Early prediction of mastery of a computerized functional skills training program in participants with mild cognitive impairment

Philip D. Harvey,[1,2] Courtney Dowell-Esquivel,[1] Justin E. Macchiarelli,[3] Alejandro Martinez,[1] Peter Kallestrup,[2] and Sara J. Czaja[2,4]

[1]*University of Miami Miller School of Medicine, Miami, FL, USA*
[2]*I-Function, Inc, Miami, FL, USA*
[3]*University of Miami, Coral Gables, FL, USA*
[4]*Weill-Cornell Medical Center, New York, NY, USA*

ABSTRACT

Background: Cognition in MCI has responded poorly to pharmacological interventions, leading to use of computerized training. Combining computerized cognitive training (CCT) and functional skills training software (FUNSAT) produced improvements in 6 functional skills in MCI, with effect sizes >0.75. However, 4% of HC and 35% of MCI participants failed to master all 6 tasks. We address early identification of characteristics that identify participants who do not graduate, to improve later interventions.

Methods: NC participants ($n = 72$) received FUNSAT and MCI ($n = 92$) participants received FUNSAT alone or combined FUNSAT and CCT on a fully remote basis. Participants trained twice a week for up to 12 weeks. Participants "graduated" each task when they made one or fewer errors on all 3–6 subtasks per task. Tasks were no longer trained after graduation.

Results: Between-group comparisons of graduation status on baseline completion time and errors found that failure to graduate was associated with more baseline errors on all tasks but no longer completion times. A discriminant analysis found that errors on the first task (Ticket purchase) uniquely separated the groups, $F = 41.40$, $p < .001$, correctly classifying 94% of graduators. An ROC analysis found an AUC of .83. MOCA scores did not increase classification accuracy.

Conclusions: More baseline errors, but not completion times, predicted failure to master all FUNSAT tasks. Accuracy of identification of eventual mastery was exceptional. Detection of risk to fail to master training tasks is possible in the first 15 minutes of the baseline assessment. This information can guide future enhancements of computerized training.

Key words: activities of daily living, cognitive training, mild cognitive impairment

Introduction

Cognitive health poses a serious challenge for the expanding elderly population, and worldwide prevalence of mild cognitive impairment (MCI) is estimated to be around 15% (Bai *et al.*, 2022). Both normative and illness-related cognitive changes can undermine the capacity to perform everyday tasks and make autonomous decisions (Beach *et al.*, 2023; Marshall *et al.*, 2017). Further, the domains of cognitive functioning affected most by normal aging

and cognitive disorders are those that are required to learn new skills and process novel information (Park & Schwarz, 2012), with research finding direct correlations between cognitive abilities and technology adoption (Czaja, *et al.*, 2006). MCI is characterized by deterioration in cognitive abilities and daily functioning (Marshall *et al.*, 2011), surpassing typical age-related decline, but not yet reaching the criteria for a dementia diagnosis. While not everyone with MCI progresses to dementia, MCI with amnestic features (aMCI) is a risk factor for developing Alzheimer's disease (AD; Albert *et al.*, 2011). The lack of an efficacious medication for cognitive impairment in MCI/early AD (Petersen *et al.*, 2005) has prompted research on the development of non-pharmacological interventions

for add-on therapy, specifically computerized cognitive (CCT) training cognitive skills.

Meta-analyses support overall efficacy of cognitive training for MCI (Hill *et al.*, 2017; Zhang *et al.*, 2019), AD in certain domains (Sherman *et al.*, 2017), and older individuals with normal cognition (NC; Lampit *et al.*, 2014). However, there are moderators of the efficacy of CCT interventions for cognition in these populations including length of training session (less than then 30 minutes is less helpful) and dose per week (More than 3 times per week had diminishing returns). Studies have also suggested that lower baseline cognition scores (Roheger, *et al.*, 2020), adding an exercise component (Gavelin *et al.*, 2021), and more structured CCT (Roheger *et al.*, 2019) lead to greater benefits. Commercially available cognitive training software was associated with wide-ranging gains in older people (Tetlow and Edwards, 2017), suggesting that specialized CCT software may not be required.

Studies in psychiatric conditions have repeatedly found that training that is titrated in difficulty, momentarily adjusted in difficulty with achievement, sustained over time, augmented by coaching, and has engaging tasks led to the greatest gains, in schizophrenia (Bowie *et al.*, 2020) and major depression (Douglas *et al.*, 2020). Remote delivery or primarily home-based computer training has had mixed results, with some reviews reporting successful training outcomes but possibly greater attrition (Best *et al.*, 2023; Douglas *et al.*, 2020) and others suggesting that home-based training is not effective (Lampit, *et al.*, 2014)

Even in studies where there were substantial cognitive gains with CCT alone and excellent near transfer to untrained cognitive skills (Edwards *et al.*, 2002), concurrent real-world functional gains were found to be limited to improved performance on previously acquired functional skills such as everyday activities (Edwards *et al.*, 2005) and driving (Ross *et al.*, 2016) and not to impact on acquisition of novel daily skills (Willis, *et al.*, 2006). Our previous study of in-person training of 6 functional skills in MCI and NC found that over 50 % of participants with NC and with MCI improved in their completion time by one standard deviation or more across the 6 skills indexed to NC baseline performance (Czaja *et al.*, 2020). However, full mastery of all six tasks was more common in the NC participants. As important as the differences in task mastery were the differences in drop-out. 32% of the MCI participants, who had lower levels of task mastery, dropped out before completing training, compared to 13% of the NC participants. Thus, the drop-out rate in participants with MCI in that study, particularly in the combined training intervention, was more than double that of the NC sample despite the substantial training gains seen in those who completed training.

