



PONT: A Protocol for Online Neuropsychological Testing

William Saban and Richard B. Ivry

Abstract

■ A major challenge for neuropsychological research arises from the fact that we are dealing with a limited resource: the patients. Not only is it difficult to identify and recruit these individuals, but their ability to participate in research projects can be limited by their medical condition. As such, sample sizes are small, and considerable time (e.g., 2 years) is required to complete a study. To address limitations inherent to laboratory-based neuropsychological research, we developed a protocol for online neuropsychological testing (PONT). We describe the implementation of PONT and provide the required information and materials for recruiting participants, conducting remote neurological evaluations, and testing patients in an automated, self-administered manner. The protocol can be easily tailored to target a broad range of patient groups, especially those who

can be contacted via support groups or multisite collaborations. To highlight the operation of PONT and describe some of the unique challenges that arise in online neuropsychological research, we summarize our experience using PONT in a research program involving individuals with Parkinson's disease and spinocerebellar ataxia. In a 10-month period, by contacting 646 support group coordinators, we were able to assemble a participant pool with over 100 patients in each group from across the United States. Moreover, we completed six experiments ($n > 300$) exploring their performance on a range of tasks examining motor and cognitive abilities. The efficiency of PONT in terms of data collection, combined with the convenience it offers the participants, promises a new approach that can increase the impact of neuropsychological research. ■

INTRODUCTION

Empirical investigations of patients with brain pathology have been central in the emergence of behavioral neurology as an area of specialization and, with its more theoretical focus, cognitive neuroscience. Information obtained through neuropsychological testing has been fundamental in advancing our understanding of the nature of brain-behavior relations (Luria, 1966) with important clinical implications (Lezak, 2000). Neuropsychological testing is not only important in evaluating the cognitive changes associated with focal and degenerative disorders of specific brain regions but has also been an essential tool in determining the contribution of targeted brain regions and particular cognitive operations. As one historical example, neuropsychological observations from the 19th century led to the classic “Broca–Wernicke–Lichtheim–Geschwind” model of language function (Poeppel, 2014; Geschwind, 1970).

Understanding the functional organization of the brain requires the use of converging methods of exploration. The spatial resolution of fMRI and the temporal resolution of EEG can reveal the dynamics of computations in local regions as well as the coordination of activity across neural networks. A limitation with these methods is that they are fundamentally correlational, revealing the relationship between cognitive events and brain regions. In contrast,

perturbation methods such as optogenetics, TMS, and lesion studies have been seen as providing stronger tests of causality.

Although neuropsychological research has been a cornerstone for cognitive neuroscience, its relative importance to the field has diminished over the past generation with the emergence of new technologies. This point is made salient when scanning the table of contents of the journals of the field: Articles using fMRI and EEG dominate the publications, with only an occasional issue including a study involving work with neurological patients. For example, in PubMed searches restricted to articles since 2011, the keyword “fMRI” yielded 409,310 hits, whereas the keyword “neuropsychological” yielded only 72,835 hits.

There are several methodological issues that constrain the availability and utility of neuropsychological research. First, the studies require access to patients who have relatively homogenous neurological pathology. This is quite challenging when drawing on patients with focal insults from stroke or tumor or when studying rare degenerative disorders. As such, the sample size in neuropsychological research tends to be small, frequently involving less than a dozen or so participants (Tang et al., 2020; Breska & Ivry, 2018; Olivito et al., 2018; Wang, Huang, Soong, Huang, & Yang, 2018; Casini & Ivry, 1999). Often, researchers have very limited access to these hard-to-reach populations, especially with rare disorders such as spinocerebellar ataxia (SCA). Second, given that the recruitment is usually from a clinic or local community, the sample may not be

University of California

representative of the population; for example, samples may be skewed if based on patients who are active in support groups. Third, given their neurological condition, the patients may have limited time and energy. Their disabilities make it challenging to recruit the participants to come to the laboratory, and for those who do, the experiments have to be tailored to avoid taxing the participants' mental capacity. As a result, it can take quite a bit of time to complete a single experiment, let alone a package of studies that might make for a comprehensive story. Hence, the progress and efficiency of each study can be very restricted.

One solution that we have pursued in our research on the cerebellum has been to take the laboratory into the field, setting up a testing room at the annual meeting of the National Ataxia Foundation. This conference is patient focused, providing people with ataxia a snapshot of the latest findings in basic research and clinical interventions as well as an opportunity to share their experiences with their peer group. For the past decade, we have set up a testing room at the conference and, over a 3-day period, been able to test about 15 people, a much more efficient way to complete a single study compared to the more traditional approach of enlisting patients with ataxia in the Berkeley community to come to the laboratory.

There remain notable limitations with taking neuropsychological research into the "field." This approach is not ideal for multiexperiment projects and entails considerable cost to send a team of researchers required to recruit the participants and coordinate the testing schedule. There are also issues concerning selection biases. Conference attendees tend to be highly motivated individuals, and those who are more severely impaired may find it difficult to meet the challenges of travel or energy to participate in a study while also attending conference events.

An alternative and simpler solution might come from the use of the Internet. In recent years, behavioral researchers in different domains of studies have developed online protocols to reach larger (Adjerid & Kelley, 2018) and more diverse (Casler, Bickel, & Hackett, 2013) populations than feasible with laboratory-based methods. Within psychology, platforms such as Amazon Mechanical Turk have been used to efficiently collect behavioral data (Crump, 2013). A number of studies have shown that the data obtained in online testing are as reliable and valid as in-person testing (Chandler & Shapiro, 2016; Casler et al., 2013; Buhrmester, Kwang, & Gosling, 2011). Clinical psychology (Chandler & Shapiro, 2016) and developmental studies (Tran, Cabral, Patel, & Cusack, 2017), which have more limited in-person access to their research population than other domains, have particularly benefited because of the unique challenges that can arise in recruiting these populations to the laboratory.

To date, online protocols have been used in a limited manner in medical (Ranard et al., 2014) and neuropsychological research, with the focus on collecting survey data (e.g., Gong et al., 2020) or patient recruitment (Hurvitz, Gross, Gannotti, Bailes, & Horn, 2020). Here, we report

an online protocol designed to recruit participants and administer behavioral tests for neuropsychological research. Although we developed this protocol to facilitate our research program involving patients with subcortical degenerative disorders, the protocol can be readily adopted for different populations. As such, it provides a valuable and efficient new approach to conduct neuropsychological research and, we hope, will contribute to the neurology, neuropsychology, and cognitive neuroscience communities.

PONT

We call the new protocol PONT, an acronym for "protocol for online neuropsychological testing." PONT entails a comprehensive package that addresses challenges involved in recruitment, neuropsychological evaluation, behavioral testing, and administration to support neuropsychological research. The protocol was approved by the institutional review board (IRB) at the University of California, Berkeley (UC Berkeley). In the initial 10-month period, we contacted 646 support group coordinators to recruit large samples of individuals with degenerative disorders of the cerebellum (SCA) or BG (Parkinson's disease [PD]). We established a workflow for online recruitment, neuropsychological assessment, behavioral testing, and follow-up. During this 10-month period, we have completed six experiments testing motor and cognitive abilities and report one of these studies in this article.

How Does PONT Work?

