



OPEN ACCESS

EDITED BY

Ayşe Altan-Atalay,
Kadir Has University, Türkiye

REVIEWED BY

Fahri Safa Cinarli,
Inonu University, Türkiye
Maria M. Johansson,
Linköping University, Sweden

*CORRESPONDENCE

Kritika Rana
✉ k.rana@westernsydney.edu.au

RECEIVED 30 May 2025

ACCEPTED 26 October 2025

PUBLISHED 17 November 2025

CITATION

Ginige JA, Rana K, Belcastro M, Gunaratne P, Lillis S, Numbers KT, Lin P-H, Nguyen K, Tanvir M and Reppermund S (2025) Technical development and user experience evaluation of the Computerised version of the Sydney Test of Activities of daily living in Memory disorders (C-STAM).
Front. Digit. Health 7:1637916.
doi: 10.3389/fdgth.2025.1637916

COPYRIGHT

© 2025 Ginige, Rana, Belcastro, Gunaratne, Lillis, Numbers, Lin, Nguyen, Tanvir and Reppermund. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Technical development and user experience evaluation of the Computerised version of the Sydney Test of Activities of daily living in Memory disorders (C-STAM)

Jeewani Anupama Ginige¹, Kritika Rana^{1,2*}, Marcus Belcastro¹, Pahan Gunaratne¹, Sam Lillis¹, Katya T. Numbers³, Ping-Hsiu Lin³, Kathy Nguyen³, Minal Tanvir³ and Simone Reppermund³

¹School of Computer, Data and Mathematical Sciences, Western Sydney University, Penrith, NSW, Australia, ²Translational Health Research Institute, Western Sydney University, Campbelltown, NSW, Australia, ³Centre for Healthy Brain Ageing, Faculty of Medicine and Health, School of Clinical Medicine, Discipline of Psychiatry and Mental Health, University of New South Wales, Sydney, NSW, Australia

The assessment of functional ability in older adults is critical for the early detection and management of age-related cognitive impairment, such as dementia and mild cognitive impairment (MCI). This paper presents a developmental and pilot user experience study of the online implementation and initial user experience evaluation of the Computerised version of the Sydney Test of Activities of daily living in Memory disorders (C-STAM), a new computerised diagnostic screening tool designed to assess functional ability in older adults with and without cognitive impairment. The technical development of the C-STAM involved translating the original STAM tasks into computerised versions, resulting in a scalable, accessible, and user-friendly tool with the potential for integration into regional and remote clinical practices. Key features of the C-STAM include automated scoring, hint options, text-to-speech functionality, a text-magnification tool, and other accessibility enhancements to accommodate diverse physical and cognitive limitations. Within the pilot study of the larger C-STAM trial, secondary feedback was sought from 30 participants ($n = 13$ with and $n = 17$ without cognitive impairment) specifically to assess the C-STAM tool's user experience (UX). This exercise did not evaluate dementia outcomes, which is planned to be reported separately. UX analysis results presented in this paper indicated a positive experience, with an average score of 5.8 ± 1.3 on a seven-point Likert scale, reflecting high perceived usability and user satisfaction. The pilot normal cognition group ($n = 17$) reported higher satisfaction, with an overall score of 6.4 ± 0.4 . In comparison, the pilot MCI group ($n = 6$) and pilot dementia group ($n = 7$) reported lower scores of 5.1 ± 1.1 and 4.9 ± 2.1 , respectively. Feedback from participants was instrumental in shaping each iteration of the tool and refining the updated version of the C-STAM that will be presented in the main validation study.

KEYWORDS

user interface design, user experience evaluation, activities of daily living, computerised assessments, older adults

1 Introduction

The assessment of functional ability in older adults is a critical component in the diagnosis and management of age-related cognitive impairment, particularly dementia and mild cognitive impairment (MCI) (1, 2). These conditions are characterised by a decline in cognitive functions such as memory, attention, and executive function, which can significantly impact an individual's ability to perform basic activities of daily living (ADL) and instrumental activities of daily living (IADL) (3). IADLs represent complex everyday activities such as handling finances, medication management, and shopping, which are essential for independent living and maintaining quality of life (4). Importantly, the main diagnostic criterion that distinguishes a diagnosis of MCI from dementia is the maintenance of functional independence in ADL and IADL (5, 6). Given the increasing prevalence of dementia due to the rising global aging population, there is a substantial burden on healthcare systems worldwide (7). There are limited disease modifying pharmacological interventions for dementia. Therefore, early detection of cognitive decline is crucial for enabling timely interventions that can slow disease progression, alleviate symptoms, and provide support for patients and carers (8, 9). Given the key role that functional ability plays in delineating MCI and dementia, there has been growing research on the development of validated tools to assist with the early detection of cognitive decline based on changes in IADL performance (10).

Traditional methods for assessing IADLs often rely on subjective reports from the person or an informant (e.g., spouse, friend or family member), which can be biased and influenced by various factors such as current mood, personality, and/or lack of insight and may not accurately reflect actual IADL performance (11–13). On the other hand, performance-based measures allow real-time standardised evaluation of IADL performance, where individuals perform a range of IADL tasks under the direct observation of the examiner (14). However, many existing performance-based IADL assessments are too time-consuming or resource-intensive for everyday practice, some requiring administration by specially trained professionals and with insufficient evidence on their psychometric properties (10, 15).

Consequently, the Sydney Test of Activities of daily living in Memory disorders (STAM) was developed and validated as a performance-based instrument to assess everyday activities in a time-efficient and reliable way (15). The STAM requires participants to perform mock versions of nine daily tasks, such as making a phone call and managing medications, that are observed, timed and subsequently scored by a trained researcher. The STAM has demonstrated good psychometric properties, including excellent specificity in distinguishing between normal cognition, MCI and dementia (15), and it has been shown to predict dementia over 4 years (16). Therefore, the STAM serves as a valuable diagnostic screening tool for assessing functional impairment and aiding in the early-stage diagnosis of dementia (15).

Despite the strengths of the STAM as a validated tool to assess IADL, there are a few inherent disadvantages in today's context. The logistical challenges associated with its administration, such as

setup time and resource requirements, may impact its efficiency in time-constrained environments. Furthermore, with technological advancements influencing daily activities, many real-life IADLs are now frequently performed on a computer, such as paying bills and shopping, making some tasks included in the STAM (e.g., looking up a number in a phone book) less relevant. In addition to improving efficiency, the transition to digital assessments enhances ecological validity by allowing tasks to be performed in formats that closely mirror contemporary daily activities, such as using digital communication platforms or online financial tools. This digital simulation provides a more authentic representation of functional performance in modern living contexts, particularly for older adults who are increasingly interacting with technology in daily life.

In response to these challenges, a computerised version of the STAM was envisioned to realistically simulate relevant IADLs and accurately assess functional performance in older people with or without cognitive impairment. The development of a scalable computerised assessment for IADLs, which builds upon the foundation of the validated STAM, allows for a streamlined and standardised functional assessment. A computerised assessment has several advantages over traditional performance-based measures, including a shorter assessment time that facilitates administration in time-restricted settings and greater cost-efficiency due to automated administration and scoring, eliminating the need for additional resources (e.g., props) or training (17). A computerised version of the STAM supports remote assessments, which have become increasingly important due to the COVID-19 pandemic, enabling clinicians to reach individuals in remote and rural areas who may lack access to specialists and face long waiting and travel times (18). Moreover, the computerised assessment provides greater measurement precision and automated data exporting and reporting, thereby enhancing accessibility, efficiency, and reliability in both clinical and research settings (17).

This work presents the online implementation efforts and initial user experience evaluation of the Computerised version of the Sydney Test of Activities of daily living in Memory disorders (C-STAM). We hypothesise that the C-STAM would demonstrate acceptable user experience (UX) across participant groups, providing preliminary evidence to support its future use in clinical and research settings. Thus, the main objective of this project was to translate, yet accurately represent, the activities and functional domains captured from the original STAM into online (computerised) versions, with the goal of developing a scalable and accessible tool to assess functional ability that is feasible and acceptable to target-end users (i.e., older adults with and without cognitive impairment) that has the potential for future integration into clinical practices.

2 Materials and methods

2.1 Development and design process

The online implementation of the nine new activities for the C-STAM was conducted with the active involvement of all stakeholders, including clinicians and subject matter experts,

lived experience consumers and carers, and clinical research and Information Technology (IT) teams. Initially, all tasks of the original STAM were reconceptualised for the C-STAM prototype using Microsoft PowerPoint slides, created by the clinical research team to communicate their ideas to subject matter experts. These prototypes were assessed by an international panel of experts in IADL, including researchers, occupational therapists, neuropsychologists, and IT specialists for feasibility, acceptability, and validity via a modified Delphi survey. The findings of the Delphi study are planned to be presented in a separate publication. The final nine items from the Delphi survey were then presented to an advisory group consisting of individuals with and without cognitive decline, as well as carers, to assess their difficulty and feasibility. Throughout this process, the IT design team worked closely with the clinical research team to explore how the original STAM activities could be effectively implemented as computerised activities and what feedback from the consumer advisory group could be effectively implemented. During this period, the IT team also began technical development and infrastructure setup for the C-STAM. This involved organising development and production server environments and selecting the appropriate software and development tools.

