autistic typically developing group in this task, replicating earlier reports (Anzulewicz et al., 2016). Greater force in hitting a target on the tablet as well as spatial targeting errors could be interpreted as a manifestation of poor motor control. Poor motor control can result from reduced use of sensory information to adjust motor behaviour and is consistent with theoretical models of sensorimotor and cognitive prediction error in autism (Van de Cruys et al., 2014)

In the domain of sensory interests, we used a tablet adaptation of a task previously associated with group differences between autistic and non-autistic children (Tavassoli et al., 2016). While the underlying mechanisms for enhanced interest in stimuli such as spinning wheels remain poorly understood, one feature shared by these stimuli is high predictability, which might be sought behaviourally as a mechanism to control sensory responsiveness or arousal. The current version of the task illustrates that autistic children show a similarly greater preference for the video of a spinning wheel, as indexed by a greater duration of looking at it compared to non-autistic children. In a phenotypic domain that is dominated by self and parent-report instruments, this task shows promise as a scalable observational measure of visual sensory interests.

Parent/Caregiver-Report and Interaction measures

The parent/caregiver-report questionnaire was based closely on a tool specific for identification of autism in an Indian context (INDT-ASD). Unsurprisingly, scores on this questionnaire significantly differed between all three groups (AS, ID, TD) in the expected direction, replicating previous reports with the original tool (Gulati et al., 2019).

The caregiver-child videos revealed substantial atypicality in both key metrics of interaction. Autistic children initiated social interactions less than the TD group did, and also trended toward fewer initiations compared to the ID group. However, we advise caution in drawing strong inferences, since the inter-rater reliability for the child initiation behaviour was moderate. Fewer synchronous responses from the caregiver were evoked in interaction with both the groups of children with neurodevelopmental conditions (AS and ID), compared to those with TD children. This result is consistent with an earlier report of reduced synchronous parent-child interactions in autistic relative to TD children (Feldman et al., 2014).

The majority of the START measures showed the expected pattern of group differences between autistic children and their TD counterparts. These data demonstrate a) the feasibility of administering a multi-domain assessment of autism-relevant phenotypic dimensions at home by non-specialist health workers, and b) the potential for scalability of this platform to other low-resource settings. However, we note the low specificity of these measures in discriminating between the AS and ID groups in the current sample. To investigate this apparent equivalence further, we re-examined each case's clinical notes, which revealed that all of the autistic participants also met the criteria for ID. This observation reflects the ground realities in India, where most autism diagnoses in children within tertiary centres are at the severe end of the spectrum, and likely to be associated with developmental delay. Additionally, a majority of the children in the ID group showed significantly elevated autistic symptoms. The phenotypic overlap in these groups likely contributed to the observed absence of group differences between AS and ID children for individual task metrics.

Notwithstanding this caveat, when combining all the measures to test their ability to discriminate the AS, ID, and the TD groups using machine learning, groups were classified with an overall accuracy of 78%, a considerable boost from the accuracy achieved by any of the measures alone. This level of classification accuracy is comparable to that achieved by machine-learning classifiers on structural brain imaging data, as well as the reliability of the autism vs other developmental conditions diagnoses by clinicians (Moon et al., 2019; Klin et al., 2000). This result highlights the advantages of a multi-measure platform that complements task performance with parent/caregiver-report to achieve greater precision in assessing autism.

To our knowledge, this is the first demonstration of a multi-measure digital platform to assess autism related symptoms by non-specialists in a low-resource setting. It adds to the growing number of international efforts toward digital assessments of autism (Dawson & Sapiro, 2019b; Mukherjee et al., 2022). The largely non-verbal nature of the tasks in the app (except the questionnaire) makes it applicable in principle to other global settings without needing significant alteration. While we found that the START battery is sensitive to detecting deviations from typical development, individual task metrics did not show a clear difference between children with ID and AS. This observation is arguably driven by the nature of our sample of children with a neurodevelopmental disorder, where all autistic children met the criteria for ID, and several of the ID children had elevated autistic features. While this level of overlap is reflective of ground realities in our target population, future validation work can focus on validating this task battery further in neurodevelopmental disorder groups with minimal symptomatic overlap.

CONCLUSION

The current study demonstrates the potential and proof of principle for a tablet-based app for assessing autistic children that can be administered by non-specialist health workers with minimal training. The app includes tasks, a questionnaire, and observational assessments of aspects of behaviour that index social, sensory, and motor function. Individual metrics from each task show a consistent pattern of differences between typically and atypically developing children. Combining the information from multiple measures within the app resulted in high classification accuracy for the three groups of children (AS, ID, TD). Future work should test this app prospectively in a large population-based study to assess the predictive validity of these measures independently, and in combination, for atypical neurodevelopmental status.

Declaration of Conflicting Interests

The authors declare no competing financial or non-financial interests.

REFERENCES

Alpern, G. D. (2007). *Developmental Profile 3, DP-3: Manual.*

Anzulewicz, A., Sobota, K., & Delafield-Butt, J. T. (2016). Toward the Autism Motor Signature: Gesture patterns during smart tablet gameplay identify children with autism. *Scientific Reports*, *6*(1), 31107. https://doi.org/10.1038/srep31107

Arora, N. K., Nair, M. K. C., Gulati, S., Deshmukh, V., Mohapatra, A., Mishra, D., Patel, V., Pandey, R. M., Das, B. C., Divan, G., Murthy, G. V. S., Sharma, T. D., Sapra, S., Aneja, S., Juneja, M., Reddy, S. K., Suman, P., Mukherjee, S. B., Dasgupta, R., … Vajaratkar, V. (2018). Neurodevelopmental disorders in children aged 2–9 years: Population-based burden estimates across five regions in India. *PLOS Medicine*, *15*(7), e1002615. https://doi.org/10.1371/journal.pmed.1002615

Arun, P., & Chavan, B. (2021). Survey of autism spectrum disorder in Chandigarh, India. *Indian Journal of Medical Research*, *154*(3), 476. https://doi.org/10.4103/ijmr.IJMR_930_19

Ausderau, K. K., Furlong, M., Sideris, J., Bulluck, J., Little, L. M., Watson, L. R., Boyd, B. A., Belger, A., Dickie, V. A., & Baranek, G. T. (2014). Sensory subtypes in children with autism spectrum disorder: latent profile transition analysis using a national survey of sensory features. *Journal of Child Psychology and Psychiatry*, *55*(8), 935–944. https://doi.org/10.1111/jcpp.12219

Ben-Sasson, A., Cermak, S. A., Orsmond, G. I., Tager-Flusberg, H., Carter, A. S., Kadlec, M. B., & Dunn, W. (2007). Extreme Sensory Modulation Behaviors in Toddlers With Autism Spectrum Disorders. *The American Journal of Occupational Therapy*, *61*(5), 584–592. https://doi.org/10.5014/ajot.61.5.584

Bishain, R., Chakrabarti, B., Dasgupta, J., Dubey, I., & Chandran, S. (2021). *Can viewer proximity be a behavioural marker for Autism Spectrum Disorder?*

Chauhan, A., Sahu, J., Jaiswal, N., Kumar, K., Agarwal, A., Kaur, J., Singh, S., & Singh, M. (2019). Prevalence of autism spectrum disorder in Indian children: A systematic review and meta-analysis. *Neurology India*, *67*(1), 100. https://doi.org/10.4103/0028-3886.253970

