Multimodal Digital Phenotyping of Behavior in a Neurology Clinic: Development of the Neurobooth Platform and the First Two Years of Data Collection

Adonay S. Nunes[1], Siddharth Patel[1], Brandon Oubre[1], Mainak Jas[2], Divya D. Kulkarni[1], Anna C. Luddy[1], Nicole M. Eklund[1], Faye X. Yang[1], Rohin Manohar[1], Nancy N. Soja[1], Katherine M. Burke[1,3], Bonnie Wong[4], Dmitry Isaev[5], Steven Espinosa[6], Jeremy D. Schmahmann[1], Christopher D. Stephen[1], Anne-Marie Wills[1], Albert Hung[1], Bradford C. Dickerson [1, 2, 4], James D. Berry[1], Steven E. Arnold[1], Vikram Khurana[7], Lawrence White[1], Guillermo Sapiro[6,8], Krzysztof Z. Gajos[9], Sheraz Khan[2], Anoopum S. Gupta[1]

[1]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[2]Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[3]Massachusetts General Hospital Institute of Health Professions, Boston, MA, USA

[4]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[5]Department of Biomedical Engineering, Duke University, Durham, NC, USA

[6]Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA

[7]Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[8]Department of Electrical and Computer Engineering, Princeton University, NJ, USA

[9]Harvard John A. Paulson School of Engineering and Applied Sciences, Allston, Massachusetts, USA

Correspondence to: Anoopum S. Gupta

101 Merrimac St Suite 300, Boston, MA 02114

agupta@mgh.harvard.edu

Abstract

Quantitative analysis of human behavior is critical for objective characterization of neurological phenotypes, early detection of neurodegenerative diseases, and development of more sensitive measures of disease progression to support clinical trials and translation of new therapies into clinical practice. Sophisticated computational modeling can support these objectives, but requires large, information-rich data sets. This work introduces Neurobooth, a customizable platform for time-synchronized multimodal capture of human behavior. Over a two year period, a Neurobooth implementation integrated into a clinical setting facilitated data collection across multiple behavioral domains from a cohort of 470 individuals (82 controls and 388 with neurologic diseases) who participated in a collective 782 sessions. Visualization of the multimodal time series data demonstrates the presence of rich phenotypic signs across a range of diseases. These data and the open-source platform offer potential for advancing our understanding of neurological diseases and facilitating therapy development, and may be a valuable resource for related fields that study human behavior.

Main Text

Neurological disorders are the leading cause of disability and the second leading cause of death worldwide[1]. There is a large unmet need in neurology for safe and effective therapies. Drug development efforts for neurodegenerative diseases have accelerated in recent years[2]. However, efficient evaluation of new therapies is limited by subjective and insensitive outcome measures, diagnostic uncertainty, incomplete patient stratification, and difficulty in evaluating therapies at the earliest stages of disease[3–5]. Quantitative characterization of digitized behavior using signal processing and machine learning approaches has the potential to address these limitations and accelerate neurology drug development[6].

Through their effects at the molecular, cellular, and circuit levels, neurodegenerative diseases progressively alter motor and cognitive behavior in life-changing and characteristic ways. These behavioral changes are carefully evaluated by the neurologist during a neurological examination[7] in order to 1) identify affected components of the nervous system, 2) make an accurate diagnosis, 3) track disease progression, and 4) evaluate response to interventions. Although powerful, clinician-performed assessments cannot capture behavioral patterns that exist beyond the threshold of human perception or that manifest during naturalistic behaviors not represented in the neurological examination. There is also information loss between what the clinician observes and the data recorded, which limits the ability to produce granular datasets to support discovery of new behavioral patterns and identification of disease subgroups. Quantitative capture of behavioral data has the potential to generate information-rich datasets and enable the use of sophisticated, data-driven computational methods to identify clinically-relevant behavioral patterns[6,8]. Adoption of standardized platforms for digital capture of behavior could lead to large scale datasets for training advanced, data-driven models that could broadly transform our understanding of neurological disease and human behavior.

There is a long history of quantitative behavioral data collection in neurological diseases. Various sensor systems have been used to record and assess eye movements using video

oculography[9–16], speech using microphones[17–21], and body movement using wearable sensors[22–28] and motion capture[29–32] systems. It has been shown that eye trackers can detect eye movement control alterations in neurodegenerative populations at early stages[11,33–35]. Computer vision-based analysis of video data obtained from consumer-grade cameras can be used to characterize and quantify gait, hand movements, and eye movements from individuals with neurological diseases[36–41]. Wearable devices have been widely used in neurological populations to quantify gait and limb movements, and to perform disease classification and estimate severity[42–46]. Computational analysis of speech has demonstrated the ability to detect early disease signs, estimate disease severity, and sensitively measure change in neurodegenerative diseases[47–56]. Finally, it has been shown that quantitative analysis of computer mouse movements in both free-living conditions[57] and during motor tasks[58–60] reveals patterns informative for detecting neurodegenerative diseases, including Alzheimer's, Parkinson's, and cerebellar ataxias. These studies have improved our understanding of how neurological conditions affect behavior. Despite the value these sensor-based systems have demonstrated in the collection and analysis of quantitative behavioral data, digital measures have not been widely adopted in clinical care or in clinical trials[23,61,62].

Reasons for the limited impact of digital measures to date are that quantitative phenotyping studies typically involve a small number of participants with a well-established diagnosis, focus on a single behavioral domain, and/or use specialized systems for data collection and analysis. There is concern that measures developed from data collected using specialized systems (i.e., custom devices and task protocols) at a single site in a small number of individuals, will not generalize to the larger population and may be too time consuming and/or complex to implement in a clinical setting or in a multi-site clinical trial. The single population, single domain approach also limits the scientific questions that can be advanced and the computational methods that can be used to discover useful disease patterns in typically heterogeneous neurological populations. For example, to adequately detect and measure behavioral patterns that reflect subtle parkinsonism and spasticity in spinocerebellar ataxia populations, conditions with overt parkinsonism and spasticity, such as idiopathic Parkinson's disease and hereditary spastic paraplegias, should be included and jointly analyzed.

To address these gaps, we developed a high-throughput, open source, and flexible platform called "Neurobooth" to support large-scale multimodal data collection. We created and deployed the first version of this system in a space embedded within the Massachusetts General Hospital outpatient neurology clinic with integration into the clinical workflow. Over a two-year period, synchronized video, audio, inertial sensor, eye tracker, computer mouse, and button-board data were collected in a standardized manner across 782 sessions from 470 unique participants with wide ranging neurological diagnoses. Here we provide an overview of the software and hardware platform and the data generated along with data visualizations. This platform is available to researchers to support quantitative behavioral phenotyping studies and the dataset can be requested. Finally, this study demonstrates the feasibility to integrate into a clinical workflow, providing an avenue for much larger scale data generation with future implementation at additional clinical locations. Integration within a clinic also makes it possible to share results with clinicians and patients at the time of the assessment to provide new information and gain feedback.

Results

System Capabilities

Neurobooth is an integrated software and hardware system for collecting quantitative behavioral data. It supports presentation of behavioral stimuli and synchronized multimodal data collection across one or more computers. Because the Neurobooth operating system (NeuroboothOS) can be distributed across multiple computers, it enables concurrent, high-throughput recording across a large number of connected devices. Neurobooth software is open source and highly customizable. It is designed to facilitate the addition of new sensors, stimuli, and study parameters, thereby supporting multiple experimental protocols and disease populations. The Neurobooth software also supports the organization, visualization, and analysis of behavioral time series data. These capabilities aim to balance flexibility (to improve and adapt the system) with standardization (to promote data aggregation and replication of results).

Dataset

The first Neurobooth system (Fig. 1) was deployed in the Massachusetts General Hospital (MGH) neurology clinic in April 2022, supported by an award from the Massachusetts Life Sciences Center. All participants provided informed consent and the research protocol was approved by the MGH Institutional Review Board (2021P000257, approved March 3, 2021). The deployment at MGH aimed to collect longitudinal digital data from a broad range of neurodegenerative disease populations, including presymptomatic individuals who carry a specific disease gene but do not yet have symptoms or definite clinical signs of disease, and individuals who have not yet received a conclusive clinical diagnosis. With MGH Neurobooth situated in the clinic, most participants with neurologic disorders participated in the study directly before or after their regularly scheduled clinic visits, enabling scalable longitudinal data collection. The MGH Neurobooth is also wheelchair accessible to support data collection in disease populations with limited mobility.

Between April 2022 and April 2024, data were collected during 782 sessions from 470 unique participants. Table 1 shows the clinical characteristics of these participants. A total of 213 of these participants have longitudinal data spanning up to 20 months (141 participated in two sessions, 45 participated in three sessions, and 27 participated in four sessions to date).

FIGURE 1 ABOUT HERE

TABLE 1 ABOUT HERE

In addition to synchronized time-series data collected from a variety of sensors (Table 2), a corpus of contextual clinical data was collected to support development of clinically meaningful digital features. These data included both clinician-performed disease rating scales that were completed during the clinic visit and a battery of patient-completed surveys capturing information about daily function and quality of life.

TABLE 2 ABOUT HERE

Behavioral Tasks

Neurobooth data collection deployed a set of tasks spanning several neuromotor and neurocognitive behavioral domains—cognition, oculomotor function, fine and gross motor function, and speech—that are broadly relevant across neurodegenerative diseases and can be performed in a limited space and time setting (see Methods for task details). Several tasks involved coordination of multiple behavioral domains, taking advantage of the synchronous multimodal data capture supported by Neurobooth. The initial task set was designed such that it could be completed by most participants in 30-40 minutes without over-taxing patients on their clinic day, and to maximize the number of people who could participate on a given day. In the sections below, we provide a brief overview of the behavioral tasks and provide a data visualization for each task that highlights a potential disease-related behavioral pattern to motivate future population-level data analysis.