After finalising the design elements of the user interface for each item for the C-STAM based on feedback from the Delphi panel and advisory group, the computerised version of all assessment items was implemented. The implementation process involved several testing iterations to refine the C-STAM activities. Feedback loops from various user groups informed each iteration. The C-STAM was first tested for feasibility and user experience (UX) in a pilot study with participants with ($n = 13$) and without ($n = 17$) cognitive impairment. There were three groups of participants in the pilot study: (i) pilot dementia group; (ii) pilot mild cognitive impairment (MCI) group, and (iii) pilot normal cognition group. Purposive sampling was utilised to recruit older adults for the pilot and main validation study via self-selection, whereby interested individuals responded to flyers and advertisements by contacting the research team via email or phone. Eligible participants were at least 60 years old, could communicate in English without an interpreter, could use a computer mouse and keyboard (basic skills), lived in Sydney, and had someone who knew them well and was in contact with them at least once a week. Pilot participants' ($n = 30$) feedback and user experience questionnaire responses were used to develop a final version for a full validation study to test the psychometric properties of the C-STAM.

2.2 Choice of technology framework

This project utilised two primary technological frameworks: Unity game development platform (19) for the front end and PHP (Hypertext Processor) Laravel web development framework for the back end (20). Unity is a cross-platform game engine widely used to develop interactive content. It supports a broad range of browsers and devices, making it an ideal choice for front-end development due to its graphical capabilities and flexibility. Despite

certain limitations associated with using Unity as a game development environment, including runtime overhead, it was selected for its flexibility and suitability for creating interactive, graphics-intensive front-end applications. PHP Laravel is a web application framework that supports accelerated development, scalability options and security, making it highly suitable for backend research data management. A limitation of this framework is that newer technologies may be available, however, it was chosen for its cost-effectiveness, implementation suitability, and overall appropriateness for the project. While other technologies, such as React for the front end and Node.js for the back end, could have been considered, Unity was chosen for its real-time interaction capabilities, and Laravel was selected for its comprehensive ecosystem and security features, which were suitable for efficient data handling and scalability in the C-STAM project.

In addition to the primary technologies, several artificial intelligence (AI) tools were integrated into the project to enhance audio and visual elements. For text-to-speech conversion, Amazon Web Services (AWS) Polly (21) and ElevenLabs (22) were utilised, providing clear and lifelike audio output. For image generation, DALL-E (23) was employed, enabling the creation of unique, non-copyrighted visual content. The use of these technologies significantly accelerated the development process, allowing for high-quality, cost-effective visuals and audio.

2.3 User experience evaluation

For the pilot study, participants completed the C-STAM as well as a brief cognitive screening test (24), an abbreviated version of an Occupational Therapy (OT) test of daily function (25), and questions about mood and recent health changes. Prior to the in-person assessment, pilot participants also completed a battery of questionnaires about their health, medical history, social and other lifestyle factors, and levels of physical activity.

When completing the C-STAM, participants' responses, keyboard strokes, mouse clicks, hints accessed, and time taken to complete each task were all recorded. Subsequently, users were asked to complete a user experience assessment questionnaire for the C-STAM tool. The questionnaire was adapted and modified from the IBM Computer Usability Satisfaction Questionnaire (CUSQ) (26) and the Post-Study System Usability Questionnaire (PSSUQ) (27), which are widely used questionnaires for assessing user satisfaction with computer systems (28). The questionnaire consisted of 18 items, each rated on a seven-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Higher scores indicate greater degrees of perceived usability or satisfaction with the C-STAM tool. The scores for the 18 items were used to calculate the overall score, as well as three subscales, as follows:

- Overall: The average scores of questions 1 to 18
- System Usefulness (SYSUSE): The average scores of questions 1 to 7
- Information Quality (INFOQUAL): The average scores of questions 8 to 13

- **Interface Quality (INTERQUAL):** The average scores of questions 14 to 17

The overall and subscale scores were compared to identify areas that needed improvement. In addition to the 18 questionnaire items, users were asked to respond to two additional open-ended questions to list the negative aspects and the most positive aspects of the C-STAM tool. One-way ANOVAs were used to examine mean differences between groups for continuous variables, with results presented as mean \pm standard deviation (SD). For variables showing significant differences in the ANOVA, Tukey HSD *post-hoc* tests were conducted to identify pairwise group differences. Categorical variables were summarised using frequencies and percentages. All analyses were performed using the latest version of R Studio (v2025.09.2 + 412), with a significance level set at $p < 0.05$. The responses to the open-ended questions were analysed using an inductive thematic approach, where two researchers independently coded the data to identify patterns and emerging themes. The coding was then compared and refined through discussion to reach consensus. The sample included a mix of older adults with normal cognition, mild cognitive impairment, and dementia, providing sufficient diversity to explore key themes related to usability, accessibility, and user experience. At the time of this writing, the main validation study is in progress; thus, user experiences concerning the main study participants are not reported here.

3 Development and design of the C-STAM

3.1 Development of the C-STAM from STAM

The development of the C-STAM marked a significant transition from the performance-based, in-person STAM that was observed and scored by an assessor to a self-completed computerised tool that is autonomously administered and scored. This process involved reimagining the original STAM assessment tasks within a digital framework, ensuring that the core functional domains were preserved while utilising the advantages of computer-administered assessments, such as standardisation, precision, and the ability to collect and score detailed response data automatically.

The original STAM consists of nine tasks assessing the following domains of function: communication (looking up and making a phone call), dressing (putting on a shirt), handling finances (paying a bill by check), managing everyday activities (preparing the check for mailing), orientation (time orientation—reading the time and setting the alarm), medication management (managing medications using a dispenser), shopping (choosing items to make a simple recipe), counting money (calculating cost and counting money), and memory (recalling activities completed) (15).

During the transition from the STAM to the C-STAM, the tasks and scenarios from the original STAM were adapted for a digital environment, ensuring they remained relevant and realistic in a computerised format. This involved redesigning

tasks to be interactive and engaging while retaining their core objectives. For example, a task that originally involved physically sorting medication was transformed into a virtual task in which users dragged and dropped medication icons into the correct compartments of a digital pillbox.

The C-STAM consists of two practice tasks and nine assessment tasks, outlined as follows:

- Practice Task 1—Making a beverage (domain: managing everyday activities)
- Practice Task 2—Setting a timer (domain: orientation)
- Task 1—Look up GP's (i.e., general practitioner) phone number (domain: communication)
- Task 2—Call GP to refill script (domain: communication)
- Task 3—BPay (i.e., bill payment) for electricity bill (domain: handling finances)
- Task 4—Online grocery shopping (domain: shopping)
- Task 5—Calculating shopping total after “special” items added (domain: counting money)
- Task 6—Dressing (domain: dressing)
- Task 7—Navigating directions (domain: managing everyday activities)
- Task 8—Managing medications (domain: medication management)
- Task 9—Memory task (domain: memory)

The scoring of the C-STAM is similar to that of the original STAM (15). Each task is scored on a 4-point scale, and points are awarded for completing the various components of each task correctly, with a point awarded for not using any hint. For each task, users receive 1 point for each correct component (maximum 4 points per item) and 0 points for an incorrect component or using a hint/s. The total score is calculated as the sum of scores for all nine tasks, with total scores ranging from 0 to 36, and higher scores indicating better performance.

3.2 Technical development and user interface design

The online implementation and user interface design of the C-STAM integrated various technologies and development principles, resulting in a comprehensive and user-friendly assessment tool. The C-STAM was developed using the Unity application, a versatile platform commonly employed in video games and interactive software development (19). The C-STAM was deployed as a web-based application using WebGL, enabling it to run on modern web browsers across various operating systems and platforms. The C-STAM assessment tool guides users through a series of interactive tasks, provides scoring based on predefined criteria, offers assistance through hints, and accommodates users with various accessibility needs (e.g., text-to-speech and text-magnifying options).

