Received: 28 June 2023 | Revised: 27 December 2023 | Accepted: 12 January 2024

DOI: 10.1002/dad2.12547

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Preliminary validation of the Virtual Kitchen Challenge as an objective and sensitive measure of everyday function associated with cerebrovascular disease

Sophia L. Holmqvist[1] | **Katie Jobson**[1] | **Dennis Desalme**[2] | **Stephanie M. Simone**[1] | **Molly Tassoni**[1] | **Moira McKniff**[1] | **Takehiko Yamaguchi**[3] | **Ingrid Olson**[1] | **Nadine Martin**[2] | **Tania Giovannetti**[1]

[1]Department of Psychology and Neuroscience, Temple University, Philadelphia, Pennsylvania, USA

[2]Department of Communication Sciences and Disorders, Eleanor M. Saffran Center for Cognitive Neuroscience, Temple University, Philadelphia, Pennsylvania, USA

[3]Department of Applied Information Engineering, Suwa University of Science, Nagano, Japan

Correspondence
Tania Giovannetti, Weiss Hall, Temple University, 1701 N. 13th St, Philadelphia, PA 19122, USA.
Email: tania.giovannetti@temple.edu

Sophia L. Holmqvist, Department of Psychology and Neuroscience, Temple University, Philadelphia, Pennsylvania, USA.
Email: sophia.holmqvist@temple.edu

Abstract

Preliminary validity of a computer-based test of everyday function (Virtual Kitchen Challenge [VKC]) was examined against brain-imaging markers of cerebrovascular disease and in contrast to conventional neuropsychological and self-report measures. Twenty community-dwelling older adults ($n = 6$ mild cognitive impairment) performed simulated breakfast and lunch tasks using a computer touchscreen (VKC). Automated measures (completion time, proportion time off screen, etc.) were computed during training and test conditions. White matter hyperintensity (WMH) volumes from brain magnetic resonance imaging and conventional measures of cognition and function also were obtained. VKC completion time and proportion time off screen improved significantly from training to test and were significantly associated with WMH volume ($r > 0.573$). VKC measures and WMH were not significantly correlated with conventional cognitive or self-report measures. The VKC holds promise as a valid measure of subtle functional difficulties in older adults that is sensitive to change and cerebrovascular pathology, highlighting its potential for clinical trials.

KEYWORDS
activities of daily living, cognitive aging, everyday action, virtual reality, white matter hyperintensities

Highlights

- Virtual Kitchen Challenge (VKC) scores showed significant improvement from training to test conditions.
- VKC scores (completion time and proportion of time off screen) were associated with a neuroimaging biomarker of brain health (white matter hyperintensities).

1 | INTRODUCTION

Assessment of everyday function is crucial for the clinical diagnosis and staging of neurodegenerative disease in older adults. Older adults and their families are most interested in the functional implications of cognitive difficulties,[1] particularly if they preclude independence. The current benchmark method of measuring functional problems includes interviews and/or questionnaires that capture only global level of functional disability but lack objectivity and standardization. The current approach is insufficient to detect subtle functional difficulties experienced by individuals with mild cognitive difficulty/decline.[2,3] Thus, there is a need for an objective assessment of mild functional difficulties for the comprehensive clinical assessment of neurodegenerative disease in older adults.

Self- and informant-report measures of function are easy to administer and score but may be limited by an unreliable reporter.[4] Cognitive difficulties as well as depression and anxiety symptoms may impact self-/informant report because of poor recall, poor insight, or biases. Informants may not always be available or knowledgeable,[5] and their reports can be influenced by demographic features and their relationship to the participant.[6] Self-/informant-report measures also lack standardization, as everyday activities and the modifications that people make to their daily activities to compensate for mild decline vary widely across people.[7–9]

Performance-based tests of everyday function address many of the limitations of self-/informant report but are generally labor intensive, requiring extensive set-up, trained human coders, and/or video recordings. Tests that are easier to administer and score may not be sufficiently complex or nuanced to capture mild difficulties.[3,10–12] The Virtual Kitchen Challenge (VKC) was created to address these obstacles in measuring subtle functional difficulties in older adults. The VKC is a standardized non-immersive virtual reality task that is easy to administer and score using a touch-screen interface. The VKC includes two tasks (breakfast, lunch) of comparable complexity that were modeled after tasks from the Naturalistic Action Test (NAT) and the Virtual Kitchen, which have been extensively researched (see Supplemental Materials in supporting information).[11,13,14]

The VKC scoring was informed by the goal-control model that posits problems in everyday activities arise due to difficulties in keeping task goals/steps active and/or in controlling the activation of goals/steps to enable efficient performance.[15] Automated VKC scores include an accomplishment score (i.e., total steps completed) and three measures of efficiency: completion time, proportion of time off screen (i.e., time spent deliberating between steps and deciding on one's next move), and the number of screen interactions (number of times the screen is touched). In an initial validation study, VKC automated efficiency scores correlated with human-coded scores of VKC, performance of the VKC tasks with real objects (NAT), and cognitive test scores.[16] Consistent with performance with real objects (NAT), healthy older adults and younger adults did not differ in accomplishment but performed significantly worse on measures of efficiency (i.e., slower, more time off screen, more screen interactions).[16] The original study did not examine whether VKC scores could detect change and did not validate the VKC against biomarkers of cognitive decline. The goal of the current study was to extend the psychometric analysis of the VKC by determining whether VKC scores (1) detect change from expected practice effects and (2) are associated with known biomarkers of cerebrovascular pathology.