Cognitive Tasks

Cognitive tasks included the digit symbol substitution test (DSST) and the multiple object tracking test (MOT). We used the validated structure from TestMyBrain.org[63] to develop the cognitive tasks in Python.

FIGURE 2 ABOUT HERE

The top row of Fig. 2 contrasts performance on a four-target, low-speed trial of MOT for a control versus a participant with corticobasal syndrome (CBS). The trajectory of gaze is shown in gray and the trajectory of the computer mouse is shown in blue, with cyan dots representing mouse clicks. The participant with CBS had less direct and longer gaze and mouse paths during the response period. One target was incorrectly selected, although their gaze and mouse trajectories indicate they considered the correct target. Though not shown in the figure, tracking gaze throughout the task could reveal different attention strategies for performing the task, which may be modified by neurological diseases[64,65].

The bottom row of Fig. 2 compares a healthy control participant versus a participant with mild cognitive impairment (MCI) performing DSST. The heat maps show the density of gaze duration over the 60-second long task. The density in the left plot shows that the gaze of the control participant was largely fixated on the location of the stimulus symbol. In contrast, the cognitively impaired participant spent a significant time gazing at the key, as is apparent from the larger spread of the density in the heat map on the right. This shows that the attention of the participant with MCI was divided between the stimulus symbol and the key. The horizontal plots below the heat maps are a time series of events over the 60-second task. Vertical blue lines indicate when a new stimulus symbol appeared. Shaded blue regions indicate the length of the response times. The control participant registered 61 responses compared to 25 responses from the MCI participant. Differences in the thickness of the blue shaded regions provide a

visual indication of the variance in response times. Both the control and the MCI participant were 100% accurate in these sessions.

Oculomotor Tasks

Oculomotor tasks included tasks that include a visual target (horizontal smooth pursuit, gaze holding, horizontal saccades, and vertical saccades) and gaze fixation without a target. Fig. 3 compares the gaze during each task between healthy controls and participants with neurologic disorders. For smooth pursuit, the individual with multiple system atrophy had saccadic pursuit, with more pronounced saccadic pursuit when tracking the object moving rightward. For horizontal saccades, the person with Gordon-Holmes syndrome had hypermetric saccades in addition to gaze-evoked nystagmus during fixations. For vertical saccades, the individual with progressive supranuclear palsy had slowed and segmented saccades and evidence of 'round the houses' sign[66] (curved trajectory apparent with concurrent movement in the horizontal direction). For gaze holding, the individual with cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS) exhibits downbeat nystagmus that varies in intensity depending on direction of gaze. For fixation without a target, the individual with spinocerebellar ataxia (SCA) type 3 (SCA-3) has vertical drift of gaze, with corrective saccades, and superimposed nystagmus.

FIGURE 3 ABOUT HERE

Motor Tasks

Motor tasks included finger-to-nose, foot tapping, rapid alternating hand movements, sit-to-stand, and the Hevelius mouse clicking task[58,59]. Fig 4 shows ten seconds of data collected during each task for a healthy control participant and an individual with a neurological disorder.

FIGURE 4 ABOUT HERE

Finger-to-nose data from an individual with spinocerebellar ataxia type 3 (SCA-3) and rapid alternating hand movements data from an individual with SCA-2 demonstrate reduced speed and variability in amplitude and rhythm compared with the healthy control participant (Fig. 4).

Foot tapping data from an individual with multiple system atrophy, parkinsonian type (MSA-P) exhibits smaller movements and variability in amplitude compared with the healthy control participant (Fig. 4).

Sit-to-stand data (lumbar sensor) from an individual with autosomal recessive cerebellar ataxia type 1 (ARCA-1) shows less regular and lower intensity movements in the dorso-ventral axis and sharp peaks in the rostro-caudal axis compared to the smoother movement profile seen in the healthy control example.

On the computer mouse task, an individual with amyotrophic lateral sclerosis (ALS) took longer to perform the task and exhibited mouse movements that were less smooth and had more target re-entries. The presence of eye tracking data alongside mouse movements and stimulus

onset provides an opportunity to assess how eye-hand coordination is affected by different motor disorders (Fig. 4).

Speech Tasks

Speech tasks included sustained phonation, diadochokinetic tasks, and passage reading[67,68]. Fig. 5 compares recorded audio and gaze (for passage reading) for healthy control participants and individuals with neurologic disorders. For sustained phonation, in comparison with the control participant, a participant with Parkinson's disease had decreasing loudness as the task progressed. For the diadochokinesis task, the individual with SCA-6 exhibited variability in amplitude and rhythm, as well as slowed consonant production. For passage reading, an individual with Friedreich's ataxia had slowed speech and longer pauses during speech production, the presence of smaller saccades, and an increased frequency of regressions in gaze during reading.

FIGURE 5 ABOUT HERE

Study Feedback

To collect participants' feedback on the Neurobooth system and their overall study experience, we administered a standardized 10-item System Usability Scale (SUS)[69] and a Study Feedback survey created by our team after each visit. A total of 581 responses were captured from 363 unique participants.

Each item in the SUS questionnaire was rated from 1 to 5 on the Likert scale (1=strongly disagree and 5=strongly agree). To interpret the score, we calculated the sum of the score contributions for each item. For items 1, 3, 5, 7 and 9 (odd-numbered, positive statements), the score contribution was the scale position minus 1. For items 2, 4, 6, 8 and 10 (even-numbered, negative statements), the score contribution was 5 minus the scale position. Next, we multiplied the total by 2.5 to obtain the overall System Usability value, which ranges from 0 (very poor perceived usability) to 100 (excellent perceived usability)[69]. Mean SUS score from Neurobooth respondents was 78.4 with a standard deviation of 16.4. This meets the benchmark score of 68.05 for digital health applications as described in the current literature[70].

In the Study Feedback survey, we assessed participants' willingness to complete the full Neurobooth assessment again in the future. The majority of responses indicated yes, they would be interested (81.1%) while some expressed willingness to complete a shorter version of the assessment (11.8%). A smaller proportion of responses indicated participants were unsure (3.5%), they would probably not be willing to participate again (2.6%), or they would definitely not participate again (1.0%). We also surveyed participants on their most and least enjoyed study tasks. 42.0% of responses indicated cognitive tasks as their most enjoyed, followed by 21.7% for motor tasks. For least enjoyed tasks, 36.7% indicated none (they enjoyed all of the activities), followed by 26.8% for speech tasks. Lastly, participants were given an optional, free-text comment box to report any neurological symptoms of theirs they felt were not captured

during testing. Participant responses, ordered from most to least frequent, included balance, gait, visual impairments (blurred vision, double vision, oscillopsia), neuropathy, weakness, handwriting, intermittent symptoms/general fluctuations throughout the day, dizziness, using stairs, reading comprehension, and lower extremity coordination during weight bearing activities.

Discussion

In the first two years of the project, data from hundreds of individuals have been collected using Neurobooth, demonstrating the feasibility of collecting high throughput, multimodal digital data from diverse disease populations in the outpatient neurology clinic setting. Data visualization at the individual level indicates the ability to capture key neurological signs in the range of populations studied. This granular and comprehensive dataset provides an opportunity to detect and quantify neurological signs, characterize early disease features and change over time, and identify the presence of neurological subgroups and subphenotypes. Developing an open source platform for standardized data collection that can be used at other institutions enables the possibility of generating even larger datasets in aggregate. These datasets, combined with sophisticated computational modeling techniques, will lead to a more complete understanding of how neurological diseases affect behavior.

Neurobooth consists of a wide array of devices, including cameras, wearable sensors, an eye tracker, microphone, computer mouse, button board, and a mobile phone (with integrated camera and microphone). These devices, along with a broad set of behavioral tasks, enable a comprehensive capture of behavior relevant across neurological conditions. In addition to supporting single-domain analyses in a larger sample than prior studies, Neurobooth provides multimodal, time-synchronized data enabling new opportunities for the identification of less obvious patterns relating to the coordination of multiple motor domains. For example, multimodal modeling[71–73] (e.g., eye movements and mouse movements during the computer mouse task; or speech, eye movements, and face movements during passage reading[74]) could enable discovery of new disease signs related to coordination of different behavioral domains that were not previously recognized or measurable by human assessors. Recent digital phenotyping efforts for screening of neurodevelopmental disorders highlights the importance of multimodal data[75,76]. The collection of contextual gold standard clinician rating scale data and patient reported measures of function enables validation of existing or newly discovered digital behavioral patterns or signs. Furthermore, the collection of longitudinal data, beginning in some individuals in the presymptomatic stage of disease, supports understanding how a digitally captured behavioral pattern begins and evolves over time. Although the diagnosis used for analysis will be the treating neurologist's diagnosis, in cases where a definitive diagnosis can only be made based on neuropathology (e.g., progressive supranuclear palsy), attempts will be made to follow individuals until autopsy when possible.

Neurobooth incorporates a broad array of devices, and also allows for the integration of new device types. The system can be extended to include devices to measure strength or to

measure physiological signals synchronously with behavior, such as respiration, electrocardiography (ECG), and electroencephalography (EEG). This flexibility makes it possible to broaden investigations while maintaining standardized data capture that allow for aggregation. Conversely, a system with a reduced set of devices can be deployed to reduce cost and allow data capture in a variety of settings. To this end, we are currently working to make it possible for a reduced version of Neurobooth to run on a single laptop and enable data collection in the home setting with consumer grade devices. In addition to device-flexibility, there is flexibility to change or remove existing behavioral tasks and to include additional tasks in a recording session. Thus the system can also be adapted to facilitate research beyond neurological populations. One goal of the platform is to maintain an open source library of behavioral tasks and device integrations that can be drawn upon by any research group or clinical team.

