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

Indu Dubey^{*1,2}, Rahul Bishain^{3*}, Jayashree Dasgupta^{4*}, Supriya Bhavnani^{4,5}, Matthew K. Belmonte^{7,1,8}, Teodora Gliga^{9,10}, Debarati Mukherjee^{4,6}, Georgia Lockwood Estrin^{4,10}, Mark H. Johnson^{10,11}, Sharat Chandran³, Vikram Patel^{4,12}, Sheffali Gulati¹³, Gauri Divan⁴, Bhismadev Chakrabarti^{1,14,15}

* equal contributions

¹ Centre for Autism, School of Psychology & Clinical Language Sciences, University of Reading, UK

² Division of Psychology, De Montfort University, UK

³ Department of Computer Science, Indian Institute of Technology Bombay, India

⁴ Child Development Group, Sangath, Bardez, Goa, India

⁵ Centre for Chronic Conditions and Injuries, Public Health Foundation of India, India

⁶ Indian Institute of Public Health-Hyderabad, Bengaluru campus, Bengaluru, India

⁷ Com DEALL Trust, Bengaluru, India

⁸ Department of Psychology, Nottingham Trent University, UK

⁹ Department of Psychology, University of East Anglia, UK

¹⁰ Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, UK

¹¹ Department of Psychology, University of Cambridge, UK

¹² Department of Global Health and Social Medicine, Harvard University, USA

¹³ Center of Excellence and Advanced Research on Childhood Neurodevelopmental Disorders, Child Neurology Division, Department of Paediatrics, All India Institute of Medical Sciences, India

¹⁴ India Autism Centre, Kolkata, India

¹⁵ Department of Psychology, Ashoka University, India

Corresponding author: Bhismadev Chakrabarti (b.chakrabarti@reading.ac.uk)

Abstract

Autism Spectrum Disorders, hereafter referred to as autism, emerge early and persist throughout life, contributing significantly to global years lived with disability. Typically, an autism diagnosis depends on clinical assessments by highly trained professionals. This high resource demand poses a challenge in resource-limited areas where skilled personnel are scarce and awareness of neurodevelopmental disorder symptoms is low. We have developed and tested a novel app, START, that can be administered by non-specialists to assess several domains of the autistic phenotype (social, sensory, motor functioning) through direct observation and parent report. N=131 children (2-7 years old; 48 autistic, 43 intellectually disabled, and 40 typically developing) from low-resource settings in the Delhi-NCR region, India were assessed using START in home settings by non-specialist health workers. We observed a consistent pattern of differences between typically and atypically developing children in all three domains assessed. 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. Parent-report further distinguished autistic from non-autistic children. Machine-learning analysis combining all START-derived measures demonstrated 78% classification accuracy for the three groups (ASD, ID, TD). Qualitative analysis of the interviews with health workers and families (N= 15) of the participants suggest high acceptability and feasibility of the app. These results provide proof of principle for START, and demonstrate the potential of a scalable, mobile tool for assessing neurodevelopmental disorders in low-resource settings.

INTRODUCTION

Autism Spectrum Disorder (ASD) is an early-onset neurodevelopmental condition with a global prevalence of ~1% and is a priority for the global mental health agenda (Baxter et al., 2015). It is estimated that India is home to ~5 million families with a child with ASD¹ (Arora et al., 2018). In India, as in many low-resource settings, most families do not receive any, let alone timely, intervention, in part because most never receive a diagnosis (Durkin et al., 2015; Krishnamurthy, 2008; Samms-Vaughan, 2014). Even those who are diagnosed experience significant delays between parents' noticing differences in behaviour and clinical diagnosis by professionals (Jain et al., 2013; Samms-Vaughan, 2014). Low community awareness about ASD 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 skilled human resources serving diagnostic services to a population of over 1.2 billion such as developmental practitioners, psychiatrists, neurologists, and psychologists (Kumar, 2011). When available, these specialists work in urban areas inaccessible to the large rural population. Second, current diagnostic approaches typically involve time-intensive, expensive and proprietary tools, greatly limiting access (Durkin et al., 2015). Most of these tools are either not adapted to local cultures or not available in local languages (Soto 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), consistent with similar evidence from high-income countries (Oono et al., 2012). In light of such evidence that task-shifting, community-based interventions can close the treatment gap, the detection gap becomes an urgent priority, highlighting the need for proactive screening for ASD. The current study aimed at developing a tool that could be used by non-specialists to assess ASD risk in low-resource settings, allowing the closing of the detection gap.