The current report comes from a study of updated skills training software. Specifically, a new version of the FUNSAT™ program was developed and tested in a randomized clinical trial, featuring fully remotely delivered cognitive and functional skills training and targeting the same 6 technology-based activities of daily living, in older adults with NC and MCI. This trial (NCT046779441) has three different pre-planned outcomes presented separately. Improvements in performance on the training simulations in errors and time to completion, Czaja *et al.* (2023) was the designated primary paper, real-world transfer of the technology-related skills assessed with ecological momentary assessment (EMA) is the second Dowell-Equivel *et al.* (2023), and near transfer to cognitive performance and far transfer to untrained functional capacity measures (Harvey, *et al.*, 2023) is the third. The study reported in this paper is a secondary analysis that was targeted at earliest possible identification of the characteristics of participants who eventually failed to develop full mastery of the 6-task training program. Identification of participants at high risk for failure to master the task could allow for the development of corrective "secondary" interventions to support training and reduce tendencies toward drop-out in participants. It seems important to identify individuals who were having challenging experiences in mastering the training tasks as rapidly as possible.

Our goal is to identify differences between participants who achieved full mastery of the training tasks, redefined in this study as completion of all subtasks within each of the 6 training tasks with no errors or two consecutive attempts with 1 error. We aimed to compare the attributes of participants who achieved full proficiency in FUNSAT, referred to as graduates, with those of nongraduates. As we were interested in very early detection of failure to graduate, we used individual differences factors to predict mastery, including baseline performance on the Montreal Cognitive Assessment (MOCA; Nasreddine *et al.*, 2005) and years of education. We also used several FUNSAT task performance characteristics as potential predictors: the number of errors and time to completion at the baseline assessments as well as training gains immediately after baseline assessments, using training gains on the first post-baseline training session as training-related predictors.

We had several hypotheses. Given the previous reports that global cognitive status and lower baseline performance predicted more training gains, we expected that lower baseline scores on the FUNSAT and possibly scores on the MOCA would predict greater gains with training. Previous studies have reported that reduced engagement in CCT

predicted reduced near transfer of training gains across populations (Harvey *et al.*, 2019), so we hypothesized that reduced training gains on the first FUNSAT training session were candidate predictors of failure to master the full set of tasks.

Methods

Overall study design

This study was a randomized controlled trial carried out at a total of fourteen community centers in South Florida and New York City. These are nonmedical community facilities attended by community residents for a variety of social and personal reasons. All recruitment was done in person, through town hall meetings and word of mouth. After initial screening, participants underwent an orientation and an in-person baseline evaluation on a fixed difficulty assessment of six functional tasks. Participants then engaged in up to 12 weeks of self-administered computer-based training at home. The study received approval from the WCG IRB, and every participant gave their signed informed consent to participate.

Participants

The study included both male and female community members over 60 years of age, without limitations based on race or ethnicity. Subjects were required to be proficient in either English or Spanish, have at least 20/60 vision, be able to read from a computer screen, and operate a touch-screen device. A neuropsychological assessment based on the Jak–Bondi criteria (Jak *et al.*, 2009) was used to determine MCI status of the participants. Based on this criteria, participants were categorized as either having normal cognitive function or falling into one of three MCI subcategories: Amnestic: deficits in two or more memory domains but not more than one in a non-memory area; non-amnestic: deficits in two non-memory cognitive areas, yet not more than one in a memory-related domain; multi-domain: deficits in two or more tests in both memory and other cognitive domains. To assess performance, normative standards were applied and impairment on any individual measure was defined as a performance of 1.0 or more standard deviations below the normative mean.

Individuals were not eligible for the study if they had a MOCA score below 18, had a reading proficiency below a 6th-grade reading level in the language in which they had selected to be assessed and trained, or could not engage in assessments conducted in English or Spanish. Participants were disqualified if they had undergone a similar

intervention in the previous year. Medical reasons for exclusion included a previous history of a serious psychiatric condition, except for depression, or histories of past neurological incidents such as seizures, brain tumors, cerebral vascular accidents, or severe traumatic brain injuries resulting in extended periods of unconsciousness.

Cognitive assessments

Data for the performance-based MCI criteria were gathered using cognitive evaluations. Assessments were conducted in the language preferred by the participants, either English or Spanish.

MONTREAL COGNITIVE ASSESSMENT (MOCA)
The MOCA evaluates cognitive abilities with scores ranging from 0 to 30 and all assessments were conducted by certified bilingual raters.

READING PERFORMANCE
English-speaking participants' literacy levels were assessed with the Wide Range Achievement Test (WRAT; Jastak, 1993), 3rd edition. Spanish speakers were assessed with the Woodcock-Munoz Language Survey, 3rd edition (WMLS-III; Woodcock *et al.*, 2017).

WECHSLER MEMORY SCALE-REVISED, LOGICAL MEMORY I AND II (ANNA THOMPSON STORY)
Participants were narrated the story and asked for an immediate recollection. After a 20-minute interval filled with other non-verbal tasks, they were then asked for a delayed recall of the original story.

BRIEF ASSESSMENT OF COGNITION (BAC): APP VERSION
The BAC evaluates cognitive domains associated with daily functioning (Keefe *et al.*, 2004). The application (Atkins *et al.*, 2017) provides these assessments via a cloud-connected tablet, simplifying administration, and ensuring consistency.

The cognitive domains assessed include the following:

- Verbal Memory; Working Memory; Motor Speed; verbal Fluency; Symbol Coding, and Executive functioning.

General procedures

The third generation of the FUNSAT™ program trains the same skills as previous generations. The skills include ATM usage, operating a ticket kiosk, Internet banking, online shopping, refilling a prescription using a telephone voice menu, and managing medication by both comprehending medication labels and organizing medications

(Supplemental Figure 1). Each task was presented in a multi-media format including text, voice, and graphic representations. Baseline assessments included a fixed difficulty (Form A) version with 6 tasks, and all subtasks were administered without training or any corrective feedback. The 6 tasks had 3–6 subtasks with sequentially increasing difficulty demands. With each error made, the original instructions would reappear in a pop-up window. If a participant made more than four errors on any one item, the software automatically moved on to the next item. Completion time and errors were collected in real time while participants completed each task, with time measured while the participant was actively engaged in the task. Participants performed the baseline assessments at the research site and then trained at home, so there was an assistant present to give encouragement in case the participant stopped participating in the assessment.