The architecture of the C-STAM comprises two main components: a front-end Unity application and a back-end PHP Laravel server. The front-end component is the interactive interface through which users engage with a series of tasks designed to assess

their functional abilities. The Unity application encompasses the bulk of the assessment's logic, including scene/task traversal, scoring mechanisms, and providing hints and narration to guide users through the assessment. On the other hand, the back-end PHP Laravel server is responsible for storing the data from the assessment and other pertinent tracking information. In line with the principle of Separation of Concern (SOC), the server is primarily focused on managing data operations (Create, Read, Update and Delete—CRUD) and does not handle any of the assessment logic (29). This clear delineation of responsibilities ensures that the Unity application can concentrate on the assessment's functionality while the Laravel server efficiently handles data management. This design choice also reduces the computational load on the server, enabling it to hold a larger capacity of incoming requests with minimal performance impact (30).

A key aspect of the user interface design is the C-STAM scoring system, which was implemented through both active and passive methods. Active scoring assigns a score when a specific event is triggered, such as clicking a button. In contrast, passive scoring calculates a score when the user moves to the next task. The design of the scoring system was focused on simplicity and modularity, with each scoring criterion linked to a specific function, making the scoring logic easy to debug, unit test, and modify. This design philosophy also prevented the embedding of scoring logic deep within other functions in the code, thereby reducing the likelihood of errors and making the system more maintainable.

Given how significantly the use of hints impacts participants' scores on the C-STAM, special consideration was provided to develop the hint features. When a user clicks the hints button, they are first asked for confirmation: "You are about to use the hints for this task. Are you sure you want to continue?" The scoring criteria were developed in such a way that if a user utilises any hints for a task, they lose a point for that particular task but are not penalised multiple times if multiple hints are used for the same task. Additionally, a point for not using hints is only awarded if the attempt has begun, which was a feature designed to ensure fairness in scoring. Conversely, if a user clicks the "Skip" button, they are prompted for confirmation to skip the task, along with an indication that they will not be able to go back. However, they still receive points depending on how far they have progressed in the task.

To accurately score the C-STAM, it was essential to track additional variables that could impact participants' scores. These variables include whether a task was skipped, if and how many of each type of hint were used, if the task attempt was begun, or whether the zoom-in or Text-to-Speech (TTS) functions were utilised. Key factors, such as the time taken to complete each task, the use of hints, and the total score, are incorporated into the scoring models, enabling adjustments to the scores as needed, making it vital to capture this information. Additionally, this detailed tracking provides a comprehensive understanding of user interactions, along with informing us about different users' needs and behaviours. Furthermore, different tasks within the assessment have specific scoring logic. For instance, in the "Call GP to refill script" task, if a user calls an incorrect number

and then the correct number, they still lose a point. Conversely, in the "Online grocery shopping" task, if a user picks the wrong item but corrects it, they still receive the point. This distinction ensures that mistakes in certain tasks, where actions are reversible, do not unfairly penalise the user, whereas irreversible mistakes in other tasks do impact the score.

Another significant aspect of the user interface design is the event collection and replaying feature, which was deemed essential for debugging issues in production environments and providing concrete evidence for bug reports. Moreover, this feature enabled the research team to review user experiences and, in some instances, help the team resolve ambiguities in participant scores when needed. The Unity app, which forms the core of the C-STAM, enabled the recording of user interactions, such as mouse clicks, keyboard inputs, and scroll actions. This data is then stored and uploaded in a single web request alongside the scoring for each assessment task, which could then be used to recreate the entire user experience, providing valuable insights into user behaviour and tool performance.

String tables provided by Unity were implemented to further streamline the user interface design of the C-STAM (31). String tables are a data structure that allows the IT team to conveniently store and access lists of strings, along with enabling easy text modification based on feedback from users and advisory groups. The localisation tables in Unity were utilised to implement narration within the C-STAM assessment, which streamlined the process of displaying text and integrating text-to-speech functionality. In the early stages of the project, any adjustments to the wording of prompts required involvement from the IT team for modification and inclusion in the subsequent releases. Therefore, an operational pipeline was developed for narration text to ensure that any necessary changes to the text could be made seamlessly after development. This pipeline enabled the research team to modify an online spreadsheet, which was accessed by the IT team while building a new release, and the changes were subsequently implemented in the Unity application, and the text-to-speech assets were updated programmatically.

3.3 Overview of the C-STAM user interface

As shown in Figure 1, users are first provided with the option of selecting between two avatars (i.e., Alex and Sam). Users are then prompted to enter full-screen mode by pressing the blue button on the bottom-right of the screen. The user interface incorporates several accessibility features, including a zoom button, a speaker button, and a hints button. Positioned on the left side of the interface is a zoom or magnifying glass icon, and activating this feature provides a magnifier for enlarging small text. The speaker icon is the Text-to-Speech button, which speaks out the on-screen text for those with vision issues or who prefer to listen instead of reading. The hints button is the lightbulb icon, which is available for users if they are stuck on a task and need some advice on what to do next. However, users are encouraged to attempt tasks without using hints. Moreover, each task within the interface includes both a "Done" and a

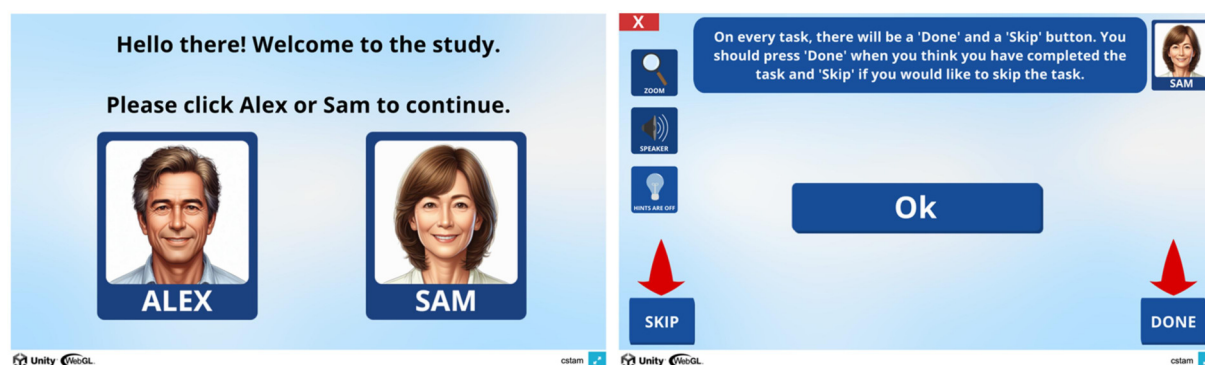


FIGURE 1

The user interface of the C-STAM along with accessibility features. Screenshot(s) from C-STAM application, UNSW Sydney (<https://www.unsw.edu.au/research/c-stam>).

“Skip” button. Users are advised to click “Done” when they believe they have completed a task and “Skip” when they would like to skip the task. The red cross button on the top left is the exit button, which users are advised to press only if they want to abandon all progress and quit the C-STAM. An iterative user-centred design process was employed for the design of the C-STAM user interface, informed by continuous feedback from pilot participants, advisory board members, and subject matter experts at multiple development stages of the C-STAM.

3.4 Design considerations for the C-STAM

Several critical considerations were essential in designing the C-STAM to ensure its effectiveness and accessibility for a diverse range of users, particularly older adults with cognitive impairment. Key considerations included integrating accessibility features to cater to users with various impairments, such as vision or hearing challenges or difficulties with thinking. These accessibility features were primarily based on the consumer advisory group and pilot participants’ feedback.

One of the key accessibility features integrated into the C-STAM is the Text-to-Speech (TTS) functionality, which is particularly important for users with vision impairments or reading difficulties who rely on audio narration (32). AWS Polly’s neural-voice TTS service employs a neural network to enhance the quality of the produced speech, which provides a more natural and comforting voice compared to traditional TTS solutions, thereby putting users at ease during the assessment (21, 33). Moreover, the use of voice synthesis AI models to replicate the voice of a voice actor and enabling the conversion of arbitrary text into realistic spoken words was essential for providing a more authentic auditory experience for users. ElevenLabs is an AI voice generator known for its ability to produce highly realistic and customisable speech from text inputs (22). The use of ElevenLabs allowed us to generate a native Australian speaker’s voice, which was aimed at further enhancing comfort, relatability and the overall user experience (22). The TTS feature in the C-STAM is used specifically for narration, rather than reading all arbitrary text on the screen as with a typical screen

reader. This narration covers approximately 80% of the text within the assessment, representing a significant step in enhancing the accessibility of the C-STAM assessment tool. The TTS functionality was especially important in tasks with lengthy instructions such as the “Look up GP’s phone number” task, along with tasks with auditory component such as “Call GP to refill script” where users dial the number, listen carefully to the options, and select the correct option to renew their prescription and select the correct GP.

