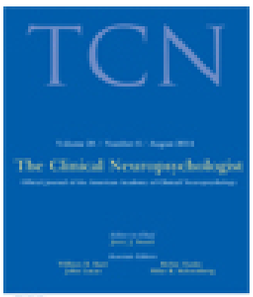

The Clinical Neuropsychologist

ISSN: 1385-4046 (Print) 1744-4144 (Online) Journal homepage: http://www.tandfonline.com/loi/ntcn20

How Do Scores on the ADAS-Cog, MMSE, and CDR-SOB Correspond?

Steve Balsis, Jared F. Benge, Deborah A. Lowe, Lisa Geraci & Rachelle S. Doody

To cite this article: Steve Balsis, Jared F. Benge, Deborah A. Lowe, Lisa Geraci & Rachelle S. Doody (2015): How Do Scores on the ADAS-Cog, MMSE, and CDR-SOB Correspond?, The Clinical Neuropsychologist, DOI: 10.1080/13854046.2015.1119312

To link to this article: http://dx.doi.org/10.1080/13854046.2015.1119312

Published online: 30 Nov 2015.

The Clinical Neuropsychologist, 2015
http://dx.doi.org/10.1080/13854046.2015.1119312

Routledge
Taylor & Francis Group

How Do Scores on the ADAS-Cog, MMSE, and CDR-SOB Correspond?

Steve Balsis[1], Jared F. Benge[2], Deborah A. Lowe[1], Lisa Geraci[1], and Rachelle S. Doody[3]

[1] Department of Psychology, Texas A & M University, College Station, TX, USA
[2] Department of Neurology, Baylor Scott & White Health, Temple, TX, USA
[3] Department of Neurology, Alzheimer's Disease and Memory Disorders Center, Baylor College of Medicine, Houston, TX, USA

Objective: Clinicians and researchers who measure cognitive dysfunction often use the Alzheimer's Disease Assessment Scale—Cognitive Subscale (ADAS-Cog), the Mini-Mental State Examination (MMSE), or the Clinical Dementia Rating scale (CDR-SOB). But, the use of different measures can make it difficult to compare data across patients or studies. What is needed is a simple chart that shows how scores on these three important measures correspond to each other. *Methods*: Using data from 1709 participants from the Alzheimer's Disease Neuroimaging Initiative and item response theory-based statistics, we analyzed how scores on each measure, the ADAS-Cog, the MMSE, and the CDR-SOB, correspond. *Results*: Results indicated multiple inflections in CDR-SOB and ADAS-Cog scores within a given MMSE score, suggesting that the CDR-SOB and ADAS-Cog are more precise in measuring the severity of cognitive dysfunction than the MMSE. *Conclusions*: This study shows how scores on these three popular measures of cognitive dysfunction correspond to each other, which is very useful information for both researchers and clinicians.

Keyword: ADAS; Alzheimer's disease; CDR; Dementia; IRT; MCI; MMSE; Psychometric; Neuropsychological assessment.

The past several decades have seen significant growth in Alzheimer's disease (AD) research. Several research and clinical priorities have emerged from this research, including maximizing the potential from clinical trial data and evaluating/identifying individuals in the earliest stages of the disease (Mitchell et al., 2012). The measurement of cognition is a critical step for identifying the early stages of the disease, and across studies and centers, various scales are employed to this end. Some of the most commonly used instruments to assess AD-related cognitive decline including global decline are the Alzheimer's Disease Assessment Scale—Cognitive Subscale (ADAS-Cog; Rosen, Mohs, & Davis, 1984), which is frequently used in pharmaceutical trials, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), which

Address correspondence to: Steve Balsis, Ph.D., Associate Professor, Department of Psychology, Texas A&M University, Mailstop 4235, College Station, TX 77843, USA. E-mail: balsis@tamu.edu

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

(*Received 31 August 2015; accepted 9 November 2015*)

is frequently used by clinicians and researchers interested in cognitive aging, and the Clinical Dementia Rating scale (CDR-SOB; Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1992; O'Bryant et al., 2008), which is commonly used in clinical trials and practice for rating severity, including in mild and prodromal stages of disease (Aisen et al., 2011).

