



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Foran, AM;Mathias, JL;Bowden, SC

Title:

Development of a Brief Screen to Detect Cognitive Impairment in Older Adults: The QuickSort

Date:

2021-02-01

Citation:

Foran, A. M., Mathias, J. L. & Bowden, S. C. (2021). Development of a Brief Screen to Detect Cognitive Impairment in Older Adults: The QuickSort. *Journal of the American Geriatrics Society*, 69 (2), pp.441-449. <https://doi.org/10.1111/jgs.16898>.

Persistent Link:

<https://hdl.handle.net/11343/276589>

Development of a brief screen to detect cognitive impairment in older adults: The QuickSort

A M Foran, M Psych (Neuro)^{1, 2}

J L Mathias, PhD¹

S C Bowden, PhD³

¹ School of Psychology, Faculty of Health & Medical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia.

² Corresponding author. Email: amie.foran@adelaide.edu.au

³ Melbourne School of Psychological Sciences, Faculty of Medicine, Dentistry, & Health Sciences University of Melbourne, Parkville, Victoria, 3010, Australia

Running head: QuickSort cognitive screen

Word Counts:

Abstract = 291

Main text = 3572

Intext tables = 3

Intext figures= 2

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/jgs.16898](https://doi.org/10.1111/jgs.16898)

Funding Sources: The study was partly funded by an Allied Health Grant from the Royal Adelaide Hospital. A M Foran was in receipt of an Australian Government Research Training Stipend.

Abstract

Background: Sorting tests detect cognitive decline in older adults who have a neurodegenerative disorder, such as Alzheimer's and Parkinson's disease. Although equally effective at detecting impairment as other cognitive screens (e.g. Mini Mental Status Examination; MMSE), sorting tests are not commonly used in this context. This study examines the QuickSort, which is a new very brief sorting test that is designed to screen older adults for cognitive impairment.

Design: Observational cohort study.

Setting: General community and inpatients, Australia.

Participants: Older (≥ 60 years) community-dwelling adults ($n=187$) and inpatients referred for neuropsychological assessment ($n=78$). A normative subsample ($n=115$), screened for cognitive and psychological disorders, was formed from the community sample.

Measurements: Participants were administered the QuickSort, MMSE, Frontal Assessment Battery (FAB) and Depression Anxiety and Stress Scale-21. The QuickSort requires people to sort 9-stimuli by color, shape and number, and to

explain the basis for their correct sorts. Sorting (range: 0-12), Explanation (range: 0-6), and Total (range: 0-18) scores were calculated for the QuickSort.

Results: The Cognitively-Healthy subsample completed the QuickSort within 2-minutes, 50% had errorless performance, and 95% had Total scores ≥ 10 . The likelihood of community-dwelling older adults and inpatients ($n=260$) being impaired on either the MMSE or FAB, or both, increased by a factor of 3.75 for QuickSort Total scores < 10 and reduced by a factor of 0.23 for scores ≥ 10 .

Conclusion: The QuickSort provides a quick, reliable and valid alternative to lengthier cognitive screens (e.g., MMSE, FAB) when screening older adults for cognitive impairment. The QuickSort performance of an older adult can be compared to a cognitively-healthy normative sample and used to estimate the likelihood they will be impaired on either the MMSE or FAB, or both. Clinicians can also use evidence-based modelling to customize the QuickSort for their setting.

1 Background

The need to screen older adults for cognitive impairment is growing as the number of people with neurodegenerative disorders increases (e.g. dementia).^{1, 2} However, clinicians' time is limited, making efficient cognitive screening imperative.³ The Mini-Mental Status Examination (MMSE) is the most commonly-used cognitive screen and is often supplemented with the Frontal Assessment Battery (FAB) to detect frontal or 'executive' deficits.^{3, 4} The MMSE and FAB each take 10 minutes to administer, with additional time for scoring.^{5, 6} Although brief, this may exceed the time available in some settings.

Sorting tests are amongst the most sensitive tests for detecting cognitive impairment,^{7, 8} but are rarely used to screen older adults. There are a number of such tests, all requiring respondents to sort stimuli according to colour, shape or number.⁹⁻¹¹ Although commonly assumed to measure 'executive' functioning, sorting tests assess multiple cognitive abilities¹²⁻¹⁴ and may, therefore, provide an alternative to the MMSE and FAB.

Of note, a recent meta-analysis found that sorting tests identify cognitive decline caused by neurodegenerative disorders of older age.¹⁵ The ability to switch categories was frequently assessed (Category score) and best detected cognitive decline in dementia, with sensitivity and specificity values both approximately 85%. These figures are comparable to those reported for the MMSE when detecting dementia in memory clinics (80% & 81%, respectively).¹³ Verbal explanations about the rule underpinning a correct sort (Explanation or Description scores) were less

common, but had the greatest sensitivity for detecting the most common neurodegenerative disorders of older age, namely Alzheimer's and Parkinsonian disorders.¹⁵

Although promising, there are multiple limitations to using existing sorting tests to screen older adults for cognitive decline, including their complexity (up to 64 stimuli) and administration time (up to 30 minutes), an inability to discontinue early when performance is intact, complicated scoring procedures, and floor effects when scoring.^{8, 11, 17} The QuickSort was developed to retain the best features of existing tests, while overcoming some of the aforementioned limitations. Specifically, the QuickSort has nine stimuli, an early discontinuation rule, a one-page record-form (reducing administration & scoring time), and a lower floor. Erroneous sorts are also explained (enabling learning) and the test can be scored even if a person is unable to complete the test or has expressive language problems.

This study was designed to evaluate the QuickSort when screening older adults for cognitive impairment by: (1) examining its user-friendliness, and inter-rater and test-retest reliability, (2) developing normative data so that the performance of older adults can be compared to their cognitively-healthy peers, and (3) evaluating its discriminant validity by assessing its ability to detect impairment on lengthier cognitive screens (MMSE, FAB) in community and clinical samples. In combination, this study will determine whether the QuickSort provides a quick, reliable and valid alternative to lengthier cognitive screens that are often used with older adults.