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 'children with ASD' interchangeably with 'autistic children' throughout the manuscript.

Mobile technologies offer a significant advantage in this effort, given their wide penetration across geographies and socioeconomic strata. The increasing availability and affordability of these devices, coupled with the steady advances in sensor technology and portable processing power, makes these technologies ideally suited for scalability. Parallel efforts to harness mobile technology to assess autism-related behavioural phenotypes have shown promise in high-resource settings (Dawson & Sapiro, 2019; Egger et al., 2018). In the current study, we aim to test a battery of tasks that index different aspects of the autistic phenotype, using a mobile device. In view of the diverse phenotypic domains associated with ASD, the mobile platform (app) includes direct assessments of the child on multiple tasks that relate to social behaviour, sensory interests, and motor function. The platform also includes assessment of parent-reported autistic features through a questionnaire, and an observational measure of parent-child interaction.

This validation exercise aims to test whether this assessment can distinguish children with neurodevelopmental disorders, and in particular autism, from their typically developing counterparts. 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 Disorder (ASD), N=48; (2) children with a diagnosis of Intellectual Disabilities (ID), N=43; and (3) typically developing (TD) children N=40 (see Table 1). The ASD and ID groups were recruited by contacting the families of children attending outpatient paediatric clinics at a tertiary clinic (All India Institute of Medical Sciences, Delhi, India). The TD group was recruited through snowballing from communities from where children with ASD and ID were recruited. - The three groups were matched for chronological age. The ASD and ID groups were matched on cognitive age using a language-adapted version of the *Developmental Profile-3 (DP3)* (Alpern, 2007). The ASD group was compared with the other two groups for the severity of autistic symptoms using a locally developed and standardised tool, the *INCLIN Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD)* validated against DSM-V criteria (Juneja et al., 2014).

Table 1: Participant characteristics

	TD	ASD	ID	F/ χ^2	p-value	Post-hoc contrasts, p-value
Chronological Age	(N=40)	(N=48)	(N=43)	F (2, 129) = 0.88	0.42	
M $\pm$ SD	4.59 $\pm$ 1.34	4.24 $\pm$ 1.22	4.56 $\pm$ 1.67			
Gender ratio (F:M)	19:21	12:36	9:34	7.99	0.02	
Cognitive age on DP3	(N=36)	(N=37)	(N=36)	F (2, 106) = 80.87	<0.0001	TD > ASD, <0.0001 TD > ID, <0.0001 ID ~ ASD, 0.19
	4.32 $\pm$ 1.49	1.49 $\pm$ 0.53	1.94 $\pm$ 0.80			
INDT-ASD	(N=37)	(N=37)	(N=39)	F (2, 110) = 109.97	<0.0001	TD < ASD, <0.0001 TD < ID, <0.0001 ID < ASD, <0.0001
	0.16 $\pm$ 0.37	17.16 $\pm$ 4.35	5.15 $\pm$ 7.51			

Tools:

The START task battery was administered on all participants alongside standardized tools of autism symptom severity and developmental level.

Screening Tool for Autism Risk using Technology (START) task battery

START 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 measuring multiple domains of development (Table 2). This choice of domains was informed by the developmental differences commonly identified in children with autism.

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 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 et al. (2011), and 2) a choice-based paradigm similar to that of Ruta et al. (2017).

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 sensitivity.

Atypical *motor* skills are commonly reported in autism (Anzulewicz et al., 2016; Ghaziuddin et al., 1994; 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.

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 Disorder (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 ASD in an Indian setting. Accordingly, the START app includes a brief questionnaire for primary caregivers 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).

[Insert Table 2 here]

[Insert Figure 1 here]

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. Note: Figures 1a and 1b have been removed from this preprint due to the policies of the medrxiv repository, and are available from the corresponding author upon request.

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-ASD questionnaire, and h) caregiver- child interaction observation.

Note: Figures 2a and 2h have been removed from this preprint due to the policies of the medrxiv repository, and are available from the corresponding author upon request.

