Developing Digital Tools for Functional Assessments that Support Clinical Needs

ANONYMOUS AUTHOR(S)

Many people face significant barriers to timely and accurate assessment of cognitive and motor function. Since traditional assessments require expert administration in clinical environments, they are inaccessible to people with mobility limitations, limited access to specialists, or those living in rural and underserved areas. While digital tools exist, they often fall short in clinical settings due to poor alignment with clinician workflows and patient needs. This paper addresses this gap by identifying challenges faced by clinicians through interviews and designing a novel digital tool–Crossroads–that tackles these challenges. Crossroads is a self-administered, touchscreen-based assessment that uses fine-finger tracking to generate detailed measures of cognitive performance. Our studies demonstrate that Crossroads is quick, reliable, and self-administered. Our work illustrates how clinician-informed design of digital tools can advance cognitive assessments, particularly for populations with accessibility and aging-related mobility concerns.

1 Introduction

People often face barriers to timely and accurate assessment of cognitive and motor function. Traditional assessments require expert administration in clinical environments and rely on rigid summative scores [6, 8, 10]. As a result, they are inaccessible to people with mobility limitations, limited access to specialists, or those living in rural and underserved areas. While digital tools exist, many remain underutilized in practice due to poor alignment with clinicians' real-world workflows and limited adaptability to patient experience.

Our research addresses this gap by grounding performance assessments of motor and cognitive function in the needs and practices of specialists. We conducted an interview study with clinicians to understand their workflow and the challenges they face. We found six key challenges that affect the accuracy and timeliness of diagnosis: 1) diagnostic dilemma, 2) subjectivity in assessment, 3) limitations of diagnostic tools, 4) disease-related, 5) logistical, and 6) interpersonal. Then, we developed Crossroads, a self-administered touchscreen-based tool that uses fine-finger tracking to assess cognitive performance. Unlike traditional assessments that rely on summative scores and visual interpretation, Crossroads tracks fine-finger trajectory data and generates continuous (e.g., speed, pauses, time taken, etc.) and episodic (e.g., errors made during the task) measures that can reveal subtle changes in cognitive performance. A user study found the tool to be brief and usable without expert supervision. The measures were reliable and correlated with clinical assessment scores. Together, these two components–clinician-informed design and the development of Crossroads–demonstrate our approach for creating clinically-relevant digital tools that can meet the needs of people with mobility concerns.

2 Related work

People often experience a decline in both cognitive and motor functions due to age or injuries [3, 14], yet assessments remain largely clinic-based, requiring in-person visits and expert supervision. This causes significant barriers for populations with mobility challenges, limited access to specialists, or geographic constraints [5]. Traditional assessments—such as the Unified Parkinson's Disease Rating Scale (UPDRS) for motor performance or the Montreal Cognitive Assessment

Manuscript submitted to ACM

(MoCA) for cognitive performance—are designed to be administered in clinical settings and often lack the flexibility needed for frequent, self-administered use at home [8, 10]. As a result, people may go undiagnosed or face delays in care, especially in the early stages of decline when intervention is most beneficial [1, 9]. Furthermore, traditional assessments provide summative scores that hide differences in strategies used by people [6, 10]. Furthermore, screening thresholds vary by demographic and contextual factors, reducing their diagnostic reliability [4, 12]. These summative scores are tied to expert interpretation, which introduces subjectivity and variability across assessments.

Novel tools–especially those supported by artificial intelligence–are underutilized due to poor alignment with clinical workflows [7, 13]. Digital tools risk creating an extra burden for clinicians when workflows are not considered in the design process [2, 15]. For example, socio-environmental factors such as poor lighting, slow internet, and additional burden on clinicians and patients influenced the performance of a deep learning model for detecting diabetic retinopathy in hospitals [2]. These challenges highlight the need for tools that not only provide clinically-meaningful data but also fit within the real-world constraints and experts' needs. Our research highlights the different challenges clinicians encounter and introduces a novel tool designed to overcome some of these issues.

3 Study 1: Specialists face several challenges in the clinical process

This study aims to identify the challenges encountered by specialists in their clinical workflow. Our research team conducted a semi-structured interview with six movement disorder specialists to understand the challenges they face in their clinical workflows. The specialists had between 5 and 20 years of experience. Five participants were based in the United States, and one had worked in both Nigeria and the United Kingdom. All interviews were conducted over Zoom. The research team recorded both audio and video during the Zoom interviews with participants' consent. The study received approval from the Institutional Review Board (IRB) of the authors' institution. We analyzed the data using an inductive thematic analysis approach. We present six main challenges identified via our study.

3.1 Clinicians face six main challenges when assessing motor performance

The six types of challenges faced by clinicians are: 1) diagnostic dilemma, 2) subjectivity in assessments, 3) limitations of diagnostic tools, 4) disease-related challenges, 5) logistical challenges, and 6) interpersonal challenges.