PONT entails five primary steps (Figure 1): (1) contacting support group leaders to advertise the project; (2) having interested individuals initiate contact with us, a requirement set by our IRB protocol; (3) conducting interactive, remote neuropsychological assessments; (4) automated

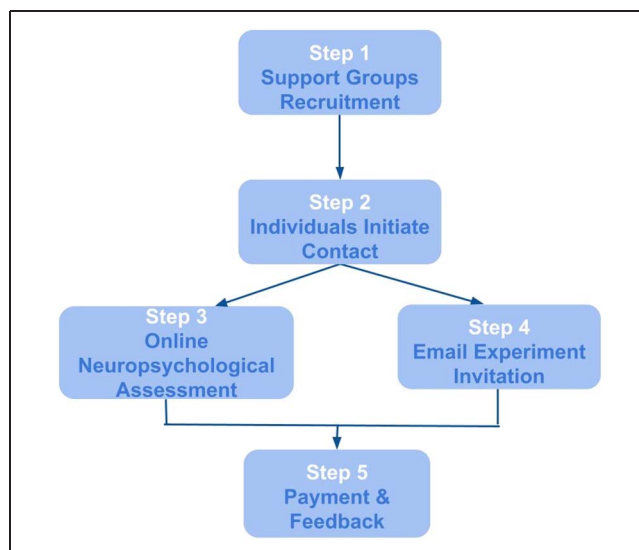


Figure 1. General PONT workflow.

administration of the experimental tasks; and (5) obtaining feedback and providing payment. A detailed, step-by-step, description of the protocol appears below and is available online (PONT-general workflow link, full URL address: osf.io/fktn9/).

Working through Support Group Leaders

Our IRB recruitment rules do not allow UC Berkeley researchers to directly contact potential participants; rather, we can provide descriptive materials with information indicating how interested individuals can contact us. Given this constraint, we sent emails to 646 contact addresses of support groups across the United States. This information was available on the National Ataxia Foundation Web site and the Parkinson and Movement Disorder Alliance Web site. The initial email provides a description of PONT and provides text that the support group administrators can pass along to their group members. Although we do not explicitly request a response from the administrators, our experience is that those who pass along our recruitment information inform us of this in a return email. The response rate is modest: After our initial email and a follow-up 2 weeks later to those who have not responded, we estimate that approximately 15% of the support groups forward our information to their members. Although this number is low, it does mean that our recruitment information has reached the membership of around 100 support groups.

This procedure is a variant of snowball sampling (Goodman, 1961), a method that has been employed to recruit patients and other hard-to-reach groups for clinical studies. Here, we modify this procedure to enlist participants via support groups and, cascading from there, word of mouth for our behavioral studies. The materials used in this phase of the recruitment process can be found online (Online Outreach Forms link, Step 1). These can be readily adopted by any research group that is targeting a population associated with support organizations and Web-based groups or when the sample might be enlisted as part of a multilaboratory collaborations.

Enrollment of the Participants

In the second step, individuals who received our flyer via their support group leaders and wished to participate could register for the study either by emailing the laboratory or filling out an online form (Online Registration Form link). In this way, the participant initiated contact with the research team, a requirement set forth in our IRB contact guidelines. Between November 2019 and October 2020, 103 individuals with SCA (age range = 24–88 years) and 133 individuals with PD (age range = 48–83 years) enrolled in the project. Reflecting the distribution of the support groups, this sample included individuals who are currently residing in 30 states, such as Hawaii, Kentucky, Ohio, California and more. Thus, our sample for PONT is

much more geographically diverse than would occur in a typical laboratory-based study.

Control participants were recruited via advertisements posted on the Craigslist Web site. The advertisement instructed interested individuals to complete our online form. For the study described below, the advertisement indicated that participation was restricted to individuals between the ages of 35 and 80 years based on our past experience in testing individuals with SCA. With hindsight, we realize it would have been more appropriate to set the range based on the ages of the patient sample. Indeed, with the large control sample available for online testing, oversampling would make it possible to do more precise age matching of controls and patients. Over a 9-month period, we have enrolled 159 individuals in the control group.

Online Interview: Demographics, Neuropsychological Assessment, and Medical Evaluation

The third step in PONT involved an online video interview that included a neurological and neuropsychological evaluation. The materials used in our interview and evaluation are provided online (Online Neuropsychological Assessment Protocol link). Although the relevant instruments will vary depending on the interests of a research group, these documents demonstrate some of the modifications we adopted over the course of pilot work with PONT (see below).

Registered individuals were contacted by email to invite them to participate in an online, live interview with an experimenter. This session provided an opportunity to describe the overall objective of the project, confirm basic demographic information, perform a short assessment of cognitive status, and, for the individuals with PD or SCA, obtain a medical history and abbreviated neurological examination. The invitation email indicated that the session would require that the participant enlist another individual (e.g., caregiver, family member) to join in for part of the session to assist with test administration and video recording. It also indicated that participation in the project would require the ability to use a computer and respond unassisted on a computer keyboard. Note that we have made sure to assess whether there are any video/audio issues before conducting each item so it will not interrupt the neuropsychological assessments. If there was any issue, we resolved it during the meeting or, in rare cases, rescheduled the session.

Approximately 240 individuals were invited for the online interview. If no response was received, a follow-up email was sent approximately 14 days later. We have received 125 affirmative responses over the past 9 months. To date, 118 interviews have been conducted (60 PD, 18 SCA, and 40 controls). As with all online interactions, we were concerned about potential hardware/software failures and ran a series of checks before starting the session. In a few instances, it was necessary to reschedule the session. Because we have limited recruitment to the United States

in our initial deployment of PONT, there were no significant challenges in terms of coordinating time zones; extending this approach internationally would require the interviewer to sometimes work at atypical hours. Note that all participants met the setting requirements.

At the start of the online session, the experimenter provided an overview of the mission of our PONT project and the goals of the interview. This introduction provided an opportunity to emphasize orally that participation is voluntary and involves a research project that will not impact the participant's clinical care or provide any direct clinical benefit. After providing informed consent, the participant completed a demographic questionnaire. The experimenter then administered a modified version of the Montreal Cognitive Assessment (MoCA) test (Nasreddine et al., 2005) as a brief evaluation of cognitive status. For control participants, the session ended with the completion of the MoCA.

The participants with PD and participants with SCA continued on to the medical evaluation phase. First, the experimenter obtained the participant's medical history, asking questions about age at diagnosis, medication and other relevant information (e.g., DBS for PD, genetic subtype if known for SCA), and a screening for other neurological or psychiatric conditions. Second, the experimenter administered a modified version of the motor section of the Unified Parkinson's Disease Rating Scale (UPDRS; Goetz, 2003) to the participants with PD and the Scale for Assessment and Rating of Ataxia (SARA; Schmitz-Hübsch et al., 2006) to the participants with SCA.

Modifications were made to these assessment instruments to make them more appropriate for online testing. For the MoCA test, we eliminated "alternating trail making" because this would require that we provide a paper copy of the task. For the UPDRS and SARA, we modified items that require the presence of a trained individual to ensure safe administration. We eliminated the "postural stability task" from the UPDRS because it requires that the experimenter apply an abrupt pull on the shoulders of the participant. We modified three items on the UPDRS ("arising from chair," "posture," and "gait"), obtaining self-reports from the participant rather than the standard evaluation by the experimenter. Similarly, we obtained self-reports of stance and gait for the SARA rather than observe the participant on these items. For the self-reports, we provided the scale options to the participant (e.g., on the SARA item for gait, 0 = *normal/no difficulty* to 8 = *unable to walk even supported*). The scores for the MoCA and UPDRS batteries were adjusted to reflect these modifications. For the online MoCA, the observed score was divided by 29 (the maximum online score) and then multiplied by 30 (the maximum score on the standard test). Hence, if a participant obtained a score of 26, the adjusted score will be $(26/29) * 30$, or 26.9. The same adjustment procedure was performed for the UPDRS. No adjustment was required for the SARA.