RESEARCH IN CONTEXT

1. **Systematic review**: Sources were reviewed from PubMed and Google Scholar. Empirical data validating efficient, objective measures of mild functional difficulties that are sensitive to change are lacking. The Virtual Kitchen Challenge (VKC) is a computer-based tests with automated scoring that is sensitive to mild functional difficulties and related to cognitive abilities in older adults. Research has not yet examined the ability of the VKC to detect change or whether VKC scores are related to neuropathology or self-reported function.
2. **Interpretation**: Significant improvement in VKC scores after practice and strong associations between VKC scores and cerebral white matter integrity suggest the VKC may be sensitive to subtle changes in functional ability and reflect brain health.
3. **Future directions**: Further study is needed to establish the full psychometric properties of the VKC, including reliability across time, predictive validity of cognitive and functional status, and validation against additional biomarkers.

Repeated practice improves performance of everyday tasks in people with dementia.[17–20] From the perspective of the goal-control framework, repeated practice may serve to strengthen activations and control of task goals, improving task accomplishment and increasing task efficiency over time. Thus, VKC performance should improve after training. If VKC scores are insensitive to expected practice effects, then its utility for longitudinal studies and clinical trials may be limited.

As stated, the VKC has not yet been validated against well-established biomarkers of cognitive impairment. White matter (WM) disease has been associated with cognitive deficits in older adults since the 1990s[21,22] and WM hyperintensities (WMH), a neuroimaging marker of small vessel cerebrovascular disease, correlate with concurrent executive dysfunction and increased risk for cognitive/functional decline,[23,24] independent of education, genetic risk, and hippocampal volume (i.e., biomarker for Alzheimer's disease).[25] WMH are present before clinical dementia symptoms,[26] and regarding everyday function, WMH volumes are associated with errors on performance-based tests reflecting poor control of goal activations (i.e., poor sequencing, distractibility) in people with dementia.[27] To our knowledge relations between WMH and performance on the VKC have not been investigated.

The current study included community-dwelling older adults who completed the VKC training and test conditions, magnetic resonance imaging (MRI) of the brain, and conventional measures of cognition and function. We hypothesized that VKC scores would improve from training to test conditions: increased accomplishment, faster completion time, fewer screen interactions, less time off screen. We also hypothesized that VKC measures reflecting lower task accomplishment, slower completion time, more screen interactions, and greater proportion of time spent off screen would be associated with greater WMH volume. Exploratory analyses examined relations between the VKC and self-reported measures of function.

2 | METHODS

2.1 | Study sample

Participants aged 60+ were recruited from a parent study (NINDCD R0101694) in Philadelphia and were classified as having healthy cognition (HC) or mild cognitive impairment (MCI) according to established criteria.[28] Participants were recruited from advertisements in the community and university-affiliated neurology clinics. Exclusion criteria were non-English speaker, current/past neurological disorder or major medical illness (e.g., dementia, traumatic brain injury, etc.), and severe sensory/motor deficits that preclude interaction with a touchscreen.

2.2 | Procedures

All procedures were approved by the Temple University Institutional Review Board. Participants completed the study in two sessions. Session 1 lasted approximately 3 hours and included informed consent, screening/demographics interview, VKC, cognitive tests, questionnaires, and blood pressure readings. Participants were encouraged to take breaks between measures. A brain MRI, lasting approximately 1 hour, occurred during a second session within 6 months of the first session.

2.3 | Virtual Kitchen Challenge

Details regarding the VKC tasks, administration, and scoring have been published previously.[3] In brief, the VKC includes two tasks (breakfast, lunch) of comparable complexity and difficulty modeled after those used in the NAT.[11,13] The VKC is administered using a non-immersive virtual reality interface developed after the Virtual Kitchen.[14] A Dell XPS 15 (9575) with a 13.5-inch PixelSense touchscreen display (3000 × 2000 resolution) was used. Participants were instructed to use their dominant hand's index finger to manipulate VKC objects.

2.3.1 | VKC movement training

All participants underwent a 5- to 10-minute training of the basic movements required to complete the VKC (e.g., pouring, stirring, spreading butter, etc.). Basic movement training was not scored.[16]

2.3.2 | VKC task training

After basic training, participants completed the breakfast and lunch task (order counterbalanced across participants). The training trial included detailed instructions (i.e., "Pack a lunch box for someone who wants a peanut butter and jelly sandwich, cookies, and a drink"). Participants were told to verbalize their actions so that errors could be corrected by the examiner immediately. Participants were allowed to ask questions and there was a back-and-forth dialogue about the tasks during the training trials.