There were challenges encountered in working with the range of disease populations and severity levels, especially given the time-constrained setting in the neurology clinic. Although Neurobooth is accessible for individuals who are wheelchair bound and would prefer to remain in their wheelchair for the session, the quality of eye tracking is sometimes reduced due to lack of adequate head support with some wheelchairs. In individuals with nystagmus or gaze holding difficulties, calibration and validation of the eye tracker can be time-consuming and occasionally unsuccessful after multiple attempts. Modifications such as inserting a cushioned headrest and adjusting the height of the computer monitor have improved eye tracking quality. The Mbient wearable inertial sensors that stream data continuously over Bluetooth to support synchronization, could at times lose connectivity resulting in lost data. Changing the location of Bluetooth receivers and automating wearable device resets prior to motor tasks improved data capture. Individuals with cognitive impairments in some cases had difficulty understanding and following task instructions, requiring repetition of the instructions or skipped tasks, and lengthening the duration of the session. In particular, the Multiple Object Tracking (MOT) task could take substantially longer for individuals with cognitive deficits to complete. An interim software update was made to automatically terminate the task after the easier trials if the task was taking extended time. Future studies focused on a reduced range of neurological populations could benefit from modifying the task set to optimize for feasibility and phenotypic capture.

In summary, Neurobooth is a flexible research platform to collect multimodal quantitative behavioral data, with the goal to advance computational investigations of human behavior. The open source platform aims to reduce barriers for collecting standardized behavioral data and allow large aggregate datasets to emerge over time. These large multimodal datasets can then fuel discovery of new disease characteristics and result in a deeper understanding for how brain diseases give rise to changes in behavior. Ongoing work aims to increase the scale of the platform by enabling the system to run the same behavioral tasks using a subset of devices connected to a single laptop computer. An additional avenue of active research is the creation of data visualizations and reports that can summarize the complex multimodal digital data and are useful for patients and clinical providers.

Methods

Neurobooth Hardware

Given space constraints (and in order to maintain a quiet and private environment), MGH Neurobooth was built within a physical booth designed to be a small meeting space for 2–4 people (Framery Q model, Framery Oy., Finland). The frame has acoustic insulation and ventilation, antimicrobial surfaces, and is wheelchair-accessible. An 80/20 aluminum rail structure was installed inside the booth to mount an adjustable computer monitor and a broad array of sensors. A sprinkler was installed in the booth ceiling to comply with hospital safety regulations and a privacy film was applied to the glass walls facing the clinic. Additional light panels were installed inside the booth to increase the luminance of recorded videos and an intercom system was installed to facilitate communication between the participant inside the booth and research staff outside of the booth. A chair with a reclining back and adjustable headrest was added to improve participant comfort and eye tracking data quality. However, the chair can be removed for those who prefer to complete their session in their wheelchair. Fig. 1 illustrates the complete physical layout of the Neurobooth.

The devices included in MGH Neurobooth support data capture across several behavioral domains (Table 2). Participants interacted with a button board to advance through task instructions and to provide responses during cognitive tasks. The computer mouse was used to capture arm movement and button clicks for a subset of tasks (i.e., MOT, Hevelius). Participant motion was captured using five wearable Mbient inertial sensors placed on the wrists, ankles, and lumbar region streaming triaxial acceleration and angular velocity. Three Intel RealSense D455 RGB-Depth cameras were located on the ceiling of the booth forming a triangle to enable estimation of body landmark positions in 3D space. High-quality eye tracking data were collected using an EyeLink Portable Duo placed below the monitor. An industrial-grade, high-speed FLIR camera was positioned above the monitor to support capture of fast facial movements. The back-facing camera of an iPhone model 12 running a custom data capture application was positioned below the eye-tracker to support capture of facial movements and eye movements[40,41]. Both a high-quality Blue Yeti microphone and the iPhone captured audio (i.e., participant speech).

Neurobooth Software

A custom Python-based library, which we call the Neurobooth Operating System (NeuroboothOS), coordinates behavioral stimulus delivery and data collection. NeuroboothOS provides a graphical user interface (GUI) for the operator to initiate and terminate data collection. Delivery of task instructions and stimuli to the participant is accomplished using PsychoPy[77], an open source Python toolbox. This design enables researchers to create and deploy behavioral tasks with relative ease. Tasks and their associated devices are organized in collections and studies.Task variables along with device and sensor parameters are defined in a set of YAML configuration files so they can be readily adjusted. A script is provided to verify that the set of files defining a study is complete and consistent. Data collected from each device is synchronized using the lab streaming layer (LSL) library[78]. Logs for each session, task, and data file are generated in a PostgreSQL database. The same database contains participants'

clinical, demographic, and patient-reported information automatically ingested from REDCap[79], an electronic data capture (EDC) system. An interactive web-based tool called the Neurobooth Explorer allows time series data to be visualized in real time after the session and can export visualizations. The time-series data figures in this manuscript were exported from the Neurobooth Explorer. All software is available on the neurobooth GitHub repository, https://github.com/neurobooth/.

Contextual Clinical Outcomes

Clinical rating scales were performed by clinical providers as part of the routine clinical appointment. For participants with ALS, the ALS Functional Rating Scale-Revised (ALSFRS-R)[80] was performed. For participants with Alzheimer's disease, mild cognitive impairment and other memory disorders, the Montreal Cognitive Assessment (MOCA) was performed[81]. For participants with a form of parkinsonism, the MDS-Unified Parkinson's Disease Rating Scale (MDS-UPDRS)[82] was performed. For participants with ataxia, the Brief Ataxia Rating Scale (BARS)[83], the Scale for Assessment and Rating of Ataxia (SARA)[84], and components of the Modified International Cooperative Ataxia Rating Scale (MICARS)[83] (arm, speech, and oculomotor sections) were performed.

We reviewed 42 patient-reported outcome measures (PROMS) used in neurodegenerative populations and selected a subset of 19 PROMS capturing information across the following domains: motor, speech, oculomotor, cognition, mood, participation in social activities, fatigue, sleep, quality of life (QoL), and environmental factors (Table 3). For the Neuro-QoL PROMS[85], the short forms were chosen wherever possible to help minimize survey burden. PROMS available in the public domain were prioritized. In addition to existing PROMs, our team also created a survey assessing the frequency of falls over the past 3, 6, and 12 months. Participants with Parkinson's Disease and ataxias completed disease-specific outcome measures, such as PROM-Ataxia[86] and the Dysarthria Impact Scale[87].

TABLE 3 ABOUT HERE

Behavioral Tasks

Cognitive Tasks

The cognitive tasks were built in Python based upon the validated structure from TestMyBrain.org[63]. For the Digit Symbol Substitution Test (DSST)[88], participants were shown six symbols arranged into three columns (Fig. 2). The columns were labeled "1", "2", and "3" such that each number was associated with two symbols. For each trial of DSST, a target symbol was shown to participants. Participants then attempted to choose the corresponding number using the button board. A new trial was immediately presented to participants once a button was pressed, with no indication of whether the selection was correct. Participants were given one minute to complete as many trials as possible. DSST assesses a large variety of cognitive domains including executive and visuoperceptual function, motor speed, attention, and associative learning.

For Multiple Object Tracking (MOT)[89], participants were shown ten black dots (Fig. 2). Three to five of these dots were targets and the rest were distractors. The color of the target dots alternated between green and black for three seconds before reverting to black. All dots then simultaneously moved for five seconds. Once the dots were stationary, participants attempted to click the targets using the computer mouse in any order. Correct clicks were displayed as green and incorrect clicks were displayed as red. If the participant did not complete their clicks within one minute, then the trial was restarted. The full MOT task consisted of three blocks of six trials each (for a total of 18 trials). Each block corresponded to the number of targets (starting with three and ending with five) and successive trials within a block linearly increased the movement speed of the dots during the animation phase. (Dots moved at approximately 40 *px/s* in the first trial and 240 *px/s* in the last trial of each chunk.) The MOT task was significantly modified in May 2024, after the two-year period detailed in this manuscript. These modifications 1) ensured a more consistent stimulus delivery across sessions, 2) improved accessibility (i.e., larger fonts and color-blind-friendly colors), and 3) implemented early-stopping criteria for study participants who had notable difficulty performing the task based on long completion times in early trials. MOT assesses visual attention[89,90].

Oculomotor Tasks

Prior to the oculomotor tasks, a five-point calibration task was performed to align the participant's gaze with the screen coordinate system. At the beginning of each oculomotor task, participants were asked to maintain their head position against the back of the chair and to keep their head still and only move their eyes.

In the horizontal smooth pursuit task, participants tracked a small circular target as it moved horizontally across the screen for 30 seconds. The circle's velocity varied in a sinusoidal manner, with zero velocity near the screen edges and maximal velocity (33.3 degrees/s) at the screen center. The task was included to assess the gain and smoothness of eye movements while tracking a slowly moving target stimulus, which can be impaired in neurological conditions[91–94].

For the horizontal saccades task, participants tracked a circular target as it instantly appeared at the left, center, and right of the screen. The vertical saccades task was similar, with the target appearing at the bottom, center, and top of the screen. The circle was stationary between transitions for a short, randomized duration (uniformly distributed between 1–1.5 s). The horizontal and vertical saccades tasks assess multiple brain areas involved in the initiation and control of ballistic eye movements[95].