The interview took around 30 min for the control participants and 40–60 min for the participants with PD and

those with SCA. We plan to repeat the interview every year for participants who remain in the pool, ensuring that the neuropsychological and neurological evaluations are up-to-date.

Completion of Online Experiments

The experiments were programmed in Gorilla Experiment Builder (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2018). For a given experiment, the participant was emailed an individual link that assigned a unique participant ID, providing a means to ensure that the data were stored in an anonymized and confidential manner (see online template, Online Experiment Invite link, Step 4). This email also included a brief overview of the goal of the study (e.g., an investigation of how PD impacts finger movements or an investigation looking at mathematical skills in SCA). The email clearly stated that the participant should click on the link when they are ready to complete the experimental task, allowing sufficient time to complete the study (30–60 min depending on the experiment). In this manner, administration of the experimental task was entirely automated, with the participant having complete flexibility in terms of scheduling.

To date, we have run six experiments with PONT, with the studies looking at various questions related to the involvement of the BG and/or cerebellum in motor learning, language, and mathematics. For the experiments involving both patient groups, we set a recruitment goal of 20 individuals per group (including controls). On the basis of similar neuropsychological studies conducted in our laboratory, we have found that sample sizes of around 12 have sufficient power to reveal group effects on sensorimotor control (and cognitive) tasks. We set a larger target for the online work given uncertainties about data quality with this approach as well as the greater availability of patient participants. In future work, the results from the online studies will provide appropriate samples to calculate, a priori, the sample size for a desired level of experimental sensitivity.

Although we set a target of 20 participants per group, the actual number varied given that recruitment was done in batches of approximately 70 emails per group. The response rate to a given batch was approximately 15%. This number was surprisingly low given that all of the invitees had indicated an interest in participating in our research program. We can only surmise that the invitees face email overload issues and, like many of us, procrastinate in responding to emails. Follow-up emails were sent after about 2 weeks, repeating the invitation and yielding similar response rates. As such, it took approximately 4–8 weeks to complete a single experiment.

We are still exploring the ideal timing for sending our experimental invitations, looking for ways to improve on the 15% response rate. There was also a considerable delay in some cases (up to a couple of months) between the time at which the participant agreed to take part in the study and when they received the email invite for a

specific study. We expect shortening this interval would increase the response rate. We have also found that participants are very likely to complete an experimental task shortly after completing the neuropsychological assessment interview (usually within 1 day). As such, we recommend that the experimental task be ready before beginning the interviews. We have also started including a survey question at the end of the experiment, asking participants how often they would like to be contacted for additional experiments.

Note, for new participants, we sent an experiment invite link right after the online interview. If the participants cannot complete it directly after the neuropsychological session, they usually complete it within a day. For participants who previously completed the neuropsychological assessment, we send an experiment within 1 year to ensure that the neuropsychological assessment is valid. On average, approximately 1 month passes between completion of the neuropsychological assessment and an experiment, but this can range from 1 day to a year.

Participant Feedback and Payment

After completing the online experiment, the participant is sent an automated thank-you note and a short online form for providing feedback about the task (e.g., rate difficulty, level of engagement). In addition, we sent the participant an email with information about payment (see template, Online Feedback Form, Step 5). Reimbursement is at \$20/hr, and the participant can opt to be paid by check sent via regular mail, via Paypal, or with an Amazon gift card. We also reimburse for the online, live interview session to further reinforce that this is a research project unrelated to their medical care. The experimenter monitors the feedback reports and, when appropriate, sends an email response to the participant.

The participant's data are downloaded from the Gorilla protocol to a secure laboratory computer. Within this secure, local environment, we have the ability to link the unique participant ID code used to access the experiment with personal information. This allows us to build a database to track the involvement of each individual in PONT as well as make comparisons across experiments.

PONT in Action: Sequence Learning in PD and SCA

To demonstrate the feasibility and efficiency of PONT, we report the results from an experiment on motor sequence learning in which we used the discrete sequence production (DSP) task. On each trial, the participant produces a sequence of four keypresses in response to a visual display. We chose to use a motor sequence learning task for two reasons. First, there exists a large literature on the involvement of the BG and the cerebellum in sequence learning (Jouen, 2013; Debas et al., 2010; Doyon et al., 2002), including studies specifically on PD and SCA (Ruitenberg et al., 2016; Gamble et al., 2014; Tremblay et al., 2010;

Spencer & Ivry, 2009; Shin & Ivry, 2003; Roy, Saint-Cyr, Taylor, & Lang, 1993). Thus, we can compare the results from our online study with published work from traditional, in-person studies. Second, the requirement that a trial consist of four, sequential responses makes this a relatively hard task in terms of response demands compared to tasks requiring a single response (e.g., two-choice RT). As such, we can evaluate the performance of PONT across the range of hardware devices used by the participants as well as evaluate participants' performance on a relatively demanding motor task.

Skill acquisition has been associated with two prominent processes: (1) memory retrieval that becomes enhanced with practice (Logan, 1988) and (2) improved efficiency in the execution of algorithmic operations (Tenison & Anderson, 2016). We modified the DSP task to look at how degeneration of the BG or cerebellum impacts these two learning processes. We compared practice benefits for repeated items (memory-based learning) with practice benefits for nonrepeated items (algorithm-based learning). Although we expected the PD and SCA groups would be slower than the controls overall, we expected sequence execution time (ET) would become faster over the experimental session for all three groups. Our primary focus was to make group comparisons of the learning benefits for the repetition and nonrepetition conditions, providing assays of memory-based and algorithm-based learning, respectively.

Previous studies have demonstrated that individuals with SCA show reduced practice benefits on both sequence and random blocks (Tzvi et al., 2017). Hence, we predicted that the SCA group would be impaired on both the repeating and nonrepeating sequences. Because the literature on motor sequence learning in PD is equivocal (Ruitenberg, Duthoo, Santens, Notebaert, & Abrahamse, 2015), we did not have an a priori prediction for this group. By using repeating and nonrepeating sequences, we hoped to be in a position to distinguish between two ways in which sequence learning might be impacted by PD.

METHODS

Participants

Drawing on the PONT participant pool, email invitations were initially sent to 85, 66, and 89 individuals in the control, SCA, and PD groups, respectively, with the differences reflecting the pool size for each group at the time of the email. The overall response rate to this first email was around 20%. Follow-up emails were sent every few weeks, and after a few rounds, we reached our goal of a minimum of 20 participants per group. Of those who initiated the study, we excluded the data of three participants (one per group) who failed to respond correctly to the attention probes (see below) and four participants (one control, two with SCA, and one with PD) who failed

Table 1. Demographic and Neuropsychological Summary of All Groups

	<i>Years of Education</i>	<i>Number of Women</i>	<i>Age (Years)</i>	<i>MoCA</i>	<i>Motor Assessment</i>
Control	17.7 ± 0.4 (14–20)	12	63.6 ± 2.3 (40–78)	27.8 ± 0.4 (21–30)	
SCA	17.1 ± 0.8 (12–22)	12	56.3 ± 2.6 (34–80)	25.8 ± 0.6 (22–29)	9.1 ± 0.9 (2–14.5)
PD	17.2 ± 0.5 (12–22)	12	65.3 ± 1.5 (50–79)	26.8 ± 0.5 (23–30)	16.2 ± 0.9 (9.7–24.8)

Mean ± *SE* (range) for each demographic and neuropsychological variable. The scores of the MoCA and UPDRS are adjusted for online administration.

to complete the experiment (either aborting the program or a loss of Internet connectivity during the session). The final sample of participants included in the analyses reported below was composed of 62 participants, 22 controls, 17 with SCA, and 23 with PD.