While each of these scales has a number of strengths and weaknesses, the fact that different measures are used across different research centers, studies, and settings creates limitations for theory and practice. For researchers, the variability in assessments used across studies limits the ability to compare results across studies or combine data sets to increase statistical power. For clinicians, the variability in assessments creates imprecision in assessing change over time or the effectiveness of various interventions.

Modern psychometric techniques can help provide statistical bridges between these various dementia staging measures. Item Response Theory (IRT), originally developed in educational assessment, is one such technique. The process underlying IRT is mathematically complex, but conceptually straightforward. The first step is to verify the assumption that all of the measures assess the same general underlying construct (see Hambleton, Swaminathan, & Rogers, 1991 for an overview of IRT methodologies). For the purposes of this article, this construct is cognitive dysfunction. The second step is to assign each participant a precise value of cognitive dysfunction based on available clinical data. In other words, because the ADAS-Cog (Rosen et al., 1984), MMSE (Folstein et al., 1975), and CDR-SOB (Hughes et al., 1982; Morris, 1992; O'Bryant et al., 2008) measures cognitive dysfunction, the observed scores on each measure can be used to estimate the underlying severity of cognitive dysfunction. This estimation is a new score known as theta. Once theta is calculated for each participant derived from multiple indicators, the raw scores on these measures can be mapped onto the ranges of theta for which they serve as indicators. Thus, similar ranges of underlying cognitive dysfunction can be "cross-walked" to different observed scores among the measures with a high degree of reliability (see Wang, Isenor, & Graybeal, 2011). Cross-walking in this regard refers to the process of equating a raw score on one measure with a raw score on a different measure. This article outlines our methodology and results and offers a usable table that shows how raw scores on these three instruments correspond in the relatively early stages of AD as defined more concretely in the participants section.

MATERIALS AND METHODS

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership. The initial goal of ADNI was to recruit 800 participants, but ADNI has been followed by two other initiatives, ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55–90, to participate in the research, consisting of older people who are cognitively healthy, people with early or late MCI, and people with early AD. Demographic information and clinical data used for this study were downloaded from the ADNI data repository (adni.loni.usc.edu) on 28 May 2014.

Participants

The analysis for this study used baseline data from 1709 participants (768 female and 941 male) enrolled across all three ADNI phases. Participants were an average of 73.70 years old (SD = 7.18) and highly educated (M = 15.93, SD = 2.85 years), and the majority identified their race as White (n = 1579, 92.4%). Other races represented include Black or African American (n = 75, 4.4%), Asian (n = 29, 1.7%), American Indian or Alaskan Native (n = 3, .2%), and Native Hawaiian or Other Pacific Islander (n = 2, .1%); 18 participants (1.1%) reported that they were more than one race, and 3 participants (.2%) did not report their race. Ten participants reported their ethnicity as Hispanic or Latino (.6%); 1643 (96.1%) indicated that they were not Hispanic or Latino, while 56 participants (3.3%) had unknown ethnicity. Baseline diagnoses represented a range of cognitive impairment: 416 (24.3%) were cognitively normal (CN), 106 (6.2%) had subjective memory complaints, 308 (18.0%) had early MCI, 561 (32.8%) had late MCI, and 318 (18.6%) had very mild AD through AD as defined just below.

According to ADNI protocol, CN participants serve as the controls and show no signs of MCI, dementia, or significant depression. CN participants have normal cognition (defined as CDR = 0, MMSE between 24 and 30, and WMS-R Logical Memory II subscale score above education-adjusted cutoffs), and neither the participants nor their study partners report memory concerns. Beginning in ADNI2, participants with significant memory concerns (SMC) were enrolled. SMC participants have normal cognition (defined the same as for CN participants), but they show significant concerns about memory (either reported by the participant, study partner, or clinician), as measured using the Cognitive Change Index, a self-report scale. MCI participants have been enrolled since ADNI1, but in ADNI-GO and ADNI-2, MCI participants, either those with early MCI or late MCI, are people who have a subjective memory concern and significant amnestic dysfunction (defined by CDR = .5 plus an abnormal score on the WMS-R Logical Memory II subscale). However, MCI subjects have sufficiently preserved functional abilities and global cognition (MMSE score between 24 and 30), such that they do not meet criteria for AD. The determination of EMCI versus LMCI is based on the severity of impairment on the WMS-R Logical Memory II subscale, a measure of delayed recall for auditory information presented in story format. The ADNI distinguishes between early MCI and late MCI based on the WMS Logical Memory score alone, and this distinction should be taken with appropriate caution as it is based on only one score.