Standardised tools to evaluate level of functioning and autism symptom severity

The *Developmental Profile 3 (DP3)* (Alpern, 2007) 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 *INCLIN Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD)* (Juneja et al., 2014) 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).

These tools were administered by trained postgraduate psychologists.

Evaluation of acceptability and feasibility of the START tool

Research assistants (1) completed a detailed observation schedule noting the environment and circumstances of each data collection, including family involvement and available resources, (2) 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, and (3) interviewed parents of participating children (TD=5, ASD=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 material.

Procedure:

Two high school graduates with no prior experience in mental health or child assessments or in using tablet computers, were recruited as non-specialist health workers. They were provided a structured two-day classroom training followed by two-day observation and supervised field training in households. Classroom training included START app setup and administration using a detailed statement of procedure. Supervised training with TD child volunteers in the classroom was followed by field training wherein they observed a researcher administer START on one TD and one atypically developing child. They then practised the data collection on two TD volunteers.

Once trained, a health worker and a research assistant visited participants' households to collect data. The health worker administered START on a Samsung SM P600 tablet while the research assistant took field notes, monitored quality of data collection and helped ensure a conducive testing

environment. Health workers used available resources within the households to create the best possible testing environment, particularly around seating and lighting. Testing was generally conducted sitting on the floor or bed.

Data analysis

Pre-set exclusion criteria were applied to the data to ensure quality. This resulted in a different number of participants' data for each task. Detailed information about the criteria used for each task, the final N for each task, and the data analysis is given below.

Preferential Looking Task: Gaze location was identified using a convolutional neural net-based algorithm, based on the gaze-capture algorithm (Bishain, 2019; Dubey, Brett, Ruta, Bishain, Chandran, Bhavnani, Belmonte, Estrin, Johnson, Gliga, & Chakrabarti, 2019; Krafka et al., 2016). Data were collected from 118 of 131 participants (TD = 40, ASD=40, ID=38). All participants met the inclusion criteria of eye detection for at least 50% of frames as well as 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 collected from 116 of 131 participants (TD = 40; ASD = 37; ID = 39). Participants who completed fewer than 50% of trials were excluded from the analysis. A trial in this context refers to an instance of the two buttons being presented (Table 2). This exclusion criterion yielded data from 104 participants (TD=39; ASD=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 collected from 125 of 131 participants (TD = 40, ASD=46, ID=39). Data were filtered for quality by removing participants who completed fewer than two of the five trials, or whose face could be detected in only 25% or fewer of the video frames. A trial was considered complete if it was not aborted. This exclusion criterion yielded data from 117 of 131 participants (TD = 37, ASD=41, ID=39) in the final analysis. Two variables were coded from this data set: a)

Time spent looking at the wheel, and b) distance of the face from the screen. The choice of these variables was based on the observations that autistic children tend to watch the spinning wheel for longer and come closer to it than their TD counterparts (Tavassoli et al., 2016). 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. 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., 2020). Distance of the head to screen has previously been shown to be related to interest (Toker et al., 2017; Witchel et al., 2020).

Motor Following Task: Data were collected from 120 of 131 participants (TD = 40, ASD=43, ID=37). Data sets were filtered for completeness by including only participants who completed two or more trials. This criterion yielded 115 participants (TD = 40, ASD=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. This measure utilized the error in the simultaneous positions of the butterfly and child's touch point. Additionally, we analyse the 'frequency gain' metric for all participants using a Fast Fourier Transformation (FFT). The trajectories of the butterfly and child's trace were resolved into multiple waves of varying amplitudes using FFT. This allowed us to analyze the closeness in the source and target motions along the frequency domain (for details see Supplementary Material).

Bubble Popping Task: Data were collected from 120 of 131 participants (TD = 40, ASD=41, ID=39). Data were included from all the participants who popped one or more balloons. Force used while popping the balloons was recorded using the getPressure parameter recorded by the Android operating system on Samsung tablets, and averaged across all balloons popped. Distance between the touch point and the centre of the balloon was calculated to estimate visuomotor targeting accuracy in approaching dynamic stimuli.

Colouring Task: Data were collected from 113 of 131 participants (TD = 40, ASD=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, ASD=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 outside or *vice versa* was counted as one crossover.