2 Method

2.1 Participants

Two samples of older adults (≥ 60 years) participated: (1) community-dwelling visitors, outpatients and volunteers at the Royal Adelaide Hospital (RAH), and (2) RAH inpatients who were consecutively referred for neuropsychological assessment (mostly relating to mental capacity), hereafter referred to as the Community ($n=187$) and Inpatient ($n=78$) samples, respectively. Participants were excluded if they (Community) or their medical team (Inpatients) reported they were acutely unwell or non-English speaking, or had red-green color blindness.

A cognitively-healthy normative subsample (hereafter named the Cognitively-Healthy subsample, $n=115$) was formed from the Community sample by excluding 72 participants who (1) were unable to complete the cognitive tests, (2) had a history indicating significant or multiple concussions, a diagnosed head injury, or an intellectual or learning disability, (3) were impaired on the MMSE (<24)¹⁸ or FAB (<11),⁶ or (4) were psychologically distressed (Depression, Anxiety and Stress Scale-21 [DASS-21] scores: depression >20 , anxiety >14 , or stress >25) (see Figure 1).¹⁹ The high exclusion rate resulted from recruiting through a large publicly-funded tertiary hospital that services a broad socio-demographic area. A similar exclusion rate has been reported for the MMSE.²⁰

2.2 Measures

Background demographic (age, gender, education, nationality) and medical information (visual & hearing disabilities, color blindness, conditions affecting cognition, e.g., head trauma, epilepsy) was recorded for each participant. Cognitive functioning was assessed using the MMSE (scored: 0-30), FAB (scored: 0-18) and Quicksort (see below), with higher scores indicating better cognition. Psychological distress was assessed using the DASS-21, with lower scores indicative of less symptomatology.²¹

The QuickSort uses nine cards, which are sorted according to three categories (colour, shape, number) over a maximum of 6 trials. The QuickSort Manual, Stimuli and Record Form are provided in Supplementary File 2 (On-line Supplementary Materials). An early discontinuation rule reduces administration time when cognition is intact. Three scores are calculated: (1) a 'Sorting' score (named the Category score in other tests), which aggregates the number of successful sorts, errors (repetition, set-loss, grouping, completion errors) and prompts during a maximum of six trials (range: 0-12); (2) an 'Explanation' score, which assesses an examinee's ability to explain the basis for their correct sorts (range:0-6); and (3) a 'Total' score, which sums the Sorting and Explanation scores (range: 0-18). Sorting scores are used when a person has problems with verbal expression, and both Sorting and Total scores can be calculated when someone fails to complete the QuickSort (incomplete trials scored zero). Repetition errors (repeated sorts using the same rule) and concrete explanations are also recorded for clinical purposes, but are not examined here.

An electronic version of the QuickSort (QuickSort-e), which generates the same scores as the original version, was developed to reduce clinician's training and scoring time, and to facilitate score-interpretation using the methods recommended by evidence-based medicine (EBM). Information regarding the participants' prior familiarity with, and comfort using, an iPad was also recorded.

2.3 Procedure

The Human Research and Ethics Committee of the RAH, South Australia, approved this project. Written informed consent was obtained according to the Declaration of Helsinki.

The QuickSort underwent initial focus group development using a convenience sample of nine clinical neuropsychologists, who provided subjective evaluations of its user-friendliness (administration, scoring & interpretation) prior to its use here. Three clinical psychologists additionally viewed and scored videos of 15 QuickSort performances (simulated impaired and actual older adults) to assess inter-rater reliability.

Participants were recruited between October 2013 and December 2017. A neuropsychologist or research assistant (post-graduate) conducted individual assessments in an office (Community sample) or bedside (Inpatients). The QuickSort was administered to (1) the Community sample, prior to the MMSE and FAB (same session), and (2) Inpatients, prior to their neuropsychological consultation, in order to blind assessors to the person's cognitive status. Forty-six

Inpatients were re-administered the QuickSort while in hospital to assess test-retest reliability.

The QuickSort-e was piloted in a subset of consecutively recruited community-dwelling participants ($n=29$) during the final stages of the study.

2.4 Data-analysis

Data was analysed using the Statistical Package for the Social Sciences²² using $p<.05$ and excluding missing data list-wise. Summary demographic, cognitive and psychological scores (means, SDs, n , %) were calculated for the Community and Inpatient samples, and the Cognitively-Healthy subsample.

A focus group of clinicians examined the user-friendliness and clinical acceptability of the QuickSort. Intraclass correlations (ICC), measuring absolute agreement (single measures), were used to assess inter-rater reliability ($n=3$ raters) and test-retest reliability ($n=46$ Inpatients), with .8 considered acceptable and .9 excellent.²³ Practice effects were indicated by differences between ICCs measuring consistency and absolute agreement.²³

Normative data for the QuickSort Total and Sorting scores were created using the Cognitively-Healthy subsample. Cumulative frequencies (base-rates) were calculated for the Total and Sorting scores, enabling the scores of older adults to be compared to their cognitively-healthy peers. Linear regressions determined whether this normative data needed to be stratified by age, education or gender. Norms were stratified if any variable accounted for >10% of the variance (small effect).²⁴

Next, the QuickSort discriminant validity was examined in terms of its ability to detect impairment on: (1) the MMSE, (2) the FAB, and (3) either the MMSE or FAB, or both. The Community and Inpatient samples were combined for this purpose, after which participants were classified into one of two Diagnostic groups: *cognitively impaired* on (1) the MMSE<24, (2) the FAB<11, and (3) either the MMSE or FAB, or both, and *non-impaired* (MMSE≥25, FAB≥12) (See Figure 1 for details). *t*-tests assessed whether the Diagnostic groups (impaired & non-impaired) were demographically comparable (age & education can independently affect cognition) and whether their QuickSort scores differed (Sorting, Explanation & Total scores, Repetition errors, Concrete responses). ANCOVAs further investigated whether demographic differences between the Diagnostic groups significantly contributed to differences in their QuickSort performance (Sorting & Total scores).