After the baseline assessments, training started. In each training session, lasting up to one hour, participants aimed to make as much progress as possible in mastering the items on individual subtasks. The program delivered training only on subtasks that had not yet been mastered. NC participants only trained with FUNSAT™ to develop normative standards for training gains. MCI participants were randomized into two groups: FUNSAT™ only or FUNSAT™ + CCT. Randomization was stratified by overall geographic area (NY vs. Miami) and sex. The FUNSAT™ program targeted development of proficiency in 6 functional tasks, with participants training 2 hours weekly for up to 12 weeks or until they achieved full mastery of all six tasks.

Those in the combined FUNSAT™ + CCT group underwent an intensive 3-week CCT training (two one-hour sessions weekly) before transitioning to FUNSAT™ for up to 9 weeks. After the 12-week period or upon mastering all tasks, participants were reevaluated using a different version of the fixed difficulty simulation administered at baseline. Follow-up evaluations took place around 30 days after completion or mastery and 3 months after that, with those results reported elsewhere. Participants were compensated $30.00 for each in-person assessment and received a bonus of $15.00 for each task mastered.

Training procedures

FUNSAT™

FUNSAT™ training was delivered through a cloud-based system on a touch-screen device with all training performed at the home of the participant. To connect to the Internet, participants had the option to use a provided hotspot or their own Wi-Fi

connection. The training protocol was adaptive, with participants receiving immediate feedback about the first error within each subtask, with additional corrective feedback being given after all subsequent errors. For example, if a participant was attempting the ATM task and entered the wrong pin, a pop-up window would appear stating "Try Again! Your ATM PIN is 1234." Following a second error, a new pop-up window would appear stating "Try Again! Remember, your PIN is 1234. Please enter 1234." A third error would prompt the participant "Try Again! Press 1, then press 2, then press 3, and then press 4. Then press ENTER." And finally, after a fourth error, each key would light up in sequence with a statement telling participants to click the corresponding key as they light up. A subtask was considered mastered if the participant completed the subtask once with no errors or twice consecutively with a maximum of one error on each attempt. Each of the tasks was considered mastered once all subtasks within a specific task were mastered. After any break from training, only the non-mastered subtasks were retrained. Training was considered complete after 12 weeks or when a participant mastered all 6 tasks, at which point the endpoint fixed difficulty assessment was delivered.

Computerized cognitive training

The BrainHQ™ "Double Decision" training exercise was selected as the CCT for the FUNSAT™ + CCT group. ACTIVE and other trials (Edwards *et al.*, 2005; Harvey *et al.*, 2019) have shown significant benefits from similar speed of processing training exercises. The exercise included two concurrent tasks where participants had to identify an item that appeared in the middle of the screen while simultaneously locating a specific stimulus among 7 others in the periphery. Participants also had the option to train up to 20% of their sessions on another BrainHQ task named "Hawk Eye" to increase variety in training.

Data analyses

The objective of the study was to contrast the characteristics of participants who successfully mastered all elements of FUNSAT prior to the end of the planned 12-week protocol, referred to as "graduates," with those who did not, known as "nongraduates." We compared the frequencies of graduation across overall site (Miami vs New York), cognitive status (MCI vs NC), racial status, \, and Latinx Ethnicity. All analyses were performed with SPSS version 28. (IBM Corporation, 2023). As we expected that poor performance on less challenging tasks would be more informative, we limited our analyses to baseline performance on the three easiest

tasks (Ticket Kiosk, ATM, and medication management) as defined by performance of the HC sample in the previous and current studies. Baseline information on completion time and errors from these three tasks was used to predict graduation status. The first analyses simply compared graduates and nongraduates in the total sample on the 6 baseline variables (3 tasks, 2 variables per task), the MOCA and education. We also examined changes from baseline to the first training session within graduators and non-graduators across all six variables to see if the changes were significant.

We used discriminant function analyses to predict graduation status (yes/no), first entering any of the 6 baseline variables that differed between groups to predict graduation status. We also used training gains (time and errors) after one training as a subsequent predictor. We used a forward entry stepwise procedure and a p value of $p < 0.05$ for a variable to enter the equation. After conducting the first analysis, we kept any predictive variables and added the time and error variables for training gains, for the first training session. After the best predictive variables were identified by the discriminant analysis, we added MOCA scores as a potential predictor. After final selection of predictors, we used ROC curve analysis to examine the area under the curve to quantify prediction of graduation status.

Results

Figure 1 presents the patient flow in the study. As can be seen in the figure, 287 participants signed a consent form and 184 were randomized, with the most common reason for not being randomized being failure to attempt to train. For randomized participants in the two cognitively defined subject groups, MCI, and NC, drop-out from training was modest. Three of 75 NC participants (4%) did not complete training, with drop-out for MCI participants in skills training only at 6 out of 51 participants (11%) and for combined training at 4 out of 52 trainees (8%).

Table 1 presents the demographic information on the participants separated by MCI status, including graduation. MCI participants had significantly less education and lower MOCA scores than NC participants but did not differ in age. There were no site, race, or training language differences in MCI status. There were slightly more Latinx participants and slightly more male participants in the MCI group than in the NC group. Chi-square tests found that MCI status was significantly associated with lower rates of graduation from all training tasks, but that ethnicity, race, location, or training language were not, all $X^2(1) < 0.46$, all $p > .50$.

As we previously reported (Dowell-Esquivel *et al.*, 2023; Czaja *et al.*, 2020), there were no site (NYC vs. Miami) differences in age, education, MOCS score, sex, and racial status due to our efforts to collect balanced samples. More participants reported Latinx Ethnicity (66%) and trained in Spanish at the Miami site (54%) than in New York (41% and 28%), $X^2(1) > 12.05$, $p < .001$.