2.3.3 | VKC test trials

The test trial for each task (breakfast/lunch) was administered immediately after the practice trial. Test trials were completed in silence and without feedback. Participants were instructed to complete the test as quickly as possible, using clear movements, and without errors. All participants completed the breakfast and lunch training and test trials.

2.3.4 | VKC-automated coding

Five automated measures were obtained from the VKC system (see Table 1) separately for training and test trials. We also computed an overall VKC composite score for exploratory analyses (see Supplemental Materials).

2.4 | Self-reported everyday function

The Functional Activities Questionnaire (FAQ[29]) includes 10 items (i.e., daily activities) that are rated according to performance over the past 4 weeks using a scale from 0 to 3 (e.g., dependent = 3; normal = 0). All items are summed with higher scores reflecting worse functioning (max = 30).

The Measurement of Everyday Cognition (E-Cog)[8] requires participants to rate their current level of functioning compared to 10 years ago on a 4-point scale (e.g., 1 = better or no change; 4 = consistently much worse). E-Cog scores for memory and executive function items were computed separately by summing item scores and dividing by the number of items completed (range 1–4). Higher scores reflect greater decline.

2.5 | MRI acquisition and processing

A Siemens 3T TrimTrio fitted with a 64-channel head coil was used for the T1- and T2-weighted fluid-attenuated inversion recovery (FLAIR) images. A high-resolution T1 magnetization-prepared rapid gradient echo (MPRAGE) sequence consisting of 176 contiguous sagittal slices (voxel size: 0.9 mm isotropic, repetition time [TR] = 1900 ms, echo time [TE] = 2.32 ms, 900 ms inversion time, 90 flip angle, pixel matrix = 256 × 256) and T2-weighted FLAIR sequence (TR = 4.2 ms,

TABLE 1 Virtual Kitchen Challenge (VKC) efficiency measures and associated descriptions.

VKC efficiency measure	Description
Accomplishment	Total steps completed during the breakfast (16 max) and lunch (16 max) tasks were summed (max = 32). Higher scores reflect greater accomplishment.
Total screen interactions	The number of instances the participant made contact with the computer screen during completion of the breakfast and lunch tasks was summed.[3] Fewer screen interactions reflect more precise and efficient movements. The minimum screen interactions are 12 for the lunch task and 18 for breakfast. More screen interactions reflect extra actions and imprecision when moving objects.
Distractor interactions	The number of times a distractor object was touched and/or moved was summed to reflect inefficient and off-task actions.
Completion time	Recording begins when the participant touches the screen after instructions and ends when the participant presses the quit button on the screen. Longer completion times reflect less efficient task performance.
Proportion of time off screen	Time spent interacting with the touch screen during the VKC was computed and subtracted from the completion time to obtain the amount of time participants spent without interacting with the screen. The time off screen was divided by completion time to control for differences in total time. Longer off-screen times could be due to multiple factors, including slower planning, difficulties locating target objects, difficulty resolving competition for object selection, and reach micro-errors.

TE = 0.039 ms, flip angle = 149 degrees, matrix size = 512 × 512, slice thickness = 2 mm with 2 mm gap) were obtained.

To quantify WMH, a cluster-based segmentation toolbox used the k-nearest neighbor (k-NN) algorithm with T1 and FLAIR images.[30] First, candidate clusters were segmented from FLAIR images with FMRIB's Automated Segmentation Tool (FAST) using default settings.[31] Co-registration with T1 scans enabled identification of the clusters' anatomical location (i.e., clusters in WM were extracted), intensity features (i.e., WMH are hyperintense in FLAIR, hypointense in T1), and size (i.e., log-transformed number of voxels; WMH size is highly variable). A supervised machine learning algorithm determined WMH versus non-WMH classification based on anatomical location, intensity, and cluster size. The average WMH of right and left hemispheres, normalized for intracranial volume, were used for analyses. WMH were quantified in mm^3 for three regions: deep WMH (DWMH), periventricular WMH (PVWMH), and whole brain WMH. All WMH volumes were log transformed due to positive skewness.

2.6 | Cognitive testing

All participants completed the Mini-Mental State Examination (MMSE)[32] to screen for dementia, an estimate of premorbid IQ (Hopkins Reading Test[33]), and widely used cognitive tests (see Table 2). Raw cognitive scores were standardized using *t* scores based on normative data from the Calibrated Neuropsychological Normative System,[34] adjusting for sex, age, education, and estimated IQ. A cognitive composite score was computed by averaging all demographically adjusted *t* scores.

2.7 | Cognitive status

Demographically adjusted cognitive *t* scores were used to classify participants as having healthy cognition or MCI according to actuarial (Jak–Bondi) neuropsychological criteria.[28,41] MCI criteria require scores > 1 standard deviation (SD) below the age-/education-/sex-adjusted mean on two tests within the same cognitive domain or on at least one measure across all five domains listed in Table 2.