A logistic regression identified the QuickSort Total cut-score that correctly classified the largest number of participants into the Diagnostic groups. A power analysis indicated that a minimum sample size of 42 was required to detect a large difference (Cohen's $d=.80$) in the QuickSort scores with 95% power and $\alpha=.05$.²⁴ In addition, the most clinically-useful QuickSort cut-score was obtained by tallying the numbers of people who scored above and below each score and who were impaired and non-impaired on (1) the MMSE, (2) the FAB, and (3) either the MMSE or FAB, or both. The CATmaker²⁵ was used to calculate sensitivity and specificity, likelihood ratios (LRs) and 95% confidence intervals (CIs). The Total cut-scores that were most clinically-useful for ruling-in (sensitivity important) or ruling-out (specificity

important) cognitive impairment were identified. LRs >1 indicate that QuickSort scores were associated with impairment on the MMSE or FAB, with LRs <1 indicating an absence of impairment on these screens.²⁶ Clinically, scores with LRs >3 or <0.3 are considered most useful because they substantially change the likelihood of the person being impaired or non-impaired, respectively.²⁷

ICCs, measuring absolute agreement (single measures), were used to assess convergent validity between the original QuickSort and the QuickSort-e.²³

3 Results

3.1 Community and Inpatient samples summary information

Table 1 provides summary demographic and test information (MMSE, FAB, QuickSort, DASS-21) for the Community ($n=187$) and Inpatient ($n=78$) samples, from which the Cognitively-Healthy normative subsample and Diagnostic groups were formed. Both samples had a mean age in their seventies and had completed approximately four years of high school. There was a slightly higher proportion of males and most participants were born in Australia. On average, the Inpatient group had poorer cognition (MMSE, FAB, QuickSort), greater psychological distress (DASS-21), and took longer to complete the QuickSort than the Community sample.

3.2 QuickSort clinical acceptability, inter-rater reliability & test-retest reliability

All clinicians in the focus group indicated that the QuickSort was user-friendly. The QuickSort Sorting, Explanation and Total scores ($n=15$ cases) provided by three independent raters were also in agreement ($ICC=1.00$), indicating that the scoring procedures are clear and have very high inter-rater reliability.

Test-retest reliability was assessed in 46 Inpatients who were readministered the QuickSort after an average 4.6 days ($SD=3.0$). The Sorting ($ICC=.75$), Explanation ($ICC=.79$) and Total ($ICC=.81$) scores all showed acceptable absolute agreement (stability) over time. These coefficients were similar to the ICCs for consistency in the Sorting, Explanation and Total scores, ($ICC= .76, .79, .81$ respectively), indicating there were minimal practice effects, even after a short interval.

3.3 QuickSort normative data

Norms for the QuickSort were created from the Cognitively-Healthy subsample ($n=115$). As seen in Table 1, this subsample closely resembled the Community sample from which it was drawn, but had slightly fewer males. On average, the Cognitively-Healthy participants completed the QuickSort in under 2-minutes. Not unexpectedly, the QuickSort scores for this subsample were skewed (see Supplementary Table S1 for interquartiles), with the average Total score approaching 16 ($max=18$). Of note, 54% ($n=62$) achieved the maximum Total score and 95% ($n=109$) scored ≥ 10 . The base-rates for the Total and Sorting score are provided in Table 2.

Linear regressions performed on both the QuickSort Total and Sorting scores of the Cognitively-Healthy subsample revealed that age ($r^2=3.4\%$ & 3.2%) and education ($r^2=8.8\%$ & 8.4%), but not gender ($r^2<0.1\%$ & 0.1%), had a significant but very small ($<10\%$) impact on performance. Thus, the test norms did not need to be demographically-stratified (see Supplementary Table S2 for analyses).

3.4 QuickSort validity: detecting impairment & non-impairment on the MMSE & FAB

As indicated above, Diagnostic groups (impaired & non-impaired) were formed from the Community and Inpatient data (total $n=260$) to assess the QuickSort ability to identify those impaired on the: (1) MMSE <24 , (2) FAB <11 , and (3) either the MMSE <24 or FAB <11 , or both.

When the Diagnostic groups were compared, the impaired groups had lower QuickSort scores (Sorting, Explanation & Total scores, Repetition errors, but not Concrete responses; see Supplementary Tables S3-S5), and took longer to complete it (excluding those only impaired on the FAB). Although significantly older and less educated than the non-impaired group, ANCOVAs revealed that the QuickSort scores of the impaired groups remained significantly lower after controlling for these demographic differences (see Supplementary Table S6). Thus, the QuickSort showed good discriminant validity.

A logistic regression identified the QuickSort Total cut-score that optimised the classification of participants as cognitively impaired versus non-impaired on

either the MMSE or FAB, or both (the only analysis that was sufficiently powered). A cut-score <4 correctly classified 84% of participants, with 43% sensitivity and 94% specificity. Scores <4 increased the likelihood of impairment on either the MMSE or FAB, or both, by a factor of 6.95 (95%CI: 3.75–8.41), but scores ≥ 4 only reduced the likelihood of impairment by a factor of 0.61 (95%CI: 0.48–0.77). The high specificity resulted from scores <4 being infrequent in the Cognitively-Healthy normative group (1.7%, Table 2).