In the fixation without a target task, participants stared at the center of a uniformly gray screen for 10 seconds. In the gaze-holding task, participants fixated on a target that moved between five different screen positions (i.e., center, down, up, right, then left). The circle was stationary in each position for 10 seconds before jumping to the next position. The fixation and gaze holding tasks were included to capture oculomotor signs such as nystagmus, saccadic intrusions, and drift[96].

Motor Tasks

For the finger-to-nose task, participants were asked to extend their arm to about 90% of their reach to touch a circular target on the monitor and then return to their nose. This action was repeated for 15 seconds for each arm. The finger-to-nose task assesses upper extremity coordination, dysmetria, speed of movement, and tremor[97].

In the repetitive foot tapping task, participants were instructed to repetitively tap their toes on the ground while maintaining heel contact with the ground. This action was performed for 10 seconds for each foot. For the rapid alternating hand movement task, participants were asked to raise their arm straight out in front of them and repetitively pronate and supinate their arm. This action was repeated for 15 seconds for each arm. These two tasks assess limb coordination and presence of bradykinesia[98].

For the sit-to-stand task, participants were asked to stand up fully from a seated position and then sit back down as quickly as possible with their arms crossed across their chest. Participants were able to use the chair's armrests if needed. The action was repeated 5 times. The sit-to-stand task assesses balance and lower extremity strength[99].

In the Hevelius mouse clicking task[58], participants were shown a single red dot on the screen and asked to click it using the mouse as quickly and as accurately as possible. After being clicked, the dot instantly moved to a new position on the screen. The task was divided into eight blocks, with each block presenting a sequence of nine dots. The size of the dot and the distance it moved on the screen remained the same within a block, but varied across blocks. It was previously observed that mouse trajectory, speed, and click characteristics were informative in quantifying ataxia and parkinsonism[58–60].

Speech Tasks

For sustained phonation, participants were asked to say "ahh" for 10 seconds. Since participants were instructed to continuously produce sound, this task was intended to measure variations in frequency and amplitude, breath patterns, and vocal fluctuations during initiation and termination of the target vowel[100]. For the diadochokinesis task, participants were asked to repeat "go go go", "la la la", "me me me", and "Pa-Ta-Ka" syllables as quickly and as many times as possible in 10 seconds. Prior to the sustained phonation and diadochokinesis tasks, participants were asked to take a deep breath and to continue the task until told to stop. These tasks are designed to capture speaking rate and vocal irregularities that may be present in many neurologic conditions[101].

For the passage reading task, participants read aloud the Bamboo passage[102,103], which is designed to include voiced consonants and phrase boundaries. Participants were asked to speak at their normal conversational volume and pace when reading the passage. Audio collected during the task can be used to assess speaking rate, phrase duration, and pauses[104]. Eye tracking allows assessment of eye movements during a naturalistic task and for assessment of eye-voice coordination[74].

Data Availability

Data can be requested by qualified researchers by visiting https://neurobooth.mgh.harvard.edu/.

Software Availability

Neurobooth software can be downloaded from our Github page: https://github.com/neurobooth/.

Funding

This work was supported by the Massachusetts Life Sciences Center, NIH R01 NS117826, Biogen, Broad Institute, Dake Family Foundation, and Massachusetts General Hospital Department of Neurology.

References

1 Feigin VL, Abajobir AA, Abate KH, Abd-Allah F, Abdulle AM, Abera SF *et al.* Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol* 2017; **16**: 877–897.

2 Cavazzoni P. The Path Forward: Advancing Treatments and Cures for Neurodegenerative Diseases. https://www.fda.gov/news-events/congressional-testimony/path-forward-advancing-treatments-and-cures-neurodegenerative-diseases-07292021#footnote1_sde0aq9. 2021.

3 Brooker SM, Edamakanti CR, Akasha SM, Kuo S-H, Opal P. Spinocerebellar ataxia clinical trials: opportunities and challenges. *Ann Clin Transl Neurol* 2021; **8**: 1543–1556.

4 Maas RPPWM, van de Warrenburg BPC. Exploring the clinical meaningfulness of the Scale for the Assessment and Rating of Ataxia: A comparison of patient and physician perspectives at the item level. *Parkinsonism Relat Disord* 2021; **91**: 37–41.

5 Petit E, Schmitz-Hübsch T, Coarelli G, Jacobi H, Heinzmann A, Figueroa KP *et al.* SARA captures disparate progression and responsiveness in spinocerebellar ataxias. *J Neurol* 2024; **271**: 3743–3753.

6 Gupta AS. Digital Phenotyping in Clinical Neurology. *Semin Neurol* 2022; **42**: 48–59.

7 Aminoff MJ. The future of the neurologic examination. *JAMA Neurol* 2017; **74**: 1291–1292.

8 Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. *Nat Biotechnol* 2015; **33**: 462–463.

9 Crutcher MD, Calhoun-Haney R, Manzanares CM, Lah JJ, Levey AI, Zola SM. Eye tracking during a visual paired comparison task as a predictor of early dementia. *Am J Alzheimers Dis Other Demen* 2009; **24**: 258–266.

10 Antoniades CA, Kennard C. Ocular motor abnormalities in neurodegenerative disorders.

Eye 2015; **29**: 200–207.

11 Sekar A, Panouillères MTN, Kaski D. Detecting Abnormal Eye Movements in Patients with Neurodegenerative Diseases - Current Insights. *Eye Brain* 2024; **16**: 3–16.

12 Anderson TJ, MacAskill MR. Eye movements in patients with neurodegenerative disorders. *Nat Rev Neurol* 2013; **9**: 74–85.

13 Garces P, Antoniades CA, Sobanska A, Kovacs N, Ying SH, Gupta AS *et al.* Quantitative oculomotor assessment in hereditary ataxia: Systematic review and consensus by the ataxia global initiative working group on digital-motor biomarkers. *Cerebellum* 2024; **23**: 896–911.

14 Chambers JM, Prescott TJ. Response times for visually guided saccades in persons with Parkinson's disease: a meta-analytic review. *Neuropsychologia* 2010; **48**: 887–899.

15 Li H, Zhang X, Yang Y, Xie A. Abnormal eye movements in Parkinson's disease: From experimental study to clinical application. *Parkinsonism Relat Disord* 2023; **115**: 105791.

16 Liu Z, Yang Z, Gu Y, Liu H, Wang P. The effectiveness of eye tracking in the diagnosis of cognitive disorders: A systematic review and meta-analysis. *PLoS One* 2021; **16**: e0254059.

17 Schalling E, Hartelius L. Speech in spinocerebellar ataxia. *Brain Lang* 2013; **127**: 317–322.

18 Boschi V, Catricalà E, Consonni M, Chesi C, Moro A, Cappa SF. Connected Speech in Neurodegenerative Language Disorders: A Review. *Front Psychol* 2017; **8**: 269.

19 Ngo QC, Motin MA, Pah ND, Drotár P, Kempster P, Kumar D. Computerized Analysis of Speech and Voice for Parkinson's Disease: A Systematic Review. *Comput Methods Programs Biomed* 2022; : 107133.

20 Vogel AP, Sobanska A, Gupta A, Vasco G, Grobe-Einsler M, Summa S *et al.* Quantitative Speech Assessment in Ataxia-Consensus Recommendations by the Ataxia Global Initiative Working Group on Digital-Motor Markers. *Cerebellum* 2023. doi:10.1007/s12311-023-01623-4.

21 Moro-Velazquez L, Gomez-Garcia JA, Arias-Londoño JD, Dehak N, Godino-Llorente JI. Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. *Biomed Signal Process Control* 2021; **66**: 102418.

22 Frost JD Jr. Triaxial vector accelerometry: a method for quantifying tremor and ataxia. *IEEE Trans Biomed Eng* 1978; **25**: 17–27.

23 Espay AJ, Bonato P, Nahab FB, Maetzler W, Dean JM, Klucken J *et al.* Technology in Parkinson's disease: Challenges and opportunities. *Mov Disord* 2016; **31**: 1272–1282.

24 Buckley E, Mazzà C, McNeill A. A systematic review of the gait characteristics associated with Cerebellar Ataxia. *Gait Posture* 2018; **60**: 154–163.

25 Vanmechelen I, Haberfehlner H, De Vleeschhauwer J, Van Wonterghem E, Feys H, Desloovere K *et al.* Assessment of movement disorders using wearable sensors during upper limb tasks: A scoping review. *Front Robot AI* 2022; **9**: 1068413.

26 Power L, Pathirana P, Horne M, Milne S, Marriott A, Szmulewicz DJ. Instrumented Objective Clinical Examination of Cerebellar Ataxia: the Upper and Lower Limb-a Review. *Cerebellum* 2022; **21**: 145–158.

27 Paredes-Acuna N, Utpadel-Fischler D, Ding K, Thakor NV, Cheng G. Upper limb intention tremor assessment: opportunities and challenges in wearable technology. *J Neuroeng Rehabil* 2024; **21**: 8.

28 Vienne-Jumeau A, Quijoux F, Vidal P-P, Ricard D. Wearable inertial sensors provide reliable biomarkers of disease severity in multiple sclerosis: A systematic review and meta-analysis. *Ann Phys Rehabil Med* 2020; **63**: 138–147.

29 Kadirvelu B, Gavriel C, Nageshwaran S, Chan JPK, Nethisinghe S, Athanasopoulos S *et al.* A wearable motion capture suit and machine learning predict disease progression in Friedreich's ataxia. *Nat Med* 2023; **29**: 86–94.