TABLE 2 Neuropsychological measures and domains assessed.

Cognitive domain	Neuropsychological measures
Attention	Trail Making Test Part A[35]
	Digit Span Forward[36]
Executive function	Trail Making Test Part B[35]
	Digit span backward[36]
Language	Letter Fluency (S and P) and Category Fluency (Animals)[34]
	Boston Naming Test 30-item version[37]
Processing speed	Salthouse Letter Comparison and Pattern Comparison[38]
Episodic memory	Hopkins Verbal Learning Test-Immediate, Delayed Recall, and Recognition trials[39]
	Brief Visual Memory Test-Revised-Immediate, Delayed Recall, and Recognition trials[40]

2.8 | Estimates of vascular risk

To characterize the sample's vascular risk, the Framingham Stroke Risk Profile score was computed for all participants based on self-report; a higher score indicates increased risk of stroke within 10 years.[42] Systolic and diastolic blood pressure (mmHg) were measured to capture pulse pressure (i.e., systolic – diastolic pressure). Elevated pulse pressure suggests increased vascular risk.[43]

2.9 | Analyses

Paired sample *t* tests assessed differences between VKC training and test conditions. Partial correlation coefficients, controlling for age, examined relations between WMH and measures of everyday function (VKC, self-reports). Parametric and non-parametric sensitivity analyses were performed and results were consistent across both approaches; thus, only parametric results are reported.

3 | RESULTS

3.1 | Participant characteristics

Twenty older adults ranging in age from 60 to 85 participated in the study. On average, participants were college educated with estimated IQ scores in the high average range (see Table 3). MMSE scores and self-reports of functioning were well within the normal range. The sample included men and women who reported either White or Black race, and six participants who met Jak–Bondi criteria for MCI. On average, the sample showed relatively low vascular risk.[44]

3.2 | VKC practice effects

As shown in Table 3, the accomplishment score was at ceiling and the distractor interaction score was at floor for both VKC test and VKC training. Therefore, accomplishment and distractor interactions were not included in statistical analyses. Paired sampled *t* tests comparing performance on the VKC training and VKC test trials revealed no significant difference in in total screen interactions ($t(19) = 0.515$, $P = 0.613$, $d = 0.115$). By contrast, completion time ($t(19) = 7.123$, $P < 0.001$, $d = 1.59$) and proportion of time off screen ($t(19) = -3.99$, $P < 0.001$, $d = -0.89$) were significantly lower (i.e., better) on the VKC test trials. Figures showing training versus test scores for each participant are included in Supplemental Materials.

3.3 | Associations among measures of everyday function

After controlling for age, none of the VKC test scores were significantly associated with self-reported measures of everyday function (FAQ, E-Cog; $rs < 0.2$). Correlation coefficients are included in Supplementary Materials.

3.4 | Associations between WMH and measures of everyday function

As shown in Table 4, after controlling for age, VKC completion time and proportion of time off screen were significantly correlated with whole brain WMH, PVWMH, and DWMH, such that slower completion times and more inefficient performance was associated with greater WMH volume. VKC total screen interactions and self-report measures (E-Cog, FAQ) were not significantly associated with WMH measures.

TABLE 3 Demographics and characteristics of the sample.

	Entire sample (*N* = 20) M (SD) or % (n)
Age, years	70.15 (6.38)
Education, years	16.68 (2.51)
Sex (% female, n)	40%, 8
Race (%, n)	
White	65%, 13
Black	35%, 7
MCI (%, n)	30%, 6
Estimated IQ	113.35 (9.69)
FAQ	1.40 (2.01)
E-Cog memory	1.78 (0.80)
E-Cog executive function	1.47 (0.73
Mini-Mental State Examination (MMSE) raw score	28.75. (2.20)
Cognitive composite	50.66 (5.33)
FSRP stroke risk percent	13.8 (9.30)
Pulse pressure, mmHg	53.02 (19.42)
Virtual Kitchen Challenge (VKC) measures	
VKC training	
Accomplishment steps completed	31.00 (1.89)
Total screen interactions	60.80 (15.61)
Total distractor interactions	0.05 (0.22)
Total completion time	230.97 (62.87)
Proportion of time off screen	0.63 (0.10)
VKC test	
Accomplishment steps completed	31.15 (1.26)
Total screen interactions	59.00 (14.39)
Total distractor interactions	0.20 (0.52)
Total completion time	158.77 (43.36)
Proportion of time off screen	0.56 (0.10)
Imaging measures	
Whole brain WMH volume (in mm^3)	2320.82 (3096.78)
Deep WMH volume (in $mm^{3)}$	329.40 (417.15)
Periventricular WMH volume (in mm^3)	1973.36 (2735.20)

Abbreviations: E-Cog, Measurement of Everyday Cognition; FAQ, Functional Assessment Questionnaire; FSRP, Framingham Stroke Risk Profile; MCI, mild cognitive impairment; SD, standard deviation; WMH, white matter hyperintensities;
Note: Cognitive Composite = average of *t* scores from all cognitive tests.