Table 1 provides demographic information for the three groups as well as the adjusted MoCA, SARA (SCA), and UPDRS (PD) scores. The SCA group was composed of 12 individuals with a known genetic subtype and five individuals with an unknown etiology (idiopathic ataxia). The mean duration since diagnosis for the SCA group was 5.9 years (range = 1 month to 25 years), and the mean age of onset was 50.4 years (range = 21–78 years). The mean duration since diagnosis for the PD group was 6.5 years (range = 1 month to 16 years), and the mean age of onset was 58.8 years (range = 45–70 years). None of the individuals in the PD group had undergone surgical intervention as part of their treatment (e.g., DBS), and all were tested while on their current medication regimen.

Procedure

The participant used their home computer and keyboard to perform the experiment. Given this and the fact that we did not control for viewing distance, the size of the stimuli varied across individuals. All of the instructions were provided on the monitor in an automated manner, with the program advancing under the participant's control.

The participant was instructed to place his or her fingers (thumbs excluded) on the keyboard, using the keys “z,” “x,” “c,” and “v” for the left hand and the keys “m,” “<,” “>,” and “?” for the right hand. Eight placeholders were displayed on the screen, with each placeholder corresponding to one of the keys of the keyboard in a spatially compatible manner (Figure 2).

We employed the DSP task (Abrahamse, Ruitenberg, Kleine, & Verwey, 2013) in which a four-element sequence is displayed at the start of the trial and the participant is asked to produce the sequence as quickly and accurately as possible. At the start of each trial, a black fixation cross appeared in the middle of a white background. After 500 msec, the fixation cross was replaced by a stimulus display that consisted of the numbers “1,” “2,” “3,” and “4,” with each number positioned over one

of the placeholders. The numbers specified the required sequence for that trial. In the example shown in Figure 2, the correct sequence required sequential keypresses with the left index finger, right pinky, right index finger, and left pinky. Note that the presentation of the stimulus sequence also served as the imperative signal.

If a keypress was not detected within 3000 msec for each element of a given sequence, the phrase “Respond faster” appeared for 200 msec. If a single response was not detected within 4000 msec or the wrong key was pressed, the trial was aborted. In the case of an erroneous keypress, a red “X” was presented above the placeholders. When the entire sequence was executed successfully, a green “√” appeared above the placeholders. The feedback screen remained visible for 500 msec, after which a fixation cross appeared for 500 msec indicating the start of the intertrial interval.

Many studies using the DSP task have participants repeatedly perform a small set of sequences (e.g., 2) and compare the ET of these trained sequences to that of untrained sequences after a learning phase (Abrahamse et al., 2013). To test memory-based and algorithm-based learning, we opted to take more continuous measures across the entire experimental session. To this end, we created two nonoverlapping categories of sequences. For memory-based learning, a set of repeating sequences was created, composed of eight 4-element sequences. For algorithm-based learning, a set of novel sequences was created, composed of 192 four-element unique sequences. The sequences for each condition were determined randomly for each participant with the constraints that each sequence required at least one keypress from

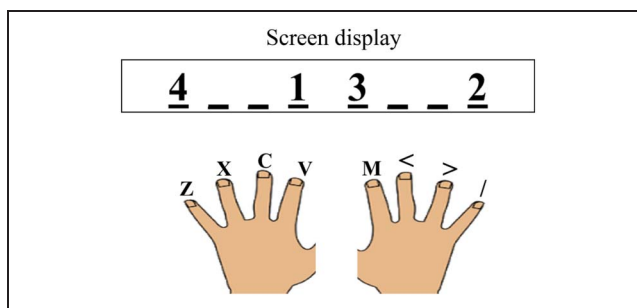


Figure 2. Each line on the screen corresponded to a finger position in a spatially compatible manner. The numbers indicate the order of the keypresses for the current trial. In this example, the participant should press the left index finger (“v”), the right pinky (“?”), the right index finger (“m”), and the left pinky (“z”).

each hand. In addition, in each sequence, a finger should not press a key more than once.

The experimental block was composed of 384 trials, divided into 24 blocks of 16 trials each. Each block included one presentation of each of the eight sequences in the repetition condition and eight of the unique sequences, with the order of the 16 sequences determined randomly within each block. A minimum 1-min break was provided after each run of four blocks, with the participant pressing the “v” key when ready to continue the experiment.

To ensure that participants remained attentive, we included five “attention probes” on the instruction pages that appeared in the start of the experiment or during the experimental block. For example, an attention probe might instruct the participant to press a specific key rather than selecting the “next” button on the screen to advance the experiment (e.g., “Do not press the ‘next’ button. Press the letter ‘A’ to continue.”). If the participant failed to respond as instructed on these probes, the experiment continued, but the participant’s results were not included in the analysis.

The experiment took approximately 45 min to complete.

RESULTS

Accuracy rates were 80%, 78%, and 78% for the control, SCA, and PD groups, respectively, $F(2, 59) = 0.198$, $p = .821$ (one-way ANOVA). Most of the errors involved an erroneous keypress (16% of all trials), and the remaining errors were because of a failure to make one of the keypresses within 4000 msec (5% of all trials). We excluded trials in which participants failed to complete the full sequence within 4000 msec (3% of the remaining trials).

Our primary dependent variable of interest was ET, the time from the first keypress to the time of the last keypress. Before turning to these data, we considered potential trade-offs in performance. First, to determine if there was a speed–accuracy trade-off, we looked at the correlation between ET and accuracy for each group. These correlations were all negative (control = $-.14$, SCA = $-.69$, PD = $-.20$), indicating that participants who made the most errors also tended to be the slowest in completing the sequence, the opposite pattern of a speed–accuracy trade-off. Second, we were concerned that participants might vary in the degree to which they prepared the series of responses before initiating the first movement. For example, a participant might opt to preplan the full sequence before making the first keypress or plan the responses in some sort of sequential manner. To look at possible trade-offs between preplanning and ET, we computed the correlation between RT (time to the first keypress) and ET (which starts at the time of the first keypress). These correlations were all positive (control = $.75$, SCA = $.57$, PD = $.51$), indicating no ET–preplanning tradeoff.