A magnification tool was integrated into the Unity application to cater to the needs of users with vision impairment (34). This tool enables users to enlarge text and other elements within the assessment, making it easier to read and engage with the content. The implementation of the magnification tool involved creating a new canvas with a User Interface (UI) image and camera, adjusting the camera settings for zoom, and ensuring that this magnified view tracks the cursor. By embedding the magnification feature into every scene as required, the tool becomes more accessible for individuals with vision challenges.

Another essential design consideration was the hints feature, which was aimed at guiding the users through the C-STAM assessment. When users seek to use the hints for a particular task, they are first asked to select which type of hints they want to see. As shown in Figure 2, two options are presented: “I don’t know how to do it” and “I don’t know what to do”. The first type of hint seeks to provide technical assistance to help the user understand which buttons to press, catering to those who may have less computer literacy and/or confidence. If a user selects this type of hint, they are shown a brief animated tutorial demonstrating how to click, drag and drop, or type using visual representations of a computer, mouse, and keyboard icons. The second type of hint seeks to assist users who may have forgotten what they were supposed to do or need clarification on the task’s objective and was designed to assist users with memory difficulties. These two distinct types of hints were designed to allow users to progress when stuck within a task without needing intervention from the interviewer or clinician.

Key design considerations in the development of the C-STAM also included aspects such as contrast, colour, vibrancy, and realism. The design focused on enhancing contrast and avoiding problematic colour combinations, such as light shades of blue and

white with minimal contrast, which can be challenging for some older users to differentiate. An essential design principle was ensuring that every interface element had a background, thus making designs independent of their backdrop and avoiding colour contrast issues. Monotone colours such as black and white were also avoided to ensure the assets appeared as realistic as possible, mimicking the challenges in everyday life accurately. Moreover, based on published research, colour vibrancy was used to create a sense of liveliness to help users feel comfortable, confident, and engaged in the assessment (35).

The use of AI-generated imagery, particularly through the DALL-E model, allowed for the efficient creation of assets for the C-STAM (23, 36). The AI-generated images were especially effective for depicting everyday objects, requiring minimal post-processing. DALL-E was particularly useful for tasks that required non-copyrighted content, such as the “Online grocery shopping” task where users were provided with images of various grocery items (e.g., Roma tomatoes, deli sliced boneless ham, sliced peaches tin, pressed orange juice) to be added to the shopping cart as part of the assessment (Figure 3). However, this approach was less effective in generating human faces and

features, such as the avatar profile images, highlighting a limitation in the current capabilities of AI-generated imagery.

Another essential design consideration was perpetual context, especially for users with cognitive impairment such as short-term memory loss. The C-STAM interface was designed to be familiar, consistent, and obvious in every scene. Clear sections were defined, and text was added to interface buttons to facilitate understanding and navigation. Avoiding the use of images or icons alone for buttons further reduced ambiguity, particularly for users with lower technical literacy (37).

3.5 Evolution of user interface designs of the C-STAM

Figures 4, 5 showcase the major evolutions in the C-STAM interface as per feedback from consumers, advisory board members, and subject matter experts from the research team, along with the participants with and without cognitive impairment from the pilot study. Initially, all tasks embedded in the C-STAM were conceptualised in Microsoft PowerPoint and subsequently implemented as a Unity application.



FIGURE 2

The hints feature of the C-STAM along with the variation of hints provided. Screenshot(s) from C-STAM application, UNSW Sydney (<https://www.unsw.edu.au/research/c-stam>).

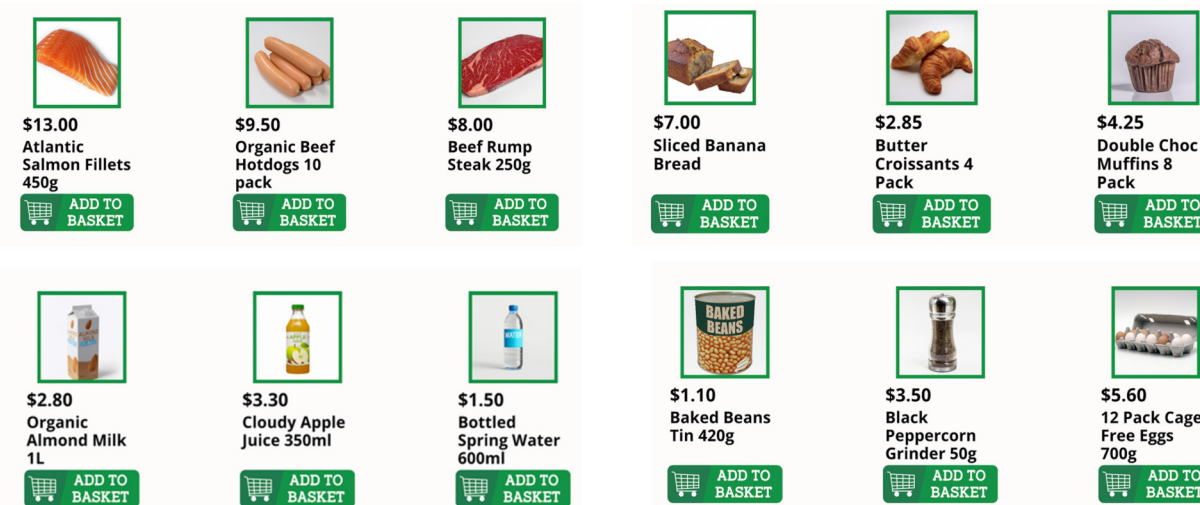


FIGURE 3

Example images of realistic grocery items generated with DALL-E (<https://openai.com/index/dall-e-2/>) and used in the “Online Grocery Shopping” task in C-STAM.

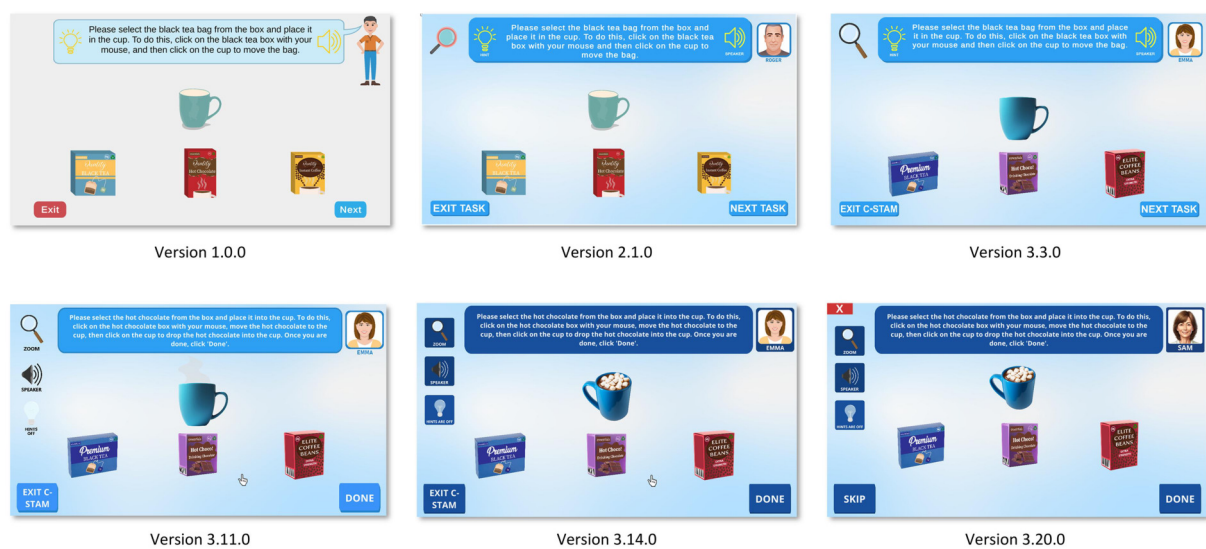


FIGURE 4

Evolution of user interface of the C-STAM. Screenshot(s) from C-STAM application, UNSW Sydney (<https://www.unsw.edu.au/research/c-stam>).

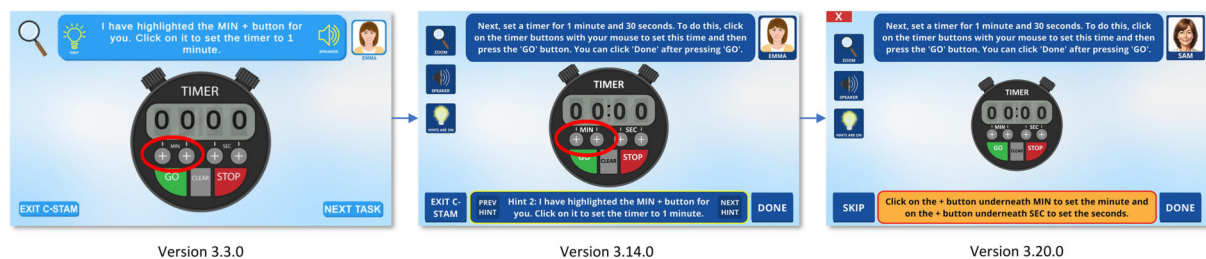


FIGURE 5

Evolution of hints in the user interface of the C-STAM. Screenshot(s) from C-STAM application, UNSW Sydney (<https://www.unsw.edu.au/research/c-stam>).