Dawson, G., & Sapiro, G. (2019a). Potential for Digital Behavioral Measurement Tools to Transform the Detection and Diagnosis of Autism Spectrum Disorder. *JAMA Pediatrics*, *173*(4), 305. https://doi.org/10.1001/jamapediatrics.2018.5269

Dawson, G., & Sapiro, G. (2019b). Potential for Digital Behavioral Measurement Tools to Transform the Detection and Diagnosis of Autism Spectrum Disorder. *JAMA Pediatrics*, *173*(4), 305. https://doi.org/10.1001/jamapediatrics.2018.5269

Divan, G., Bhavnani, S., Leadbitter, K., Ellis, C., Dasgupta, J., Abubakar, A., Elsabbagh, M., Hamdani, S. U., Servili, C., Patel, V., & Green, J. (2021). Annual Research Review: Achieving universal health coverage for young children with autism spectrum disorder in low- and middle-income countries: a review of reviews. *Journal of Child Psychology and Psychiatry*, *62*(5), 514–535. https://doi.org/10.1111/jcpp.13404

Dubey, I., Brett, S., Ruta, L., Bishain, R., Chandran, S., Bhavnani, S., Belmonte, M. K., Estrin, G. L., Johnson, M., Gliga, T., & Chakrabarti, B. (2022). Quantifying preference for social stimuli in young children using two tasks on a mobile platform. *PLOS ONE*, *17*(6), e0265587. https://doi.org/10.1371/journal.pone.0265587

Dubey, I., Ropar, D., & Hamilton, A. (2015). Measuring the value of social engagement in adults with and without autism. *Molecular Autism*, *6*, 35. https://doi.org/10.1186/s13229-015-0031-2

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., Lee, L., Rattazzi, A., Stapel-Wax, J., Stone, W. L., Tager-Flusberg, H., Thurm, A., Tomlinson, M., & Shih, A. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. *Autism Research*, *8*(5), 473–476. https://doi.org/10.1002/aur.1575

Egger, H. L., Dawson, G., Hashemi, J., Carpenter, K. L. H., Espinosa, S., Campbell, K., Brotkin, S., Schaich-Borg, J., Qiu, Q., Tepper, M., Baker, J. P., Bloomfield, R. A., & Sapiro, G. (2018). Automatic emotion and attention analysis of young children at home: a ResearchKit autism feasibility study. *Npj Digital Medicine*, *1*(1), 20. https://doi.org/10.1038/s41746-018-0024-6

Feldman, R., Golan, O., Hirschler-Guttenberg, Y., Ostfeld-Etzion, S., & Zagoory-Sharon, O. (2014). Parent–child interaction and oxytocin production in pre-schoolers with autism spectrum disorder. *British Journal of Psychiatry*, *205*(2), 107–112. https://doi.org/10.1192/bjp.bp.113.137513

Forti, S., Valli, A., Perego, P., Nobile, M., Crippa, A., & Molteni, M. (2011). Motor planning and control in autism. A kinematic analysis of preschool children. *Research in Autism Spectrum Disorders*, *5*(2), 834–842. https://doi.org/10.1016/j.rasd.2010.09.013

Ghaziuddin, M., & Butler, E. (1998). Clumsiness in autism and Asperger syndrome: a further report. *Journal of Intellectual Disability Research*, *42*(1), 43–48. https://doi.org/10.1046/j.1365-2788.1998.00065.x

Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., Le Couteur, A., Leadbitter, K., Hudry, K., Byford, S., Barrett, B., Temple, K., Macdonald, W., & Pickles, A. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *The Lancet*, *375*(9732), 2152–2160. https://doi.org/10.1016/S0140-6736(10)60587-9

Gulati, S., Kaushik, J. S., Saini, L., Sondhi, V., Madaan, P., Arora, N. K., Pandey, R. M., Jauhari, P., Manokaran, R. K., Sapra, S., Sharma, S., Paul, V. K., & Sagar, R. (2019). Development and validation of DSM-5 based diagnostic tool for children with Autism Spectrum Disorder. *PLOS ONE*, *14*(3), e0213242. https://doi.org/10.1371/journal.pone.0213242

Hedger, N., Dubey, I., & Chakrabarti, B. (2020). Social orienting and social seeking behaviors in ASD. A meta analytic investigation. *Neuroscience and Biobehavioral Reviews*, *119*. https://doi.org/10.1016/j.neubiorev.2020.10.003

Hedger, N., Haffey, A., McSorley, E., & Chakrabarti, B. (2018). Empathy modulates the temporal structure of social attention. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1893), 20181716. https://doi.org/10.1098/rspb.2018.1716

Juneja, M., Mishra, D., Russell, P. S. S., Gulati, S., Deshmukh, V., Tudu, P., Sagar, R., Silberberg, D., Bhutani, V. K., Pinto, J. M., Durkin, M., Pandey, R. M., Nair, M., & Arora, N. K. (2014). INCLEN diagnostic tool for autism spectrum disorder (INDT-ASD): Development and validation. *Indian Pediatrics*, *51*(5), 359–365. https://doi.org/10.1007/s13312-014-0417-9

Klin, A., Lang, J., Cicchetti, D. V, & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders*, *30*(2), 163–167. https://doi.org/10.1023/a:1005415823867

Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., & Torralba, A. (2016). Eye Tracking for Everyone. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2176–2184. https://doi.org/10.1109/CVPR.2016.239

Kumar, A. (2011). Mental health services in rural India: challenges and prospects. *Health*, *03*(12), 757–761. https://doi.org/10.4236/health.2011.312126

Manjiviona, J., & Prior, M. (1995). Comparison of Asperger syndrome and high-functioning autistic children on a Test of Motor Impairment. *Journal of Autism and Developmental Disorders*, *25*(1), 23–39. https://doi.org/10.1007/BF02178165

Minhas, A., Vajaratkar, V., Divan, G., Hamdani, S. U., Leadbitter, K., Taylor, C., Aldred, C., Tariq, A., Tariq, M., Cardoza, P., Green, J., Patel, V., & Rahman, A. (2015). Parents' perspectives on care of children with autistic spectrum disorder in South Asia – Views from Pakistan and India. *International Review of Psychiatry*, *27*(3), 247–256. https://doi.org/10.3109/09540261.2015.1049128

Moon, S. J., Hwang, J., Kana, R., Torous, J., & Kim, J. W. (2019). Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: Systematic review and meta-analysis of brain magnetic resonance imaging studies. *JMIR mental health*, *6*(12), e14108.