Table 2 presents the scores for the 6 baseline completion time and error variables across graduation status, education and MOCA scores, as well as training gains from baseline to the first training session. We used t-tests to compare the graduates and nongraduates on the baseline task performance variables, MOCA scores, and education. As seen in the table, nongraduates made more baseline errors, had baseline slower performance, lower MOCA scores, and less education than the graduates. Effect sizes for the differences were all $d = 0.84$ or larger. As the variance estimates appeared to be potentially unbalanced, we performed F tests for homogeneity of variance. Only 1 was significant, ATM baseline errors. When we used the Mann-Whitney U test to confirm the results of the t-tests, all 6 tests were significant, all $U > 455$, all $z > 4.51$, all $p < .001$. We performed similar analyses (data not shown) for the difference of graduates and nongraduates within the MCI participants alone. All 6 t-tests were statistically significant, with graduates performing better (all $t > 2.41$, all $p < .022$).

At the bottom of the table, we present change scores from baseline to first training session. For graduates, all changes in completion time and errors were significant at $p < 0.001$, with effect sizes of $d = 0.33$ or larger. For the non-graduators, two of the variables did not change significantly from baseline to the first training session: time and errors on the medication management test. The effect sizes for group differences at baseline were uniformly larger, across all 6 measures, than the effect size for trial 1 training changes. Thus, all 6 baseline performance (time and error) variables and all trial 1 training gains were considered for use in the multivariate analyses.

Table 3 presents the results of the discriminant function analyses with the 6 baseline time and error variables. As can be seen in the top of the table, only baseline errors on the Ticket Kiosk Task entered the discriminant function, $p < .001$. This analysis yielded correct overall classification based on graduation status of 85%, while correctly identifying 94% of the graduates

When we entered the training gains after 1 training session as a predictor of graduation status, including both changes in errors and time to completion, none of the variables entered the discriminant function, all $F < 1.84$, all $p > .18$.
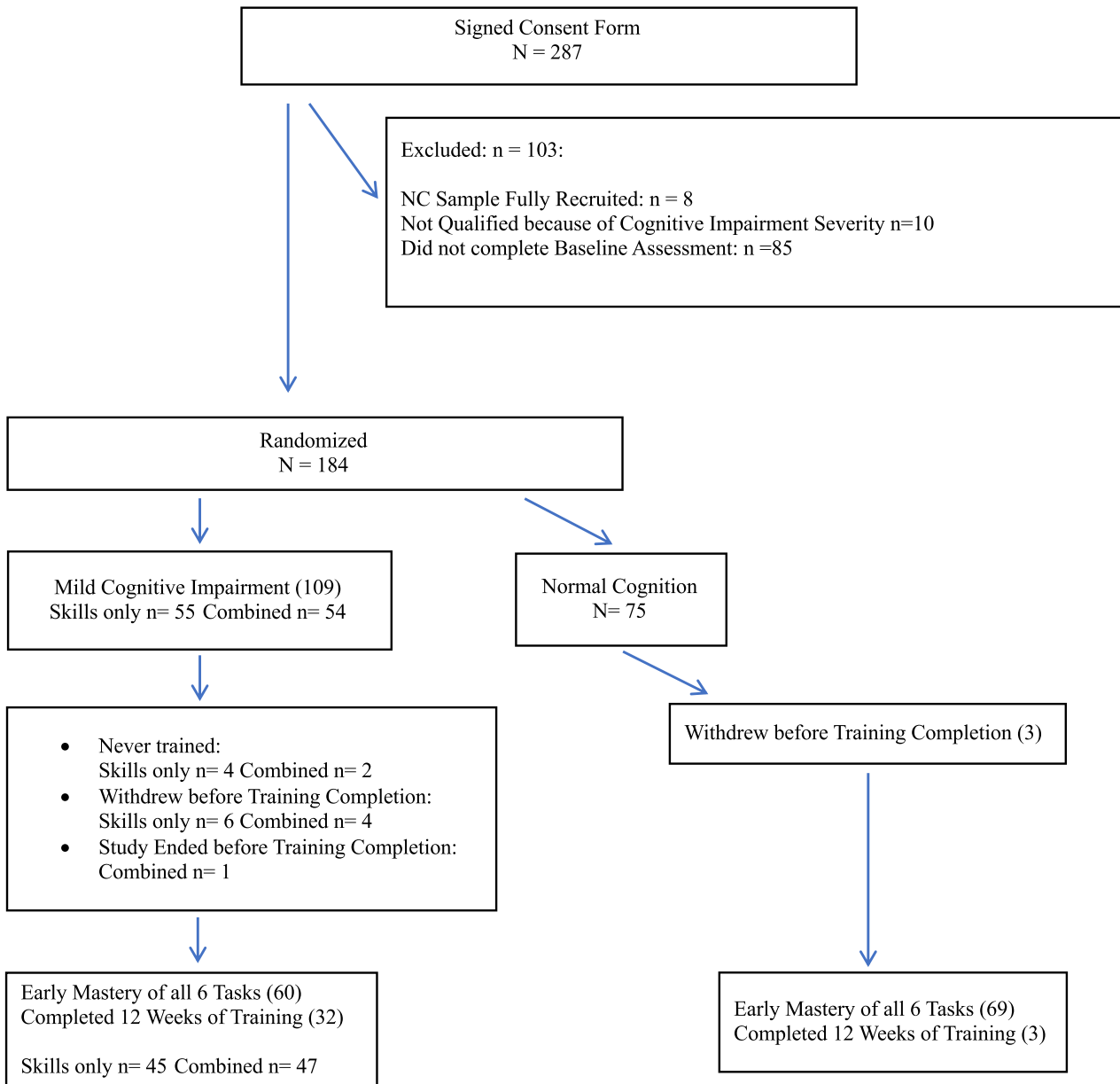
Figure 1. CONSORT diagram of participant flow in the study.

