Journal of the International Neuropsychological Society (2023), **29,** 704–714
doi:10.1017/S1355617722000662

Research Article

Cultural influence on clock drawing test: A systematic review

Giorgia Maestri[1], Alessia Nicotra[1], Simone Pomati[1], Marco Canevelli[2], Leonardo Pantoni[3] and Ilaria Cova[1]

[1]Neurology Unit, Luigi Sacco University Hospital, Milan, Italy, [2]Department of Human Neuroscience, Sapienza University, Rome, Italy and [3]Stroke and Dementia Lab, "Luigi Sacco" Department of Biomedical and Clinical Sciences, University of Milan, Milano, Italy

Abstract

Objective: Dementia among migrants is an emerging phenomenon worldwide and the development of neuropsychological tests sensitive to cultural differences is increasingly regarded as a priority. The Clock Drawing Test (CDT) is one of the most used screening tools for the detection of cognitive decline. Nevertheless, there is still a debate about its adoption as a cross-cultural assessment. **Methods:** To identify cultural variables influencing performance at CDT, we performed a systematic review of literature on three databases of all studies considering the role of at least one of the following: (1) language; (2) education; (3) literacy; (4) acculturation; and (5) ethnicity. **Results:** We extrapolated 160 analyses from 105 studies. Overall, an influence of cultural determinants on performance at CDT was found in 127 analyses (79.4%). Regarding specific cultural factors, 22 analyses investigated the effect of ethnicity on CDT scores, reporting conflicting results. Only two scoring systems turned out to be sufficiently accurate in a multicultural population. Language influenced performance in only 1 out of 8 analyses. A higher level of education positively influenced test performance in 118 out of 154 analyses (76.6%), and a better quality of education in 1 analysis out of 2. A negative influence of illiteracy on CDT performance emerged in 9 out of 10 analyses. Acculturation affected performances at CDT in 1 out of 2 studies. **Conclusions:** Based on the present findings, caution is needed when using CDT in a multicultural context, even if it requires limited linguistic competence.

Keywords: cross-cultural; neuropsychology; cognitive screening; cognitive impairment; ethnicity; executive functions

(Received 9 January 2022; final revision 3 June 2022; accepted 8 August 2022; First Published online 25 November 2022)

Introduction

International migration, with migrants defined as people who have moved across an international border, is an increasingly important phenomenon-worldwide. Due to population aging, dementia among migrants is becoming an emerging public health concern. In 2017, the number of migrants aged 65 years or older living in Europe was 6.5 million. Nearly 400.000 dementia cases were estimated in this population (Canevelli et al., 2019).

Aiming to develop diversity-sensitive models of diagnosis and care, neuropsychologists must keep in mind the impact of culture on performances in cognitive tests. Culture represents the set of learned traditions and living styles shared by the members of a society and includes the ways of thinking, feeling, and behaving (Harris, 1983). Ardila (Ardila, 2020) identified different cultural variables influencing human behavior in a neuropsychological context such as language, the quality and the degree of formal education, and the pattern of abilities and values developed as a consequence of the cultural background (such as familiarity with a one-to-one relationship, background authority of the examiner, the concept of "best performance"). These considerations led to the question of whether tests commonly used in neuropsychological assessment are free from cultural biases.

The Clock Drawing Test (CDT) is one of the most widely used cognitive tests. The subject is presented with white paper with instructions to draw a clock. In the free-drawn method, the subject is asked to draw a clock from memory. In the pre-drawn method, the request consists in drawing the numbers in the clock face and setting the hands at a fixed time. Another version requires only setting the hands at a fixed time on a pre-drawn clock complete with contour and numbers. Several different CDT scoring methods have been developed, including quantitative and qualitative systems, but no consensus exists regarding which scoring method is the most accurate (Spenciere et al., 2017).

The CDT requires the use of many mental skills: comprehension of the request of the examiner; memory to remember the instruction to set the hands at a fixed time once the clock face is complete; executive functions to coordinate the planning, organization, and simultaneous processing (including corrections and inhibition of incorrect responses such as perseveration); visual-perceptual and visual-motor abilities to internally represent the clock, to translate the mental representation into a motor program, and to monitor the output; linguistic competence for the graphomotor representation of numbers (Freedman et al., 1994). Neuroanatomical regions involved in performing the CDT include both cortical (dorsolateral prefrontal cortex, frontal, and parietal lobes) and subcortical structures (thalamus, caudate, and corpus callosum) (Eknoyan et al., 2012; Supasitthumrong et al., 2019). Due to the various cognitive functions and the underlying neuroanatomical areas, the CDT is considered a cognitive screening tool,

Corresponding author: Ilaria Cova, email: ilaria.cova@unimi.it

Cite this article: Maestri G., Nicotra A., Pomati S., Canevelli M., Pantoni L., & Cova I. (2023) Cultural influence on clock drawing test: A systematic review. *Journal of the International Neuropsychological Society*, **29:** 704–714, https://doi.org/10.1017/S1355617722000662