Mukherjee, D., Bhavnani, S., Lockwood Estrin, G., Rao, V., Dasgupta, J., Irfan, H., Chakrabarti, B., Patel, V., & Belmonte, M. K. (2022). Digital tools for direct assessment of autism risk during early childhood: A systematic review. *Autism*, 136236132211331. https://doi.org/10.1177/13623613221133176

Mukherjee, S. B., Malhotra, M. K., Aneja, S., Chakraborty, S., & Deshpande, S. (2015). Diagnostic accuracy of Indian Scale for Assessment of Autism (ISAA) in chidren aged 2–9 years. Indian Pediatrics, 52(3), 212–216. https://doi.org/10.1007/s13312-015-0608-z

Patra, S., & Kar, S. K. (2021). Autism spectrum disorder in India: a scoping review. *International Review of Psychiatry*, *33*(1–2), 81–112. https://doi.org/10.1080/09540261.2020.1761136

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for Geometric Patterns Early in Life as a Risk Factor for Autism. *Archives of General Psychiatry*, *68*(1), 101. https://doi.org/10.1001/archgenpsychiatry.2010.113

Posar, A., & Visconti, P. (2018). Sensory abnormalities in children with autism spectrum disorder. *Jornal de Pediatria*, *94*(4), 342–350. https://doi.org/10.1016/j.jped.2017.08.008

Rahman, A., Divan, G., Hamdani, S. U., Vajaratkar, V., Taylor, C., Leadbitter, K., Aldred, C., Minhas, A., Cardozo, P., Emsley, R., Patel, V., & Green, J. (2016). Effectiveness of the parent-mediated intervention for children with autism spectrum disorder in south Asia in India and Pakistan (PASS): a randomised controlled trial. *The Lancet Psychiatry*, *3*(2), 128–136. https://doi.org/10.1016/S2215-0366(15)00388-0

Raw, R. K., Wilkie, R. M., Culmer, P. R., & Mon-Williams, M. (2012). Reduced motor asymmetry in older adults when manually tracing paths. *Experimental Brain Research*, *217*(1), 35–41. https://doi.org/10.1007/s00221-011-2971-x

Rudra, A., Belmonte, M. K., Soni, P. K., Banerjee, S., Mukerji, S., & Chakrabarti, B. (2017). Prevalence of autism spectrum disorder and autistic symptoms in a school-based cohort of children in Kolkata, India. *Autism Research*, *10*(10), 1597–1605. https://doi.org/10.1002/aur.1812

Ruta, L., Famà, F. I., Bernava, G. M., Leonardi, E., Tartarisco, G., Falzone, A., Pioggia, G., & Chakrabarti, B. (2017). Reduced preference for social rewards in a novel tablet based task in young children with Autism Spectrum Disorders. *Scientific Reports*, *7*(1), 3329. https://doi.org/10.1038/s41598-017-03615-x

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, *10*(1), 91–103. https://doi.org/10.1007/BF02408436

Tavassoli, T., Bellesheim, K., Siper, P. M., Wang, A. T., Halpern, D., Gorenstein, M., Grodberg, D., Kolevzon, A., & Buxbaum, J. D. (2016). Measuring Sensory Reactivity in Autism Spectrum Disorder: Application and Simplification of a Clinician-Administered Sensory Observation Scale. *Journal of Autism and Developmental Disorders*, *46*(1), 287–293. https://doi.org/10.1007/s10803-015-2578-3

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*(4), 649–675. https://doi.org/10.1037/a0037665

Vats, P., Juneja, M., & Mishra, D. (2018). Diagnostic Accuracy of International Epidemiology Network (INCLEN) Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD) in Comparison with Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5). *Indian Pediatrics*, *55*(6), 485–487. https://doi.org/10.1007/s13312-018-1339-8

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, *15*(5), 778–790. https://doi.org/10.1002/aur.2696

Table 1: Participant characteristics

	TD	AS	ID	F/χ2	p-value	Post-hoc contrasts, p-value
Chronological Age M ±SD	(N=40) 4.59 ±1.34	(N=48) 4.24 ±1.22	(N=43) 4.56 ±1.67	F (2, 129) = 0.88	0.42	
Gender ratio (F:M)	19:21	12:36	9:34	7.99	0.02	
Cognitive age on DP3	(N=36) 4.32 ±1.49	(N=37) 1.49 ±0.53	(N=36) 1.94 ±0.80	F (2, 106) = 80.87	<0.001	TD > AS, <0.001 TD > ID, <0.001 ID ~ AS, 0.19
INDT-ASD	(N=37) 0.16 ±0.37	(N=37) 17.16 ±4.35	(N=39) 5.15 ±7.51	F (2, 110) = 109.97	<0.001	TD < AS, <0.001 TD < ID, <0.001 ID < AS, <0.001

Table 2: Description of the tasks included in the START task battery

Task	Relevant References	Task details	Stimuli	Procedure	Dependent variable
			Domain: Social		
Preferential Looking Task (Figure 2a)	(Chakrabarti et al., 2017; Dubey et al., 2022; Hedger et al., 2018)	• Pairs of social and nonsocial videos were presented. • Size: each image covered half the screen in landscape mode leaving approximately 0.5cm between the images. • No inter-trial interval or central fixation • Counterbalanced presentation across the two sides of the screen. • Eight trials; total duration ~60 seconds.	• Social: Four videos of children looking towards the camera and smiling. • Nonsocial: Four videos of spinning washing machines or garden wind-fans.	• Tablet position: Upright on a stand (Figure 1a). • Setup: Child's and tablet's position were adjusted for a) camera alignment, b) light, c) distance, and d) stability. • Task: Child was instructed to look at the tablet screen and to keep their head still. • Child's face was video-recorded using the front camera while s/he looked at the stimuli. • Rule to proceed: Once setup completed, wait till the trial was completed	• Proportion of looking time to social stimuli

Button Task (Figure 2b)	(Dubey et al., 2022; Ruta et al., 2017)	• Two buttons were presented at random locations on the screen. • One button, when touched, showed the social video, and the other showed a non-social video. • The association between buttons and stimuli videos were counterbalanced between participants. • Eight trials; total duration ~120 seconds	• Social: a video of a child swimming and waving underwater • Non-social: a video of a dynamic geometric pattern.	• Tablet position: Flat on the table (Figure 1b) with a soft frame underneath. • Demonstration: child was shown that each button plays a video of either a child or a dynamic pattern. • Task: Child chose a button to play the linked video on each trial. • Rule to proceed: If the child successfully touched one button to start a video.	• Proportion of choices made to look at social stimulus
Domain: Sensory Sensitivity					
Wheel Task (Figure 2c)	(Tavassoli et al., 2016)	• A video of a black and white wheel presented on the screen. • A red button was presented in the right lower corner. This button could be pressed at any time to end the trial.	• A 15 second video of a black and white wheel spinning to create a visual illusion.	• Tablet position: upright on a stand (Figure 1a). • Setup: Child was positioned in front of the screen. • Demonstration: Child was shown the task and instructed that they could press the red button any time they	• Looking time at the video, calculated as a proportion of duration for which the wheel videos were played, divided by the sum of the maximum duration for which the videos could be played if the

		• Five trials each lasting maximum 15 seconds. • Total duration 75 seconds		wanted to stop looking at the wheel. • Rule to proceed: If the child looked at the wheel carefully or tried to stop it by touching the red button. • Child's face was video-recorded.	'terminate' button was not pressed • Minimum and maximum distance between the child's face and the tablet.
		Domain: Motor			
Motor Following Task (Figure 2d)	(Raw et al., 2012)	• A butterfly flying across the screen. • Random trajectories for the butterfly were generated with variable velocities in both x and y axes • Counterbalancing: the butterfly flew from left to right on two trials and right to left on the other two. • Four trials; total duration ~120 seconds.	• An image of a colourful butterfly over a background of a green field.	• Tablet position: Flat on the table with a soft protective cover underneath (Figure 1b). • Demonstration: Child was shown how to follow the butterfly across the screen by keeping the index finger of their dominant hand on top of the butterfly. • Task: Child followed the butterfly's trajectory with their index finger. Child's trajectory was displayed in real-time. • Rule to proceed: If the child was able to follow the butterfly making	• Spatio-temporal error, jerk, and weighted frequency gain for X and Y axes.

				about ⅓ of its trajectory in the demonstration trial (maximum 3 attempts).	
Bubble Popping Task (Figure 2e)	(Anzulewicz et al., 2016)	• A series of bubbles were presented on the screen floating up and down in a straight line parallel to the Y axis. • Six trials showing increasing number (1 to 6) of bubbles on each trial. • Total duration: ~36 seconds.	• Images of bubbles on a colourful background of an underwater scene.	• Tablet position: Flat on the table with a soft protective cover underneath (Figure 1b). • Demonstration: Child was shown how to pop the bubbles and given a chance to do the same. • Task: Child was instructed to pop the bubbles as quickly as they could using one finger. • Rule to proceed: If the child popped two bubbles over a maximum of 3 demonstrations.	• Distance of touch from the center of the bubble • Force of touching the tablet.