Figure 4 illustrates the evolution of the user interface of the C-STAM, taking “Practice Task 1—Making a beverage” as a reference. Significant changes are evident from Version 1.0.0 to 2.1.0, including alterations in the colours within the interface, integration of the magnification tool, and the inclusion of the avatar profile image (i.e., Roger). Version 3.3.0 showcases further refinement of images (e.g., cup and beverage boxes) and icons (e.g., magnification tool). In Version 3.11.0, the hints and speaker buttons were relocated from the task description panel to the left of the interface (below the magnification tool) to consolidate accessibility features. Notable changes from Version 3.11.0 to 3.14.0 include adjustments to the colour profile for enhanced contrast (i.e., light blue to dark blue) and modifications to images for a more realistic appearance (e.g., marshmallows within the cup). Finally, essential modifications in Version 3.20.0 include the incorporation of a more realistic profile image of the avatar (without appearing ageist), and the change of avatar names to include gender-neutral names (i.e., Roger and Emma changed to Alex and Sam). Figure 5 further demonstrates the evolution of the hints feature in the user interface, using “Practice Task 2—Setting a timer” as a reference, with hints replacing the task description in Version 3.3.0, the hints panel appearing at the bottom

of the screen in Version 3.14.0, and the hints panel being further highlighted for clarity in Version 3.20.0. Some challenges were encountered during development, including ensuring accessibility for users with visual or cognitive impairments, creating realistic AI-generated imagery for avatars and objects, and designing fair scoring logic for tasks with reversible and irreversible actions. These were addressed through iterative testing, integration of TTS and magnification features, consultation with consumer advisory groups, and continuous refinement of scoring criteria to ensure usability.

4 User experience evaluation of the C-STAM

4.1 Quantitative results of user experience evaluation

A total of 30 participants completed the user experience assessment questionnaire as part of the feasibility testing of the C-STAM tool in a pilot study involving individuals with and without cognitive impairment. Of the 30 participants, 23.3%

($n = 7$) were in the pilot dementia group, 20.0% ($n = 6$) were in the pilot mild cognitive impairment (MCI) group, and 56.7% ($n = 17$) were in the pilot normal cognition group. As shown in Table 1, the average age of participants was 71.7 ± 5.7 years, and 50% ($n = 15$) were female. The majority of participants identified as Caucasian (90%; $n = 27$) and had a tertiary education (80%; $n = 24$).

The summary of results from the user experience assessment questionnaire is presented in Figure 6. When considering all 30 participants across groups, the overall average scores for System Usefulness (SYSUSE), Information Quality (INFOQUAL), and Interface Quality (INTERQUAL) were 5.6 ± 1.5 , 5.9 ± 1.3 , and 6.1 ± 1.3 , respectively, resulting in an overall score of 5.8 ± 1.3 . Higher scores across all categories were reported by the pilot

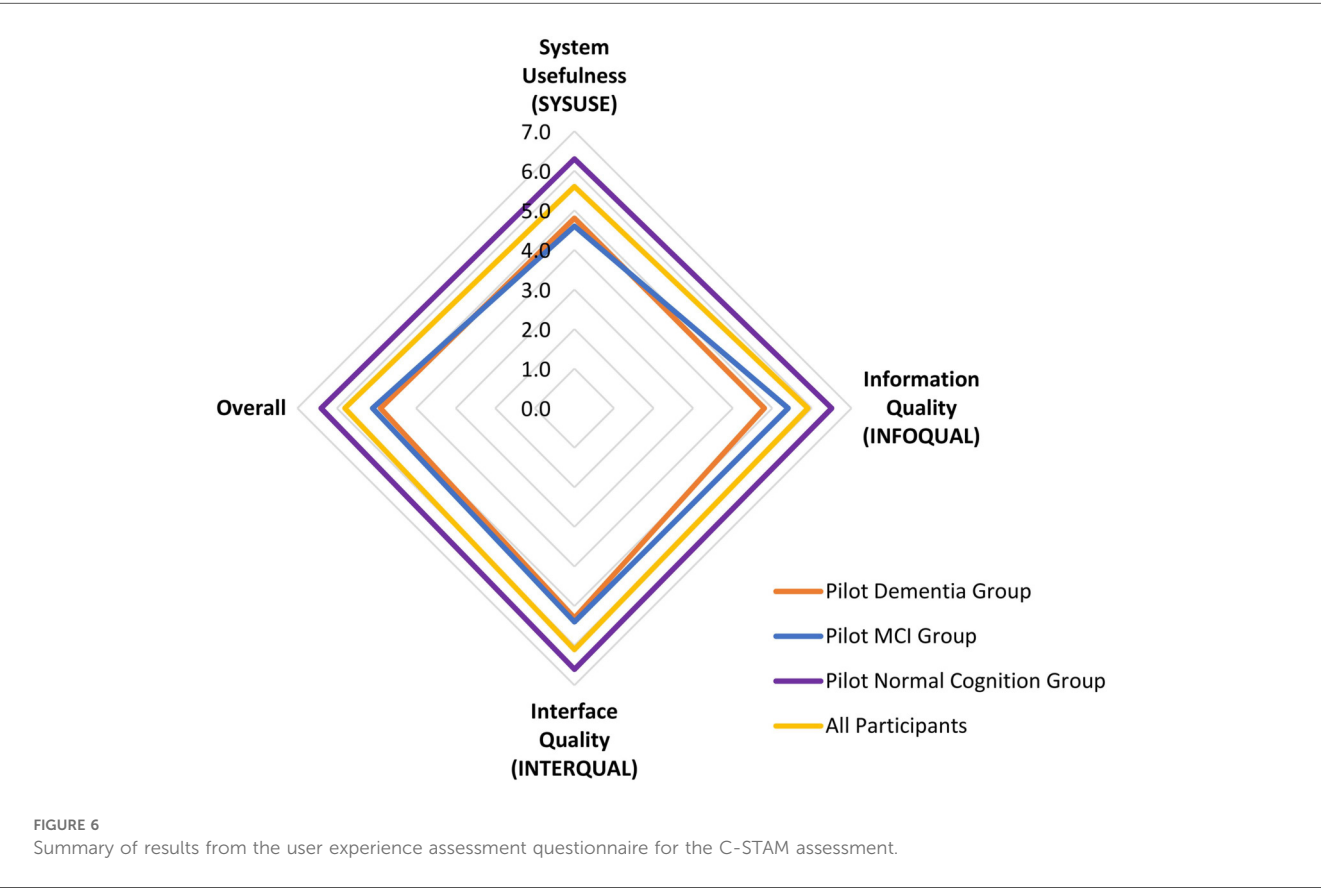
normal cognition group (SYSUSE: 6.3 ± 0.6 ; INFOQUAL: 6.5 ± 0.4 ; INTERQUAL: 6.6 ± 0.5), with an overall score of 6.4 ± 0.4 . In comparison, the pilot MCI group had a lower overall score (5.1 ± 1.1) and across categories (SYSUSE: 4.6 ± 1.4 ; INFOQUAL: 5.4 ± 1.3 ; INTERQUAL: 5.4 ± 0.8). The pilot dementia group had the lowest overall score (4.9 ± 2.1), as well as across most categories (SYSUSE: 4.8 ± 2.3 ; INFOQUAL: 4.8 ± 2.2 ; INTERQUAL: 5.3 ± 2.4).

One-way ANOVA indicated significant differences between groups for all categories (SYSUSE: $p = 0.013$; INFOQUAL: $p = 0.009$; INTERQUAL: $p = 0.031$; overall: $p = 0.008$). *post-hoc* Tukey HSD tests showed that differences were most pronounced between the Pilot Dementia Group and Pilot Normal Cognition

TABLE 1 Characteristics of study participants ($n = 30$).

Characteristics	All participants ($n = 30$)	Pilot dementia group ($n = 7$)	Pilot mild cognitive impairment group ($n = 6$)	Pilot normal cognition group ($n = 17$)	p -value ^a
Age (in years)	71.7 ± 5.7	74.1 ± 4.1	73.7 ± 3.2	70.0 ± 6.4	0.170
Sex					0.043
Male	15 (50%)	5 (71.4%)	5 (83.3%)	5 (29.4%)	
Female	15 (50%)	2 (28.6%)	1 (16.6%)	12 (70.6%)	
Caucasian ethnicity	27 (90%)	7 (100%)	5 (83.3%)	15 (88.2%)	0.765
Completed tertiary education	24 (80%)	6 (85.7%)	5 (83.3%)	13 (76.5%)	0.999

^aGroup comparison was conducted using one-way ANOVA for age, and Fisher’s exact for sex, ethnicity and education.