30 de los Reyes-Guzmán A, Dimbwadyo-Terrer I, Trincado-Alonso F, Monasterio-Huelin F, Torricelli D, Gil-Agudo A. Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review. *Clin Biomech* 2014; **29**: 719–727.

31 Klucken J, Barth J, Kugler P, Schlachetzki J, Henze T, Marxreiter F *et al.* Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease. *PLoS One* 2013; **8**: e56956.

32 Delrobaei M, Memar S, Pieterman M, Stratton TW, McIsaac K, Jog M. Towards remote monitoring of Parkinson's disease tremor using wearable motion capture systems. *J Neurol Sci* 2018; **384**: 38–45.

33 Gorges M, Pinkhardt EH, Kassubek J. Alterations of eye movement control in neurodegenerative movement disorders. *J Ophthalmol* 2014; **2014**: 658243.

34 Haque RU, Pongos AL, Manzanares CM, Lah JJ, Levey AI, Clifford GD. Deep Convolutional Neural Networks and Transfer Learning for Measuring Cognitive Impairment Using Eye-Tracking in a Distributed Tablet-Based Environment. *IEEE Trans Biomed Eng* 2021; **68**: 11–18.

35 Pavisic IM, Pertzov Y, Nicholas JM, O'Connor A, Lu K, Yong KXX *et al.* Eye-tracking indices of impaired encoding of visual short-term memory in familial Alzheimer's disease. *Sci Rep* 2021; **11**: 8696.

36 Fleyeh H, Westin J. Extracting Body Landmarks from Videos for Parkinson Gait Analysis. In: *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. ieeexplore.ieee.org, 2019, pp 379–384.

37 Nachum R, Jackson K, Duric Z, Gerber L. A Novel Computer Vision Approach to Kinematic Analysis of Handwriting with Implications for Assessing Neurodegenerative Diseases. *Conf Proc IEEE Eng Med Biol Soc* 2021; **2021**: 1309–1313.

38 Sapiro G, Hashemi J, Dawson G. Computer vision and behavioral phenotyping: an autism case study. *Current Opinion in Biomedical Engineering* 2019; **9**: 14–20.

39 Nunes AS, Kozhemiako N, Stephen CD, Schmahmann JD, Khan S, Gupta AS. Automatic

Classification and Severity Estimation of Ataxia From Finger Tapping Videos. *Front Neurol* 2021; **12**: 795258.

40 Chang Z, Chen Z, Stephen CD, Schmahmann JD, Wu HT, Sapiro G *et al.* Accurate detection of cerebellar smooth pursuit eye movement abnormalities via mobile phone video and machine learning. *Sci Rep* 2020; **10**: 1–10.

41 Azami H, Chang Z, Arnold SE, Sapiro G, Gupta AS. Detection of Oculomotor Dysmetria From Mobile Phone Video of the Horizontal Saccades Task Using Signal Processing and Machine Learning Approaches. *IEEE Access* 2022; **10**: 34022–34031.

42 Balakrishnan A, Medikonda J, Namboothiri PK, Manik, Natarajan A. Role of Wearable Sensors with Machine Learning Approaches in Gait Analysis for Parkinson's Disease Assessment: A Review. *Engineered Science* 2022; **19 (September 2022)**: 5–19.

43 Bo F, Yerebakan M, Dai Y, Wang W, Li J, Hu B *et al.* IMU-Based Monitoring for Assistive Diagnosis and Management of IoHT: A Review. *Healthcare (Basel)* 2022; **10**. doi:10.3390/healthcare10071210.

44 Phan D, Nguyen N, Pathirana PN, Horne M, Power L, Szmulewicz D. A Random Forest Approach for Quantifying Gait Ataxia With Truncal and Peripheral Measurements Using Multiple Wearable Sensors. *IEEE Sens J* 2020; **20**: 723–734.

45 Oubre B, Daneault J-F, Whritenour K, Khan NC, Stephen CD, Schmahmann JD *et al.* Decomposition of Reaching Movements Enables Detection and Measurement of Ataxia. *Cerebellum* 2021. doi:10.1007/s12311-021-01247-6.

46 Lee J, Oubre B, Daneault J-F, Stephen CD, Schmahmann JD, Gupta AS *et al.* Analysis of Gait Sub-Movements to Estimate Ataxia Severity using Ankle Inertial Data. *IEEE Trans Biomed Eng* 2022; **PP**. doi:10.1109/TBME.2022.3142504.

47 Tsanas A, Little MA, McSharry PE, Spielman J, Ramig L. Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Trans Biomed Eng* 2012; **59**: 1264–1271.

48 Hlavnička J. *Automated Analysis of Speech Disorders in Neurodegenerative Diseases*. 2019.https://search.proquest.com/openview/f8423f1b44d58c0b1137a74c2b51da1f/1?pq-origsite=gscholar&cbl=2026366&diss=y.

49 Rusz J, Tykalová T, Salerno G, Bancone S, Scarpelli J, Pellecchia MT. Distinctive speech signature in cerebellar and parkinsonian subtypes of multiple system atrophy. *J Neurol* 2019; **266**: 1394–1404.

50 Takakura Y, Otsuki M, Sakai S, Tajima Y, Mito Y, Ogata A *et al.* Sub-classification of apraxia of speech in patients with cerebrovascular and neurodegenerative diseases. *Brain Cogn* 2019; **130**: 1–10.

51 Vogel AP, Magee M, Torres-Vega R, Medrano-Montero J, Cyngler MP, Kruse M *et al.* Features of speech and swallowing dysfunction in pre-ataxic spinocerebellar ataxia type 2. *Neurology* 2020; **95**: e194–e205.

52 Kashyap B, Pathirana PN, Horne M, Power L, Szmulewicz D. Quantitative Assessment of Speech in Cerebellar Ataxia Using Magnitude and Phase Based Cepstrum. *Ann Biomed*

Eng 2020; **48**: 1322–1336.

53 Song J, Lee JH, Choi J, Suh MK, Chung MJ, Kim YH *et al.* Detection and differentiation of ataxic and hypokinetic dysarthria in cerebellar ataxia and parkinsonian disorders via wave splitting and integrating neural networks. *PLoS One* 2022; **17**: e0268337.

54 Vieira FG, Venugopalan S, Premasiri AS, McNally M, Jansen A, McCloskey K *et al.* A machine-learning based objective measure for ALS disease severity. *NPJ Digit Med* 2022; **5**: 45.

55 Isaev DY, Vlasova RM, Di Martino JM, Stephen CD, Schmahmann JD, Sapiro G *et al.* Uncertainty of Vowel Predictions as a Digital Biomarker for Ataxic Dysarthria. *Cerebellum* 2023. doi:10.1007/s12311-023-01539-z.

56 Vattis K, Oubre B, Luddy AC, Ouillon, Jessey S, Eklund NM, Stephen CD, Schmahmann JD, Nunes AS, Gupta, AS. Sensitive quantification of cerebellar speech abnormalities using deep learning models. *IEEE Access* 2024; **12**: 62328–62340.

57 White RW, Murali Doraiswamy P, Horvitz E. Detecting neurodegenerative disorders from web search signals. npj Digital Medicine. 2018; **1**. doi:10.1038/s41746-018-0016-6.

58 Gajos KZ, Reinecke K, Donovan M, Stephen CD, Hung AY, Schmahmann JD *et al.* Computer Mouse Use Captures Ataxia and Parkinsonism, Enabling Accurate Measurement and Detection. *Mov Disord* 2020; **35**: 354–358.

59 Pandey V, Khan NC, Gupta AS, Gajos KZ. Accuracy and reliability of at-home quantification of motor impairments using a computer-based pointing task with children with ataxia-telangiectasia. *ACM Trans Access Comput* 2023; **16**: 1–25.

60 Eklund NM, Ouillon J, Pandey V, Stephen CD, Schmahmann JD, Edgerton J *et al.* Real-life ankle submovements and computer mouse use reflect patient-reported function in adult ataxias. *Brain Communications* 2023; : fcad064.

61 Dorsey ER, Venuto C, Venkataraman V, Harris DA, Kieburtz K. Novel methods and technologies for 21st-century clinical trials: a review. *JAMA Neurol* 2015; **72**: 582–588.

62 Artusi CA, Imbalzano G, Sturchio A, Pilotto A, Montanaro E, Padovani A *et al.* Implementation of Mobile Health Technologies in Clinical Trials of Movement Disorders: Underutilized Potential. *Neurotherapeutics* 2020; **17**: 1736–1746.

63 Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychon Bull Rev* 2012; **19**: 847–857.

64 Fehd HM, Seiffert AE. Looking at the center of the targets helps multiple object tracking. *J Vis* 2010; **10**: 19.1–13.

65 Norton DJ, Nguyen VA, Lewis MF, Reynolds GO, Somers DC, Cronin-Golomb A. Visuospatial Attention to Single and Multiple Objects Is Independently Impaired in Parkinson's Disease. *PLoS One* 2016; **11**: e0150013.

66 Quinn N. The 'round the houses' sign in progressive supranuclear palsy. *Ann Neurol* 1996; **40**: 951.

67 Kent RD, Kent JF, Rosenbek JC. Maximum performance tests of speech production. *J Speech Hear Disord* 1987; **52**: 367–387.

68 Fletcher SG. Time-by-count measurement of diadochokinetic syllable rate. *J Speech Hear Res* 1972; **15**: 763–770.

69 Brooke J. SUS-A Quick and Dirty Usability Scale. In: Jordan PW, Thomas B, Weerdmeester BA, Mcclelland IL (eds). *Usability Evaluation in Industry*. Taylor & Francis: London, 1996, pp 189–194.