Colouring Task (Figure 2f)		• Simple outline figures and a colour palette were presented for colouring in. • Two trials; maximum duration 240 seconds.	• Line figures of a flower and a butterfly. • Colour palette presented at the left lower corner with red, yellow, blue, and green colours to choose from.	• Tablet position: Flat on the table with a soft protective cover underneath (Figure 1b). • Demonstration: Child was shown how to touch the palette to activate a colour and move the finger to colour the figure. The child was given a chance to practise before starting the task. • Task: Child was asked to freely colour the figure. • Rule to proceed: If the child was able to make a stroke and go to the color palette to pick a colour.	• Count of events of crossing over the outlines of the target figure
Domain: Parent/Caregiver Report and Interaction					
START Questionnaire	(Gulati et al., 2019; Vats et al., 2018)	• 14 binary choice (yes, no) items focussed on exploring the early signs of autsim such as poor eye-contact,	• Items adapted from INCLEN-INDT-ASD, M-CHAT, and ICD classification system.	• Tablet position: Held by health worker. • Items were read aloud by the health worker.	• Each item is scored 1 or 0. Six items are coded as yes = 1, and no = 0. Eight items marked with (R) are reverse coded. The sum of scores is

		limited social interactions etc. • Language used: Hindi			calculated across all items to index the severity of autistic symptoms.
Parent/Care giver- Child Interaction (Figure 2g)	Parent Child Interaction protocol from the Duke Center for Autism and Brain Development (personal communicati-on)	• A set of toys was presented to the caregiver and child to play with.	• Toys to facilitate pretend play e.g. kitchen set, pretend food, cars, dolls, etc. Toys with lights and sounds to elicit sensory sensitivity. Toys to facilitate interaction and verbal outputs e.g. toy phone, toy mobiles.	• Tablet position: The tablet was held by the health worker to record a video of the child & caregiver • Video was recorded for ten minutes.	• Synchronous responses from caregiver • Initiation of interactions by child
All tasks on the app can be aborted using a unique gesture of sequential tapping on the corners of the tablet PC. All health workers were trained on the use of this gesture, with the instruction to abort a task in case a child did not engage with the tablet.					

Table 3: Group comparison using GLM for all measures from the START app.

Task	Dependent Variable	TD Mean (SD)	AS Mean (SD)	ID Mean (SD)	F	η2p	p	Post Hoc comparison p value
Domain: Social								
Preferential Looking Task	**Social preference**	n = 40 M = .59 (.09)	n = 40 M= .52 (.12)	n = 38 M = .53 (.09)	F (2,115) = 5.996	0.09	0.003	TD > AS = 0.005 TD > ID = 0.025 AS ~ ID = 1.00
Button Task	**Social choice**	n = 39 M = .47 (.24)	n = 27 M = .52 (.28)	n = 38 M = .52 (.21)	F (2,101) = .638	0.01	.53	NA
Domain: Sensory								
Wheel Task	**Looking at the wheel**	n = 37 M = 0.46 (0.37)	n = 41 M = 0.73 (0.33)	n = 39 M = 0.66 (0.36)	F (2,114) = 6.24	0.10	0.003	TD < AS = 0.003 TD < ID = 0.039 AS ~ ID = 1.000
Domain: Motor								
Motor Following task	**RMSE**	n = 40 M = 203.80 (97.95)	n = 40 M = 591.27 (283.63)	n = 35 M = 404.70 (216.74)	F (2,112) = 32.93	0.37	<0.001	TD < AS < 0.001 TD < ID < 0.0001 AS > ID < 0.001
	FFT X Axis	n = 40 M = 1.53 (.40)	n =36 M = 2.01 (.76)	n = 34 M = 2.02 (.72)	F (2,107) =7.21	0.12	0.001	TD < AS = 0.005 TD < ID = 0.005 AS ~ ID = 1.000
	FFT Y Axis	n = 40 M = 10.33 (9.12)	n = 36 M= 25.28 (15.42)	n = 34 M = 20.27 (11.51)	F (2,107) =14.87	0.22	<0.001	TD < AS <0.001 TD < ID = 0.002 AS ~ ID = 0.268
	Jerk	n = 40 M= .06 (.13)	n = 40 M = .05 (.13)	n = 35 M = .05 (.13)	F (2,112) =.053	0.001	0.948	
Bubble Popping Task	**Force**	n = 40 M = 0.07 (0.01)	n = 41 M = 0.09 (0.02)	n = 39 M = 0.08 (0.02)	F (2,117) =8.49	0.13	<0.001	TD < AS = <0.0001 TD ~ ID = 0.140 AS ~ ID = 0.122

	Distance on X axis	n = 40 M= 45.20 (11.78)	n = 41 M = 88.97 (57.34)	n = 39 M = 66.81 (22.49)	F (2,117) =14.54	0.20	<0.001	TD < AS = <0.001 TD < ID = 0.029 AS > ID = 0.023
	Distance on Y axis	n = 40 M = 54.95 (13.59)	n = 41 M = 87.02 (43.17)	n = 39 M = 75.01 (25.03)	F (2,117) =11.76	0.17	<0.001	TD < AS <0.001 TD < ID = 0.011 AS ~ ID = 0.229
Colouring Task	**Crossing over**	n = 37 M = 23.05 (16.98)	n = 29 M = 56.40 (29.43)	n = 27 M = 49.81 (28.31)	F (2,90) =16.95	0.27	<0.001	TD < AS = <0.001 TD < ID <0.001 AS ~ ID = 0.972
Domain: Parent/Caregiver report and Interaction								
Parent/Caregiver-Child Interaction	**Caregiver: Synchronous response**	n = 32 M = .33 (.22)	n = 35 M = .15 (.11)	n = 33 M = .20 (.14)	F (2,97) =11.46	0.19	<0.001	TD > AS = <0.0001 TD > ID = 0.004 ID ~ AS =0.607
	Child: Initiation	n = 32 M = .48(.24)	n = 35 M = .23 (.24)	n = 33 M = .40 (.23)	F (2,97) =9.94	0.17	<0.001	TD > AS = <0.0001 TD ~ ID = 0.516 AS < ID = 0.011
START Questionnaire	**Total Score**	n = 40 M = 1.03 (1.31)	n = 48 M = 5.06 (2.26)	n = 43 M = 3.09 (2.32)	F (2,128) =44.06	0.41	<0.001	TD < AS <0.001 TD < ID <0.001 AS > ID <0.001

Table 4: *Machine learning results. The overall classification accuracy for the best combination of feature vectors is listed. Refer to Figure 4 for corresponding Feature Vector IDs. So1: Button Task, So2: Preferential Looking task, Se1:Wheel task, Mo1: Motor Following Task, Mo5, Mo7: Bubble-popping task, Ob1: Parent Child Interaction, Ob2: Questionnaire responses*