Participants were diagnosed with AD if they met National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD (McKhann et al., 1984). At the time of the ADNI diagnosis, AD participants demonstrated deficits in their global cognition, abnormal memory functioning (based on scores on the WMS-R Logical Memory II subscale), and showed significant concerns with memory (reported by the participant, study partner, or clinician). Finally, participants were excluded if they had a history of significant neurologic disease (including multi-infarct dementia and subdural hematoma). Detailed information about the ADNI inclusion and exclusion criteria can be found online, http://www.adniinfo.org/scientists/ADNIStudyProcedures.aspx.

Measures

The ADAS-Cog (Rosen et al., 1984) is a cognitive assessment with raw scores that range from 0 to 70, where higher scores indicate greater cognitive dysfunction. The MMSE (Folstein et al., 1975) is a brief screening instrument for cognitive dysfunction with raw scores that range from 0 to 30, where lower scores indicate greater cognitive dysfunction. The CDR-SOB (Hughes et al., 1982; Morris, 1992; O'Bryant et al., 2008) is a clinician-rated staging method for cognitive dysfunction and functional ability. It requires an individual severity rating on each of six domains (or boxes): memory, orientation, judgment and problem-solving, community affairs, and activities of daily living. Under the traditional scoring method, an algorithm is used to produce global scores that indicate five stages of cognitive dysfunction: 0 = no dementia, .5 = questionable dementia, 1 = mild dementia, 2 = moderate dementia, and 3 = severe dementia (Morris, 1992). Alternatively, the Sum of Boxes scoring approach (CDR-SOB) yields scores from 0 to 18 by summing the CDR domain scores (see O'Bryant et al., 2008 for a description of how to stage using CDR-SOB). We used CDR-SOB in the current study to give a more fine grained assessment of the correspondence between the CDR measure and the ADAS-Cog and MMSE. In the present sample, participants had the following scores (MMSE: $M = 27.24$, $SD = 2.60$; ADAS-Cog: $M = 10.59$, $SD = 6.42$; CDR-SOB: $M = 1.53$, $SD = 1.66$). Ranges appear in Table 1.

Statistical analyses

IRT assumes unidimensionality of the data, so we conducted exploratory factor analyses to test the unidimensionality. Specifically, we tested a three-variable model, where each measure (i.e., the ADAS-Cog, MMSE, and CDR-SOB) served as the variables. Then, we determined whether the first factor was sufficiently large relative to the second and third factors to assume unidimensionality. We then conducted IRT analyses in Multilog (Thissen, 1991) using Samejima's graded response model (Samejima, 1969) to determine how the ADAS-Cog, MMSE, and CDR-SOB each function across the range of latent cognitive dysfunction. Although it can be considered atypical to use full measures as "items" in IRT, it is a statistically coherent practice within an IRT framework. We also used Multilog to map raw scores on each of these measures to corresponding ranges of latent AD-related cognitive dysfunction, or theta. Finally, we developed a "crosswalk" table that details how raw scores on each of these three measures correspond to one another based on their relationship to theta.

RESULTS

Findings indicated that data from the ADAS-Cog, MMSE, and CDR-SOB were adequately unidimensional for IRT analyses. Indeed, the results of an exploratory factor analysis with maximum likelihood estimation were nearly perfectly unidimensional. The first-to-second eigenvalue ratio was 2.45:.28 or 8.73:1. The third eigenvalue also was .28, and revealed no significant drop-off from the second eigenvalue. Our ratio of first-to-second eigenvalue, 8.73:1, far exceeded the 3:1 ratio standard which is commonly cited for sufficient unidimensionality (see Embretson & Reise, 2000). This is not