70 Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S *et al.* System Usability Scale benchmarking for digital health apps: Meta-analysis. *JMIR MHealth UHealth* 2022; **10**: e37290.

71 Xu P, Zhu X, Clifton DA. Multimodal learning with Transformers: A survey. *IEEE Trans Pattern Anal Mach Intell* 2023; **45**: 12113–12132.

72 Venugopalan J, Tong L, Hassanzadeh HR, Wang MD. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci Rep* 2021; **11**: 3254.

73 Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review. *Brief Bioinform* 2022; **23**: bbab569.

74 Terao Y, Tokushige S-I, Inomata-Terada S, Miyazaki T, Kotsuki N, Fisicaro F *et al.* How do patients with Parkinson's disease and cerebellar ataxia read aloud? -Eye-voice coordination in text reading. *Front Neurosci* 2023; **17**: 1202404.

75 Perochon S, Di Martino JM, Carpenter KLH, Compton S, Davis N, Eichner B *et al.* Early detection of autism using digital behavioral phenotyping. *Nat Med* 2023; **29**: 2489–2497.

76 Krishnappa Babu PR, Di Martino JM, Aiello R, Eichner B, Espinosa S, Green J *et al.* Validation of a Mobile App for Remote Autism Screening in Toddlers. *NEJM AI* 2024; **1**: AIcs2400510.

77 Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H *et al.* PsychoPy2: Experiments in behavior made easy. *Behav Res Methods* 2019; **51**: 195–203.

78 Kothe C, Medine D, Boulay C, Grivich M, Stenner T. Lab Streaming Layer (LSL) - A software framework for synchronizing a large array of data collection and stimulation devices. 2014.https://github.com/sccn/labstreaminglayer/.

79 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; **42**: 377–381.

80 Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B *et al.* The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci* 1999; **169**: 13–21.

81 Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I *et al.* The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005; **53**: 695–699.

82 Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT *et al.* Movement

Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. *Mov Disord* 2007; **22**: 41–47.

83 Schmahmann JD, Gardner R, MacMore J, Vangel MG. Development of a brief ataxia rating scale (BARS) based on a modified form of the ICARS. *Mov Disord* 2009; **24**: 1820–1828.

84 Schmitz-Hübsch T, du Montcel ST, Baliko L, Berciano J, Boesch S, Depondt C *et al.* Scale for the assessment and rating of ataxia: development of a new clinical scale. *Neurology* 2006; **66**: 1717–1720.

85 Cella D, Lai J-S, Nowinski CJ, Victorson D, Peterman A, Miller D *et al.* Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. *Neurology* 2012; **78**: 1860–1867.

86 Schmahmann JD, Pierce S, MacMore J, L'Italien GJ. Development and Validation of a Patient-Reported Outcome Measure of Ataxia. *Mov Disord* 2021; **36**. doi:10.1002/mds.28670.

87 Vogel A, Graf L, Chan J, Hepworth G, Bade M, Synofzik M. Development and validation of the dysarthria impact scale. *International Congress on Ataxia Research* 2022.

88 Singh S, Strong RW, Jung L, Li FH, Grinspoon L, Scheuer LS *et al.* The TestMyBrain Digital Neuropsychology Toolkit: Development and psychometric characteristics. *J Clin Exp Neuropsychol* 2021; **43**: 786–795.

89 Pylyshyn ZW, Storm RW. Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat Vis* 1988; **3**: 179–197.

90 Jaeger J. Digit Symbol Substitution Test: The Case for Sensitivity Over Specificity in Neuropsychological Testing. *J Clin Psychopharmacol* 2018; **38**: 513–519.

91 White OB, Saint-Cyr JA, Tomlinson RD, Sharpe JA. Ocular motor deficits in Parkinson's disease. II. Control of the saccadic and smooth pursuit systems. *Brain* 1983; **106 (Pt 3)**: 571–587.

92 Das VE, Leigh RJ. Visual-vestibular interaction in progressive supranuclear palsy. *Vision Res* 2000; **40**: 2077–2081.

93 Zee DS, Yee RD, Cogan DG, Robinson DA, Engel WK. Ocular motor abnormalities in hereditary cerebellar ataxia. *Brain* 1976; **99**: 207–234.

94 Fletcher WA, Sharpe JA. Smooth pursuit dysfunction in Alzheimer's disease. *Neurology* 1988; **38**: 272–277.

95 John Leigh R, Zee DS. *The Neurology of Eye Movements*. Oxford University Press, 2015.

96 Thurtell MJ, Leigh RJ. Nystagmus and saccadic intrusions. *Handb Clin Neurol* 2011; **102**: 333–378.

97 Gagnon C, Mathieu J, Desrosiers J. Standardized finger-nose test validity for coordination assessment in an ataxic disorder. *Can J Neurol Sci* 2004; **31**: 484–489.

98 Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating

Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008; **23**: 2129–2170.

99 Csuka M, McCarty DJ. Simple method for measurement of lower extremity muscle strength. *Am J Med* 1985; **78**: 77–81.

100 Gerratt BR, Kreiman J, Garellek M. Comparing Measures of Voice Quality From Sustained Phonation and Continuous Speech. *J Speech Lang Hear Res* 2016; **59**: 994–1001.

101 Wang Y-T, Kent RD, Duffy JR, Thomas JE. Analysis of diadochokinesis in ataxic dysarthria using the motor speech profile program. *Folia Phoniatr Logop* 2009; **61**: 1–11.

102 Yunusova Y, Graham NL, Shellikeri S, Phuong K, Kulkarni M, Rochon E *et al.* Profiling Speech and Pausing in Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD). *PLoS One* 2016; **11**: e0147573.

103 Wang Y-T, Green JR, Nip ISB, Kent RD, Kent JF. Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatr Logop* 2010; **62**: 297–302.

104 Barnett C, Green JR, Marzouqah R, Stipancic KL, Berry JD, Korngut L *et al.* Reliability and validity of speech & pause measures during passage reading in ALS. *Amyotroph Lateral Scler Frontotemporal Degener* 2020; **21**: 42–50.

105 Baylor C, Yorkston K, Eadie T, Kim J, Chung H, Amtmann D. The Communicative Participation Item Bank (CPIB): item bank calibration and development of a disorder-generic short form. *J Speech Lang Hear Res* 2013; **56**: 1190–1208.

106 Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971; **9**: 97–113.

107 Whiteneck GG, Harrison-Felix CL, Mellick DC, Brooks CA, Charlifue SB, Gerhart KA. Quantifying environmental factors: a measure of physical, attitudinal, service, productivity, and policy barriers. *Arch Phys Med Rehabil* 2004; **85**: 1324–1335.

108 Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B *et al.* The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007; **45**: S3–S11.

109 Sloane M, Ball K, Owsley C, Bruni J, Roenker D. The visual activities questionnaire: Developing an instrument for assessing problems in everyday visual tasks. *Noninvasive Assessment of the Visual System* 1992. doi:10.1364/navs.1992.sub4.

110 Guy W. *ECDEU assessment manual for psychopharmacology*. U.S. Department of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs, 1976.

Figures

Figure 1. The Neurobooth, located in the MGH neurology clinic.