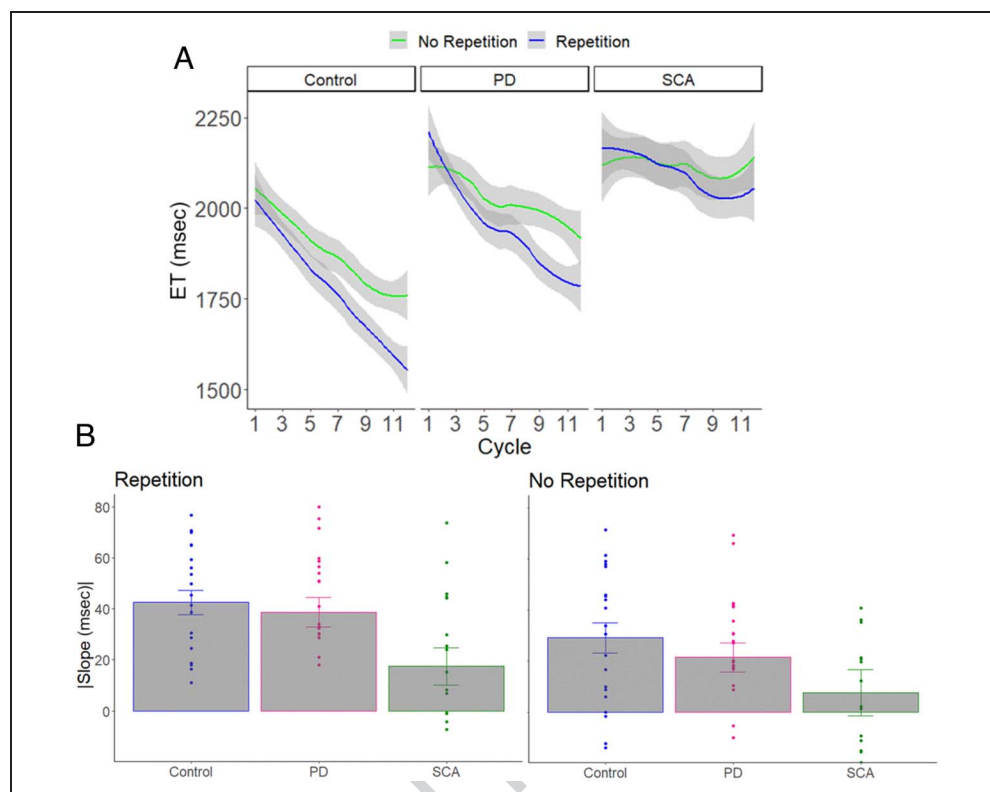
Figure 3 shows ET as a function of cycle of learning and group, with separate functions for the no-repetition and repetition conditions. Across all conditions, the SCA group was 362.8 msec slower than the control group in executing the sequences ($SE = 164.2$, $p = .031$). The PD group was 167.6 msec slower than the control group, but this effect was not significant ($SE = 151.6$, $p = .274$). To statistically evaluate learning, we employed a linear mixed-effect model with the factors Group, Repetition condition, and Cycle as well as Participant as a random factor (R software, lme4 library; Bates, Mächler, Bolker, & Walker, 2015). For the Cycle variable, we averaged across pairs of blocks to collapse the 24 blocks into 12 cycles. We also included years of education, age, and MoCA score as covariates in the model.

The control participants got faster over the experimental session with an overall slope of -35.8 msec/cycle ($SE = 1.9$, $p < .0001$). These participants showed significant learning in both the no-repetition condition (-29.4 msec/cycle, $SE = 2.8$, $p < .0001$) and the repetition condition (-41.5 msec/cycle, $SE = 2.5$, $p < .0001$). Similar to what has been observed with the DSP task in young adults (Jouen, 2013), the improvement was more pronounced for the repetition condition than for the no-repetition condition (difference in slope: -11.8 msec/cycle, $SE = 3.7$, $p = .001$).

The patient groups also got faster over the experimental session with an overall slope of -29.2 msec/cycle for the PD group ($SE = 1.9$, $p < .0001$) and -13.1 msec/cycle for the SCA group ($SE = 2.3$, $p < .0001$). The improvement was significant in both the no-repetition condition (PD: -20.5 msec/cycle, $SE = 2.7$, $p < .0001$; SCA: -7.8 msec/cycle, $SE = 3.3$, $p = .019$) and the repetition condition (PD: -37.1 msec/cycle, $SE = 2.6$, $p < .0001$; SCA: -18.4 msec/cycle, $SE = 3.0$, $p < .0001$). Similar to the control participants, this magnitude of the improvement was more pronounced for the repetition condition than for the no-repetition condition (difference in slope: PD = -16.8 msec/cycle, $SE = 3.7$, $p = .001$; SCA = -10.3 msec/cycle, $SE = 4.5$, $p = .023$).

We next compared the learning effects for each patient group to the control group. Both patient groups showed less improvement than the controls in the no-repetition condition (SCA vs. control: difference in slope = 21.6 msec/cycle, $SE = 4.19$, $p < .0001$; PD vs. control: difference in slope = 9.1 msec/cycle, $SE = 3.8$, $p = .017$). However, only the SCA group was impaired in the repetition condition, showing less improvement than the control group (23.2 msec/cycle, $SE = 4.1$, $p < .0001$); the comparison between the control and PD groups was not significant (4.1 msec/cycle, $SE = 3.7$, $p = .275$). In a comparison of the two patient groups, the SCA group showed less improvement than the PD group on the no-repetition condition (12.6 msec/cycle, $SE = 4.23$, $p = .002$) and the repetition condition (18.7 msec/cycle, $SE = 3.94$, $p < .001$). In terms of the covariates, there were no significant effects of education (18.5 msec/year, $SE = 67.9$,

Figure 3. (A) ET as a function of cycle of learning, repetition condition, and group. Error bars = 95% confidence level. (B) The slope for each individual (dots) in each repetition condition. Error bars = *SEM*.



$p = .786$), age (114.2/year, $SE = 71.1$, $p = .113$), or MoCA score (-122.6 /point, $SE = 74.2$, $p = .104$).

Taken together, the results demonstrate the viability of using an automated, online protocol to examine sequence learning in neurological populations. As expected, the two patient groups were slower than the control group overall, although this effect was only significant for the SCA group. More importantly, the results show distinct patterns of sequence learning impairments in the two patient groups, similar to what has been observed in laboratory studies (Ruitenberg et al., 2015, 2016; Gamble et al., 2014; Tremblay et al., 2010; Spencer & Ivry, 2009; Shin & Ivry, 2003; Molinari et al., 1997; Roy et al., 1993). Interestingly, the SCA group was impaired in both the repetition and no-repetition conditions, a pattern suggestive of impairment in both memory-based and algorithm-based learning. In contrast, the PD group was only significantly impaired in the no-repetition condition, a pattern suggestive of impairment in algorithm-based learning.

Sequence production and learning has been the focus of many previous studies involving both of these patient groups. We are unaware of any studies with either patient group using the DSP task with repetition versus no-repetition manipulation, precluding direct, cross-experiment comparisons. Most of the prior work has been done with the serial RT task (Tremblay et al., 2010; Muslimović, Post, Speelman, & Schmand, 2007; Shin & Ivry, 2003), a method in which sequence learning is assessed by comparing blocks of trials in which a sequence of length n repeats in a cyclic manner (e.g., $n = 8$ and a

block involves 10 cycles) to blocks in which the stimuli are selected at random. As such, sequence learning is operationalized as reductions in RT for a repeating sequence relative to improvements on the random blocks.

Using the repetition benefit measure, individuals with cerebellar pathology, either from degeneration or focal lesions, consistently exhibit a pronounced learning deficit (Tzvi et al., 2017; Molinari et al., 1997), one that within the present framework, would be attributed to memory-based learning. One exception is a study (Spencer & Ivry, 2009) that observed normal learning when the responses were made directly to the stimulus positions. It may be either that these “direct” cues support the formation of distinct memory associations not impacted by cerebellar pathology (e.g., direct stimulus–response mappings, not requiring an intermediary mapping operation) or that the direct cues provide a boost to memory formation. One could consider the random blocks as a long “nonrepeating” sequence. For example, it was demonstrated that individuals with SCA showed reduced practice benefits on both sequence and random blocks (Tzvi et al., 2017), similar to the deficits we observed for both repeating and nonrepeating sequences, respectively.