3.1.1 Diagnostic dilemma. Diagnostic dilemma refers to the uncertainty in arriving at a clear diagnosis. For instance, clinicians face difficulty in determining a clear diagnosis for older adults, whose symptoms may overlap with aging.

> "But if someone who is 70, 80, or 90 starts developing signs of Parkinsonism, distinguishing what's the disease from the effects of normal aging can become more difficult." (P4)

Diagnostic criteria are limited and do not account for individual differences and the unique experiences of each patient.

3.1.2 Subjectivity in assessments. Diagnoses of subtle motor performance–like tremor amplitude, slowness, or rigidity–rely on subjective interpretation, which can vary among clinicians, leading to inconsistent diagnoses.

> "...our Inter-rater reliability is not great depending on the disease... Diagnosis A from one movement specialist, diagnosis B from a second one,..., and I, finally, as the last stop, giving diagnosis D. That's a very typical situation." (P5)

3.1.3 Limitations of diagnostic tools. Most diagnostic tools produce helpful clues, but do not help with early detection. Most diagnoses are based on observable symptoms, which often appear after substantial neurological damage.

"By the time that Parkinson's disease is clinically diagnosed by someone like me in the office, 40 to 60% of dopamine-producing cells at the base of the brain... are already lost." (P5)

Diagnostic tools are not always accurate or reliable. Some provide rich data, but are time-consuming and labor-intensive.

3.1.4 Disease-related challenges. The unpredictable nature of symptoms within the same individual and across patients is a common issue. The variability between and within patients makes diagnosis and tracking difficult.

"...there's no two Parkinson's patients that are the same. And some people might not have your classical symptoms and still have Parkinson's." (P1)

3.1.5 Logistical challenges. Logistical challenges include a shortage of specialists; limited time during clinic visits; and long delays between follow-up appointments. These challenges can lead to misdiagnosis and a delay in treatment. Long intervals between appointments reduce the opportunity to adjust treatment plans based on changes in symptoms.

"We're so busy right now. . . I barely have the capacity to see my own patients, who often go longer than they need to before their next visit." (P1)

3.1.6 Interpersonal challenges. Clinicians rely on patients (and often family members/caregivers) to explain their symptoms, provide an accurate medical history, and other details. This is especially challenging for patients with cognitive problems or limited awareness of their condition.

"So I think part of the challenges is that so much of it is based on history, and not all of our patients are very good historians, or very aware of symptoms. . . Certainly, we have a patient population that deals with cognitive issues, and so getting a reliable history can be difficult" (P6)

3.2 Tools informed by clinical needs can overcome some challenges faced by specialists

Some of the challenges faced by specialists can be addressed with tools that produce fine-granular data, do not require visual interpretation that causes subjective assessments, take a few minutes to complete, and can potentially be used remotely to collect frequent data, as seen in Table 1.

4 Study 2: Digital tools can be designed to overcome some of these challenges

We contribute a novel tool *Crossroads* that overcomes some of the challenges faced by clinicians. Crossroads uses fine-finger tracking on a touchscreen device to assess cognitive performance.

Table 1. Challenges faced by clinicians can be overcome using appropriately designed digital tools.

Clinician Need	**How Tools Can Addresses It**
Diagnostic uncertainty	Tracks fine-grained error and trajectory data that go beyond summative scores
Subjective assessments	Provides concrete measures that do not require visual interpretations
Shortage of specialists	Self-administered: Does not require an expert
Time-constrained workflows	Takes <2 minutes to complete
Long intervals between appointments	Potential remote use for frequent, longitudinal data
Communication gaps with patients	Visualizing error profiles can aid shared understanding.

4.1 The Crossroads System

Crossroads' design derives from the traditional Trail Making Test (TMT), which is a key part of many cognitive assessments. The user is asked to alternatively drag their finger to a dot labeled with a number and then a dot labeled with a letter (Figure 1). The user needs to reach the final dot without lifting their finger while alternating between numbers and letters in ascending order. Each drag (number-to-letter or letter-to-number) is considered to be a single trial. The task is organized into 8 blocks. Each block consists of 10 trials (1-A-2-B-3-C-4-D-5-E-6).

Crossroads uses episodic measures (recorded for specific events or occurrences that happen infrequently) and continuous measures (calculated from the finger trajectory when the user's finger is in contact with the screen) to assess cognitive performance. Unlike traditional assessments that provide limited insights beyond summative scores, Crossroads uses speed-time profiles to better understand errors made by users. Crossroads categorizes errors into mistakes and slips. Mistakes occur when the user touches the incorrect dot. Slips occur when the user initiates movement towards an incorrect dot but changes direction to reach the correct dot. The speed-time profile of the slip trajectory is analyzed to capture features like initiation time, detection time, correction time, slip duration, and confidence gain. Crossroads measures continuous measures like time taken, speed, acceleration, jerk, and pause duration.