In our final discriminant analysis, presented at the bottom of Table 3, we added MOCA scores to the original baseline variables as an additional predictor of graduation status. Interestingly, MOCA scores entered the analysis at a very significant level but did not displace ticket task baseline errors as the primary discriminator. Classification accuracy was improved by 2% overall, with detection accuracy for nongraduates increased by 2% and detection accuracy for graduates unaffected.

Figure 2 presents the ROC curve analysis for graduation status. Using ticket task baseline errors as the predictor, the area under the curve (AUC) was 0.83, with a standard error of measurement of .042. The p value for significance test was $p < .001$;

and the 95% Confidence interval for the AUC was 0.75–.92.

Discussion

In a well-characterized sample of participants with NC and MCI, nearly all NC participants and the majority of those with MCI fully mastered 3 different functional skills training tasks. Prediction of those who did not manifest full mastery suggested that errors on the very first, and easiest, of the fixed difficulty pre-training simulations, the Ticket Kiosk Task, was a substantial predictor of eventual mastery. Further, errors on the task, not completion

Table 1. Demographic and descriptive information on participants

	MILD COGNITIVE IMPAIRMENT $N = 92$		NORMAL COGNITION $N = 72$			
	M	SD	M	SD	t	p
Age	71.68	6.40	71.17	6.36	−0.50	0.65
MOCA score	22.45	3.21	27.15	1.37	11.50	<0.001
Years of education	13.29	3.94	15.56	2.52	4.18	<0.001

SITE	MILD COGNITIVE IMPAIRMENT $N = 92$ N (%)	NORMAL COGNITION $N = 72$ N (%)	X^2	p
Miami	38 (41%)	32 (43%)	1.21	0.55
NY	54 (59%)	40 (57%)		
Sex			6.00	0.05
Male	18 (20%)	7 (8%)		
Female	74 (80%)	65 (92%)		
Race			7.95	0.44
White	36 (36%)	34 (47%)		
Black	28 (30%)	14 (18%)		
Other/	28 (30%)	24 (35%)		
More than 1/None				
Ethnicity			8.20	0.04
Latinx	52 (57%)	34 (46%)		
Non-Latinx	40 (43%)	38 (54%)		
Training language			2.64	0.27
English	50 (55%)	48 (67%)		
Spanish	42 (35%)	24 (33%)	16.13	0.001
Training gradua- tion	60 (65%)	69 (96%)		
MCI classification				
Amnestic	14 (15%)			
Multi-domain	38 (41%)			
Non-amnestic	40 (43%)			

time nor training gains after initiation of training, were the best predictor of eventual mastery. Adding MOCA scores as a predictor did not change the proportion of graduating cases identified.

With the high levels of graduation, the incremental prediction was not numerically substantial, because if everyone was designated as a graduate, 78% of the classifications would be correct. The improvement is statistically significant across two analysis strategies. However, the finding that error scores on the very first task are the best predictor of this incremental prediction provides pragmatic information about how it is possible to rapidly identify failure to master the task.

Given that MCI and more severe cognitive challenges lead to disability, the availability of training that can lead to mastery of functionally relevant technology-related tasks may be a treatment advance. In our previous study with in-person training (Czaja *et al.*, 2020), we have found that

drop-out from training, although minimal compared to pharmacological interventions, can handicap global training outcomes. Drop-out rates in this study for participants with MCI were less than half that seen in the previous intervention. In the current fully remote version of the training simulations, identification as early as possible of possible challenges to completion could allow the developers to modify the tasks increase efficiency of training, further reduce drop-out, and attenuate experiences of frustration on the part of potential participants and their families. Given the high levels of mastery of MCI participants, we do not see any reason that participants with slightly more severe impairments could not receive some benefit from training.

The origin of high early error rates and eventual failure to master all tasks cannot be clearly identified from these data. Poor motivation seems unlikely as a cause of failure to master the tasks, because the participants who were identified as not mastering all

Table 2. Baseline and first training session scores on potential predictors as a function of mastery of all training tasks

	GRADUATED		DID NOT GRADUATE $N = 129$		HOMOGENEITY OF VARIANCE $N = 35$				
	M	SD	M	SD	t	p	d	F	p
Baseline time to completion									
Ticket test	999.23	409.29	1448.19	573.17	4.73	<0.001	1.02	3.49	0.07
ATM test	1382.23	670.53	2389.28	1214.38	5.86	<0.001	1.28	21.41	<0.001
Medication management	1067.01	649.75	1635.32	649.75	3.85	<0.001	0.84	0.69	0.41
Baseline Errors									
Ticket test	15.42	8.92	27.96	10.49	6.34	<0.001	1.37	1.09	0.30
ATM test	24.74	20.49	50.80	14.11	5.92	<0.001	1.24	0.36	0.55
Medication management	31.78	24.70	53.28	22.35	3.91	<0.001	0.87	1.26	0.26
MOCA	24.63	3.23	21.37	2.95	4.86	<0.001	1.45		
Years of education	12.65	3.85	14.72	3.40	2.76	0.006	1.02		

Training gains on the first learning trial: change from baseline to training trial 1

	GRADUATED					DID NOT GRADUATE				
Completion time	M	SD	T	p	d	M	SD	t	p	d
Ticket	224.26	285.78	8.84	<0.001	0.79	259.65	472.76	2.80	0.01	0.55
ATM	336.36	414.78	9.07	<0.001	0.81	465.48	159.70	2.91	0.008	0.58
Medication management	266.64	576.65	5.21	<0.001	0.46	154.44	741.09	1.04	0.31	0.21
Errors										
Ticket	5.83	6.88	9.54	<0.001	0.85	5.88	12.22	1.45	0.021	0.48
ATM	9.71	15.88	6.87	<0.001	0.61	8.64	14.15	3.05	0.005	0.61
Medication management	17.10	51.64	3.73	<0.001	0.33	21.60	55.62	1.94	0.064	0.39

tasks continued training until the end of the study. It is possible that reduced experience with technology-related tasks was associated with high error rates on the first simulation. It is also possible that the characteristics of the fixed difficulty assessment, where the task challenges are generally hierarchical in difficulty may lead to participants "getting behind" and never catching up. Also, some requirements for successful performance of the task, such as the need to orient to the touch-screen and correctly execute responses, are not trained by the current version of the software. Other fixed difficulty functional capacity assessments, such as the Virtual Reality Functional Capacity Assessment Task (VRFCAT; Keefe *et al.*, 2016), have a formal orientation training program that preceded the task itself. However, the VRFCAT does not have a remote delivery option, so eventually having both remote delivery and a formal training period would be the optimal development. In the FUNSAT fixed difficulty stimulations, participants have only 4 opportunities to complete each item before it is designated as failed and a progression takes place.