There may be situations where detecting cognitive impairment (sensitivity) takes precedence over ruling it out (specificity), or vice versa. Sensitivity, specificity and LR (& 95% CIs) statistics are therefore provided for every QuickSort Total score when predicting impairment on the MMSE, FAB, and either the MMSE or FAB, or both (see Table 3). These data were also calculated for Sorting scores (see Supplementary Table S7), but should only be used when a person has problems with verbal expression because the Total score had larger positive and smaller negative LRs.

In clinical settings where ruling-in (sensitivity) and ruling-out (specificity) cognitive impairment are both important, a cut-score <10 may be preferable because sensitivity increases to 78%, with 82% specificity (see Table 3). LRs indicate that scores <10 increase the likelihood that a person is impaired by a factor of 3.18 (95%CI: 2.43–4.17), 3.63 (95%CI: 2.74–4.72), 3.75 (95%CI: 2.81–4.98) on the MMSE, FAB, and either the MMSE or FAB, or both, respectively. Alternatively, scores of ≥ 10 reduce the likelihood that a person is impaired by a factor of 0.24

(95%CI: 0.11–0.49), 0.16 (95%CI: 0.06–0.39) and 0.23 (95%CI: 0.13–0.41) on the MMSE, FAB, and either the MMSE or FAB, or both, respectively. Therefore, a cut-score of <10 may prove more useful because it can both rule-in *and* rule-out impairment on lengthier cognitive screens.

As with most tests, there is greater certainty surrounding very high or low scores. For example, Total scores <2 increase the likelihood of impairment on either the MMSE or FAB, or both, by a factor of 9.26 (95%CI: 2.96–28.75) and Total scores ≥ 17 reduce the likelihood of impairment by a factor of 0.16 (95%CI: 0.06–0.41). Thus, the most informative way to interpret any QuickSort score is to use the associated LR.

According to EBM, the local prevalence of impairment (pre-test probability, which can be estimated from published research or a clinical audit) should be taken into consideration when estimating a patient's probability of impairment on the MMSE or FAB.^{28, 29} As seen in Figure 2, the likelihood of impairment on either the MMSE or FAB, or both, can be calculated for two patients who score 5 on the QuickSort (see <6 cut-score LR+=5.45, Table 3, right panel), but are seen in different clinical settings: one with a pre-test probability of impairment of 20% (solid line) and the other with a 50% pre-test probability (dotted line). Lines from these two different pre-test probabilities (left Y-axis) through the LR+ of 5.45 (centre Y-axis), yield post-test probabilities (right Y-axis) of 58% and 85%, respectively.

3.5 QuickSort-e preliminary findings

Author Manuscript

Twenty-nine consecutive community-dwelling participants were additionally administered the QuickSort-e. Everyone reported being very comfortable using the iPad, despite 45% having not previously used an iPad. The QuickSort-e took slightly longer to administer than the original version ($M = 3.08$, $SD = 1.84$), but this is offset by automatic scoring. The two versions of the QuickSort had satisfactory to good convergent validity (Sorting ICC=.72, Explanation ICC=.86, Total ICC=.84).

4. Discussion

The QuickSort is a brief new cognitive screen that is designed to identify cognitive impairment in clinical settings where resources are very limited. The QuickSort assesses sorting ability, which deteriorates as a consequence of multiple neurodegenerative disorders,¹⁵ and seeks to improve on existing sorting tests. Specifically, the QuickSort is quicker to administer and score, simpler for older adults, provides a larger range of scores, and can be used even when a person is unable to complete the test or has expressive language problems. This study examined whether the QuickSort provides a fast, reliable and valid screen for older adults that can be used as an alternative to two of the most common, but lengthier, cognitive screens: the MMSE and FAB.

Although simple from a test-taker's perspective, the QuickSort requires clinicians to follow detailed instructions, including specific prompts for different types of errors and early discontinuation when cognition is intact. Despite these complexities, clinicians reported that the QuickSort was user-friendly; a finding that was supported by its good inter-rater reliability. Test-retest data additionally

indicated that the QuickSort provides stable scores in an inpatient setting and is not impacted by practice effects. Preliminary findings for the QuickSort-e suggest that it is also very brief, even when participants were unfamiliar with an iPad, and it generates scores that are comparable to those of the original version.

The normative data for the QuickSort was based on a subgroup of cognitively- and psychologically-healthy older adults, most of whom completed the task within two minutes, with over half achieving a perfect score. Cumulative frequencies for the Sorting and Total scores enable clinicians to evaluate whether a person's performance is common or unusual, relative to their cognitively-healthy peers. Age, education and gender did not significantly affect QuickSort performance, eliminating the need for demographically-adjusted norms.

Discriminant validity was evaluated by comparing people who were impaired with those who were not impaired on the MMSE or FAB. The impaired group had significantly lower QuickSort scores than the non-impaired group, a finding that was not attributable to the former being older and less educated. The QuickSort Total cut-score that correctly classified the largest number of people as impaired on either the MMSE or FAB, or both, was <4, however sensitivity (43%) was sacrificed for specificity (94%). A cut-score <10 may therefore be preferred in clinical settings because it can be used to rule-in (82% sensitivity) and rule-out (78% specificity) impairment (see Table 3).

Although useful, single cut-scores fail to utilize the information provided by low and high scores.²⁶ LRs help to address these problems. For example, QuickSort

scores <2 were not seen in cognitively-healthy older adults and increase the likelihood of impairment on either the MMSE or FAB, or both, by a factor of 9.26 (95% CI: 2.96 – 28.75). Conversely, QuickSort scores ≥ 17 , which were common in cognitively-healthy adults (63%), reduce the likelihood of impairment by a factor of 0.16.

The use of a nomogram or on-line calculator³¹ to estimate a person's post-test probability of impairment – based on their QuickSort score and the prevalence of impairment in that clinical setting – further enhances its clinical utility. For example, a patient who gets a QuickSort score of five in a clinical setting where approximately 50% of patients are cognitively impaired has an 85% likelihood that they will be impaired on lengthier cognitive screens, suggesting they need to undergo further investigations into the presence of cognitive decline.