4.2 Study

An in-lab user study tested the feasibility of Crossroads' design to assess cognitive performance. Our study aimed to answer the following question: 1) What insights do episodic measures provide? 2) Do continuous measures demonstrate high test-retest reliability across the blocks? 3) Do some measures demonstrate a high correlation with scores on traditional assessments? 4) What did participants think of Crossroads?

The user study consisted of 11 participants who reported having no cognitive or motor impairments. Each participant received a $20 Amazon gift card for participating in the study. The participants were asked to 1) use Crossroads, 2) complete a digital version of the Montreal Cognitive Assessment (MoCA)–a traditional assessment, 3) take part in a short interview, and 4) complete a survey. Consecutive participants were counter-balanced across using Crossroads (or MoCA) before the other. Speed vs time profiles were analyzed to better understand slips made by each participant. Differences in speed values during a slip among participants grouped by MoCA scores were assessed using the Kruskal-Wallis test. The Intraclass Correlation Coefficient (ICC) scores quantified the reliability of continuous measures. Pearson correlation

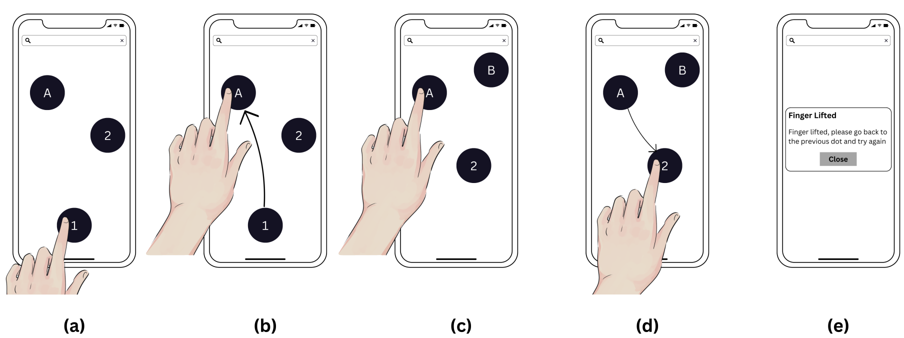

Fig. 1. Crossroads displays three dots labeled with numbers or letters. Users drag their fingers from one dot to another while alternating between numbers and letters in ascending order. Users restart from the previous dot if they lift their finger or touch an incorrect dot.

computed validity between all measures and MoCA scores. Acceptability of Crossroads was studied using participants' interview responses.

4.3 Results

4.3.1 Crossroads provides a rich understanding of errors. Participants with the same MoCA scores demonstrated different numbers of slips and mistakes. Analysis of the speed vs time profile demonstrated that the median speed during a slip increased with higher MoCA scores: participants with lower MoCA scores moved slower and for longer durations (Figure 2). A Kruskal-Wallis test indicated that there was a significant difference in speed across the five MoCA scores (26 to 30), χ^2 (4, N = 5) = 9.48, p = < .001. A moderate correlation was found between MoCA scores and detection time.

4.3.2 Crossroads produces reliable measures. Median jerk (0.94) and median time (0.9) showed excellent reliability (ICC > 0.9). Nine other measures showed good reliability (0.90 > ICC > 0.75) when the median of three blocks was analyzed (Table 1 in the supplementary).

4.3.3 Measures from Crossroads show strong correlation to MoCA scores. Six continuous measures strongly correlated (between 0.7 and 0.89) with MoCA scores (Table 2 in the supplementary).

4.3.4 Acceptability of Crossroads. Using Crossroads took a median of 1 minute 23 seconds. All participants used Crossroads without needing any help from the research team. Four of eleven participants mentioned that they enjoyed the task.

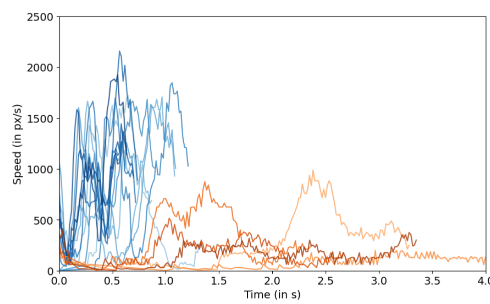

Fig. 2. Speed vs time profiles of individual slips demonstrate varying performance. People with MoCA score = 26 (orange) moved slower and for longer duration, while people with MoCA score = 30 (blue) moved faster and for shorter duration.