TABLE 4 Partial correlation coefficients between measures of everyday function and cognition and WMH controlling for age.

		White matter hyperintensities		
		Whole brain	Periventricular	Deep
		r (*P* value)	*r* (*P* value)	*r* (*P* value)
Virtual Kitchen Challenge test trials	Total completion time	0.575* (0.012)	0.557* (0.016)	0.505* (0.033)
	Proportion of time off screen	0.612* (0.007)	0.596* (0.009)	0.533* (0.023)
	Total screen interactions	0.205 (0.415)	0.167 (0.509)	0.328 (0.184)
Self-report questionnaires	FAQ	0.161 (0.523)	0.170 (0.499)	0.106 (0.677)
	E-Cog memory	0.087 (0.731)	0.120 (0.634)	−0.185 (0.462)
	E-Cog executive function	0.011 (0.967)	0.044 (0.862)	−0.319 (0.196)

*$P < 0.05$.
Abbreviations: E-Cog, Measurement of Everyday Cognition; FAQ, Functional Assessment Questionnaire; WMH, white matter hyperintensities.

TABLE 5 Partial correlation coefficients between functional measures and cognitive composite score controlling for age.

		Cognitive composite score *r* (*P* value)
Virtual Kitchen Challenge test trials	Total completion time	−0.185 (0.449)
	Proportion of time off screen	−0.008 (0.974)
	Total screen interactions	−0.238 (0.327)
Self-report questionnaires	FAQ	−0.536* (0.018)
	E-Cog memory	−0.226 (0.352)
	E-Cog executive function	−0.279 (0.248)

*$P < 0.05$.
Abbreviations: E-Cog, Measurement of Everyday Cognition; FAQ, Functional Assessment Questionnaire.

3.5 | Associations between measures of everyday function and cognitive tests

All correlation coefficients were in the expected direction, such that worse everyday function on all measures was associated with a lower cognitive composite score (Table 5). However, only the correlation between the FAQ and the cognitive composite score was statistically significant.

4 | DISCUSSION

This study investigated the validity of the VKC as a sensitive measure of subtle changes and cerebrovascular pathology in older adults. VKC performance scores were compared across two conditions (training vs. test) and examined in relation to biomarkers of cerebrovascular disease (i.e., WMH). The VKC showed significant improvement in efficiency scores after practice, such that older adults completed the VKC more quickly (shorter completion time) and with less deliberation and effort (lower proportion of time off screen). Completion time and proportion of time off screen also were strongly positively related to WMHs. Other VKC measures, namely, number of screen interactions and accomplishment steps, did not differ significantly from training to test conditions and were not strongly related to WMHs. Thus, VKC measures of completion time and the proportion of time off screen hold most promise as objective, performance-based measures of everyday function that may be sensitive to change and reflect cerebropathology in older adults with very mild functional difficulties.

Performance on VKC measures of completion time and proportion of time spent off screen significantly improved from training to test conditions while accomplishment steps and number of screen interactions showed only minimal improvement in our sample. The accomplishment score was affected by ceiling effects, as even participants with MCI completed all the task steps during the practice trials, a finding that has been demonstrated on conventional performance-based tasks that use real objects.[45] Thus, the VKC task accomplishment score may be more useful when evaluating participants with greater functional difficulties (e.g., dementia). By contrast, the number of screen interactions required to complete the VKC was not affected by ceiling effects. On average participants required approximately 60 interactions to complete the VKC, whereas the VKC may be completed in as few as 30 screen interactions by experts. Younger adults

(M age = 20) have been shown to complete the VKC in an average of 45 (SD = 14) interactions.[16] It is possible that the number of times older participants interact with the screen to complete the VKC may reflect a variety of non-cognitive factors, including computer touch-screen skills/experience, peripheral motor abilities (e.g., arthritis), or idiosyncrasies related to personal preferences or personality traits. For example, participants may have differed regarding their interactions in completing specific task steps depending on personal preference, such as spreading the peanut butter. Although they were instructed to perform the task precisely, efficiently, and quickly, participants also differed in performing clean-up actions, like replacing lids on jars, etc. More work is needed to understand screen interactions, but our results suggest that completion time and proportion of time off screen were sensitive to change and hold promise for future longitudinal studies and clinical trials.

VKC measures of completion time and proportion of time off screen also were significantly associated with PVWMH, DWMH, and total WMHs, highlighting the potential of the VKC as an objective test of mild functional difficulties that may be used to identify people at risk for future cognitive decline in a variety of clinical and research settings. There were no meaningful differences between PVWMH and DWMH in relation to the VKC. Past studies have suggested that PVWMH may be more strongly associated with cognition and dementia risk, but studies of everyday function have demonstrated both PVWMH and DWMH are each associated with different aspects of function in people with dementia (e.g., informant reports vs. specific errors on performance-based tests).[27] More work is needed to map regional WMHs to specific functional outcomes with special attention to the potential moderating role of age and cognitive status.