Feature Vector ID combination providing the best accuracy	Mean Classification Accuracy (AS)	Mean Classification Accuracy (ID)	Mean Classification Accuracy (TD)	Mean Overall Classification Accuracy	Mean proportion % of subjects across different groups (AS:ID:TD)
Social: So1, So2 + **Sensory: Se1** + **Motor: Mo1, Mo5, Mo7** + **Observation: Ob1, Ob2**	61.61%	78.23%	86.40%	78.02%	23:30:47

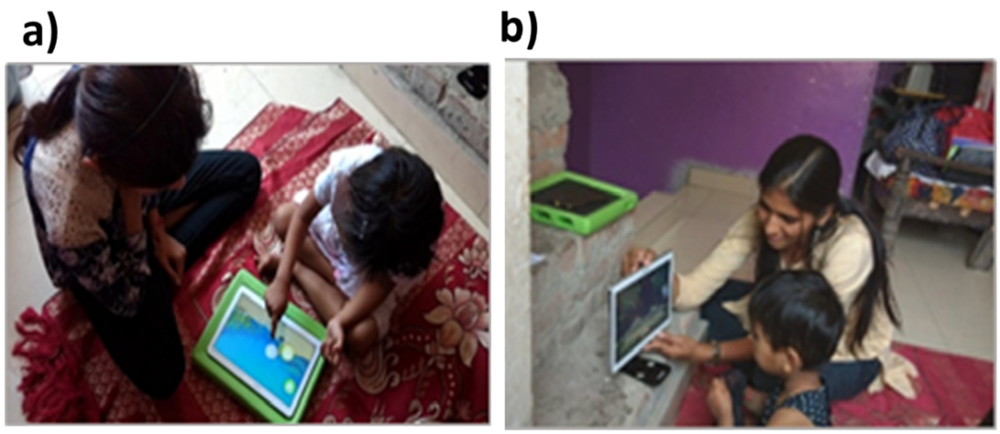


Figure 1: Administration of the START task battery in field settings. a) Tablet positioned upright for preferential looking task, and wheel task; b) Tablet positioned flat on a surface with a frame underneath for the button task, motor following task, bubble popping task, and colouring task. Health workers and primary caregiver (in case of children under 18 years of age), provided written informed consent for publication of the images in this figure.

327x139mm (96 x 96 DPI)

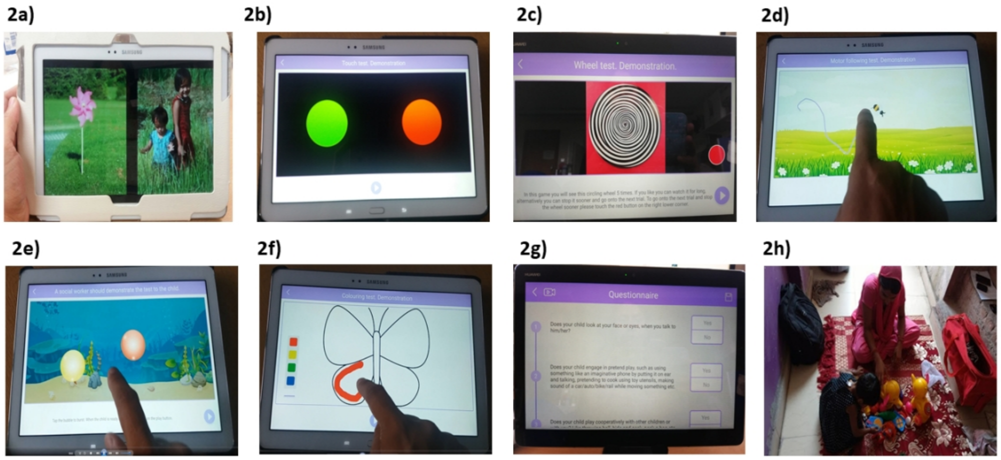


Figure 2: Sample screenshots from the a) preferential looking task, b) button task, c) wheel task, d) motor following task, e) bubble popping task, f) colouring task, g) START questionnaire, and h) caregiver-child interaction observation. The primary caregiver provided written informed consent for publication of the image 2h showing themselves and the child in this figure.

336x153mm (96 x 96 DPI)

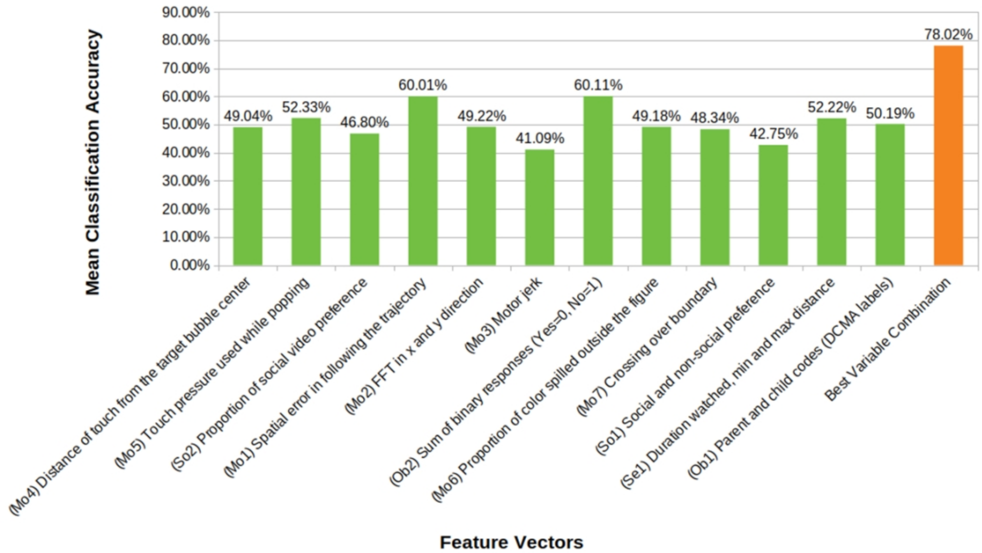


Figure 3: Mean classification accuracies of the feature vectors taken from the eight START tasks. The figure also represents the most accurate classification achieved by a combination of these features. (Prefixes on x-axis in parentheses refer to corresponding feature IDs). Some feature vectors are multidimensional amalgams of several different measures within a task. Chance level classification accuracy is 33.3%.

338x190mm (96 x 96 DPI)

Supporting Information

Using mobile health technology to assess childhood autism in low-resource community settings in India: an innovation to address the detection gap

Contents

1.1 Dependent variable calculation for the Motor following task

Root Mean Square Error (RMSE) in the motor following task was calculated using the following formula

$$\mathrm{RMSE} = \sqrt{\sum \frac{(x_{pred} - x_{ref})^2 + (y_{pred} - y_{ref})^2}{N}}$$

where x_{pred} and y_{pred} are the participant's finger position on x and y axes on the screen while x_{ref} and y_{ref} are the corresponding positions for the butterfly. N indicates the total number of recorded data points for a test attempt.