Parent/Caregiver- Child Interaction: Data were included from 100 of 131 participants (TD = 32; ASD = 35, ID = 33). The video-recording of the session was coded using the Dyadic Communication Measure for Autism, a standardised framework for coding dyadic interaction between parents and children by three independent coders (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. Child initiations were scored as a percentage of the total number of attempts to interact with the caregiver, which included a) initiating interactions, b) responding to caregiver, and c) other behaviours. Similarly, caregiver responses were scored as the percentage of synchronous interactions within the total number of interactions. 13% of the videos were coded by all the three coders and used to calculate intra-class correlation (ICC) using a 2-way mixed-effects model, based on 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-ASD Questionnaire: Data were collected from all 131 participants (TD = 40, ASD = 48, ID = 43). The items were scored as binary responses. The summed score indicates the number of 'red flag' signs of ASD.

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 Supplementary Tables 2 and 3.

Machine-learning analysis

This analysis applied a data-driven technique to combine the information from the several measures within START so as to optimise discrimination between the three groups (ASD, ID, TD). In this analysis, 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. First, each feature was evaluated independently for its accuracy in classifying individuals into the three groups. In addition, all possible (non-singleton) combinations of the features were explored to determine a set of features optimised for classification. This combinatorial exploration identifies a suite of tasks that yields the most accurate classification (see Supplementary material 1.2 for details).

Acceptability and feasibility

Interviews were conducted with the non-specialist health workers and caregivers to evaluate feasibility and acceptability of the START task battery in home settings.

Environmental conditions for data capture and nature and frequency of disruptions during the assessment were assessed from the observation schedule used by the research assistant. All interviews were audio-recorded, transcribed and then translated to English and cross-checked for accuracy of translation. In-depth interviews were qualitatively analysed using thematic analysis. Researchers JD and DV independently familiarised themselves with the data by re-reading transcripts, following which the emerging codes were highlighted. Codes were then organised into themes and subthemes using NVIVO. Data were synthesised and triangulated using an iterative process of consensus validation between researchers.

RESULTS

Results are presented below in three sections: a) group comparisons in the START measures, b) group classification accuracy using machine-learning analysis, and c) feasibility and acceptability of START administration by non-specialists in households.

a) 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 children with neurodevelopmental disorders (ASD and ID) looked at the social stimuli less than the TD group did. However, no such group difference was seen in the active choice task (button task). In the sensory domain, an effect of group membership was noted in time looking at the spinning wheel: Children with ASD and ID looked at the spinning wheel longer than their TD counterparts did. In the motor domain, neurodevelopmentally disordered 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 children with ASD endorsed higher numbers of items from the START-ASD questionnaire than parents of either ID or TD children. An overall inspection of Table 3 suggests a consistent pattern of group difference for neurodevelopmental disorder and TD groups. However the two neurodevelopmental disorder groups (ASD and ID) are not clearly discriminable from most of the measures (the questionnaire data and visuomotor RMS error being notable exceptions).

[Insert Table 3 here]

b) Machine-learning analysis

The classification accuracies, the sample sizes for each group and other details as determined in the machine-learning analysis are provided in Supplementary Table 1. Based on these results, the Motor task (RMSE in following the trajectory) was the most promising independent task with 60% overall classification accuracy into three groups (TD, ID, ASD), superior to a random chance classification accuracy of approximately 33%. This discrimination accuracy is at par with that reported by the questionnaire measure (Figure 4).

Figure 4: Bar graphs showing comparison of mean classification accuracies of the feature vectors taken from the eight tasks included in the START battery. The figure also represents the best classification accuracy achieved by a combination of these features. (Prefixes on x-axis in brackets refer to corresponding feature IDs). Note that some features are designed as multidimensional vectors with different measures along its dimensions

While the overall classification accuracy for each of the tasks is relatively weak, we observe a significant improvement when various combinations of the corresponding features are analysed together. The predictive power of the weaker features is significantly boosted and we report a much stronger overall classification accuracy of 78% for the three groups of ASD, ID, and TD (See Table 4). The final set of features comprised three motor / visuomotor features from three tasks- RMS error

in 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.