Note that the literature on motor sequence learning in PD is equivocal (for a review, see Ruitenberg et al., 2015). Previous studies have produced mixed results that may suggest that the involvement of the BG is selective and depends on the specific learning conditions, medication state, or symptom severity (Ruitenberg et al., 2015, 2016; Shin & Ivry, 2003). For example, in studies utilizing a

modified version of the serial RT task, patients with PD were less efficient in learning random nonrepeating sequences but had no impairment in learning repeated sequences (Ruitenberget al., 2016; Tremblay et al., 2010; Muslimović et al., 2007). Thus, in line with our results, whereas we observed deficits in the algorithm-based learning, memory-based learning may be spared in PD.

DISCUSSION

In this article, we describe a novel PONT. PONT was designed to take advantage of features that have motivated many experimental psychologists to move to online studies over the past decade, yet it is tailored to address the unique demands of patient-based research. The main focus of this article was to outline the key steps required for this approach, with our report of the results from a sequence learning included to provide a concrete example of the application of PONT.

PONT was developed to address five main challenges facing researchers who conduct neuropsychological studies. Perhaps the most profound limiting factor that this valuable approach addresses is access to a targeted population (see Figure 4), be it a disorder with a relatively high prevalence rate in the population (e.g., 0.3% for PD) or one that is rare (e.g., 0.003% for SCA). Not only can it be difficult to recruit these individuals, but the patients themselves may have limited time and energy given their neurological condition. As a result, it can take a long time to complete a study (e.g., 1–2 years), a problem that is magnified for a project requiring multiple experiments. Indeed, neuropsychological articles tend to involve only a single patient group and a single experiment. PONT provides a protocol that makes data collection extremely efficient. In the first 10 months after the launch of PONT, we were able to complete six studies, with each having around 20 participants per group. Moreover, as our patient participant pool grows, we should be in a position to run studies with much larger sample sizes, something that is relatively rare in the cognitive neuroscience literature but desirable for looking at factors underlying heterogeneity in performance within a

particular group (e.g., because of genetic subtype, pattern of pathology).

Second, recruitment for laboratory-based experiments usually comes from a local community. As such, the sample may be small and not representative of the population (see Figure 4). For example, our PD sample from the Berkeley community tends to be highly educated compared to the general PD population in the United States. In addition, if the study requires a relatively homogenous pathology or genetic subtype (e.g., a particular variant of SCA), the sample size is likely to be small (e.g., eight participants). By casting the recruitment net across the entire country (or, as we envision, internationally), PONT will be ideal for running experiments with larger and more diverse samples that will better represent the general population.

Third, although we developed PONT to facilitate our research program on subcortical contributions to cognition, this protocol can be readily adopted for different populations. It can be readily adopted for disorders such as Alzheimer’s disease, stroke, and dyslexia, which are associated with support groups or online social networks. We anticipate recruitment will be more challenging for any individual laboratory when these points of contact are absent. However, PONT can provide a common test protocol to support multilaboratory collaborative operations, an alternative way to increase the participant pool that may be especially useful for studies aimed at more targeted populations (e.g., medulloblastoma, nonfluent aphasia). As such, PONT provides a valuable approach to conduct neuropsychological research to explore a wide range of brain regions and neurological disorders.

An important benefit of PONT is that it is user-friendly for the participants. The participants complete the experiment at home, choosing a time that fits into their personal schedule. As we have learned from our experience over the past three decades working with participants with SCA and those with PD, arranging transportation to and from the laboratory can present a major obstacle, especially when the travel is to participate in a study that offers no direct clinical benefit to the participants. The participants also seem to enjoy the challenge of their “assignment,” although they, like our college-age participants, may complain about the repetitive nature of a task when hundreds of trials are required to obtain a robust data set. Table 2 provides a sample of the comments we have received in our postsession surveys. Although we had embarked on the PONT project before the onset of the COVID-19 pandemic, the timing was fortuitous, allowing us to continue and ramp up our neuropsychological testing, allowing vulnerable populations to participate in studies in a safe manner. PONT can serve as a “defense system” for neuropsychological research development—protecting research development even when patients’ accessibility is more limited than ever.

Finally, PONT provides a considerable benefit in terms of reduced economic and environmental costs. Setting

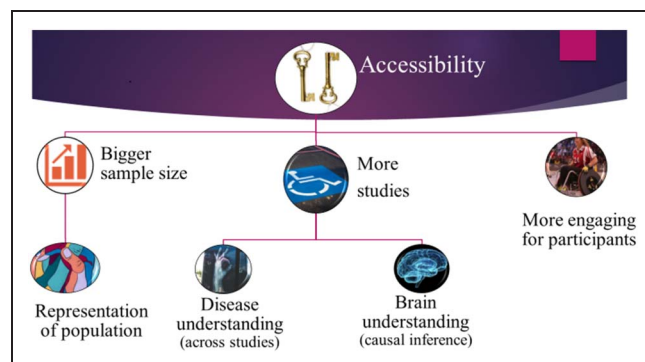


Figure 4. Advantages of the online protocol for neuropsychological research.

Table 2. A Sample of the Participants' Feedback After Completing the Online Experiments

<i>Participant</i>	<i>Feedback</i>	<i>Type</i>
1	"There were a few that I struggled with but over-all it helped me practice my mental math. Thank you."	Positive
2	"I tried so hard to do them faster each round and just wanted to get them right which caused me to push the wrong key several times. Competing against myself was good for my brain. This old lady thanks you for the challenge."	Positive
3	"I thought this was fun. A little repetitive but enjoyable."	Positive
4	"Thanks for inviting me. Would like to know more about the study outcomes."	Positive
5	"At times I knew the correct answer and pressed the wrong finger. I feel that there was a learning curve in my responses."	Neutral
6	"I needed a blowing the nose break."	Neutral
7	"What's to say? my eyes got tired"	Negative
8	"I had issues with my computer—twice I inadvertently brought up an unwanted window and had to click to get rid of it, and during the last set I lost the whole page, had to click on the link the email and it took me to where I believe I left off."	Negative
9	"Test is too long."	Negative

aside the hard-to-calculate benefits from reduced travel and reductions in personnel time, the savings are substantial when only considering participant costs. For our in-laboratory studies, we typically conduct two 35-min experiments, sandwiched around a 30-min break. When we factor in compensation for travel time, the cost tends to come to around \$100/session. The same amount of data can be obtained for approximately \$40 over two PONT sessions, with the added benefit that potential fatigue effects are reduced given that each session is run on a different day.

Although there are many advantages to this online approach, there are also some notable limitations that need to be taken into consideration. In terms of participant recruitment, PONT is limited to populations that have consistent access to quality Internet connections, a limitation in terms of inclusiveness. We also recognize there

will be variation across institutions and countries in terms of ethical constraints. Our protocol conforms to the guidelines of UC Berkeley where the main issues of concern are that participants freely choose to participate and the data are preserved in a manner to ensure anonymity and minimize risk to data breaches. We anticipate the requirements will differ in other locales. We are now developing collaborations with researchers in other countries and will update the online materials as this work unfolds. We anticipate our efforts here will identify new challenges for taking PONT international, beyond the requirements of translating the materials into different languages.