Additionally, we analysed the 'frequency gain' metric for all participants using a Fast Fourier Transformation (FFT). For this the trajectories of the cursor and finger motion are resolved into multiple waves of varying amplitudes using FFT. This allows us to analyze the closeness in the source and target motions along both axes by observing them in the frequency domain. This is achieved by calculating the average gain in amplitude for the source motion in the vicinity of each target frequency. The average gain in amplitude for the finger motion in the neighbourhood of each frequency bin represents the accuracy of the finger in copying the cursor trajectory, approaching unity in case of high degree of correspondence between the two trajectories. More specifically, $G_f$the gain at a given frequency f, is calculated as

$$G_f = \frac{U_{fm}}{B_f}$$

Here, B_f is the amplitude for the cursor's (butterfly) motion at frequency$f$; and U_{fm} is the average amplitude of the user's (finger) motion in the vicinity of frequency f (a neighborhood of three frequency bins including f is used). It is given as

$$U_{fm} = \frac{U_{f_{-1}} + U_f + U_{f_{+1}}}{3}$$

Here, $U_f$represents the amplitude of user's motion at frequency f. The subscripts f_{-1} and f_{+1} represent frequency bins adjacent to frequency bin f. Target motion is pre-determined and hence its amplitude is not approximated. The child is assumed to be following this motion and hence U_f is approximated at the central bin at frequency f.

In addition, Jerk, the change in acceleration per time, was derived as the third-order differential of the participant's distance along their trajectory with respect to time.

1.2 Machine learning-based data analysis:

Per current best practices, the machine learns from a subset of the labelled data and creates a classification model; the model is then subjected to the remaining unseen data to determine the desired accuracy. Classification accuracies have been reported for a specific group as:

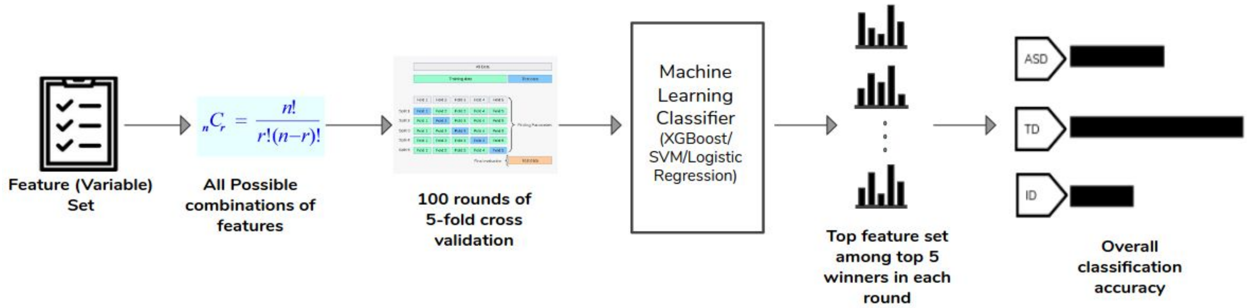
$$C_{ASD} = \frac{P_{ASD}}{N_{ASD}}$$

where, P_{ASD} is the number of children correctly classified as autistic, N_{ASD} is the actual number of autistic children, and thus the ratio C_{ASD} is the classification accuracy for the autistic group (reported in %). C_{ID} and C_{TD} have been similarly derived from the pairs P_{ID}, N_{ID} and P_{TD}, N_{TD}, respectively. Finally, the overall classification accuracy $C_{Overall}$ is determined as:

$$C_{Overall} = \frac{P_{ASD} + P_{ID} + P_{TD}}{N_{ASD} + N_{ID} + N_{TD}}$$

A 5-fold cross-validation scheme has been followed to minimize any bias and variance which might be introduced by the relatively small size of our dataset (a deep neural net-based model may not be feasible).

Figure S1: Analyzing all possible combinations of features for classification accuracy in a 100-round polled scheme



To determine the set of features from the combinatorial set (all possible combinations of features were evaluated) with the best overall accuracy, a polling scheme consisting of 100 rounds was used as shown in supplementary Figure S1. In each round, a 5-fold cross-validation experiment was performed for all possible sets or combinations of features. The top 5 winning combinations of features, based on the overall classification accuracy, were noted for each round. Finally, the most frequently occurring feature combination was declared as the overall winner and its average accuracy, across the 100 rounds, is reported. This polling scheme ameliorates any undesirable effect of outliers and the cherry-picking of favourable results.

Figure S2: Task completion percentage for each START task for each of the three groups of children (AS, ID and TD)

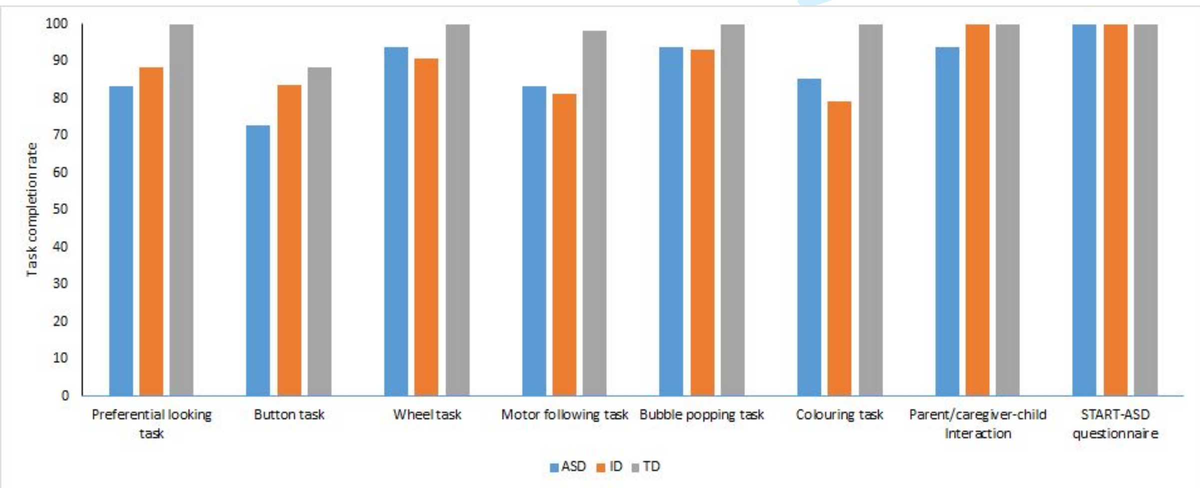


Table S1: Results from Machine Learning analysis. Overall classification accuracy for each feature vector (dependent variable) is listed.

Task	Dependent Variable ID	Dependent Variable	Mean Overall Classification Accuracy (%)	Mean proportion % of subjects across different groups (AS:ID:TD)
Preferential Looking Task	(1)	Social preference	42.75	34:32:34
Button Task	(2)	Social choice	46.80	29:35:36
Wheel Task	(3)	Proportion of looking at the wheel, Minimum distance, Maximum distance	52.22	36:33:31
Motor Following Task	(4.1)	RMSE	60.01	34:32:34
	(4.2)	FFT X Axis, FFT Y Axis	49.22	32:31:37
	(4.3)	Jerk	41.09	34:31:35
Bubble Popping Task	(5.1)	Distance on X Axis, Distance on Y Axis	49.04	34:33:33
	(5.2)	Force	52.33	34:33:33
Colouring Task	(6.1)	Crossing over	48.34	30:30:40
	(6.2)	Proportion of color spilt out to the total area inside figure	49.18	33:30:37
Parent/Caregiver- Child Interaction	(7)	Caretaker: Synchronous response, Child: Initiation	50.19	34:33:33
START Questionnaire	(8)	Total Score	60.11	37:33:30