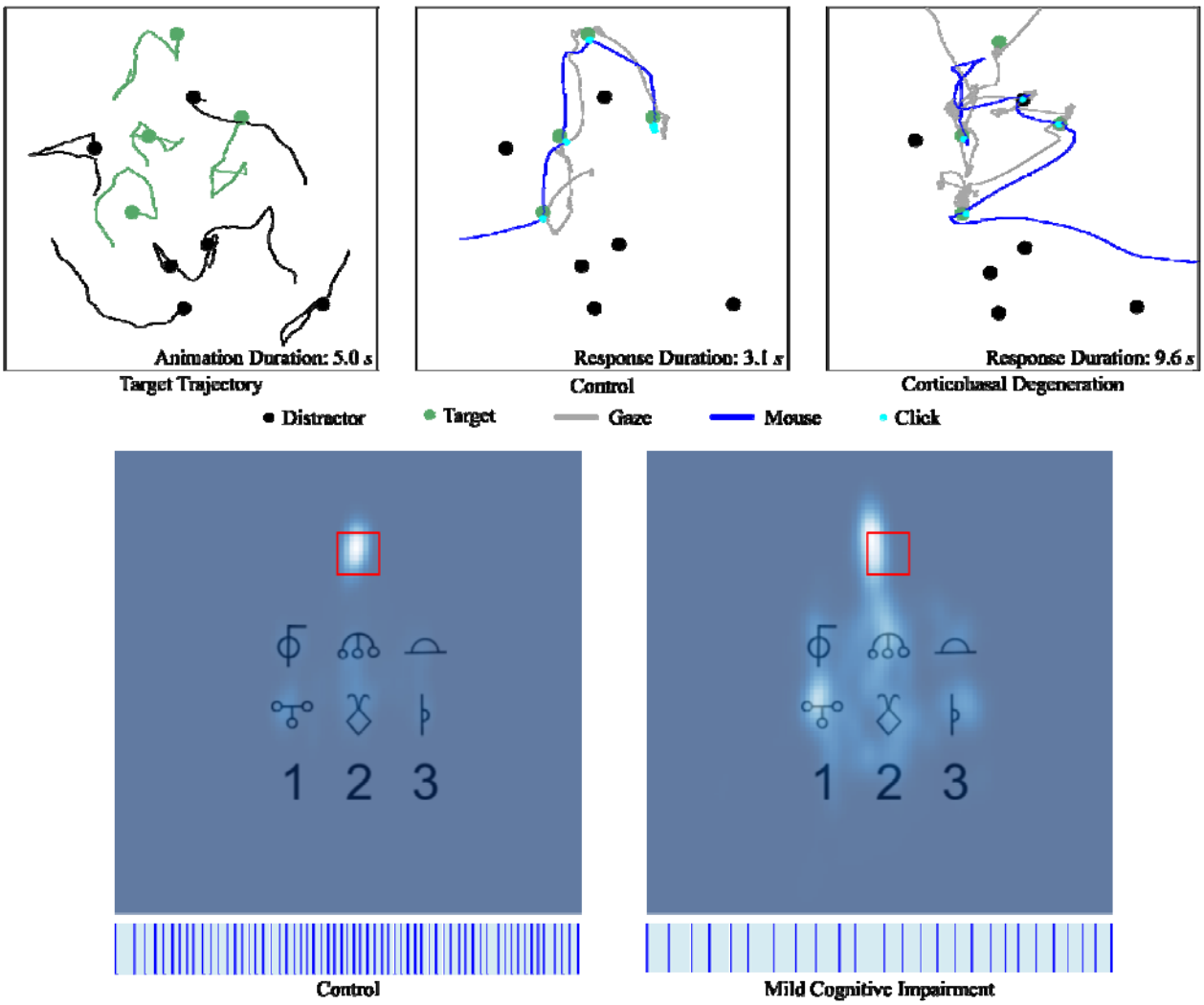

Figure 2. Illustrations of Multiple Object Tracking (MOT) and the Digital Symbol Substitution Test (DSST). (**top**) One trial of MOT with four targets and six distractors. The left panel displays the trajectory of each circle, with the dot indicating the final position. The middle and right panels display gaze and mouse trajectories from a control and participant with corticobasal syndrome, respectively. (**bottom**) Heat maps showing the relative gaze duration during DSST. The red square indicates the location of the target symbol. Lighter colors represent a longer time looking at a given region. The left panel shows a control and the right panel shows a participant with mild cognitive impairment. The vertical blue lines at the bottom represent the presentation of symbols during the one minute task duration, with shaded blue regions indicating the response time before the button press.

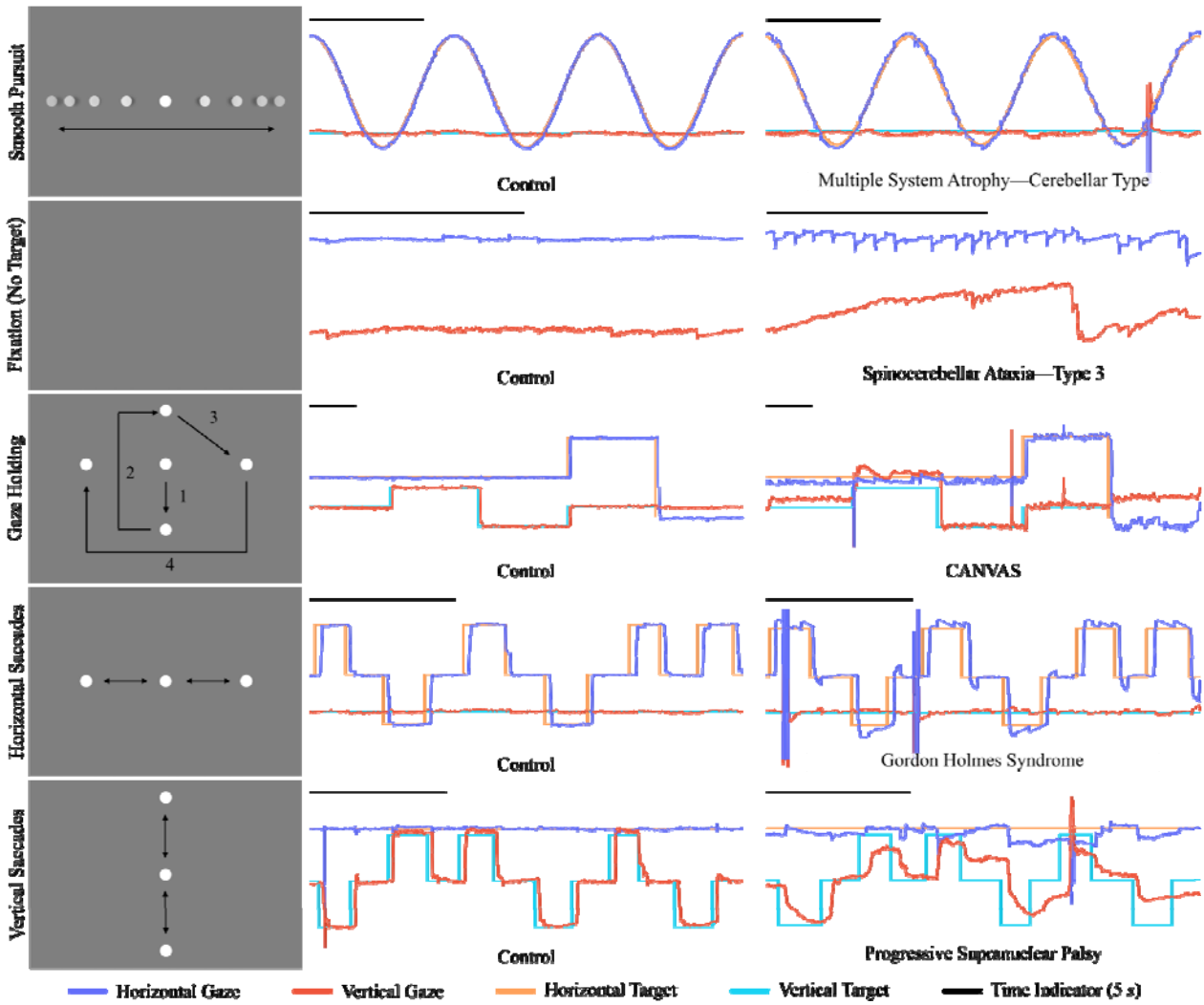

Figure 3. Illustrations of oculomotor tasks accompanied by example gaze time-series. Arrows and numbers indicate the motion of the visual target and are not shown to participants. Only a single visual target is displayed throughout each task. All control time-series were obtained from a single data collection session (i.e., from a single participant). Horizontal and vertical gaze positions are plotted on the y-axis with time represented on the x-axis. The y-axis values representing horizontal gaze increase with eye movement to the far right of the screen, whereas y-axis values representing vertical gaze increase with eye movement to the bottom of the screen.

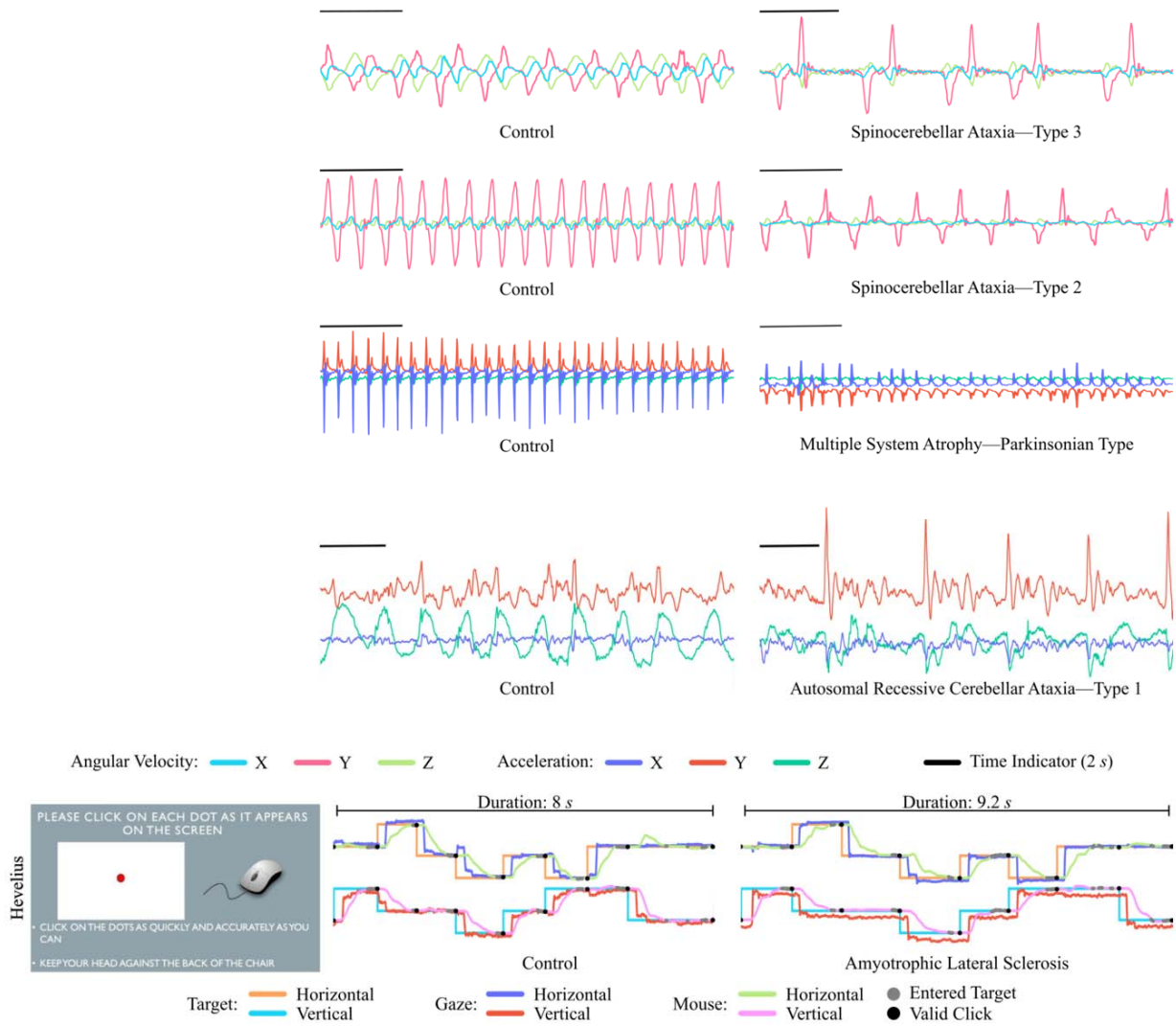

Figure 4. Illustrations of motor tasks. The finger-to-nose, alternating hand, foot tapping, and sit-to-stand tasks are accompanied by example triaxial inertial time-series. For the foot tapping task, the ankle sensor was oriented differently for the two displayed time-series. For the Hevelius computer mouse task, the example data are represented in screen coordinates. Horizontal screen coordinates increase from left to right. Vertical screen coordinates increase from top to bottom. The target moved immediately after a valid mouse click. It was possible for the mouse to enter the target multiple times prior to a valid click.