The VKC measures were developed to reflect the cognitive processing demands of everyday tasks, including the rapid and controlled activation of task goals, which may be disrupted by WM disease. Importantly, unlike other performance-based measures, the VKC obtains measures of task efficiency immediately in an automated fashion without the need for video recording and/or human coding. Although the VKC is more time consuming than questionnaires, questionnaire measures of function were not associated with WMH in our sample. In mild stages of impairment, questionnaire measures may be too blunt to identify slowed and inefficient everyday task processing abilities that are not yet at the level of severity to preclude completion of everyday tasks, limit independent function, or raise concern regarding decline.

Although relations between the VKC and standard questionnaires of everyday function and tests of cognition were not the focus of this study, it is worth noting that VKC measures were not correlated with the cognitive composite or questionnaires of everyday functioning. The cognitive composite score may not have been sensitive enough to capture mild cognitive difficulties in our sample of older adults, as the mean cognitive composite score demonstrated that our sample was relatively healthy and unimpaired. Our findings contrast with an earlier study of the VKC that showed strong and significant relations (r values between 0.35 and 0.71) between cognitive test scores (memory and executive function) and VKC measures (e.g., completion time, number of screen interactions, proportion of time off screen) in a sample that included both younger and older adults.[16] It is possible that within a sample of only older adults, the range of cognitive test scores was too restricted to identify relations between cognitive tests and VKC measures. We do not think our use of a global cognitive composite obscured relations with specific cognitive abilities, because correlation coefficients for individual test scores (see supporting information) also were non-significant. Given the discrepancies between the VKC and cognitive measures, longitudinal studies are needed to determine which measures (VKC vs. traditional cognitive tests) are best for detecting early cognitive/functional difficulties that predict future decline.

Questionnaires, such as the FAQ and E-Cog, may not be as sensitive for capturing mild cognitive difficulties in completing everyday tasks. Past studies report that performance-based measures and self-report measures of everyday functioning are only weakly related.[46] Questionnaires ask about functioning in one's natural environment, where task demands may vary widely across people. Task demands are known to impact performance and cognitive complaints. For example, younger adults report more memory complaints than older adults, despite hundreds of experiments with objective tests demonstrating age effects on memory performance.[47] This age paradox may be explained by different task demands across the lifespan, varying levels of business, and the use of compensatory strategies.[47] Thus, objective performance measures and questionnaires likely capture different aspects of everyday function. Performance-based measures, such as the VKC, assess the cognitive processes that are important for performing everyday tasks in a standardized manner that enables objective comparisons across participants and over time.

The sample size, though racially diverse, was small, reducing statistical power to detect small to medium effects. Replication with a larger sample is needed. Furthermore, participants were characterized as "mild cognitive impairment" based on only cognitive test scores and not a comprehensive clinical assessment, which is a limitation. Longitudinal study is needed to determine predictive validity of the VKC measures for identifying meaningful cognitive and functional decline, including conversion to MCI or dementia. WMH volume has been linked to executive dysfunction and processing/motor speed, which could explain the specific relations with VKC measures. VKC completion time may measure processing speed, motor speed, or a combination (e.g., psychomotor speed). Although participants with motor difficulties were not recruited and the VKC was designed to evaluate the cognitive aspects of everyday function, future work must disambiguate the effects of motor ability level on the VKC measures and evaluate the implications of the different motor demands of the VKC versus real, everyday activities. Despite the study limitations, the results support the validity of the VKC, particularly measures of completion time and proportion of time off screen, as sensitive to subtle changes in functional abilities and to mild functional difficulties associated with cerebrovascular disease in older adults without dementia. Thus, the VKC holds promise as an efficient measure of everyday function in a variety of settings.

ACKNOWLEDGMENTS

This project was funded by an Administrative Supplement to NINDCD R0101694 (to NM). TG, SLH, MT, and MM were supported by R01AG062503 and R21AG060422 (to TG).

CONFLICTS OF INTEREST STATEMENT

The authors have no conflicts of interest to report. Author disclosures are available in the supporting information.

CONSENT STATEMENT

All human subjects provided informed consent.