Table 1. Corresponding scores across three measures of cognitive dysfunction

ADAS-Cog	MMSE	CDR-SOB
0	30	0
1	30	0
2	30	0
3	30	0
4	30	0
5	30	0
6	29	0
7	29	.5
8	29	.5
9	28	1
10	28	1
11	27	1.5
12	27	1.5
13	26	2
14	26	2.5
15	26	2.5
16	25	3
17	25	3
18	24	3.5
19	24	4
20	24	4
21	23	4.5
22	23	4.5
23	23	4.5
24	22	5
25	22	5
26	22	5.5
27	22	5.5
28	21	5.5
29	21	6
30	21	6
31	21	6
32	21	6.5
33	21	6.5
34	20	6.5
35	20	6.5
36	20	6.5
37	20	6.5

Notes: ADAS-Cog, Alzheimer's Disease Assessment Scale—Cognitive Subscale. MMSE, Mini-Mental State Examination. CDR-SOB, Clinical Dementia Rating scale—Sum of the Boxes.

to say that these measures assess only one aspect of cognitive dysfunction. Rather, this result suggests that all of the cognitive processes that these measures assess covary quite strongly in this sample, and strongly enough to justify the application of unidimensional IRT procedures.

The IRT analyses revealed that the measures indicated cognitive dysfunction in the expected direction. Scores on the CDR-SOB and ADAS-Cog were higher at greater

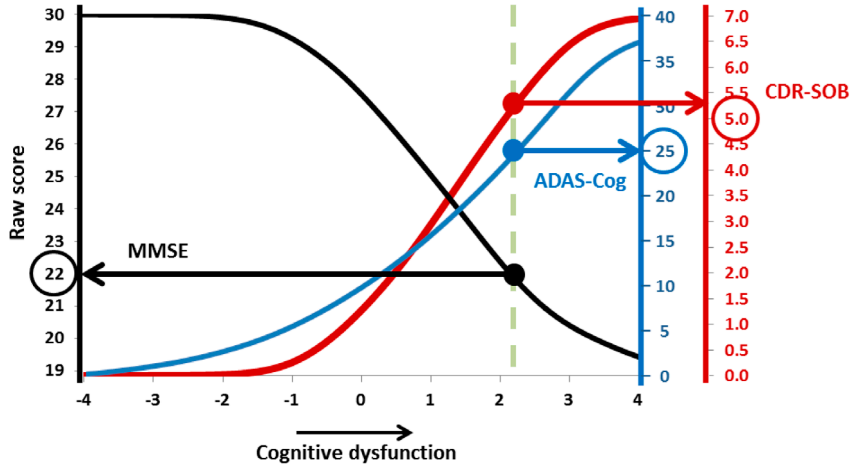

Figure 1. The correspondence among three measure of cognitive dysfunction in Alzheimer's disease.
Notes: ADAS-Cog, Alzheimer's Disease Assessment Scale—Cognitive Subscale. MMSE, Mini-Mental State Examination. CDR-SOB, Clinical Dementia scale—Sum of the Boxes.

levels of cognitive dysfunction. In contrast, scores on the MMSE were lower at greater levels of cognitive dysfunction (see Figure 1). Next, expected scores on the three measures at each level of cognitive dysfunction were compared. Results indicated multiple inflections in CDR-SOB and ADAS-Cog scores within a given MMSE score (see Table 1), suggesting that the CDR-SOB and ADAS-Cog are more precise in measuring the severity of cognitive dysfunction than the MMSE. Figure 1 illustrates how an MMSE score of 22 is equivalent to an ADAS-cog score of 25 and a CDR-SOB score of 5.0. Table 1 summarizes the correspondence among each raw score in the relatively mild range of cognitive dysfunction (ADAS-cog = 0 through 37, CDR-SOB = 0 through 6.5, MMSE = 20 through 30) across the three measures. Note that a threshold procedure ensured that the relationship among the scores was computed consistently. For example, a raw score on the ADAS-Cog of 28 would need to be fully reached before the raw score on the chart switched from 28 to 29.

DISCUSSION

This study examined the relationship among raw scores on three commonly used measures of AD severity related to cognitive dysfunction. Using IRT, we were able to determine the correspondence among scores from each of the three measures (Table 1 provides a useful catalog for raw score equivalence across the measures). This table is designed to allow clinicians and investigators to compare patients or research participants who were assessed using one measure to those who were assessed using another measure. Results are consistent with previous research in that they replicate the well-documented general correspondence among these measures. Higher scores on the ADAS-Cog and the CDR-SOB are associated with lower scores on the MMSE, and vice versa. The associations detailed in the table we present are important because they

show how individual raw scores correspond among the measures. Likewise, the table shows the meaningful model-based floor and ceiling effects of these measures across the continuum of cognitive dysfunction.