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 (ASD)	Mean Classification Accuracy (ID)	Mean Classification Accuracy (TD)	Mean Overall Classification Accuracy	Mean proportion % of subjects across different groups (ASD: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

c) Feasibility and Acceptability

High completion rates (>70%) were obtained in all task measures collected (see Figure 5). 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.

Figure 5: Task completion rate for each START task for each of the three groups of children (ASD, ID and TD)

Triangulation of data from the observation schedule and in-depth interviews (see Supplementary Table 4) with non-specialist health workers revealed that although conducting a child developmental assessment in households posed unique challenges such as limitations of space, variations in lighting and background noise and interruptions by family members, the health workers were able to identify suitable conditions to optimise data capture, particularly through cooperation with family members. Health workers specified that they found the detailed written standard operating protocols and scripts provided to them during training useful in guiding assessments. Interestingly, while some design elements in the app facilitated its smooth administration, other innovations, such as webcam video-based eye-tracking, led to confidentiality concerns in parents. App-based assessment seems to have high acceptability for children, who engaged well with the health workers, actively played the “games” on the tablet and enjoyed its child-friendly design elements. As expected, TD children were better able to engage with the team and the task than were the children with neurodevelopmental disorders, but parents of children with neurodevelopmental disorders reported that their child was comfortable during these assessments. Results also show that parents found START acceptable but questioned the credibility of an app-based assessment of child development.

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 disorders. This mobile platform was found to be both feasible for delivery by non-specialists in a community-based setting in homes and acceptable to all users including community health workers, parents and children. Workers were able to collect quantitative, multidimensional data in cooperation with families; this experience is consonant with those of others working with profoundly autistic children (McKinney et al., 2021). We find strong evidence for group differences on the majority of the measures between children with and without neurodevelopmental disorders, and comparatively little evidence for differences between the two groups of children with neurodevelopmental disorders (ASD and ID).

Task measures

The task measures focused on social, sensory, and motor functioning. Significant group differences between the TD and neurodevelopmental disorder groups were noted in tasks within each of these three domains. Specifically in the social domain, greater attention to social over non-social rewards was noted in TD children, using a preferential looking task. This pattern of results is consistent with reports on similar paradigms applied in laboratory settings, using standard infra-red eye trackers (Hedger et al., 2020). This observation supports the view that preferential looking toward social stimuli can be a proxy for social features of autism, and is not a proxy for general cognitive ability. It is also consistent with the results of a recent meta-analysis that showed no effect of IQ on preferential looking toward social stimuli (Hedger et al., 2020). Reduced attention to social stimuli has been widely noted in autistic children and adults, and has been suggested to be predictive of autistic symptomatology in later childhood (Bacon et al., 2020).

In contrast to the preferential looking task, where participants view the stimuli passively, the button task involves making an instrumental choice. Contrary to expectations, the button task did not show a difference between the three groups. This absence of any 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). In the original report, the contingencies between the button and the

associated stimuli (social/non-social) were not explicitly told to the child. In contrast, the current administration involved the child being clearly shown how one of the buttons was associated with the video of a child while the other was associated with a set of moving shapes. This change might have reduced the child's motivation to explore on his/her own, leading to roughly equal numbers of button presses for the social or the non-social stimulus. It also is possible that this apparent inconsistency might be driven by cultural and/or ascertainment-related differences in this sample from a tertiary clinic in urban India, in contrast to those of the Italian sample on which the first report was based.

Strong group differences were noted in all task measures of motor function. The TD group performed more accurately than both the ASD and ID groups in the motor following task, as indexed by lower spatial errors (RMSE). However, no group differences were noted in the jerk profile in this dynamic visuomotor task, as have previously been shown in adults with autism using a gross motor task (Cook et al., 2013) and in autistic children using a visuomotor task of touchscreen movements to static targets (Weisblatt et al., 2019). Convergent findings indicating poorer visuomotor control in autistic children compared to the TD 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 ASD group used significantly greater force than the TD group in this task, replicating similar results reported using a different tablet-based task in children with ASD (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 closed-loop motor control, wherein autistic users are less able to decelerate their action on time to hit a precise location on the tablet screen. Poor motor control can result from a reduced use of sensory information to adjust motor behaviour as an online process (Haswell et al., 2009), and is consistent with longstanding theoretical models of sensorimotor and cognitive prediction error in autism (Courchesne & Allen, 1997; Van de Cruys et al., 2014; Sinha et al., 2014; Palmer et al., 2017).