ORCID

Sophia L. Holmqvist https://orcid.org/0000-0001-7303-3338

REFERENCES

1. Marcum ZA, Rosenberg D, Barnes DE, Yaffe K, Larson EB. engaging patients to design the systematic multi-domain Alzheimer's risk reduction trial (SMARRT) intervention: findings from a web-based survey. *J Alzheimers Dis Rep*. 2020;4:255-260. doi:10.3233/ADR-200210
2. Farias ST, Giovannetti T, Payne BR, et al. Self-perceived difficulties in everyday function precede cognitive decline among older adults in the ACTIVE study. *J Int Neuropsychol Soc*. 2018;24:104-112.
3. Schmitter-Edgecombe M, McAlister C, Weakley A. Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: the day-out task. *Neuropsychology*. 2012;26(5):631-641. doi:10.1037/a0029352
4. Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Subjective cognitive complaints contribute to misdiagnosis of mild cognitive impairment. *J Int Neuropsychol Soc*. 2014;20(8):836-847.
5. Department of Health and Human Services. *A profile of older Americans*: 2021.
6. 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(5):503-514.
7. 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(9):827-834. doi:10.1002/gps.1367
8. Farias ST, Mungas D, Reed BR, et al. The measurement of Everyday Cognition (ECog): scale development and psychometric properties. *Neuropsychology*. 2008;22(4):531-544. doi:10.1037/0894-4105.22.4.531
9. Snyder PJ, Kahle-Wrobleski K, Brannan S, et al. Assessing cognition and function in alzheimer's disease clinical trials: do we have the right tools? *Alzheimers Dement*. 2014;10(6):853-860. doi:10.1016/j.jalz.2014.07.158
10. Rotenberg S, Ruthralingam M, Hnatiw B, et al. Measurement properties of the multiple errands test: a systematic review. *Arch Phys Med Rehabil*. 2020;101(9):1628-1642. doi:10.1016/j.apmr.2020.01.019
11. Giovannetti T, Libon DJ, Buxbaum LJ, Schwartz MF. Naturalistic action impairments in dementia. *Neuropsychologia*. 2002;40(8):1220-1232. doi:10.1016/s0028-3932(01)00229-9
12. McDougall GJ, Becker H, Vaughan PW, Acee TW, Delville CL. The revised direct assessment of functional status for independent older adults. *Gerontologist*. 2010;50(3):363-370. doi:10.1093/geront/gnp139
13. Schwartz MF, Segal M, Veramonti T, Ferraro M, Buxbaum LJ. The Naturalistic Action Test: a standardised assessment for everyday action impairment. *Neuropsychol Rehabil*. 2002;12(4):311-339. doi:10.1080/09602010244000084
14. Allain P, Foloppe DA, Besnard J, et al. Detecting everyday action deficits in Alzheimer's disease using a nonimmersive virtual reality kitchen. *J Int Neuropsychol Soc*. 2014;20(5):468-477. doi:10.1017/S1355617714000344
15. Giovannetti T, Mis R, Hackett K, Simone SM, Ungrady MB. The goal-control model: an integrated neuropsychological framework to explain impaired performance of everyday activities. *Neuropsychology*. 2021;35(1):3. doi:10.1037/NEU0000714
16. Giovannetti T, Yamaguchi T, Roll E, et al. The Virtual Kitchen Challenge: preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*. 2019;26(6):823-841. doi:10.1080/13825585.2018.1536774
17. Farina E, Fioravanti R, Chiavari L, et al. Comparing two programs of cognitive training in Alzheimer's disease: a pilot study. *Acta Neurol Scand*. 2002;105(5):365-371. doi:10.1034/j.1600-0404.2002.01086.x
18. Josephsson S, Bäckman L, Borell L, Bernspång B, Nygård L, Rönnberg L. Supporting everyday activities in dementia: an intervention study. *Int J Geriatr Psychiatry*. 1993;8(5):395-400. doi:10.1002/gps.930080505
19. Josephsson S, Bäckman L, Borell L, Nygård L, Bernspång B. Effectiveness of an intervention to improve occupational performance in dementia. *Occup Ther J Res* . 1995;15(1):36-49. doi:10.1177/153944929501500103
20. Avila R, Bottino CM, Carvalho IA, Santos CB, Seral C, Miotto EC. Neuropsychological rehabilitation of memory deficits and activities of daily living in patients with Alzheimer's disease: a pilot study. *Braz J Med Biol Res*. 2004;37(11):1721-1729. doi:10.1590/s0100-879X2004001100018
21. Libon DJ, Swenson RA, Malamut BL, Scanlon M, Coslett HB, Sands LP. Periventricular white matter alterations, dementia, and Binswanger's disease. *Dev Neuropsychol*. 1993;9(2):87-102. doi:10.1080/87565649109540546
22. Libon DJ, Price CC, Davis Garrett K, Giovannetti T. From Binswanger's disease to leuokoaraiosis: what we have learned about subcortical vascular dementia. *Clin Neuropsychol*. 2004;18(1):83-100. doi:10.1080/13854040490507181
23. Bangen KJ, Preis SR, Delano-Wood L, et al. Baseline white matter hyperintensities and hippocampal volume are associated with conversion from normal cognition to mild cognitive impairment in the Framingham offspring study. *Alzheimer Dis Assoc Disord*. 2018;32(1):50-56. doi:10.1097/WAD.0000000000000215
24. Brickman AM, Zahodne LB, Guzman VA, et al. Reconsidering harbingers of dementia: progression of parietal lobe white matter hyperintensities predicts Alzheimer's disease incidence. *Neurobiol Aging*. 2015;36(1):27-32. doi:10.1016/J.NEUROBIOLAGING.2014.07.019
25. Puzo C, Labriola C, Sugarman MA, et al. Independent effects of white matter hyperintensities on cognitive, neuropsychiatric, and functional decline: a longitudinal investigation using the National Alzheimer's Coordinating Center Uniform Data Set. *Alzheimers Res Ther*. 2019;11(1):1-13. doi:10.1186/S13195-019-0521-0/FIGURES/1
26. D'arbeloff T, Elliott ML, Knodt AR, et al. White matter hyperintensities are common in midlife and already associated with cognitive decline. *Brain Commun*. 2019;1(1):fcz041. doi:10.1093/BRAINCOMMS/FCZ041
27. Par KH, Lee JY, Na DL, et al. Different associations of periventricular and deep white matter lesions with cognition, neuropsychiatric symptoms, and daily activities in dementia. *J Geriatr Psychiatry Neurol*. 2011;24(2):84-90. doi:10.1177/0891988711402351
28. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry*. 2009;17(5):368-375. doi:10.1097/JGP.0B013E31819431D5