A limitation of this study relates to the sample sizes. Test-retest reliability was assessed using a small convenience sample of inpatients and now needs to be evaluated in a community-dwelling sample. Administration time was only recorded for a subset of participants and needs to be assessed further. Although the QuickSort normative sample is larger than those originally reported for the MMSE and FAB,^{32, 33} larger normative datasets, stratified by age and education, are now available for these measures,^{20, 34-37} suggesting the QuickSort norms should be expanded. Lastly, cumulative frequencies above and below cut-scores were reported because the small samples precluded multiple level LRs and interpretation of stand-alone scores.

Future research should examine patients who have more overt cognitive impairments, such as adults with dementia. The QuickSort-e also needs to be developed further by integrating decision-making algorithms that instantaneously provide post-test probabilities in response to clinical questions that are relevant to specific settings (e.g., the likelihood the patient is cognitively impaired, has a neurodegenerative disorder, or will be readmitted to hospital). Lastly the QuickSort-e may be suitable for telehealth assessments, which are in much greater demand as a result of COVID-19, but additional reliability and validity studies are needed to support this type of use.

Overall, the QuickSort assesses the cognitive decline associated with various neurodegenerative disorders of older age.¹⁵ It provides a quick, easy, reliable and valid cognitive screen that is suitable for use in busy clinical settings. Importantly, the QuickSort provides a viable alternative to lengthier screens, such as the MMSE and FAB. Clinicians are encouraged to customize it to their clinical setting, using post-test calculators, to improve the accuracy and efficiency of their cognitive screening.

5 Acknowledgements

5.1 Conflict of interest

The authors have no conflicts of interest to disclose.

5.2 Author contributions

A M Foran designed and created the QuickSort, and was responsible for study inception and execution, data collection and coding, data analysis and reporting, and final manuscript preparation. J L Mathias supervised all aspects of this research, particularly the study design, data analysis, and preparation of both the QuickSort manual and study manuscript. S C Bowden also contributed throughout all stages of the study, and was particularly involved in providing statistical advice and oversight.

5.3 Sponsors role

The research was partially funded by a Royal Adelaide Hospital Allied Health Grant, which assisted with participant recruitment and data collection. The first author was supported by an Australian Government Research Training Stipend.

References

1. Ecomonics A. Dementia epidemic must be 'front of mind'. *Australian Nursing Journal*. 2009;**17**: 24.
2. Prince MJ, Wimo A, Guerchet M, Ali G-C, Wu Y-T, Prina M. World Alzheimer Report 2015: The Global Impact of Dementia: an analysis of prevalence, incidence, cost and trends. London: <https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf>: Alzheimer's Disease International & King's College, London, 2015.
3. Robinson L, Tang E, Taylor J-P. Dementia: timely diagnosis and early intervention. *BMJ : British Medical Journal*. 2015; **350**.
4. Lee DY, Kim JW, Seo EH, Sohn BK, Park SY, Woo JI. Improvement of dementia screening accuracy of Mini-Mental State Examination (MMSE) by education-adjustment and supplementation of Frontal Assessment Battery (FAB) performance. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. 2012;**8**: P415-P415.
5. Sheehan B. Assessment scales in dementia. *Therapeutic Advances in Neurological Disorders*. 2012;**5**: 349-358.
6. Slachevsky A, Villalpando JM, Sarazin M, Hahn-Barma V, Pillon B, Dubois B. Frontal Assessment Battery and differential diagnosis of frontotemporal dementia and Alzheimer disease. *Archives of Neurology*. 2004;**61**: 1104-1107.
7. Reitan R. *The Halstead-Reitan Neuropsychological Test Battery: theory and clinical interpretation*. 2 ed. Tucson, USA: Neuropsychology Press, 1993.
8. Delis D, Squire L, Bihrlle A, Massman P. Componential analysis of problem-solving ability: performance of patients with frontal lobe damage and amnesic patients on a new sorting test. *Neuropsychologia*. 1992;**30**: 683-697.
9. Goldstein KH, Scheerer M. Abstract and concrete behaviour: an experimental study with special tests. *Psychologic Monograph*. 1941;**53**: 239.
10. Weigl E. Zur Psychologie sogenannter Abstraktionsprozesse. *Psychol* 103:2-45. 1927;[translated in *J Abnorm Soc Psychol* **36**: 3-33].
11. Berg EA. A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology*. 1948;**39**: 15.
12. Floyd RG, Bergeron R, Hamilton G, Parra GR. How do executive functions fit with the Cattell-Horn-Carroll Model? some evidence from a joint factor analysis of the Delis-Kaplan Executive Function System and the Woodcock-Johnson III tests of cognitive abilities. *Psychology in the Schools*. 2010;**47**: 721-738.
13. Schneider WJ, McGrew KS. The Cattell-Horn-Carroll Theory of cognitive abilities. In: Flanagan DP, McDonough EM, eds. *Contemporary intellectual assessment: Theories, tests and issues*, 4 edn. New York: Guilford Press, 2018.
14. Hobson P, Meara, J., & Taylor, C. The Weigl Colour-Form Sorting Test: a quick and easily administered bedside screen for dementia and executive dysfunction. *Int J Geriatr Psychiatry*. 2007;**22**: 909-915.
15. Foran AM, Mathias JL, Bowden SC. Effectiveness of sorting tests for detecting cognitive decline in older adults with dementia and other common neurodegenerative disorders: a meta-analysis. under final review.
16. Mitchell AJ. A meta-analysis of the accuracy of the Mini-Mental State Examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*. 2009;**43**: 411-431.