In the domain of sensory sensitivity, we used a tablet adaptation of a task previously associated with group differences between children with and without ASD (Tavassoli et al., 2016). Greater preference for specific types of sensory stimuli such as spinning wheels has been widely observed in autism (Sasson et al., 2011). While the underlying mechanisms for enhanced interest in such stimuli remain poorly understood, one feature shared by these stimuli is high predictability. We observed an

expected pattern of group differences in the current sample, with autistic children spending longer time looking at the spinning wheel with illusory contours compared to the TD children. As such, this result demonstrates generalisability of the original observations with a real spinning wheel, and points to the potential for a scalable assay of visual sensory sensitivity.

Parent/Caregiver-Report and Interaction measures

The parent/caregiver-report questionnaire was based closely on a tool specific for identification of ASD in an Indian context (INDT-ASD). Unsurprisingly, scores on this questionnaire significantly differed between all three groups (ASD, ID, TD) in the expected direction, replicating previous reports with the original tool (Gulati et al., 2019; Juneja et al., 2014).

The caregiver-child interactive episode revealed substantial atypicality in both key metrics of interaction. Autistic children initiated social interaction less than the TD group did, and also trended toward fewer initiations compared to the ID group. Fewer synchronous responses from the caregiver were evoked in interaction with ASD and ID children, 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 in home settings. 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. Two caveats do need to be noted. First, there was little evidence for specificity of these measures in discriminating between the ASD and ID groups. To investigate this apparent equivalence further, we re-examined each case's clinical notes, which revealed two insights: a) all of the ASD cases also met criteria for ID. This observation reflects the ground realities in India, where most ASD diagnoses in children within tertiary centres are at the severe end of the spectrum, and likely to be associated with developmental delay. b) A majority of the children in the ID group showed significantly elevated autistic symptoms. The phenotypic overlap in these groups could

therefore have contributed to the observed lack of group differences between ASD and ID children. Second, all children in the atypical neurodevelopmental groups came from a tertiary clinic, which would have introduced a bias toward more severe levels of dysfunction, leading to greater group differences and also, as aforementioned, potentially to inconsistencies with previous results such as the absence of group difference on the instrumental social choice button task.

These caveats notwithstanding, we combined all the measures to test their ability to discriminate the ASD, ID, and the TD groups using machine-learning. This analysis revealed an overall accuracy of 78% in classifying the three groups, a considerable boost from the accuracy achieved by any of the measures alone. This result highlights the advantages of a multi-measure platform that combines task performance with parent/caregiver-report to achieve greater precision in assessing autism.

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 task, questionnaire, and observational assessments of aspects of behaviour that index social, sensory, and motor function. Individual metrics from each task showed 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 (ASD, ID, TD). Future work should test this app prospectively in a large population-based study to assess the predictive validity of each of these measures independently, and in combination, for atypical neurodevelopmental status. The eventual goal for this platform is its incorporation into a routine health surveillance system to help identify children with- or at risk of- neurodevelopmental disorders.

Acknowledgements

This work was funded by a Medical Research Council Global Challenge Research Fund grant to the START consortium (PI: Chakrabarti; Grant ID: MR/P023894/1). The authors express their gratitude to Geraldine Dawson for sharing the protocol for parent-child interaction, to Therapy Box UK for app development, and to all families and health workers who took part in this study.