Of particular note, we replicated the structure among these three measures via a confirmatory factor analysis using MPlus software (Muthen & Muthen, 2007) in the database maintained at Baylor College of Medicine's Alzheimer's Disease and Memory Disorders Center (n = 859), (see Doody et al., 2005). Results confirmed equivalent factor loadings between the samples. Notably, the differences between the samples did not exceed the required magnitude to impose additional constraints as recommended by simulations (Cheung & Rensvold, 2002). Confirmatory fit indices between the constrained (CFI = .99) and unconstrained (CFI = 1.00) models did not differ more than .01. This simple analysis helps to further the replicability of the findings.

A clinician, for example, who is treating a client and has assessed the client using the MMSE would be able to use our table to understand the patient's score in relation to published literature using the ADAS-Cog, for example. Or, if a clinician had an MMSE score for the patient, but did not have the required collateral source to arrive at a CDR-SOB score, the clinician could use our table to estimate the client's CDR-SOB scores and to arrive at a dementia stage (see O'Bryant et al., 2008, for how to stage dementia using a CDR-SOB score). The clinical applications extend beyond these two examples, of course, into the clinical trial literature where the ADAS-Cog is ubiquitous, yet these other measures are not always used. Our findings provide a way for clinicians or researchers to directly apply clinical trial findings using the ADAS-Cog to their own patients, who might have been assessed using the MMSE or CDR-SOB scores.

It is important to note that these main findings are based on just one sample, although the structural correspondence among the measures was replicated in a second sample, and the ADNI sample is one of the best characterized and largest samples in Alzheimer's research. That said, the ADNI sample is skewed in terms of education and race, and one should exercise caution when applying this table to underrepresented minority groups. Also, participants had relatively mild cognitive dysfunction, and this study does not address scores in the range of severe cognitive dysfunction. While this is a limitation of the study, it is also a potential strength because most of the relevant comparisons among these measures will be made for individuals and studies that address mild cognitive dysfunction.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

Aisen, P. S., Gauthier, S., Ferris, S. H., Saumier, D., Haine, D., Garceau, D., … Sampalis, J. (2011). Tramiprosate in mild-to-moderate Alzheimer's disease–a randomized, double-blind, placebo-controlled, multi-centre study (the Alphase Study). *Archives of Medical Science, 7*, 102–111.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233–255.

Doody, R. S., Pavlik, V., Massman, P., Kenan, M., Yeh, S., Powell, S., … Chan, W. (2005). Changing patient characteristics and survival experience in an Alzheimer's disease center patient cohort. *Dementia and Geriatric Cognitive Disorders, 20*, 198–208.

Embretson, S., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British Journal of Psychiatry, 140*, 566–572.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*, 939–939.

Mitchell, S. L., Black, B. S., Ersek, M., Hanson, L. C., Miller, S. C., Sachs, G. A., … Morrison, R. S. (2012). Advanced dementia: State of the art and priorities for the next decade. *Annals of Internal Medicine, 156*, 45–51.

Morris, J. C. (1992). The Clinical Dementia Rating (CDR): Current vision and scoring rules. *Neurology, 43*, 2412–2414.

Muthen, L. K., & Muthen, B. O. (2007). *MPlus user's guide* (6th ed.). Los Angeles, CA: Muthen & Muthen.

O'Bryant, S. E., Waring, S. C., Cullum, C. M., Hall, J., Lacritz, L., Massman, P. J., … Doody, R. (2008). Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: A Texas Alzheimer's research consortium study. *Archives of Neurology, 65*, 1091–1095.

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *American Journal of Psychiatry, 141*, 1356–1364.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. Psychometric Monograph (No. 17). Richmond, VA: Psychometric Society.

Thissen, D. (1991). *Multilog user's guide, version 6.3*. Mooresville, IN: Scientific Software.

Wang, H., Isenor, A., & Graybeal, J. (2011). Moving between standards (crosswalking). In *The MMI guides: Navigating the world of marine metadata*. Retrieved October 8, 2015, from http://marinemetadata.org/guides/mdatastandards/crosswalks