Group, particularly for Information Quality ($p = 0.011$) and overall score ($p = 0.015$) as shown in Table 2.

Figure 7 illustrates the variations in participants' responses to the 18 items of the user experience assessment questionnaire for the C-STAM, stratified by group, highlighting differences in feedback on the system's functionality and user experience across participant categories.

4.2 Qualitative results of user experience evaluation

Participant feedback on the C-STAM was collected as free-text responses and analysed thematically. Based on the responses of the 30 participants regarding the positive and negative aspects of the C-STAM tool, two key themes were identified: user interface and usability, and accessibility and adaptability. Direct quotations from the participants are presented to illustrate the identified themes and labelled as per the participant type and ID (e.g., *Pilot MCI group, Participant 1*).

4.2.1 User interface and usability

Many pilot participants ($n = 14$) commented on the simplicity and clarity of the C-STAM tool, which made it easier to use and understand.

"Very simple interface" (Pilot normal cognition group, Participant 13)

"It was clear and easy to navigate" (Pilot normal cognition group, Participant 16)

"Quite intuitive.... easy to understand" (Pilot MCI group, Participant 6)

"Simple language" (Pilot dementia group, Participant 4)

A few participants ($n = 6$) indicated that the features of the C-STAM tool, such as hints and audio, were pivotal for having a positive experience with the user interface and usability.

"The most positive aspect of the C-STAM tool was ... hints" (Pilot dementia group, Participant 5)

"It was great having the audio. The done button was consistent on each page. The font and size of the writing was big enough to read without having to adjust" (Pilot normal cognition group, Participant 17)

However, some participants ($n = 5$) indicated having difficulties with unclear instructions and challenges in navigating or understanding certain aspects of the system.

"Some instructions were not completely specific, especially finding Dr Gary's phone number" (Pilot normal cognition group, Participant 5)

"Had to read instructions a few times to understand how to do it. But that is probably my comprehension skills" (Pilot normal cognition group, Participant 11)

"Early instructions were a bit vague. Maybe a bit of a pre-primer [introduction] at the beginning, e.g., There are multiple components to answer these. Take your time to understand all components of a task" (Pilot dementia group, Participant 3)

A few participants ($n = 5$) also found the design of the C-STAM to be challenging, especially those who were less familiar with computers or specific tasks within the C-STAM.

"Harder at first if not used to using mechanics of computer" (Pilot MCI group, Participant 2)

"Those not familiar with online shopping and computers would find it difficult. So, I guess a trial would help to clarify the questions for users" (Pilot normal cognition group, Participant 14)

4.2.2 Accessibility and adaptability

Most participants ($n = 19$) commented that they found the C-STAM tool to be easy to follow and use and appreciated its

TABLE 2 User experience scores across pilot groups and *post-hoc* comparisons.

	System usefulness (SYSUSE)	Information quality (INFOQUAL)	Interface quality (INTERQUAL)	Overall
All Participants	5.6 ± 1.5	5.9 ± 1.3	6.1 ± 1.3	5.8 ± 1.3
Pilot Dementia Group	4.8 ± 2.3	4.8 ± 2.2	5.3 ± 2.4	4.9 ± 2.1
Pilot MCI Group	4.6 ± 1.4	5.4 ± 1.3	5.4 ± 0.8	5.1 ± 1.1
Pilot Normal Cognition Group	6.3 ± 0.6	6.5 ± 0.4	6.6 ± 0.5	6.4 ± 0.4
<i>p</i> -value ^a	0.013	0.009	0.031	0.008
Post-hoc Tests (Tukey HSD) ^b				
Group 1 vs. Group 2	0.961	0.586	0.960	0.918
Group 1 vs. Group 3	0.053	0.011	0.061	0.015
Group 2 vs. Group 3	0.036	0.151	0.117	0.062

^aOne-way ANOVAs were used to test for mean differences between groups.

^bTukey HSD *post-hoc* tests were conducted for significant findings; group 1 was the pilot dementia group, group 2 was the pilot MCI group, and group 3 was the pilot normal cognition group.

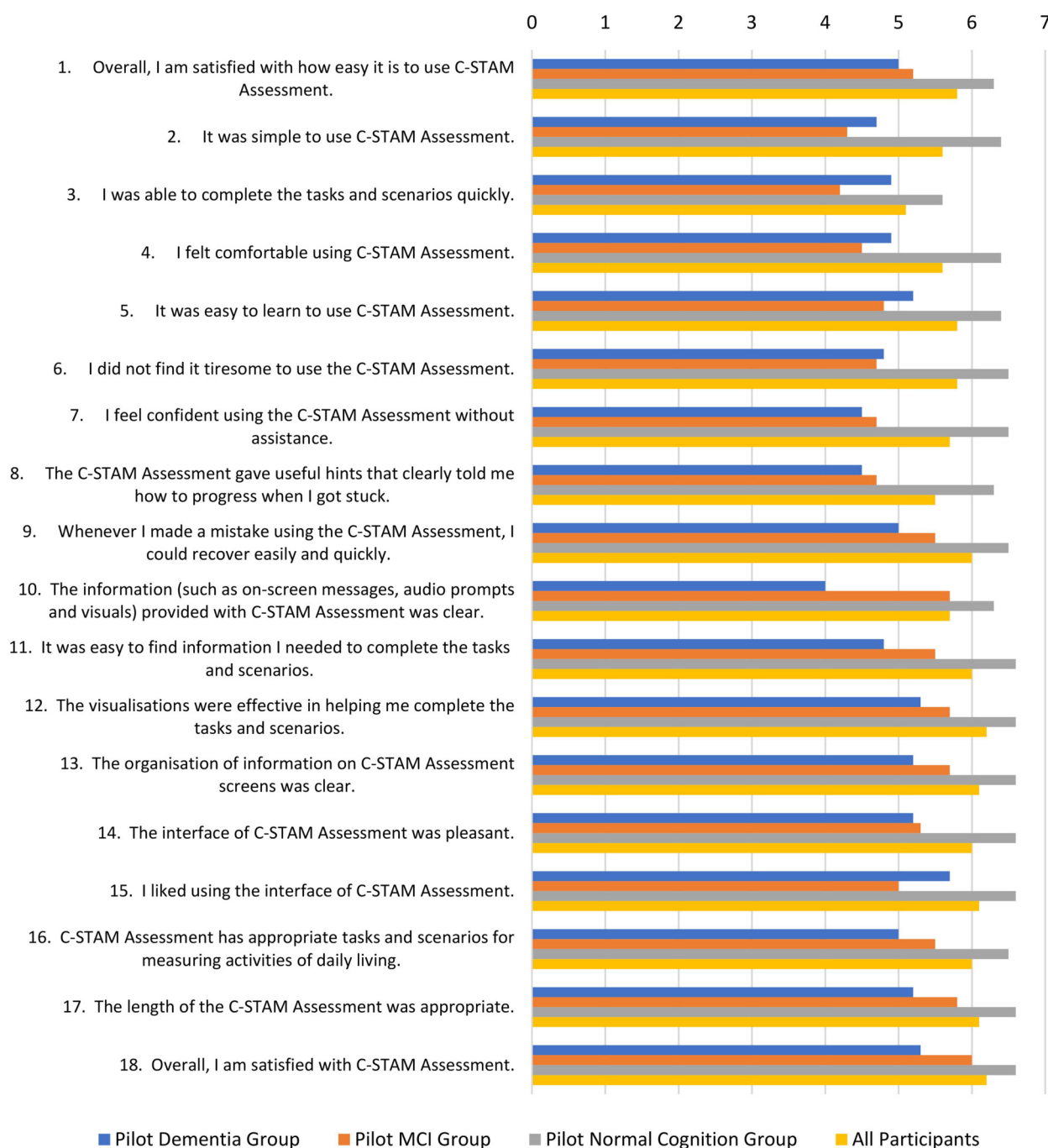


FIGURE 7

Participants' responses on the 18 items of the user experience assessment questionnaire for the C-STAM assessment.

accessible design. They noted that once started, the system was straightforward, making it accessible for both users and assessors.

“Was easy to follow instructions and clear” (Pilot dementia group, Participant 7)

“Easy to follow once started” (Pilot MCI group, Participant 2)

“Easy for assessor to use” (Pilot normal cognition group, Participant 3)

A few participants ($n = 4$) also expressed enthusiasm about the potential of the C-STAM as a more engaging and reliable diagnostic screening tool compared to traditional tasks for cognitive assessments, with suggestions provided on enhancing the effectiveness of the C-STAM tool.