It is worth noting that the FUNSAT is fully modular, and any combination of training simulations can be administered to participants. Since

errors on all the tasks were greater in non-graduators, if a protocol was targeting only ATM banking, for instance, high levels of errors on that simulation also discriminated eventual graduators and those who did not.

The limitations of the study include inability to subdivide participants with MCI based on Jak–Bondi subtypes because we did not stratify at the time of selection. Racial and ethnic status is not balanced across the MCI overall subgroups and fewer participants overall trained in Spanish than in English. Failing to achieve mastery of all tasks does not mean that training gains were not substantial in general (see Table 2) or that real-world transfer did not occur in that subset of participants.

Although training gains in the FUNSAT across simulations were previously reported to be similar across different racial, ethnic, language, educational, and baseline cognitive factors (Dowell-Esquivel *et al.*, 2023), there were still a subset of participants, generally limited to those with MCI, who did not fully master the training tasks. The fact that these participants can be identified very early on, and through error rates at baseline rather than reduced early training gains (which requires completing the full baseline assessment before training

Table 3. Results of discriminant analyses predicting mastery status from errors and completion time

MODEL 1	WILKS LAMBDA	PILLAIS' APPROX. F	p
Ticket errors	0.78	43.35	<0.001
Ticket time	0.78	0.01	0.99
ATM errors	0.77	2.36	0.13
ATM time	0.76	3.82	0.06
Medication management errors	0.78	0.73	0.39
Medication management time	0.78	0.00	0.99

CLASSIFICATION ACCURACY

	PREDICTED			
	NON-GRADUATOR		GRADUATOR	
ACTUAL	N	%	N	%
Non-graduator	18	51	17	49
Graduator	8	6	121	94

Total Correct: 85%

MODEL 2: MOCA ADDED	WILKS LAMBDA	PILLAIS' APPROX. F	p
Ticket errors	0.78	41.84	<0.001
MOCA	0.76	23.94	<0.001
Ticket time	0.76	0.00	0.99
ATM errors	0.75	0.81	0.37
ATM time	0.74	2.96	0.09
Medication management errors	0.76	0.09	0.76
Medication management time	0.75	0.23	0.63

Classification accuracy

	PREDICTED			
	NON-GRADUATOR		GRADUATOR	
ACTUAL	N	%	N	%
Non-graduator	20	57	15	43
Graduator	8	6	121	94

Total correct: 86%

starts), suggests that targeting these participants with task-based interventions may be possible. Formal training for orientation to the task demands, possible alternative assessment strategies, and modification of training strategies including more opportunities to pass easier items, smaller incremental training units, or more feedback might reduce the learning challenges. Given the general absence of previous successful computerized skills training interventions targeting this population, a 65% success rate for full mastery with training of 6 technology-related functional skills for participants with MCI seems substantial. The importance of these training gains is underscored by the results of the previous papers from this study showing the following: (1). greater proportionate gains on training tasks for MCI participants than NC (Czaja et al., 2023); (2). real-world transfer of performance of the trained functional skills task to the real-world environment, in both MCI and NC samples (Dowell-Esquivel et al., 2023); and (3) training gains in cognition and functional capacity that were statistically significant with, effect sizes greater than $d = 0.75$ for MCI participants (Harvey et al., 2023). The fact that drop-out on the part of MCI participants was reduced by 30% through adjustments in training delivery and standards for mastery suggests that a goal of eliminating failure to master all tasks through alternations in training delivery does not seem to be an unrealistic goal.
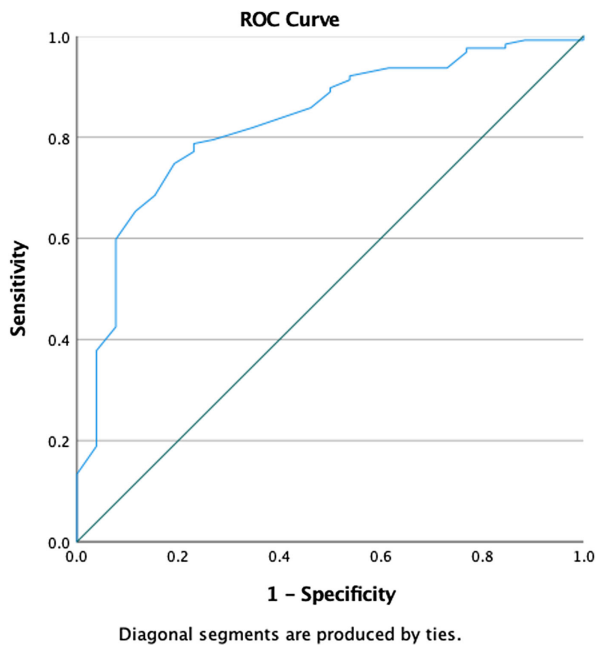
Figure 2. Receiver operating characteristic analysis yielding area under the curve for identification of participants who mastered all tasks ticket kiosk errors.