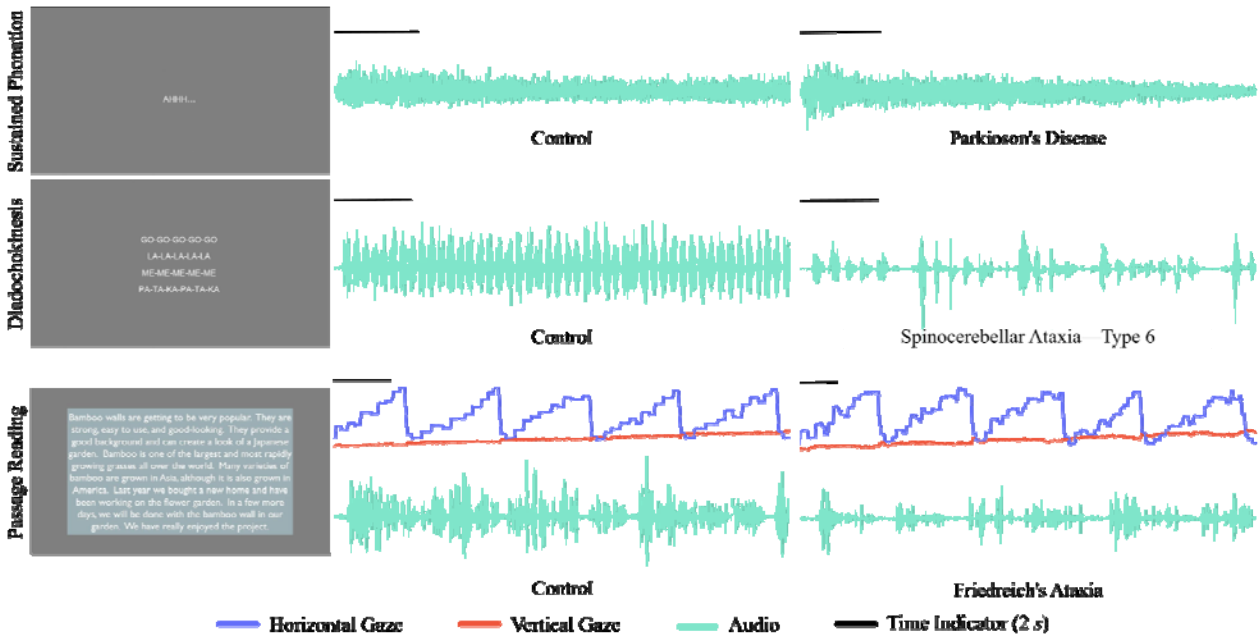

Figure 5. Illustrations of the speech tasks - sustained phonation (top), diadochokinesis tasks (middle), and passage reading (bottom). For diadochokinesis tasks, only a single line of centered text is displayed per task. The illustrated diadochokinesis data are from the “Go-Go-Go” task. For passage reading, the y-axis values representing horizontal gaze (blue curve) increase with eye movement to the far right of the screen, whereas y-axis values representing vertical gaze (red curve) increase with eye movement to the bottom of the screen.

Tables

Table 1. Participant demographics and clinical diagnoses.

Abbreviations: ARSACS—Autosomal recessive spastic ataxia of Charlevoix-Saguenay; CANVAS—Cerebellar Ataxia, Neuropathy, and Vestibular Areflexia Syndrome

Participant diagnosis	*N*	Age in decades	Female %
Control	82	2nd to 8th	70
Alzheimer's disease	5	5th to 8th	60
Amyotrophic lateral sclerosis	4	5th to 7th	25
ARSACS	3	1st to 2nd	100
Atypical parkinsonism	9	5th to 8th	56
CANVAS	11	3rd to 7th	45
Corticobasal syndrome	4	7th	75
Fragile X-associated tremor/ataxia syndrome	4	5th to 7th	50
Friedreich's ataxia	14	2nd to 7th	79
Frontotemporal dementia	2	7th	100
Hereditary spastic paraplegias	10	4th to 7th	50
Huntington's disease	1	6th	0
Lewy body dementia	3	6th to 8th	100
Mild cognitive impairment	5	3rd to 8th	80
Multiple system atrophy - cerebellar type	8	6th to 7th	88
Multiple system atrophy - parkinsonian type	4	5th to 7th	25
Other neurological diagnosis	36	3rd to 9th	53
Other ataxia	61	2nd to 8th	34
Parkinson's disease	132	4th to 8th	33
Primary progressive aphasia	7	5th to 7th	29
Progressive supranuclear palsy	14	6th to 8th	79
Spinocerebellar ataxia 1	2	3rd to 6th	100
Spinocerebellar ataxia 2	8	2nd to 7th	63
Spinocerebellar ataxia 3	24	2nd to 7th	71
Spinocerebellar ataxia 6[†]	13	5th to 8th	62
Spinocerebellar ataxia 8	3	3rd to 5th	67
Spinocerebellar ataxia 27B	1	5th	0

[†]One subject with spinocerebellar ataxia 6 was concurrently diagnosed with spinocerebellar ataxia 8.

Table 2. Data collection devices in the initial Neurobooth deployment. Data were captured from all devices during most tasks, with exception of computer mouse data, which were only captured during tasks that used the mouse as an input device. Data were not collected from the eye tracker or iPhone during the motor tasks because the arm holding the eye tracker and iPhone was repositioned during these tasks to create room for task performance.

Device name	Data Modality	Sampling rate (Hz)	No. Devices
Intel D455	RGB-Depth Video	60	3
FLIR Camera	RGB (Infrared)	196	1
EyeLink Portable Duo	Binocular Gaze	1000	1
Mbient IMU	6-axis Inertial	100	5
iPhone 12	RGB Video, Audio	150–240*	1
Microphone Blue Yeti	Audio	44,100	1
Computer Mouse	Mouse movement and clicks	—	1
Button Pad	User Input	—	1

* The optimal iPhone sampling rate is 240 Hz. Though some videos achieve this rate, lower rates were observed in practice. The iPhone application has been updated multiple times to reduce dropped frames.

Table 3. Descriptions of each collected patient reported outcome measure (PROM).

PROM	Items	Brief Description
PROM-Ataxia*	70	Assesses physical function, activities of daily living, and mental function in individuals with cerebellar ataxia [86].
Dysarthria Impact Scale	23	Assesses how speech affects everyday life [87]
Falls Questionnaire	8	Assesses the number of falls and near falls over the past 3, 6 and 12 months (Created for Neurobooth)
Communicative Participation Item Bank- General	10	Assesses how much participant's condition affects how they communicate on an average day [105]
Edinburgh Handedness Inventory	15	Asks whether the right or left hand is used for various everyday tasks [106]
Craig Hospital Inventory of Environmental Factors	10	Assesses how various environmental factors act as barriers to participation as an active member of society [107]
PROMIS 10	10	Assesses physical, mental and social health [107,108]
Visual Activities Questionnaire	13	Assesses vision-related quality of life and function in everyday activities [109]
Patient Global Impressions scale	1	Assesses patient impression of overall clinical change since their last participation [110]
Neuro-QoL (Quality of Life in Neurological Disorders)		
Upper Extremity Function (Fine Motor, ADL)	8	Assesses the difficulty of performing fine motor tasks [85]
Lower Extremity Function (Mobility)	8	Assesses mobility difficulties [85]
Cognition Function	8	Assesses the difficulty of thinking, reading, concentrating and planning [85]
Ability to Participate in Social Roles and Activities	8	Assesses how often participants kept up with social responsibilities over the past week [85]
Anxiety	8	Assesses how often participants felt symptoms of anxiety over the past week [85]
Depression	8	Assesses how often participants felt symptoms of depression over the

		past week [85]
Emotional and Behavioral Dyscontrol	8	Assesses how often participants felt unable to control their emotions over the past week [85]
Positive Affect and Well-Being	9	Assesses how often participants felt that their life was satisfying and hopeful “lately” [85]
Fatigue	8	Assesses how often participants felt tired completing a range of activities over the past week [85]
Sleep Disturbance	8	Assesses how often participants had trouble getting adequate sleep over the past week [85]

*****PROM-Ataxia was collected in controls and individuals who are seen by a movement disorders and/or ataxia clinical provider.