“Overall, I enjoyed the experience and hope that a reliable testing tool for assessors will be an outcome. It will be a bit more interesting than the current pen and paper tasks”. (Pilot normal cognition group, Participant 1)

“I liked this tool and could see how it would be good for cognitive assessments... if this test is to be repeated annually, please change the scenarios so they are not the same”. (Pilot dementia group, Participant 2)

However, some participants ($n = 6$) expressed concerns about the accessibility of the C-STAM and its adaptability to different user needs. Some of the issues reported were difficulties with mouse and keyboard functionality and lengthy assessment time.

“I was somewhat startled that my finger presses on the pad and mouse did not work as I am used to on my Apple laptop pad” (Pilot normal cognition group, Participant 1)

“Unfamiliar keyboard ... strange logic” (Pilot dementia group, Participant 4)

“Quite long assessment especially for a person with dementia” (Pilot normal cognition group, Participant 3)

4.3 Discussion of findings of user experience evaluation

The user experience evaluation of the C-STAM tool provided valuable insights into its usability, functionality, and overall user satisfaction as a diagnostic screening tool for assessing functional ability in older adults with and without cognitive impairment. The findings revealed both the strengths and limitations of the C-STAM tool from the perspectives of diverse user groups, including individuals with normal cognition, MCI, and dementia. It is important to note that cultural and linguistic factors may influence user interactions, particularly for users from non-English speaking backgrounds, which could affect accessibility and the generalisability of the findings. It is essential to note that participants may have interacted with different versions of the C-STAM tool during the study, and some limitations identified may have already been addressed in subsequent iterations of the tool. Feedback from participants in this pilot study was crucial in shaping each iteration of the tool and played a vital role in refining the updated version of the C-STAM that is currently being tested in the main validation study.

This pilot study demonstrated a generally positive user experience with the C-STAM tool, with an overall average score of 5.8 ± 1.3 on the seven-point Likert scale, reflecting a high level of perceived usability and user satisfaction. This finding aligns with existing literature that emphasises the importance of co-design principles to ensure accessibility and engagement, particularly for digital tools targeting older adults (38, 39).

When stratified by group, the pilot normal cognition group reported the highest levels of satisfaction (6.4 ± 0.4), followed by the pilot MCI group (5.1 ± 1.1) and pilot dementia group (4.9 ± 2.1). The greater variability observed in the pilot dementia group may be attributed to the inclusion of individuals at varying stages of dementia, whose differing cognitive and functional abilities likely contributed to a wide range of user experiences with the tool. Given the exploratory and descriptive nature of this pilot study, these differences should be interpreted cautiously and can only be taken as preliminary evidence of the tool's ability to differentiate cognitive status. Moreover, given that individuals with cognitive impairment often encounter barriers when engaging with complex digital tasks, the relatively lower scores and feedback from the MCI and dementia groups led to the recognition of areas for further refinement (e.g., hints showing how to complete task components), particularly for users with cognitive impairment (40). While the validation study is still underway, the differences in user experiences among the three groups provide a pre-emptive indication that the C-STAM tool may differentiate between individuals with and without cognitive impairment. This also emphasises the need for additional supports such as stepwise guidance and enhanced hints in future versions. Nonetheless, these findings provide preliminary insights into how the tool might be tailored to support different cognitive profiles.

The findings from the three subscales of the user experience evaluation provided a detailed understanding of the positive features and areas for improvement of the C-STAM tool. The System Usefulness (SYSUSE) subscale revealed high perceived practicality and effectiveness (average score: 5.6 ± 1.5), with most participants finding the tool effective and easy to use. However, some noted that the assessment could feel lengthy and tiresome, particularly for individuals with cognitive impairment. This feedback underscores the importance of considering assessment length as a factor that may impact test burden and cognitive load for older adults (41). The Information Quality (INFOQUAL) subscale received positive feedback on the clarity and relevance of the information provided (average score: 5.9 ± 1.3). However, some participants expressed difficulty understanding certain tasks (e.g., the navigating directions task). These reported challenges led to refinements in subsequent iterations of the tool, with a focus on providing clear and straightforward information while designing digital tools for older adults (42). The Interface Quality (INTERQUAL) subscale received high ratings for the design and layout of the C-STAM interface (average score: 6.1 ± 1.3). Most participants emphasised the simplicity and intuitive nature of the interface, although some reported challenges with the device itself (e.g., keyboard and mouse functionality). These findings suggest the potential for a web-based solution compatible with mobile devices, which highlights an essential area of research as the future of cognitive assessment is likely to involve mobile applications for smartphones and tablets (43). These findings are also indicative of potential opportunities for integrating the C-STAM into routine clinical workflows, once validated through the main trial, for example as a complementary digital assessment

alongside traditional screening tools or to reduce reliance on resource-intensive performance-based evaluations.

There are certain limitations to this study. The reliance on perceived subjective measures of user experience may introduce response bias, including social desirability bias, as participants may overstate positive experiences. The pilot study involved a relatively small sample of individuals with and without cognitive impairment ($n = 30$), which may limit the generalisability of the findings. Additionally, the majority of participants identified as Caucasian and had a tertiary education, which may restrict the applicability of the findings to other demographic groups, such as non-English speaking populations for whom language translation may be necessary, and individuals with limited or no prior computer experience. Furthermore, participants' prior familiarity with digital devices may have confounded usability ratings, potentially inflating scores for those comfortable with technology and underestimating challenges for less experienced users. Another limitation is the lack of objective performance metrics, such as error types, or time to complete assessments, which would complement subjective evaluations and provide a more comprehensive understanding of usability and functional engagement. These factors accentuate the importance of future studies examining cross-cultural applicability and addressing language, literacy, and computer literacy barriers and the importance of conducting larger-scale studies with more diverse populations, and some of these gaps will be addressed in the main validation study.

5 Conclusion

This paper presents the online implementation and initial user experience evaluation of the C-STAM, a new computerised diagnostic screening tool currently under validation, designed to assess functional ability in older adults with and without cognitive impairment. This article outlines the evolution of the design of the C-STAM, the implementation of design into practice, and the iterative feedback processes that informed its development. The user experience evaluation, conducted with 30 pilot participants with and without cognitive impairment, demonstrated a generally positive user experience and high satisfaction levels. These findings suggest that a fully computerised, self-administered assessment of functional ability is both feasible and acceptable for older adults with and without cognitive impairment. Furthermore, the principles of co-design played a critical role in ensuring that the C-STAM tool was effectively tailored to meet the needs of its target end users. The validation study will address these factors in a larger, more diverse sample to assess usability, reliability, and generalisability.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Human Research Ethics Committee (HREC) of the University of New South Wales (HC210967). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JG: Investigation, Validation, Project administration, Supervision, Conceptualization, Writing – review & editing, Software, Writing – original draft. KR: Writing – original draft, Formal analysis, Writing – review & editing, Visualization, Software, Validation, Methodology, Investigation. MB: Investigation, Software, Writing – review & editing. PG: Writing – review & editing, Software, Investigation. SL: Writing – review & editing, Investigation, Software. KTN: Data curation, Investigation, Validation, Methodology, Supervision, Writing – review & editing. P-HL: Data curation, Investigation, Writing – review & editing. KN: Investigation, Project administration, Data curation, Writing – review & editing. MT: Writing – review & editing, Data curation, Investigation. SR: Resources, Validation, Funding acquisition, Project administration, Writing – review & editing, Supervision, Methodology, Investigation, Data curation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is funded by a National Health & Medical Research Council/Dementia Ageing and Aged Care Mission (MRFF) grant 2007541.