Conflicts of interest

Peter Kallestrup is CEO of i-Function, Inc. Sara J. Czaja is Co-Chief Scientific Officer of i-Function, Inc. Courtney Dowell-Equivel, Justin Macchiarelli, and Alejandro Martinez have no competing interests. Philip D. Harvey is Co-Chief Scientific Officer of i-Function, Inc. He has other interests unrelated to the content of this paper: Consulting fees or travel reimbursements from Alkermes, Boehringer Ingelheim, Karuna Therapeutics, Merck Pharma, Minerva Neurosciences, and Sunovion Pharma in the past year. He receives royalties from the Brief Assessment of Cognition in Schizophrenia (Owned by WCG Endpoint Solutions, Inc. and contained in the MCCB). He is Scientific Consultant to EMA Wellness, Inc.

Source of funding

Description of author(s)' roles

PD Harvey: Designed and supervised the study. Analyzed data and wrote and edited the manuscript.

P Kallestrup: Designed and supervised the study. Wrote and edited the manuscript.

J Macchiarelli: Collated and organized data, wrote first draft of the manuscript.

A Martinez: Collated and organized data, wrote first draft of the manuscript.

C Dowell-Esquivel: Conceptualized the specific substudy, collated data, and wrote first draft of the manuscript.

S Czaja: Designed and supervised the study. Wrote and edited the manuscript.

Acknowledgements

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S10416102240 00115

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia*, 7(3), 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

Atkins, A. S., Tseng, T., Vaughan, A., Twamley, E. W., Harvey, P., Patterson, T., Narasimhan, M., & Keefe, R. S. E. (2017). Validation of the tablet-administered brief assessment of cognition (BAC app). *Schizophrenia Research*, 181, 100–106. https://doi.org/10.1016/j.schres.2016.10.010

Bai, W., Chen, P., Cai, H., Zhang, Q., Cheung, T., Jackson, T., Sha, S., Xiang, Y.-T., & Su, Z. (2022). Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: A meta-analysis and systematic review of epidemiology studies. *Age andAageing*, 51(8), afac173. https://doi.org/10.1093/ageing/afac173

Beach, S. R., Czaja, S. J., & Schulz, R. (2023). Novel methods for assessment of vulnerability to financial exploitation (FE). *Journal of Elder Abuse and Neglect*, 35(4-5), 1–23. https://doi.org/10.1080/08946566.2023.2281672

Best, M. W., Romanowska, S., Zhou, Y., Wang, L., Leibovitz, T., Onno, K. A., Jagtap, S., & Bowie, C. R. (2023). Efficacy of remotely delivered evidence-based psychosocial treatments for schizophrenia-spectrum disorders: A series of systematic reviews and meta-

analyses. *Schizophrenia Bulletin,* 49(4):973–986. doi: 10.1093/schbul/sbac209.

Bowie, C. R., Bell, M. D., Fiszdon, J. M., Johannesen, J. K., Lindenmayer, J.-P., McGurk, S. R., Medalia, A. A., Penadés, R., Saperstein, A. M., Twamley, E. W., Ueland, T., & Wykes, T. (2020). Cognitive remediation for schizophrenia: An expert working group white paper on core techniques. *Schizophrenia Research*, 215, 49–53. https://doi.org/10.1016/j.schres.2019.10.047

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (CREATE). *Psychology and Aging*, 21(2), 333–352. https://doi.org/10.1037/0882-7974.21.2.333

Czaja, S. J., Kallestrup, P., & Harvey, P. D. (2023). The efficacy of a home-based functional skills training program for older adults with and without a cognitive impairment. *Innovations in Aging*.

Czaja, S. J., Kallestrup, P., Harvey, P. D., & Pak, R. (2020). Evaluation of a novel technology-based program designed to assess and train everyday skills in older adults. *Innovation in Aging*, 4(6), igaa052. https://doi.org/10.1093/geroni/igaa052

Douglas, K., Jordan, J., Inder, M., Crowe, M., Mulder, R., Lacey, C., Beaglehole, B., Bowie, C., & Porter, R. (2020). Cognitive remediation for outpatients with recurrent mood disorders: A feasibility study. *Journal of Psychiatric Practice*, 26(4), 273–283. https://doi.org/10.1097/PRA.0000000000000487

Douglas, K. M., Milanovic, M., Porter, R. J., & Bowie, C. R. (2020). Clinical and methodological considerations for psychological treatment of cognitive impairment in major depressive disorder. *BJPsych Open*, 6(4), e67. https://doi.org/10.1192/bjo.2020.53

Dowell-Esquivel, C., Czaja, S. J., Kallestrup, P., Depp, C. A., Saber, J. N., & Harvey, P. D. (2023). Computerized cognitive and skills training in older people with mild cognitive impairment: Using ecological momentary assessment to index treatment-related changes in real-world performance of technology-dependent functional tasks. *The American Journal of Geriatric Psychiatry*, S1064-7481(23), 00463–00473. https://doi.org/10.1016/j.jagp.2023.10.014

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). Transfer of a speed of processing intervention to near and far cognitive functions. *Gerontologia*, 48(5), 329–340. https://doi.org/10.1159/000065259

Edwards, J. D., Wadley, V. G., Vance, D. E., Wood, K., Roenker, D. L., & Ball, K. K. (2005). The impact of speed of processing training on cognitive and everyday performance. *Aging and Mental Health*, 9(3), 262–271. https://doi.org/10.1080/13607860412331336788

Gavelin, H. M., Dong, C., Minkov, R., Bahar-Fuchs, A., Ellis, K. A., Lautenschlager, N. T., Mellow, M. L., Wade, A. T., Smith, A. E., Finke, C., Krohn, S., & Lampit, A. (2021). Combined physical and cognitive training for older adults with and without cognitive impairment: A systematic review and network meta-analysis of randomized controlled trials. *Ageing Research Reviews*, 66, 101232. https://doi.org/10.1016/j.arr.2020.101232

Harvey, P. D., Balzer, A. M., & Kotwicki, R. J. (2019). Training engagement, baseline cognitive functioning, and cognitive gains with computerized cognitive training: A cross-diagnostic study. *Schizophrenia Research: Cognition*, 19, 100150. https://doi.org/10.1016/j.scog.2019.100150

Harvey, P. D., Zayas-Bazan, M., Tibiriçá, L., Kallestrup, P., & Czaja, S. J. (2023). *Improvements in performance based measures of cognition and functional capacity after computerized functional skills training in older people with mild cognitive impairment and healthy comparators.* Psychiatry Research.