Table S2: Alternative test statistics when assumption of homogeneity of variance is violated

Task	Variable	Levene's test p value	Robust test: Welch	Robust test: Brown-Forsyth	Post-hoc contrasts (Games-Howell), p-value
Domain: Social					
Preferential looking task	Social looking	.216	NA	NA	NA
Button task	Social preference	.145	NA	NA	NA
Domain: Sensory					
Wheel task	Duration	.179	NA	NA	Na
	Maximum distance	.002	6.64, *p* = .002	7.58, *p* = .001	TD < AS, .002 TD ~ ID, .183 AS ~ ID, .080
	Minimum distance	.049	20.20, *p* <.0001	17.73, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS ~ ID, .629
Domain: Motor					
Motor following task	RMSE	<.0001	41.08, *p* <.0001	32.86, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS > ID, .005
	FFTx	.024	9.64, *p* <.0001	6.92, *p* = .002	TD < AS, .004 TD < ID, .003 AS ~ ID, .999
	FFTy	.002	16.40, *p* <.0001	14.55, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS ~ ID, .277
	Jerk	.749	NA	NA	NA

Bubble popping task	Force	.050	9.57, *p* <.0001	8.50, *p* <.0001	TD < AS, <.0001 TD ~ ID, .086 AS ~ ID, .152
	Distance X	<.0001	23.23, *p* <.0001	14.93, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS ~ ID, .065
	Distance Y	<.001	17.33, *p* <.0001	11.97, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS ~ ID, .283
Colouring task	Crossover errors	.021	19.99, *p* <.0001	15.60, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS ~ ID, .672
Domain: Parent report/observation					
Parent/ Caregiver-Child Interaction	Caretaker's synchronous interaction	.001	9.21, *p* <.0001	11.15, *p* <.0001	TD < AS, <.0001 TD < ID, .013 AS ~ ID, .246
	Child's initiation	.624	NA	NA	NA
START Questionnaire	Sum score	.006	56.82, *p* <.0001	45.71, *p* <.0001	TD < AS, <.0001 TD < ID, <.0001 AS > ID, <.0001

Table S3: Kruskal-Wallis test for group comparison for tasks where assumption of normality is violated

Task	**Variable**	**Kolmogorov-Smirnov test**	**X^2**	**df**	***p***	**$\mathcal{E}^2$**	**Dwass-Steel-Critchlow-Flinger pairwise contrasts, p-value**
Domain: Social							

Preferential looking task	Social looking	0.06, p =.73	NA	NA	NA	NA	NA
Button task	Social preference	0.11, p = 0.16	NA	NA	NA	NA	NA
Domain: Sensory							
Wheel task	Proportion duration of watching	0.19, p <.001	10.35	2	0.006	0.09	TD < AS, 0.006 TD > ID, 0.047 AS ~ ID, 0.773
Domain: Motor							
Motor following task	RMSE	0.10, p = 0.167	NA	NA	NA	NA	NA
	FFTx	0.09, p = 0.403	NA	NA	NA	NA	NA
	FFTy	0.09, p = 0.278	NA	NA	NA	NA	NA
	Jerk	0.36, p <.001	6.83	2	0.033	0.06	TD > AS, 0.050 TD ~ ID, 0.139 AS ~ ID, 0.548
Bubble popping task	Force	0.09, p = 0.335	NA	NA	NA	NA	NA
	Distance X	0.17, p = 0.002	37.46	2	<.0001	0.31	TD < AS, <.001 TD < ID, <.001 AS ~ ID, 0.316
	Distance Y	0.13, p = 0.031	25.76	2	<.0001	0.22	TD < AS, <.001 TD < ID, <.001 AS ~ ID, 0.869
Colouring task	Crossover errors	0.10, p = 0.277	NA	NA	NA	NA	NA
Domain: Parent report/observation							
Parent/ Caregiver-Child	Caretaker's synchronous interaction	0.09, p = 0.394	NA	NA	NA	NA	NA

Interaction	Child's initiation	0.11, $p = 0.189$	NA	NA	NA	NA	NA
Questionna-ire	Sum score	0.17, $p = 0.001$	61.23	2	<.0001	0.47	TD < AS, <.001 TD < ID, <.001 AS > ID, <.001

Table S4: Summary of themes and subthemes emerging from the interview of health workers and parents.

Topic	Theme, sub-theme, quotation
Health worker's experience using the START app	Facilitators to smooth administration: Statement of Procedure, script, and app design elements *We have been given words [script]– if we speak them as it is, we remain confident.* (health worker 1) *When the game finishes, a small dialogue box appears on the screen that this game is finished and we press the arrow button to go to the next game. It helps a lot. We get to know that have to go to next [game].* (health worker 2) *If a child didn't take interest in wheel task, we switched to button or butterfly task and so on…and if a child is not at all interested in playing on the tablet then we used to record PCI. [we would say] "It's fine if you don't want to play on a tablet. See! We have got toys for you, let's play with them".* (health worker 1)
Challenges faced by health workers during data collection and strategies adopted to overcome them	Suitability of household environment for data capture
	Sub-theme 1: Availability of space in households *Some families had a single room house – they were living and eating in the same room. In these cases, adult family members used to go out while we made siblings sit in a corner.* (health worker 1)
	Sub-theme 2: Disruptions by family members *We ensured that … no other family members except mother-child are in the room.* (health worker 2)
	Sub-theme 3: Disruptions in the testing environment

	If an air cooler was on then we requested them (parents) to switch it off or if a phone was ringing in the room then we gestured to them to put it on silent mode. (health worker 2)
	Engaging atypical children *There is a difference between a normal child and child with problems. A normal child engages with us quickly but a child with problems might not be comfortable in sitting with us.* (health worker 2)
	Confidentiality concerns *Consenting video has really helped in giving a clear picture to the families (about the assessment)...Also, families had concerns – will these games cause any harm to the child and will the video be uploaded to any website or shown publicly.* (health worker 2)
Acceptability to children	Interest in digital devices *Nowadays children like laptops or tablets if you make them play on it, they like it. It could be any game.* (Father, ID child) *He was interested and accordingly the assessment proceeded smoothly.* (Father, autistic child)
	App design elements *It (START task battery) was appropriate for them. Otherwise the child gets bored and runs away.* (Mother, autistic child) *He liked bursting bubbles and colouring* (Father, autistic child)
	Health worker engagement *She (health worker) was able to understand how to deal with the child.* (Mother, autistic child) *The health worker was doing it nicely – she was explaining to the child quite well.* (Mother, ID child)
Acceptability to parents	Overall high acceptability of the app *It was nice but she (child) wasn't so successful in games (wasn't able to play well). She is quite young so accordingly it was fine.* (Mother, ID child)
	Scepticism of apps as a valid assessment of child development *Suppose any child has been identified and a highly qualified doctor from your team explains it to them then they would feel that their child actually requires it (intervention)…How will they get convinced through an app? Obviously, they would need a doctor.* (Father, autistic child).

1.3 START Questionnaire

Instructions for the health workers: Please read out the items to the caregiver and ask them to choose from the options.