29. Teng E, Becker BW, Woo E, Knopman DS, Cummings JL, Lu PH. Utility of the functional activities questionnaire for distinguishing mild cognitive impairment from very mild Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2010;24(4):348-353. doi:10.1097/WAD.0B013E3181E2FC84
30. Warfield SK, Kaus M, Jolesz FA, Kikinis R. Adaptive, template moderated, spatially varying statistical classification. *Med Image Anal*. 2000;4(1):43-55. doi:10.1016/S1361-8415(00)00003-7
31. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging*. 2001;20(1):45-57. doi:10.1109/42.9064
32. Folstein MF, Folstein SE, McHugh PR. mini-mental state. *J Psychiatr Res*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6
33. Schretlen DJ, Winicki JM, Meyer SM, Testa SM, Pearlson GD, Gordon B. Development, psychometric properties, and validity of the hopkins adult reading test (HART). *Clin Neuropsychol*. 2009;23(6):926-943. doi:10.1080/13854040802603684
34. Schretlen DJ, Testa SM, Pearlson GD, Staff P. *Calibrated neuropsychological normative system software portfolio (CNNS-SP)*. Psychological Assessment Resources; 2010.
35. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. *Percept Mot Skills*. 1958;8(7):271. doi:10.2466/pms.8.7.271-276
36. Wechsler D. *Wechsler memory scale-revised*. Psychological Corporation; 1987.
37. Goodglass H, Kaplan E. The assessment of aphasia and related disorders. *Lea & Feibinger;*. 1983.
38. Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. *Dev Psychol*. 1991;27(5):763-776. doi:10.1037/0012-1649.27.5.763
39. Brandt J, Ralph. Hopkins Verbal Learning Test– Revised. 2001.
40. Benedict RHB, Groninger L, Schretlen D, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: studies of normal performance, reliability, and, validity. *Psychol Assess*. 1996;8(2):145-153. doi:10.1037/1040-3590.8.2.145
41. Edmonds EC, Delano-Wood L, Clark LR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimers Dement*. 2015;11(4):415-424. doi:10.1016/J.JALZ.2014.03.005
42. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: a risk profile from the Framingham Study. *Stroke*. 1991;22(3):312-318. doi:10.1161/01.STR.22.3.312
43. Nation DA, Preis SR, Beiser A, et al. Pulse pressure is associated with early brain atrophy and cognitive decline: modifying effects of APOE4. *Alzheimer Dis Assoc Disord*. 2016;30(3):210. doi:10.1097/WAD.0000000000000127
44. Glynn RJ, Chae CU, Guralnik JM, Taylor JO, Hennekens CH. Pulse pressure and mortality in older people. *Arch Intern Med*. 2000;160(18):2765-2772. doi:10.1001/ARCHINTE.160.18.2765
45. Giovannetti T, Bettcher BM, Brennan L, et al. Characterization of everyday functioning in mild cognitive impairment: a direct assessment approach. *Dement Geriatr Cogn Disord*. 2008;25(4):359-365. doi:10.1159/000121005
46. Schmitter-Edgecombe M, Parsey C, Cook DJ. Cognitive correlates of functional performance in older adults: comparison of self-report, direct observation, and performance-based measures. *J Int Neuropsychol Soc*. 2011;17(5):853-864. doi:10.1017/S1355617711000865
47. Tassoni MB, Drabick DAG, Giovannetti T. The frequency of self-reported memory failures is influenced by everyday context across the lifespan: implications for neuropsychology research and clinical practice. *Clin Neuropsychol*. 2023;37(6):1115-1135. doi:10.1080/13854046.2022.2112297

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Holmqvist SL, Jobson K, Desalme D, et al. Preliminary validation of the Virtual Kitchen Challenge as an objective and sensitive measure of everyday function associated with cerebrovascular disease. *Alzheimer's Dement*. 2024;16:e12547. https://doi.org/10.1002/dad2.12547