As with all online research, there is always concern about how to verify the participant's status. We do think there are various safeguards that make it very likely the information is accurate. First, the patients reach out to us after they get the information about the project from the coordinator of their support group; thus, they are part of a support group for a specific disorder of interest (to us). Second, we conduct a live video session. This provides us with an opportunity to confirm that their medical information is in line with typical disease-related data (e.g., medication type [e.g., L-dopa] and scores in motor tests [e.g., UPDRS]). Our experience is that the patients take very seriously the accuracy of the information they provide. Many patients possess and reference meticulous records of their prescriptions.

On the procedural front, online experiments are likely to be less standardized across participants. A major source of variability comes about because participant uses his or her own computer and response device. These are, of course, standardized for in-laboratory studies, including positioning the participants such that the visual displays are near identical. It may be possible to impose some degree of control on viewing angle by specifying how the participants position themselves or using simple calibration methods; for example, the size of the stimuli might be adjusted based on having the participant adjust a display line to match a standard distance (e.g., the width of a credit card). Nonetheless, we expect it will be challenging to control the visual angle of a display when the participants are self-administering the experimental task.

Online experiments are also unlikely to be well suited for all domains of study. For example, given the variability in hardware, it would be difficult to run experiments that require precise timing, for instance, studies in which stimuli are presented for a short duration followed by a mask. Moreover, we have avoided studies that use auditory stimuli because computer speakers are so variable and can be of poor quality, and it is not possible use tactile or olfactory stimuli.

Most important, the use of an automated, self-administered system likely comes with a cost in terms of data quality. In our in-person studies, it is not uncommon for a participant to misunderstand the instructions

or need extra practice when they find a particular task difficult. These situations can be readily identified and addressed when an experimenter is present. Moreover, a participant can get discouraged if they find the task too challenging or have an error rate higher than their personal “standard.” An encouraging experimenter can help ensure motivation remains high. We incorporate various feedback messages to keep the participants’ morale and motivation high, but an online system will be less flexible than in-person testing (which could be seen as a positive for maintaining test uniformity). We imagine there will be some experiments that may require the virtual presence of an experimenter. We have also conducted postexperiment live check-ins with some participants to both maintain a human touch and get feedback on how we can make the experience more enjoyable and beneficial. Inspired by the feedback, we have started providing periodic newsletters to the participants, describing recent findings from both basic and translational studies (mailchi.mp/ab940c63fa5c/newsletter-by-the-cognac-lab-at-uc-berkeley-the-neuroscience-of-ataxia-3556742?e=ae8c913a37).

Data quality may also be compromised when testing is self-administered given that the environment is likely conducive to distraction (e.g., from a TV in the background or attention-grabbing text alerts). Although the COVID-19 pandemic precluded a direct comparison of in-person and online testing with PONT, the results from the sequence learning study seem similar to those published from in-person experiments. Several studies have made more systematic comparisons of online and in-person experiments (Barnhoorn, Haasnoot, Bocanegra, & van Steenbergen, 2014; Simcox & Fiez, 2014). Overall, the results are encouraging in that, although the data may be noisier, the general patterns are similar. We note that this is especially true with large sample sizes, something that is unlikely to be true in many neuropsychological studies. One recent review provides recommendations for improving data quality that are certainly appropriate for PONT (Grootswagers, 2020): Keep the experiments as short as possible, provide reasonable compensation, and make the tasks as engaging as possible.

The emergence of online protocols has provided behavioral scientists with the opportunity to conduct studies that involve large and diverse samples, while being cost efficient. The PONT project described in this article describes how this general approach can be adopted to meet the challenges associated with neuropsychological testing. Currently, our work is limited to individuals with ataxia and PD, allowing us to expand our research program on subcortical contributions to cognition. The materials provided with this article can be readily adopted by researchers working with any patient population, especially when recruitment can be conducted via support groups, via Web-based groups, or through collaborations across multiple laboratories/clinics. In terms of the latter, we see PONT as a fertile tool to support multinational collaborative research operations. We expect PONT will

significantly increase the sample size, the number of studies conducted, and the overall pace of neuropsychological research. As such, it offers a powerful tool for this field, one that has and will continue to yield fundamental insights into brain–behavior relationships.

Acknowledgments

We thank Sharon Binoy, Nandita Radhakrishnan, Rachel Woody, Sidra Seddiqee, and Sravya Borra for their assistance.

Reprint requests should be sent to William Saban, Department of Psychology and Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, or via e-mail: williamsaban@gmail.com.

Author Contributions

William Saban: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing—Original draft; Writing—Review & editing. Richard B. Ivry: Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Writing—Original draft; Writing—Review & editing.

Funding Information

This research was supported by funding from the National Institutes of Health (NS116883).

Diversity in Citation Practices

A retrospective analysis of the citations in every article published in this journal from 2010 to 2020 has revealed a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were $M(\text{an})/M = .408$, $W(\text{oman})/M = .335$, $M/W = .108$, and $W/W = .149$, the comparable proportions for the articles that these authorship teams cited were $M/M = .579$, $W/M = .243$, $M/W = .102$, and $W/W = .076$ (Fulvio et al., *JoCN*, 33:1, pp. 3–7). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article’s gender citation balance.

REFERENCES

- Abrahamse, E. L., Ruitenberg, M. F. L., de Kleine, E., & Verwey, W. B. (2013). Control of automated behavior: Insights from the discrete sequence production task. *Frontiers in Human Neuroscience*, 7, 82. <https://doi.org/10.3389/fnhum.2013.00082>, PubMed: 23515430