5 Clinician needs-informed and remote-friendly digital tools can support assessments of motor and cognitive performance

Overall, Crossroads produces fine-grained data that does not require visual interpretation, takes less than two minutes to complete, and can potentially be used remotely to collect frequent data. These features of Crossroads help it to overcome some challenges faced by specialists. Our work demonstrates the potential of clinician needs-informed digital tools to improve accessibility of assessments. Crossroads addresses some of the challenges identified in the interview study, including subjectivity in evaluation, diagnostic uncertainty, and time-constrained workflows. These findings can expand assessment reach among populations with limited mobility or clinic access. However, the study has limitations. Our sample consisted of normative, tech-literate adults; future work could evaluate usability and reliability among people with cognitive or motor impairments, as well as those with limited digital experience. Estimating clinical scores from tracked measures (like Hevelius [11]) will enhance Crossroads' relevance in healthcare settings and for remote use. Future work can develop interpretable regression models that estimate standard cognitive scores (like MoCA scores) from Crossroads' measures. Future longitudinal at-home studies could collect self-reports on contextual factors and administer the tool under varying conditions to study its generalizability.

Our findings emphasize that accessibility in digital health extends beyond interface design—it requires aligning with the real-world experiences of stakeholders, including patients and clinicians. Designing for real-world use is essential for tools to achieve meaningful impact in clinical practice and to effectively support the needs of the user.

References

[1] Halima Amjad, David L Roth, Orla C Sheehan, Constantine G Lyketsos, Jennifer L Wolff, and Quincy M Samus. 2018. Underdiagnosis of dementia: an observational study of patterns in diagnosis and awareness in US older adults. *Journal of general internal medicine* 33 (2018), 1131–1138.

[2] Emma Beede, Elizabeth Baylor, Fred Hersch, Anna Iurchenko, Lauren Wilcox, Paisan Ruamviboonsuk, and Laura M Vardoulakis. 2020. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–12.

[3] Sandra A Black and Ronald D Rush. 2002. Cognitive and functional decline in adults aged 75 and older. *Journal of the American Geriatrics Society* 50, 12 (2002), 1978–1986.

[4] Nicole Carson, Larry Leach, and Kelly J Murphy. 2018. A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *International journal of geriatric psychiatry* 33, 2 (2018), 379–388.

[5] R Turner Goins, Kimberly A Williams, Mary W Carter, S Melinda Spencer, and Tatiana Solovieva. 2005. Perceived barriers to health care access among rural older adults: a qualitative study. *The Journal of Rural Health* 21, 3 (2005), 206–213.

[6] Edith Kaplan. 1988. The process approach to neuropsychological assessment. *Aphasiology* 2, 3-4 (1988), 309–311.

[7] Saif Khairat, David Marc, William Crosby, Ali Al Sanousi, et al. 2018. Reasons for physicians not adopting clinical decision support systems: critical analysis. *JMIR medical informatics* 6, 2 (2018), e8912.

[8] P1 Martínez-Martín, A Gil-Nagel, L Morlan Gracia, J Balseiro Gómez, J Martinez-Sarries, F Bermejo, and Cooperative Multicentric Group. 1994. Unified Parkinson's disease rating scale characteristics and structure. *Movement disorders* 9, 1 (1994), 76–83.

[9] Loren Mowszowski, Jennifer Batchelor, and Sharon L Naismith. 2010. Early intervention for cognitive decline: can cognitive training be used as a selective prevention technique? *International psychogeriatrics* 22, 4 (2010), 537–548.

[10] Ziad S Nasreddine, Natalie A Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L Cummings, and Howard Chertkow. 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53, 4 (2005), 695–699.

[11] Vineet Pandey, Nergis C Khan, Anoopum S Gupta, and Krzysztof Z Gajos. 2023. Accuracy and reliability of at-home quantification of motor impairments using a computer-based pointing task with children with ataxia-telangiectasia. *ACM Transactions on Accessible Computing* 16, 1 (2023), 1–25.

[12] Janice M Ranson, Elżbieta Kuźma, William Hamilton, Graciela Muniz-Terrera, Kenneth M Langa, and David J Llewellyn. 2019. Predictors of dementia misclassification when using brief cognitive assessments. *Neurology: Clinical Practice* 9, 2 (2019), 109–117.

[13] Venkatesh Sivaraman, Leigh A Bukowski, Joel Levin, Jeremy M Kahn, and Adam Perer. 2023. Ignore, trust, or negotiate: understanding clinician acceptance of AI-based treatment recommendations in health care. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[14] William C Walker and Treven C Pickett. 2007. Motor impairment after severe traumatic brain injury: A longitudinal multicenter study. *Journal of Rehabilitation Research & Development* 44, 7 (2007).

[15] Liuping Wang, Zhan Zhang, Dakuo Wang, Weidan Cao, Xiaomu Zhou, Ping Zhang, Jianxing Liu, Xiangmin Fan, and Feng Tian. 2023. Human-centered design and evaluation of AI-empowered clinical decision support systems: a systematic review. *Frontiers in Computer Science* 5 (2023), 1187299.