Hill, N. T. M., Mowszowski, L., Naismith, S. L., Chadwick, V. L., Valenzuela, M., & Lampit, A. (2017). Computerized cognitive training in older adults with mild cognitive impairment or dementia: A systematic review and meta-analysis. *The American Journal of psychiatry*, 174(4), 329–340. https://doi.org/10.1176/appi.ajp.2016.16030360

IBM Corporation (2023). *Statistical Package for the Social Sciences (SPSS) version 28.* IBM Corporation.

Jak, A. J., Bondi, M. W., Delano-Wood, L., Wierenga, C., Corey-Bloom, J., Salmon, D. P., & Delis, D. C. (2009). Quantification of five neuropsychological approaches to defining mild cognitive impairment. *The American Journal of Geriatric Psychiatry*, 17(5), 368–375. https://doi.org/10.1097/JGP.0b013e31819431d5

Jastak, S. (1993). *Wide-range achievement test* (3rd ed.) Wide Range, Inc.

Keefe, R. S., Goldberg, T. E., Harvey, P. D., Gold, J. M., Poe, M. P., & Coughenour, L. (2004). The brief assessment of cognition in schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research*, 68(2-3), 283–297. https://doi.org/10.1016/j.schres.2003.09.011

Keefe, R. S. E., Davis, V. G., Atkins, A. S., Vaughan, A., Patterson, T., Narasimhan, M., & Harvey, P. D. (2016). Validation of a computerized test of functional capacity. *Schizophrenia Research*, 175(1-3), 90–96. https://doi.org/10.1016/j.schres.2016.03.038

Lampit, A., Hallock, H., Valenzuela, M., & Gandy, S. (2014). Computerized cognitive training in cognitively healthy older adults: A systematic review and meta-analysis of effect modifiers. *PLOS Medicine*, 11(11), e1001756.

Marshall, G. A., Aghjayan, S. L., Dekhtyar, M., Locascio, J. J., Jethwani, K., Amariglio, R. E., Johnson, K. A., Sperling, R. A., & Rentz, D. M. (2017). Activities of daily living measured by the harvard automated phone task track with cognitive decline over time in non-demented elderly. *Journal of Prevention of Alzheimer's Disease*, 4(2), 81–86. https://doi.org/10.14283/jpad.2017.10

Marshall, G. A., Rentz, D. M., Frey, M. T., Locascio, J. J., Johnson, K. A., Sperling, R. A., & Alzheimer's Disease Neuroimaging Initiative (2011). Executive function and instrumental activities of daily living in mild cognitive impairment and Alzheimer's disease. *Alzheimer's and Dementia*, 7(3), 300–308. https://doi.org/10.1016/j.jalz.2010.04.005

Nasreddine, Z. S., Phillips, N. A., Bédirian, Vérie, Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The montreal cognitive assessment, MoCA: A brief screening

tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Park, D., & Schwarz, N. (2012). *Cognitive aging: A primer.* Psychology Press.

Petersen, R. C., Thomas, R. G., Grundman, M., Bennett, D., Doody, R., Ferris, S., Galasko, D., Jin, S., Kaye, J., Levey, A., Pfeiffer, E., Sano, M., van Dyck, C. H., & Thal, L. J. (2005). Vitamin E and donepezil for the treatment of mild cognitive impairment. *New England Journal of Medicine*, 352(23), 2379–2388. https://doi.org/10.1056/NEJMoa050151

Roheger, M., Kalbe, E., Corbett, A., Brooker, H., & Ballard, C. (2020). Lower cognitive baseline scores predict cognitive training success after 6 months in healthy older adults: Results of an online RCT. *International Journal of Geriatric Psychiatry*, 35(9), 1000–1008. https://doi.org/10.1002/gps.532

Roheger, M., Kessler, J., & Kalbe, E. (2019). Structured cognitive training yields best results in healthy older adults, and their ApoE4 state and baseline cognitive level predict training benefits. *Cognitive and Behavioral neurology*, 32(2), 76–86. https://doi.org/10.1097/WNN.0000000000000195

Ross, L. A., Edwards, J. D., O'Connor, M. L., Ball, K. K., Wadley, V. G., & Vance, D. E. (2016). The transfer of cognitive speed of processing training to older adults' driving mobility Across 5 Years. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 71(1), 87–97. https://doi.org/10.1093/geronb/gbv022

Sherman, D. S., Mauser, J., Nuno, M., & Sherzai, D. (2017). The efficacy of cognitive intervention in mild cognitive impairment (MCI): A meta-analysis of outcomes on neuropsychological measures. *Neuropsychology Review*, 27(4), 440–484. https://doi.org/10.1007/s11065-017-9363-3

Tetlow, A. M., & Edwards, J. D. (2017). Systematic literature review and meta-analysis of commercially available computerized cognitive training among older adults. *Journal of Cognitive Enhancement*, 1(4), 559–575.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., Morris, J. N., Rebok, G. W., Unverzagt, F. W., Stoddard, A. M., Wright, E., & ACTIVE Study Group, for the (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA*, 296(23), 2805–2814. https://doi.org/10.1001/jama.296.23.2805

Woodcock, R. W., Alvarado, C. G., Ruef, M., & Shrank, R. (2017). *Woodcock-Muñoz language survey* (Third ed.) Riverside.

Zhang, H., Huntley, J., Bhome, R., Holmes, B., Cahill, J., Gould, R. L., Wang, H., Yu, X., & Howard, R. (2019). Effect of computerized cognitive training on cognitive outcomes in mild cognitive impairment: A systematic review and meta-analysis. *BMJ Open*, 9(8), e027062. https://doi.org/10.1136/bmjopen-2018-027062