17. Beglinger LJ, Unverzagt FW, Beristain X, Kareken D. An updated version of the Weigl discriminates adults with dementia from those with mild impairment and healthy controls. *Archives of Clinical Neuropsychology*. 2008;**23**: 149-156.
18. Hancock P, Larner AJ. Test your memory test: diagnostic utility in a memory clinic population. *International Journal of Geriatric Psychiatry*. 2011;**26**: 976-980.
19. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. In: Psychology Foundation. Sydney, 1995.
20. Grigoletto F, Zappalà G, Anderson DW, Lebowitz BD. Norms for the Mini-Mental State Examination in a healthy population. *Neurology*. 1999;**53**: 315.
21. Lovibond SH, Lovibond PF. *Manual for the Depression Anxiety Stress Scales*. 2nd ed. Sydney: Psychology Foundation, 1995.
22. IBM. SPSS Statistics for Windows. **Volume Version 25.0** Armonk, NY: IBM Corp, Released 2017.
23. Bowden SC, Finch S. When is a test reliable enough and why does it matter? In: Bowden SC, ed. *Neuropsychological assessment in the age of evidence-based practice* NY: Oxford University Press, 2017.
24. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates, 1988.
25. Badenoch DS, D; Straus, S; Ball, C; Dawes, M. CATmaker. In: 1.1 v, ed. Oxford: Centre for Evidence-based Medicine, 2004.
26. Deeks JA, D. Diagnostic tests 4: likelihood ratios. *British Medical Journal*. 2004;**329**: 168.
27. McGee S. Teaching Evidence-Based Physical Diagnosis: Six Bedside Lessons. *Southern medical journal*. 2016;**109**: 738.
28. Straus SE, Pattani R, Veroniki A. *Evidence-based medicine: how to practice and teach EBM*: Elsevier, 2019.
29. Bunnage M. How Do I Know When a Diagnostic Test Works? In: Bowden SC, ed. *Neuropsychological Assessment in the Age of Evidence-Based Practice: Diagnostic and Treatment Evaluations National Academy of Neuropsychology: Series on Evidence-Based Practice* OUP, USA.: Oxford University Press, 2017.
30. Fagan TJ. Nomogram for Bayes's Theorem. *New England Journal of Medicine*. 1975;**293**: 257.
31. Kohn MA, Jarrett MS, Senyak J. Post-test probability calculator. Clinical and Translational Science Institute, 2020.
32. Folstein M, Folstein S, McHugh P. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975;**12**: 189-198.
33. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology*. 2000;**55**: 1621.
34. Sakuma N, Ura C, Miyamae F, et al. Distribution of Mini-Mental State Examination scores among urban community-dwelling older adults in Japan. *International Journal of Geriatric Psychiatry*. 2017;**32**: 718-725.
35. Li H, Jia J, Yang Z. Mini-Mental State Examination in elderly Chinese: a population-based normative study. *Journal of Alzheimer's disease : JAD*. 2016;**53**: 487.
36. Freitas S, Simões MR, Alves L, Santana I. The relevance of sociodemographic and health variables on MMSE normative data. *Applied Neuropsychology: Adult*. 2015;**22**: 311-319.

37. Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based Norms for the Mini-Mental State Examination by age and educational level. *JAMA*. 1993;**269**: 2386-2391.

Figure 1. Participant flow chart

¹MMSE <24; ²MMSE ≥24 ($n = 1$ in the Community sample and $n = 6$ in the Inpatient sample were intact on the MMSE, but impaired on the FAB); ³FAB <11; ⁴FAB ≥11 ($n = 1$ in the Community sample and $n = 14$ in the Inpatient sample were intact on the FAB, but impaired on the MMSE); ⁵MMSE ≥24 and FAB ≥11; ⁶impaired on either the MMSE <24 or FAB <11, or both

Figure 2. Nomogram showing the post-test probability of impairment on the MMSE or FAB, or both

Nomogram showing the post-test probability of impairment on the MMSE or FAB, or both, for a person with a QuickSort Total score of 5 and $LR+ = 5.54$ (see Table 3).

Example 1 (yellow line) shows a clinical setting where the pre-test probability (left Y-axis) is estimated to be 20%, resulting in an estimated post-test probability of approximately 58% (right Y-axis).

Example 2 (red line) shows a situation where the pre-test probability is estimated to be 50% resulting in a post-test probability of approximately 85%.

Adapted from Fagan (1975) Nomogram for Bayes's Theorem

Author Manuscript

Table 1. Summary demographic and test data for the Community and Inpatient samples, and Cognitively-Healthy normative subsample

	Community sample (<i>n</i> = 187)				Inpatient sample (<i>n</i> = 78)				Cognitively-healthy normative subsample (<i>n</i> = 115)			
	<i>n</i>	%	<i>M</i>	<i>SD</i>	<i>n</i>	%	<i>M</i>	<i>SD</i>	<i>n</i>	%	<i>M</i>	<i>SD</i>
Age	187		71.3	7.48	78		74.6	8.76	115		71.2	7.70
Education (years)	187		11.5	2.96	72		11.1	3.88	115		11.7	2.84
Gender												
Male	98	52%			47	60%			53	46%		
Female	89	48%			31	40%			62	54%		
Nationality												
Australian ¹	121	65%			54	69%			72	63%		
European	57	30%			20	26%			36	31%		
Asian	3	2%			0	0%			3	3%		
Other	6	3%			2	3%			4	4%		
MMSE ²												
Total score (range: 0-30)	187		28.2	1.68	71		23.7	3.87	115		28.5	1.29
Impaired: MMSE <24	3	2%			31	40%			0	0%		
FAB ³												
Total score (range: 0-18)	187		15.1	2.59	69		11.8	3.09	115		15.6	1.99
Impaired: FAB <11	11	6%			23	29%			0	0%		
MMSE and FAB												
Impaired: either or both screens	12 ⁶	6%			39	50%			0	0%		
QuickSort												
Sorting score (range: 0-18)	187		9.4	3.53	78		4.2	3.83	115		10.4	2.66
Explanation score (range: 0-9)	187		4.9	1.59	78		2.6	2.22	115		5.4	1.14
Total score (range: 0-18)	187		14.3	4.80	78		6.8	5.88	115		15.8	3.37
Total score <10	34	18%			58	74%			6	5%		
Administration time ⁴	26		2 _{min} 8 _s	1 _{min} 17 _s	17		4 _{min} 92 _s	2 _{min} 37 _s	19		1 _{min} 43 _s	51 _s
Repetition errors (range: 0-5)	187		0.5	0.96	78		1.7	1.51	115		0.3	0.53