Acknowledgments

We would like to extend our gratitude to all programmers who have dedicated their valuable time for the online implementation of the C-STAM. We thank all investigators on the project and the advisory board members for their valuable input.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

References

1. Martyr A, Clare L. Awareness of functional ability in people with early-stage dementia. *Int J Geriatr Psychiatry*. (2018) 33:31–8. doi: 10.1002/gps.4664
2. Greenaway MC, Duncan NL, Hanna S, Smith GE. Predicting functional ability in mild cognitive impairment with the dementia rating Scale-2. *Int Psychogeriatr*. (2012) 24:987–93. doi: 10.1017/S1041610211002717
3. Marshall GA, Rentz DM, Frey MT, Locascio JJ, Johnson KA, Sperling RA. Executive function and instrumental activities of daily living in mild cognitive impairment and Alzheimer's disease. *Alzheimers Dement*. (2011) 7:300–8. doi: 10.1016/j.jalz.2010.04.005
4. Gold DA. An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *J Clin Exp Neuropsychol*. (2012) 34:11–34. doi: 10.1080/13803395.2011.614598
5. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. (2004) 256:183–94. doi: 10.1111/j.1365-2796.2004.01388.x
6. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment. *J Intern Med*. (2004) 256:240–6. doi: 10.1111/j.1365-2796.2004.01380.x
7. Nichols E, Steinmetz JD, Völset SE, Fukutaki K, Chalek J, Abd-Allah F, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. *Lancet Public Health*. (2022) 7:e105–e25. doi: 10.1016/S2468-2667(21)00249-8
8. Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. *Br Med J*. (2015) 350:h3029. doi: 10.1136/bmj.h3029
9. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the lancet commission. *Lancet*. (2020) 396:413–46. doi: 10.1016/s0140-6736(20)30367-6
10. Wesson J, Clemson L, Brodaty H, Reppermund S. Estimating functional cognition in older adults using observational assessments of task performance in complex everyday activities: a systematic review and evaluation of measurement properties. *Neurosci Biobehav Rev*. (2016) 68:335–60. doi: 10.1016/j.neubiorev.2016.05.024
11. Farias ST, Mungas D, Jagust W. Degree of discrepancy between self and other-reported everyday functioning by cognitive status: dementia, mild cognitive impairment, and healthy elders. *Int J Geriatr Psychiatry*. (2005) 20:827–34. doi: 10.1002/gps.1367
12. Hackett K, Mis R, Drabick DAG, Giovannetti T. Informant reporting in mild cognitive impairment: sources of discrepancy on the functional activities questionnaire. *J Int Neuropsychol Soc*. (2020) 26:503–14. doi: 10.1017/s1355617719001449
13. Numbers K, Jang S, Brodaty H, Sachdev PS, Draper B, Reppermund S. Instrumental activities of daily living by subjective and objective measures: the impact of depression and personality. *Front Aging Neurosci*. (2022) 14:829544. doi: 10.3389/fnagi.2022.829544
14. Goldberg TE, Koppel J, Keehlisen L, Christen E, Dreses-Werringloer U, Conejero-Goldberg C, et al. Performance-based measures of everyday function in mild cognitive impairment. *Am J Psychiatry*. (2010) 167:845–53. doi: 10.1176/appi.ajp.2010.09050692
15. Reppermund S, Birch RC, Crawford JD, Wesson J, Draper B, Kochan NA, et al. Performance-based assessment of instrumental activities of daily living: validation of the Sydney test of activities of daily living in memory disorders (STAM). *J Am Med Dir Assoc*. (2017) 18:117–22. doi: 10.1016/j.jamda.2016.08.007
16. Jang S, Numbers K, Lam BCP, Sachdev PS, Brodaty H, Reppermund S. Performance-based vs informant-reported instrumental activities of daily living in predicting dementia. *J Am Med Dir Assoc*. (2022) 23:1342–7.e9. doi: 10.1016/j.jamda.2021.09.020

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

17. Romero-Ayuso D, Castellero-Perea Á, González P, Navarro E, Molina-Massó JP, Funes MJ, et al. Assessment of cognitive instrumental activities of daily living: a systematic review. *Disabil Rehabil*. (2021) 43:1342–58. doi: 10.1080/09638288.2019.1665720
18. Geddes MR, O'Connell ME, Fisk JD, Gauthier S, Camicioli R, Ismail Z. Remote cognitive and behavioral assessment: report of the Alzheimer society of Canada task force on dementia care best practices for COVID-19. *Alzheimers Dement*. (2020) 12:e12111. doi: 10.1002/dad2.12111
19. Unity Technologies. Unity. (2024). Available online at: <https://unity.com/> (Accessed May 29, 2025).
20. Laravel Holdings. Laravel. (2024). Available online at: <https://laravel.com> (Accessed May 29, 2025).
21. Amazon Web Services. Amazon Polly. (2024). Available online at: <https://aws.amazon.com/polly/> (Accessed May 29, 2025).
22. ElevenLabs. Generative Voice AI. (2024). Available online at: <https://elevenlabs.io> (Accessed May 29, 2025).
23. OpenAI. DALL-E. (2024). Available online at: <https://openai.com/index/dall-e-2/> (Accessed May 29, 2025).
24. Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR. The Addenbrooke's cognitive examination revised (ACE-R): a brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatry*. (2006) 21:1078–85. doi: 10.1002/gps.1610
25. Fisher AG. *Performance Assessment of Self-care Skills (PASS)*. Assessment of Motor and Process Skills. Fort Collins, CO: Three Star Press (1997).
26. Lewis JR. IBM Computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum Comput Interact*. (1995) 7:57–78. doi: 10.1080/10447319509526110
27. Lewis JR. Psychometric evaluation of the PSSUQ using data from five years of usability studies. *Int J Hum Comput Interact*. (2002) 14:463–88. doi: 10.1080/10447318.2002.9669130
28. Vlachogianni P, Tselios N. Perceived usability evaluation of educational technology using the post-study system usability questionnaire (PSSUQ): a systematic review. *Sustainability*. (2023) 15:12954. doi: 10.3390/su151712954
29. Rashid A, editor. *A Hybrid Approach to Separation of Concerns: The Story of SADES. Metalevel Architectures and Separation of Crosscutting Concerns*. Berlin, Heidelberg: Springer Berlin Heidelberg (2001).
30. Ortin F, Lopez B, Perez-Schofield JBG. Separating adaptable persistence attributes through computational reflection. *IEEE Software*. (2004) 21:41–9. doi: 10.1109/MS.2004.52
31. Unity Technologies. String Tables. (2024). Available online at: <https://docs.unity3d.com/Packages/com.unity.localization@1.1/manual/StringTables.html#top> (Accessed May 29, 2025).
32. Isewon I, Oyelade J, Oladipupo O. Design and implementation of text to speech conversion for visually impaired people. *Int J Appl Inform Syst*. (2014) 7:25–30. doi: 10.5120/ijais14-451143
33. Trueba JL, Klimkov V. Neural text-to-speech makes speech synthesizers much more versatile. *Amazon Science*. (2019). Available online at: <https://www.amazon.science/blog/neural-text-to-speech-makes-speech-synthesizers-much-more-versatile> (Accessed May 29, 2025).
34. Bittner AK, Yoshinaga PD, Shepherd JD, Kaminski JE, Malkin AG, Chun MW, et al. Acceptability of telerehabilitation for magnification devices for the visually impaired using Various approaches to facilitate accessibility. *Transl Vis Sci Technol*. (2022) 11:4. doi: 10.1167/tvst.11.8.4
35. Ho C-H, Hou K-C. Exploring the attractive factors of app icons. *KSII Trans Int Inform Syst (TIIS)*. (2015) 9:2251–70. doi: 10.3837/tiis.2015.06.016

36. Vayadande K, Bhemde S, Rajguru V, Ugile P, Lade R, Raut N, editors. AI-based image generator web application using OpenAI's DALL-E system. In: *2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET)*; 2023 Nov 23-24; Karnataka, India. Karnataka: IEEE (2023). p. 1-5. Available online at: <https://ieeexplore.ieee.org/document/10420306>
37. Chaudry BM, Connelly KH, Siek KA, Welch JL, editors. Mobile interface design for low-literacy populations. In: *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*; Miami, FL, United States. New York, NY: Association for Computing Machinery (2012). p. 91-100. doi: 10.1145/2110363.2110377
38. Omaghomi TT, Elufioye OA, Akomolafe O, Anyanwu EC, Daraojimba AI. Health apps and patient engagement: a review of effectiveness and user experience. *World J Adv Res Rev.* (2024) 21:432-40. doi: 10.30574/wjarr.2024.21.2.0476
39. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. *J Med Internet Res.* (2022) 24:e40192. doi: 10.2196/40192
40. Henni SH, Maurud S, Fuglerud KS, Moen A. The experiences, needs and barriers of people with impairments related to usability and accessibility of digital health solutions, levels of involvement in the design process and strategies for participatory and universal design: a scoping review. *BMC Public Health.* (2022) 22:35. doi: 10.1186/s12889-021-12393-1
41. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. *BMC Med Inform Decis Mak.* (2022) 22:317. doi: 10.1186/s12911-022-02064-5
42. Nordin S, Sturge J, Ayoub M, Jones A, McKee K, Dahlberg L, et al. The role of information and communication technology (ICT) for older adults' decision-making related to health, and health and social care services in daily life—a scoping review. *Int J Environ Res Public Health.* (2021) 19:151. doi: 10.3390/ijerph19010151
43. Dieciuc M, Zhang S, Gray N, Lustria MLA, Carr DC, Charness N, et al. Terms of engagement: understanding the motivations, preferences, and attitudes of older adults toward mobile cognitive assessment and training. *Gerontologist.* (2024) 64:gnad048. doi: 10.1093/geront/gnad048