References

- Alpern, G. D. (2007). *Developmental Profile 3, DP-3: Manual*.
- Anzulewicz, A., Czahaj, D., Delafield-Butt, J. T., Jamolkowicz, P., Mnich, M., Paciorek, A. W., & Others. (2016). Using Tablet-Based Gameplay for the Identification of Autism-Related Movement Patterns. *IMFAR Abstract Book*, 225, 2016.
- 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.
- 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.
- Ausderau, K., Sideris, J., Furlong, M., Little, L. M., Bulluck, J., & Baranek, G. T. (2014). National survey of sensory features in children with ASD: factor structure of the sensory experience questionnaire (3.0). *Journal of Autism and Developmental Disorders*, 44(4), 915–925.
- Bacon, E. C., Moore, A., Lee, Q., Carter Barnes, C., Courchesne, E., & Pierce, K. (2020). Identifying prognostic markers in autism spectrum disorder using eye tracking. *Autism: The International Journal of Research and Practice*, 24(3), 658–669.
- Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, 45(3), 601–613.
- 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: Official Publication of the American Occupational Therapy Association*, 61(5), 584–592.
- Bishain, R. (2019, May 2). An Open-Source, Computing Platform-Agnostic, Calibration-Free Preferential Gaze Detection Approach for Social Preference Assessment. *INSAR 2019 Annual Meeting*. <https://insar.confex.com/insar/2019/webprogram/Paper31575.html>
- Chakrabarti, B., Haffey, A., Canzano, L., Taylor, C. P., & McSorley, E. (2017). Individual differences in responsivity to social rewards: Insights from two eye-tracking tasks. *PloS One*, 12(10), e0185146.
- Cook, J. L., Blakemore, S.-J., & Press, C. (2013). Atypical basic movement kinematics in autism spectrum conditions. *Brain: A Journal of Neurology*, 136(Pt 9), 2816–2824.
- Courchesne, E., & Allen, G. (1997). Prediction and preparation, fundamental functions of the cerebellum. *Learning and Memory*, 4, 1-35.
- Dawson, G., & Sapiro, G. (2019). Potential for Digital Behavioral Measurement Tools to Transform the Detection and Diagnosis of Autism Spectrum Disorder. *JAMA Pediatrics*, 173(4), 305–306.
- Dubey, I., Brett, S., Ruta, L., Bishain, R., Chandran, S., Bhavnani, S., Belmonte, M., Estrin, G. L., Johnson, M. H., Gliga, T., & al., E. (2019). *Social is not always special: Evidence for task specificity in measuring social preference in young children*. <https://doi.org/10.31234/osf.io/3un5e>
- Dubey, I., Brett, S., Ruta, L., Bishain, R., Chandran, S., Bhavnani, S., Belmonte, M., Estrin, G. L., Johnson, M.

- H., Gliga, T., & Chakrabarti, B. (2019). Social is not always special: Evidence for task specificity in measuring social preference in young children. In *PsyArXiv*.
- Dubey, I., Ropar, D., & Hamilton, A. (2015). Measuring the value of social engagement in adults with and without autism. *Molecular Autism*, 6, 35.
- Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., Lee, L.-C., 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: Official Journal of the International Society for Autism Research*, 8(5), 473–476.
- 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., Jr, & Sapiro, G. (2018). Automatic emotion and attention analysis of young children at home: a ResearchKit autism feasibility study. *NPJ Digital Medicine*, 1, 20.
- 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. In *British Journal of Psychiatry* (Vol. 205, Issue 2, pp. 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.
- Ghaziuddin, M., Butler, E., Tsai, L., & Ghaziuddin, N. (1994). Is clumsiness a marker for Asperger syndrome? *Journal of Intellectual Disability Research: JIDR*, 38 (Pt 5), 519–527.
- 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., & PACT Consortium. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *The Lancet*, 375(9732), 2152–2160.
- 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.
- Haswell, C. C., Izawa, J., Dowell, L. R., Mostofsky, S. H., & Shadmehr, R. (2009). Representation of internal models of action in the autistic brain. *Nature Neuroscience*, 12(8), 970–972.
- Hedger, N., Dubey, I., & Chakrabarti, B. (2020). Social orienting and social seeking behaviors in ASD. A meta analytic investigation. *Neuroscience and Biobehavioral Reviews*, 119, 376–395.
- Hedger, N., Haffey, A., McSorley, E., & Chakrabarti, B. (2018). Empathy modulates the temporal structure of social attention. *Proceedings. Biological Sciences / The Royal Society*, 285(1893), 20181716.
- Jain, R., Juneja, M., & Sairam, S. (2013). Children With Developmental Disabilities in India. In *Journal of Child Neurology* (Vol. 28, Issue 4, pp. 455–460). <https://doi.org/10.1177/0883073812447685>
- 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. K. C., Arora, N. K., & INCLEN Study Group. (2014). INCLEN Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD): development and validation. *Indian Pediatrics*, 51(5), 359–365.
- Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., & Torralba, A. (2016). Eye Tracking for Everyone. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Krishnamurthy, V. (2008). A clinical experience of autism in India. *Journal of Developmental and Behavioral Pediatrics: JDBP*, 29(4), 331–333.
- Kumar, A. (2011). Mental health services in rural India: challenges and prospects. In *Health* (Vol. 03, Issue 12, pp. 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.
- McKinney, A., Hotson, K. L., Ahmed, Z. B., Días, C., Ben Shalom, D., Weisblatt, E. J. L., Foster, J., Villar, S. S., Murphy, S., & Belmonte, M. K. (2021). Overcoming hurdles to intervention studies with autistic children with profound communication difficulties and their families. *Autism*, 25, in press. doi:

- 10.1177/1362361321998916
- 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. In *International Review of Psychiatry* (Vol. 27, Issue 3, pp. 247–256). <https://doi.org/10.3109/09540261.2015.1049128>
- Oono, I. P., McConachie, H., & Honey, E. J. (2012). Parent-mediated early intervention for young children with autism spectrum disorders (ASD). In *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.cd009774>
- Palmer, C. J., Lawson, R. P., & Hohwy, J. (2017). Bayesian approaches to autism: Towards volatility, action, and behavior. *Psychological Bulletin*, 143(5), 521–542.
- 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), 101–109.
- Posar, A., & Visconti, P. (2018). Sensory abnormalities in children with autism spectrum disorder. *Jornal de Pediatria*, 94(4), 342–350.
- 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.
- 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. Experimentelle Hirnforschung. Experimentation Cerebrale*, 217(1), 35–41.
- Robins, D. L., & Dumont-Mathieu, T. M. (2006). Early screening for autism spectrum disorders: update on the modified checklist for autism in toddlers and other measures. *Journal of Developmental and Behavioral Pediatrics: JDBP*, 27(2 Suppl), S111–S119.
- 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.
- Samms-Vaughan, M. E. (2014). The status of early identification and early intervention in autism spectrum disorders in lower- and middle-income countries. In *International Journal of Speech-Language Pathology* (Vol. 16, Issue 1, pp. 30–35). <https://doi.org/10.3109/17549507.2013.866271>
- Sasson, N. J., Elison, J. T., Turner-Brown, L. M., Dichter, G. S., & Bodfish, J. W. (2011). Brief report: Circumscribed attention in young children with autism. *Journal of Autism and Developmental Disorders*, 41(2), 242–247.
- 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.
- Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42), 15220–15225.
- Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism: The International Journal of Research and Practice*, 19(6), 646–661.
- Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the Screening Tool for Autism in Two-Year-Olds (STAT) for children under 24 months: an exploratory study. *Autism: The International Journal of Research and Practice*, 12(5), 557–573.
- 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.
- Toker, D., Lallé, S., & Conati, C. (2017). Pupillometry and Head Distance to the Screen to Predict Skill Acquisition During Information Visualization Tasks. *Proceedings of the 22nd International Conference*

- on *Intelligent User Interfaces*, 221–231.
- 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.
- Vats, P., Juneja, M., & Mishra, D. (2018). Diagnostic Accuracy of International Epidemiology Network (INCLIN) Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD) in Comparison with Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5). In *Indian Pediatrics* (Vol. 55, Issue 6, pp. 485–487). <https://doi.org/10.1007/s13312-018-1339-8>
- Weisblatt, E. J. L., Langensiepen, C. S., Cook, B., Días, C., Plaisted Grant, K., Dhariwal, M., Fairclough, M. S., Friend, S. E., Malone, A. E., Varga-Elmiyeh, B., Rybicki, A., Karanth, P., & Belmonte, M. K. (2019). A tablet computer-assisted motor and language skills training programme to promote communication development in children with autism: Development and pilot study. *International Journal of Human-Computer Interaction* 35(8):643-665. doi: 10.1080/10447318.2018.1550176
- Witchel, H. J., Santos, C. P., Ackah, J. K., Chockalingam, N., & Westling, C. E. I. (2020). Proxemics of screen mediation: engagement with reading on screen manifests as diminished variation due to self-control, rather than diminished mean distance from screen. *Electronic Journal of Communication*, 29(3 & 4). http://sro.sussex.ac.uk/id/eprint/89749/1/Witchel_proxemics_ElecJComm_accepted_2019-12-08.pdf