Concrete responses (range: 0-3)	187	0.4	0.68	78	0.2	0.50	115	0.4	0.74
DASS-21 ⁵									
Depression (range: 0-21)	187	2.3	3.36	4	5.5	2.05	115	2.3	3.35
Anxiety (range: 0-21)	187	2.6	2.87	4	6.0	5.60	115	2.6	2.83
Stress (range: 0-21)	187	4.3	3.97	4	3.5	3.70	115	4.3	3.88

¹n=1 person identified as Indigenous in each of the Community and Inpatient samples, ²MMSE = Mini Mental Status Examination, ³FAB = Frontal Assessment Battery, ⁴min = minutes, s = seconds ⁵DASS-21 Depression, Anxiety and Stress Scale-21

Table 2. Cumulative frequency (base-rates) for the QuickSort Total & Sorting scores in the Cognitively-Healthy normative subsample ($n = 115$)

Total	QuickSort cut-score	Cumulative Frequency ¹
	<3	0
	<4	1.7
	<5	1.7
	<6	2.6
	<7	4.3
	<8	4.3
	<9	5.2
	<10	5.2
	<11	7.0
	<12	9.6
	<13	12.2
	<14	21.7
	<15	24.3
	<16	36.5
	<17	37.4
	<18	46.1
	18	53.9
Sorting		
	<1	0
	<2	0.9
	<3	2.6
	<4	4.3
	<5	4.3
	<6	7.0
	<7	8.7
	<8	16.5
	<9	16.5
	<10	32.2
	<11	32.2
	<12	32.2
	12	67.8

¹percentage of the normative sample with scores below the cut score

Table 3. Diagnostic data for QuickSort Total scores, when predicting impairment on the MMSE (left columns), FAB (centre columns) and either the MMSE or FAB, or both (right columns)

QuickSort Total cut-score	Impaired MMSE (<24) (n = 258)				Impaired FAB (<11) (n = 256)				Impaired on either the MMSE or FAB, or both (n = 260)			
	Se	Sp	LR+	LR-	Se	Sp	LR+	LR-	Se	Sp	LR+	LR-
<1	0.06	0.98	3.29 (0.63 – 17.30)	0.96 (0.88 – 1.04)	0.15	0.99	16.32 (3.30 – 80.82)	0.86 (0.75 – 0.99)	0.10	0.99	9.80 (2.05 – 51.31)	0.91 (0.83 – 1.00)
<2	0.15	0.97	5.49 (1.77 – 17.01)	0.88 (0.76 – 1.01)	0.21	0.97	7.62 (2.72 – 21.32)	0.82 (0.69 – 0.97)	0.18	0.98	9.26 (2.96 – 28.75)	0.84 (0.74 – 0.96)
<3	0.38	0.95	7.79 (3.80 – 15.95)	0.65 (0.50 – 0.85)	0.32	0.96	8.98 (3.89 – 20.72)	0.70 (0.56 – 0.89)	0.35	0.96	9.29 (4.25 – 20.01)	0.67 (0.55 – 0.83)
<4	0.44	0.92	5.49 (3.07 – 9.83)	0.61 (0.45 – 0.82)	0.44	0.92	5.44 (3.04 – 9.74)	0.61 (0.45 – 0.82)	0.43	0.94	6.95 (3.75 – 8.41)	0.61 (0.48 – 0.77)
<5	0.53	0.89	4.14 (2.92 – 7.72)	0.53 (0.37 – 0.76)	0.50	0.88	4.27 (2.61 – 6.99)	0.57 (0.40 – 0.80)	0.51	0.91	5.60 (3.38 – 9.30)	0.54 (0.41 – 0.72)
<6	0.62	0.85	4.19 (2.78 – 6.33)	0.45 (0.29 – 0.69)	0.56	0.86	4.00 (2.57 – 6.23)	0.51 (0.35 – 0.75)	0.63	0.89	5.45 (3.55 – 8.41)	0.42 (0.29 – 0.61)
<7	0.62	0.81	3.22 (2.21 – 4.69)	0.47 (0.31 – 0.73)	0.74	0.82	4.19 (2.95 – 5.93)	0.32 (0.18 – 0.56)	0.65	0.84	4.10 (2.82 – 5.95)	0.42 (0.29 – 0.61)
<8	0.76	0.79	3.57 (2.61 – 4.88)	0.30 (0.16 – 0.55)	0.76	0.80	3.77 (2.74 – 5.20)	0.30 (0.16 – 0.54)	0.77	0.82	4.32 (3.11 – 6.01)	0.29 (0.17 – 0.47)
<9	0.82	0.74	3.18 (2.43 – 4.17)	0.24 (0.11 – 0.49)	0.88	0.76	3.63 (3.63 – 4.72)	0.16 (0.06 – 0.39)	0.82	0.78	3.75 (2.81 – 4.98)	0.23 (0.13 – 0.41)
<10	0.82	0.74	3.18 (2.43 – 4.17)	0.24 (0.11 – 0.49)	0.88	0.76	3.63 (2.74 – 4.72)	0.16 (0.06 – 0.39)	0.82	0.78	3.75 (2.81 – 4.98)	0.23 (0.13 – 0.41)
<11	0.82	0.73	3.02 (2.32 – 3.94)	0.24 (0.12 – 0.50)	0.88	0.74	3.44 (2.66 – 4.40)	0.16 (0.06 – 0.40)	0.82	0.77	3.52 (2.67 – 4.63)	0.23 (0.13 – 0.42)
<12	0.82	0.70	2.75 (2.14 – 3.55)	0.25 (0.12 – 0.52)	0.91	0.72	3.26 (2.53 – 4.13)	0.12 (0.04 – 0.36)	0.84	0.74	3.27 (2.52 – 4.23)	0.21 (0.11 – 0.40)
<13	0.82	0.68	2.56 (2.00 – 3.28)	0.26 (0.13 – 0.54)	0.91	0.70	3.02 (2.41 – 3.79)	0.13 (0.04 – 0.37)	0.84	0.72	1.65 (1.36 – 2.09)	0.31 (0.16 – 0.61)
<14	0.82	0.61	2.12 (1.69 – 2.66)	0.24 (0.14 – 0.60)	0.91	0.63	2.47 (2.02 – 3.02)	0.14 (0.05 – 0.41)	0.84	0.64	2.83 (2.83 – 0.24)	0.24 (0.13 – 0.46)
<15	0.85	0.57	1.97 (1.60 – 2.42)	0.26 (0.11 – 0.59)	0.91	0.58	2.18 (1.81 – 2.62)	0.15 (0.05 – 0.45)	0.86	0.60	2.15 (1.76 – 2.62)	0.23 (0.11 – 0.46)