No	Items in English	Option
1(R)	Does your child look at your face or eyes, when you talk to him/her?	Yes/no
2(R)	Does your child engage in pretend play, such as using something like an imaginative phone by putting it on ear and talking, pretending to cook using toy utensils, making sound of a car/auto/bike/rail while moving something etc.	Yes/no
3(R)	Does your child play cooperatively with other children or with you? Like throwing ball, hide and seek, peek-a-boo etc.	Yes/no
4	Does your child get disturbed by usual sound or light? Such as getting annoyed by the sound of the kitchen utensils and trying to close the ears with hands/fingers, not able to bear the sound of the vehicles, unable to bear the fairy/festival lights, gets irritable by the sharp light of the bulb, etc. (Social worker please ask the opposite behaviour too, such as does the child like loud sounds or sharp lights? He/she watches bright lights by going close to them and/or listen to the radio / TV by sticking ears to them?)	Yes/no
5(R)	Does your child imitate you? Like making gesture for "bye-bye" or hello, or wearing a scarf or bag like you?	Yes/no
6	Does your child get annoyed with cloth tags, woollen or tight cloths, toothbrushes, socks etc. Or does he like rubbing some items / cloth on his body repeatedly even if it results in scratches.	Yes/no
7(R)	Is your child able to use language according to his/her age? Like adding words to make sentence "let's go out", or to answer you correctly and asking questions "what is that?", "when are we going?" etc.	Yes/no

8	Does your child call himself by his/her name like "Vivek will eat food".	Yes/no
9(R)	Does your child show you the things he/she likes by pointing fingers to them?	Yes/no
10	Does your child repeat any kind of movement frequently? Like constantly making flapping/wriggling movement with his hands/fingers, constantly moving the body back and forth while sitting, constantly moving the head or body in unusual manner, etc.	Yes/no
11(R)	Does your child look at you / responds when called by name?	Yes/no
12	Does your child repeat certain voices, such as the sharp (high pitched) meaningless sounds, repeating your spoken words without context or meaning, repeating any sound heard on TV/redio/computer meaninglessly?	Yes/no
13(R)	Does your child come to you and show you when he/she has done something good?	Yes/no
14	Does your child play oddly with toys? Such as instead of using them meaningfully he/she just lines them up, or instead of running the toy car he spends long time looking at its wheels, smells or rubs toys on his body.	Yes/no

Scoring: The items with the (R) indicate reverse scoring i.e. a score of 1 is given for "No" and for the other items a score of 1 is given for "Yes". Then the sum is calculated to get the severity of autistic symptoms.

1.4 Observation schedule used by the research assistant

Child code
Date of assessment
General observations
Specific observations
Observation coding

	Low	Medium	High
Mother factors			
Interest in the visit	1	2	3
Favourable reaction to tablet	1	2	3
Distractions with other duties	1	2	3
Distractions with other family members	1	2	3
Child factors			
Exposure to smartphone/tablet	1	2	3
Child's interest in the assessment	1	2	3
Did parent have to help child engage with the assessment			(Y/N)
Ability to swipe	1	2	3
Ability to tap	1	2	3
Total time of engagement (establishment of rapport with the child)			Minutes

	Low	Medium	High
Environment factors			
List all family members present (other than mother/father)			
Number of times siblings disrupted assessment			times
Number of times other family members disrupted assessments			times
Any other types of disruption to assessment			times
Interest of other family members in the assessment	1	2	3
Level of noise in assessment room	1	2	3
Level of light in assessment room	1	2	3

Type of lighting Natural Artificial (Bulb) Artificial (Tubelight) Torch

Assessor factors	Low	Medium	High
How well was assessment process explained to mother	1	2	3
How well was the child engaged by assessor	1	2	3
How well did the assessor judge mood of the child	1	2	3
How well did the assessor administer the devices	1	2	3
How well did the assessor manage the family	1	2	3

Eye tracking	Low	Medium	High
What was the arrangement			

table/chair floor/table floor/chair 2 chairs bed/chair others (specify)

How difficult was it to get a suitable arrangement	1	2	3
Did the child need to sit in the mother's lap			(Y/N)
Was mother's face coming in the parameters screen			(Y/N)
Was mother prompting the child during the assessment			(Y/N)
Did the child try to touch the tablet during eye tracking			(Y/N)
Was there a need to move to another task and then back			(Y/N)
Time taken to calibrate the parameters	immediately	within 3 minutes	not at all
Did the child disengage from the task			(Y/N)
Was the task aborted			(Y/N)

Mention reason:

Wheel task			
No of demos needed	1	2	3
Did the child understand the task			(Y/N)
Did the child lose interest during play mode			(Y/N)
Was the task aborted			(Y/N)
Mention reason:			
Button task			
No of demos needed	1	2	3
Did the assessor have to hold the child's hand during demo			(Y/N)
Did the child understand the task			(Y/N)
Did the child lose interest during play mode			(Y/N)
Was the task aborted			(Y/N)
Mention reason:			
Butterfly task			
No of demos needed	1	2	3
Did the assessor have to hold the child's hand during demo			(Y/N)
Did the child understand the task			(Y/N)
Did the child lose interest during play mode			(Y/N)
Was the task aborted			(Y/N)
Mention reason:			
Bubbles task			
No of demos needed	1	2	3
Did the assessor have to hold the child's hand during demo			(Y/N)
Did the child understand the task			(Y/N)
Did the child lose interest during play mode			(Y/N)
Was the task aborted			(Y/N)
Mention reason:			
Colouring			
No of demos needed	1	2	3
Did the assessor have to hold the child's hand during demo			(Y/N)
Did the child understand the task			(Y/N)
Did the child lose interest during play mode			(Y/N)
Was the child pressing too hard i.e. colour not coming			(Y/N)
Was the task aborted			(Y/N)
Mention reason:			
PCI observations			
Questionnaire observations			

1.5 Interview schedule used to evaluate acceptability of the assessment from caregivers/families

The purpose of the in-depth interview (IDI) with mothers of children completing the START assessment was to understand the acceptability and feasibility of using START in Delhi households. Permission to audio record the interview was taken prior to the interview. If the parent was uncomfortable with audio recording, permission for note taking during the interview was sought. The following information was provided to the mother to guide the interview process:

"Thank you for meeting me today and for participating in our study. We had visited you at your home to carry out a tablet assessment that we are developing. I would like to understand more about your experience of the assessment by asking you a few questions. I am interested in knowing *your opinions*/suggestions and you can refuse to answer any question in case you feel uncomfortable. Could we begin?

Experience of the consenting process

You were approached by a health worker who explained the purpose of the study and requested for a time when she could visit you at home.

- Could you describe how you felt when you were approached by the health worker?
- Could you describe any immediate concerns you had about the assessment/home visit?

Experience during the visit

I would like to know more about your experience during our visit.

What did you like about the assessment? What did you dislike?

Probes: Time duration of the visit, comfort with a tablet assessment, comfort with video recording.

What did your family think about the assessment?

Child engagement with START:

What was your child's reaction to the health worker visiting them and during the assessment?

Probe: what do you think might be the reasons that he/she enjoyed/did not enjoy our visit?

What are your suggestions to make this more enjoyable for other children in the future?

Health worker training:

What did you think of the way the tablet was administered?

Would you have liked the health worker to do anything differently?

Probe: Was the health worker sensitive to your child's requirements/needs during the assessment?

Could you describe any concerns you had during the assessment process? How did the health worker address these?

Scaling up

In the future, we would like to carry out this assessment with more children at their homes.

How would other families like yours respond?

What are your suggestions so that most families would be happy to participate?

I would like to thank you on behalf of our team for taking the time out to not only be a part of our work but also for speaking with me today. Your feedback is very important for us to understand how we can make this experience better for families with young children."