- Adjerid, I., & Kelley, K. (2018). Big data in psychology: A framework for research advancement. *American Psychologist*, 73, 899–917. <https://doi.org/10.1037/amp0000190>, PubMed: 29469579
- Barnhoorn, J. S., Haasnoot, E., Bocanegra, B. R., & van Steenbergen, H. (2014). QRTEngine: An easy solution for running online reaction time experiments using Qualtrics. *Behavior Research Methods*, 47, 918–929. <https://doi.org/10.3758/s13428-014-0530-7>, PubMed: 25407763
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67. <https://doi.org/10.18637/jss.v067.i01>
- Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences, U.S.A.*, 115, 12283–12288. <https://doi.org/10.1073/pnas.1810596115>, PubMed: 30425170
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk. *Perspectives on Psychological Science*, 6, 3–5. <https://doi.org/10.1177/1745691610393980>, PubMed: 26162106
- Casini, L., & Ivry, R. B. (1999). Effects of divided attention on temporal processing in patients with lesions of the cerebellum or frontal lobe. *Neuropsychology*, 13, 10–21. <https://doi.org/10.1037/0894-4105.13.1.10>, PubMed: 10067771
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29, 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, 12, 53–81. <https://doi.org/10.1146/annurev-clinpsy-021815-093623>, PubMed: 26772208
- Debas, K., Carrier, J., Orban, P., Barakat, M., Lungu, O., Vandewalle, G., et al. (2010). Brain plasticity related to the consolidation of motor sequence learning and motor adaptation. *Proceedings of the National Academy of Sciences, U.S.A.*, 107, 17839–17844. <https://doi.org/10.1073/pnas.1013176107>, PubMed: 20876115
- Doyon, J., Song, A. W., Karni, A., Lalonde, F., Adams, M. M., & Ungerleider, L. G. (2002). Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proceedings of the National Academy of Sciences, U.S.A.*, 99, 1017–1022. <https://doi.org/10.1073/pnas.022615199>, PubMed: 11805340
- Gamble, K. R., Cummings, T. J., Lo, S. E., Ghosh, P. T., Howard, J. H., & Howard, D. V. (2014). Implicit sequence learning in people with Parkinson's disease. *Frontiers in Human Neuroscience*, 8, 563. <https://doi.org/10.3389/fnhum.2014.00563>, PubMed: 25136303
- Geschwind, N. (1970). The organization of language and the brain. *Science*, 170, 940–944. <https://doi.org/10.1126/science.170.3961.940>, PubMed: 5475022
- Gong, Y., Chen, Z., Liu, M., Wan, L., Wang, C., Peng, H., et al. (2020). *Mental health of spinocerebellar ataxia patients during COVID-19 pandemic: A cross-sectional study*. 1–17. <https://doi.org/10.21203/rs.3.rs-40489/v1>
- Goodman, L. A. (1961). Snowball sampling. *Annals of Mathematical Statistics*, 32, 148–170. <https://doi.org/10.1214/aoms/1177705148>
- Grootswagers, T. (2020). A primer on running human behavioural experiments online. *Behavior Research Methods*, 52, 2283–2286. <https://doi.org/10.3758/s13428-020-01395-3>, PubMed: 32291730
- Hurvitz, E. A., Gross, P. H., Gannotti, M. E., Bailes, A. F., & Horn, S. D. (2020, February 1). Registry-based research in cerebral palsy: The cerebral palsy research network. *Physical Medicine and Rehabilitation Clinics of North America*, 31, 185–194. <https://doi.org/10.1016/j.pmr.2019.09.005>, PubMed: 31760991
- Jouen, A.-L. (2013). Discrete sequence production with and without a pause: The role of cortex, basal ganglia, and cerebellum. *Frontiers in Human Neuroscience*, 7, 492. <https://doi.org/10.3389/fnhum.2013.00492>, PubMed: 24065900
- Lezak, M. D. (2000). Nature, applications, and limitations of neuropsychological assessment following traumatic brain injury. In A. L. Christensen & B. P. Uzzell (Eds.), *International handbook of neuropsychological rehabilitation* (pp. 67–79). Boston: Springer. https://doi.org/10.1007/978-1-4757-5569-5_4
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527. <https://doi.org/10.1037/0033-295X.95.4.492>
- Molinari, M., Leggio, M. G., Solida, A., Ciorra, R., Misciagna, S., Silveri, M. C., et al. (1997). Cerebellum and procedural learning: Evidence from focal cerebellar lesions. *Brain*, 120, 1753–1762. <https://doi.org/10.1093/brain/120.10.1753>, PubMed: 9365368
- Muslimović, D., Post, B., Speelman, J. D., & Schmand, B. (2007). Motor procedural learning in Parkinson's disease. *Brain*, 130, 2887–2897. <https://doi.org/10.1093/brain/awm211>, PubMed: 9365368
- Olivito, G., Lupo, M., Iacobacci, C., Clausi, S., Romano, S., Masciullo, M., et al. (2018). Structural cerebellar correlates of cognitive functions in spinocerebellar ataxia type 2. *Journal of Neurology*, 265, 597–606. <https://doi.org/10.1007/s00415-018-8738-6>, PubMed: 29356974
- Poeppl, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology*, 28, 142–149. <https://doi.org/10.1016/j.conb.2014.07.005>, PubMed: 25064048
- Ranard, B. L., Ha, Y. P., Meisel, Z. F., Asch, D. A., Hill, S. S., Becker, L. B., et al. (2014, July 11). Crowdsourcing—Harnessing the masses to advance health and medicine, a systematic review. *Journal of General Internal Medicine*, 29, 187–203. <https://doi.org/10.1007/s11606-013-2536-8>, PubMed: 23843021
- Roy, E. A., Saint-Cyr, J., Taylor, A., & Lang, A. (1993). Movement sequencing disorders in parkinson's disease. *International Journal of Neuroscience*, 73, 183–194. <https://doi.org/10.3109/00207459308986668>, PubMed: 8169053
- Ruitenbergh, M. F. L., Duthoo, W., Santens, P., Notebaert, W., & Abrahamse, E. L. (2015). Sequential movement skill in Parkinson's disease: A state-of-the-art. *Cortex*, 65, 102–112. <https://doi.org/10.1016/j.cortex.2015.01.005>, PubMed: 25681652
- Ruitenbergh, M. F. L., Duthoo, W., Santens, P., Seidler, R. D., Notebaert, W., & Abrahamse, E. L. (2016). Sequence learning in Parkinson's disease: Focusing on action dynamics and the role of dopaminergic medication. *Neuropsychologia*, 93, 30–39. <https://doi.org/10.1016/j.neuropsychologia.2016.09.027>, PubMed: 27702597
- Schmitz-Hübsch, T., Du Montcel, S. T., Baliko, L., Berciano, J., Boesch, S., Depondt, C., et al. (2006). Scale for the assessment and rating of ataxia: Development of a new clinical scale. *Neurology*, 66, 1717–1720. <https://doi.org/10.1212/01.wnl.0000219042.60538.92>, PubMed: 16769946
- Shin, J. C., & Ivry, R. B. (2003). Spatial and temporal sequence learning in patients with Parkinson's disease or cerebellar lesions. *Journal of Cognitive Neuroscience*, 15, 1232–1243. <https://doi.org/10.1162/089892903322598175>, PubMed: 14709239
- Simcox, T., & Fiez, J. A. (2014). Collecting response times using Amazon Mechanical Turk and Adobe Flash. *Behavior*

- Research Methods*, 46, 95–111. <https://doi.org/10.3758/s13428-013-0345-y>, PubMed: 23670340
- Spencer, R. M. C., & Ivry, R. B. (2009). Sequence learning is preserved in individuals with cerebellar degeneration when the movements are directly cued. *Journal of Cognitive Neuroscience*, 21, 1302–1310. <https://doi.org/10.1162/jocn.2009.21102>, PubMed: 18752399
- Tang, Z. C., Chen, Z., Shi, Y. T., Wan, L. L., Liu, M. J., Hou, X., et al. (2020). Central motor conduction time in spinocerebellar ataxia: A meta-analysis. *Aging*, 12, 25718–25729. <https://doi.org/10.18632/aging.104181>, PubMed: 33232267
- Tran, M., Cabral, L., Patel, R., & Cusack, R. (2017). Online recruitment and testing of infants with Mechanical Turk. *Journal of Experimental Child Psychology*, 156, 168–178. <https://doi.org/10.1016/j.jecp.2016.12.003>, PubMed: 28088051
- Tremblay, P. L., Bedard, M. A., Langlois, D., Blanchet, P. J., Lemay, M., & Parent, M. (2010). Movement chunking during sequence learning is a dopamine-dependant process: A study conducted in Parkinson's disease. *Experimental Brain Research*, 205, 375–385. <https://doi.org/10.1007/s00221-010-2372-6>, PubMed: 20680249
- Tzvi, E., Zimmermann, C., Bey, R., Münte, T. F., Nitschke, M., & Krämer, U. M. (2017). Cerebellar degeneration affects cortico-cortical connectivity in motor learning networks. *Neuroimage: Clinical*, 16, 66–78. <https://doi.org/10.1016/j.nicl.2017.07.012>, PubMed: 28761810
- Wang, R. Y., Huang, F. Y., Soong, B. W., Huang, S. F., & Yang, Y. R. (2018). A randomized controlled pilot trial of game-based training in individuals with spinocerebellar ataxia type 3. *Scientific Reports*, 8, 1–7. <https://doi.org/10.1038/s41598-018-26109-w>, PubMed: 29777115

Uncorrected Proof