<16	0.88	0.48	1.70 (1.43 – 2.03)	0.24 (0.10 – 0.62)	0.94	0.50	1.87 (1.60 – 2.18)	0.12 (0.03 – 0.46)	0.90	0.51	1.85 (1.57 – 2.18)	0.19 (0.08 – 0.45)
<17	0.88	0.46	1.65 (1.39 – 1.96)	0.25 (0.10 – 0.64)	0.97	0.49	1.90 (1.65 – 2.19)	0.06 (0.01 – 0.42)	0.92	0.50	1.84 (1.57 – 2.15)	0.16 (0.06 – 0.41)
<18	0.97	0.40	1.61 (1.43 – 1.82)	0.07 (0.01 – 0.51)	0.97	0.40	1.62 (1.43 – 0.51)	0.07 (0.01 – 0.51)	0.98	0.43	1.71 (1.51 – 1.93)	0.05 (0.01 – 0.32)

MMSE = Mini Mental Status Examination; FAB = Frontal Assessment Battery; Se = sensitivity; Sp = specificity; LR+ = positive likelihood ratio (& 95% Confidence intervals), LR- = negative likelihood ratio (& 95% Confidence Intervals); LR are calculated from raw frequencies as reported from the CATmaker

Supplementary Files

Supplementary File 1

Supplementary Table Title	Page N ^o
Supplementary Table S1. Demographic & test data for the Cognitively-Healthy normative subsample, with additional data showing the distribution of scores	1
Supplementary Table S2. Results of linear regression analyses examining the influence of demographic variables on the QuickSort Total and Sorting scores in the normative subsample	2
Supplementary Table S3. Summary demographic & tests scores for the impaired and non-impaired Diagnostic groups, formed using MMSE scores ($n = 258$)	3
Supplementary Table S4. Summary demographic & tests scores for the impaired and non-impaired Diagnostic groups, formed using the FAB ($n = 256$)	4
Supplementary Table S5. Summary demographic & tests scores for the impaired and non-impaired Diagnostic groups, formed using either the MMSE or FAB, or both ($n = 260$)	5
Supplementary Table S6. ANCOVA investigating the influence of age & education when predicting impairment on the MMSE, FAB, and either the MMSE or FAB, or both	6
Supplementary Table S7. Diagnostic data for QuickSort Sorting scores, when predicting impairment on the MMSE (left columns), FAB (centre columns) and either the MMSE or FAB, or both (right columns)	7

Supplementary File 2

QuickSort Manual and Test Stimuli

Color Charge Form Journal of the American Geriatrics Society

Upload this form to your submission OR return it to:
Emily Hammond / JAGS Editorial Office
jags@jgeditorial.com

Dear Author:
If you have supplied color figures in your submission, the first color figure in print is free of cost. Any additional figures you wish to print in color will incur a fee of \$500 each. Please indicate how EACH of your figures should be handled in the table below.

INSTRUCTIONS:	Figure No.	Color in PRINT (waive fee)	Color in PRINT (charge \$500 per)*	Color ONLINE ONLY (no charge)	Black/White ONLINE and PRINT
<p>For EVERY figure in your paper, an X should be provided in the appropriate row.</p> <p>Place an 'X' in column 1 for a maximum of ONE figure.</p> <p>Place an 'X' in column 2 if you want to pay for additional color printing. The total in this column should be entered into the Quantity field below.</p>	1				X
	2	X			
	3				
	4				
	5				
	TOTAL:				

*If you indicate ANY figures in this column, the contact information below MUST be provided.

Please Complete & Please Print

Component(s) in Article	Charge(s)	Quantity	Estimated Charge(s) (in USD)
Color Figure Charge Standard	\$500 per figure	1	\$ 0
			\$ 0 = Total Estimated

Name:	Amie May Foran
Manuscript No.:	JAGS-1638-C1-Jul-20.R1
Manuscript Title:	Development of a brief screen to detect cognitive impairment in older adults: ^{The} QuickSort
Address:	School of Psychology & Medical Sciences Faculty of Health University of Adelaide, South Australia, 5005, Australia
Phone Number:	+ 61 402453978
Fax Number:	N/A
Email Address:	amie.foran@adelaide.edu.au

Verify your agreement to pay by signing and dating below. Wiley will send an invoice to the email address you've provided:

Signature:  Date: 4/10/20