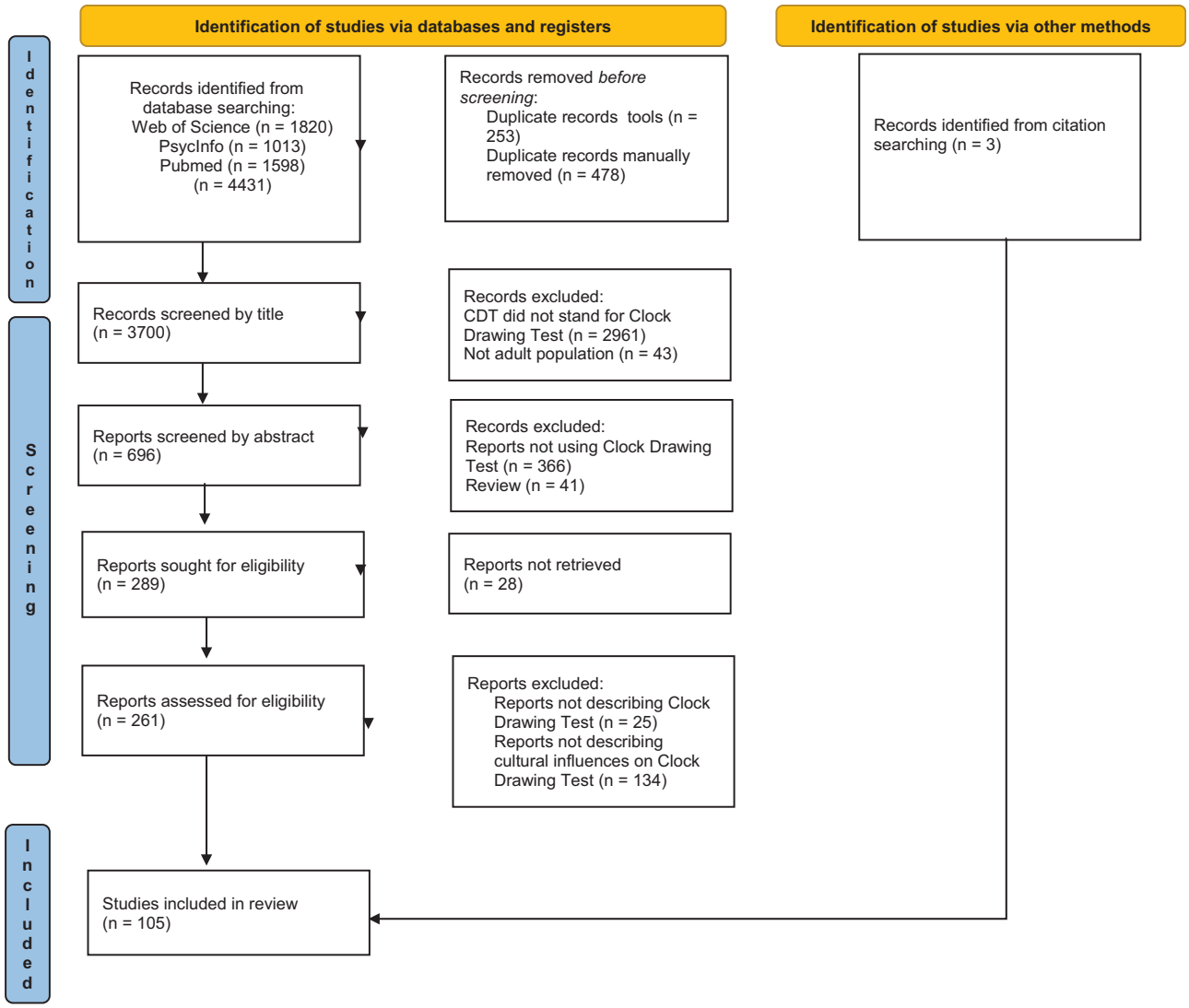

Figure 1. PRISMA 2020 flow diagram of the systematic review.

providing a measure of the overall cognitive performance of the individual (Ehreke et al., 2010; Shulman, 2000).

However, the use of the CDT as a screening test in a cross-cultural context is still debated (Franzen et al., 2020). Given the ease of administration and the limited linguistic competence required, the CDT may be deemed appropriate to support a culture fair assessment of the individual's global cognitive functioning (Parker & Philp, 2004). As proof of this, the CDT has been included as a subtest of the European Cross-Cultural Neuropsychological Test Battery (CNTB) (Nielsen et al., 2018) which demonstrated cross-cultural diagnostic properties for the evaluation of dementia in targeted minority and majority populations.

The present systematic review aimed to summarise the available evidence on the impact of the most commonly available and measurable cultural variables on CDT performance. Special attention was paid to the language used for the administration of the test, education (considering both the level and the quality of education), illiteracy (i.e., the absence of formal education or the inability to read and write), the level of acculturation (i.e., cultural modification of a group by adopting certain values and practices of a culture that is not originally their own) and to ethnicity (intended as any human grouping that shares common racial, cultural, and linguistic characteristics).

Methods

Study design

This systematic review was conducted and reported in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols 2015 (PRISMA-P 2015) (Shamseer et al., 2015) guidelines.

Search strategy

A literature search of original articles was conducted on three comprehensive medical databases (Web of Science, PsycInfo, and PubMed) from their respective dates of inception up to March 2022.

A targeted search was based on predefined search terms and used various Boolean terms to build the various algorithms. The search identified key concept combinations which can be described as follows: (*"Clock drawing test"* OR *"clock"* OR *"CDT")* AND

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Figure 2. Worldwide distribution of papers describing effects of cultural variables on clock drawing test performance. The size of the bubbles corresponds to the number of studies conducted in each country.

("cultur"* OR *"educat*"* OR "*norm**" OR *"ethnic*"* OR *"illiter*"* OR *"languag*").* These words were translated into specific search fields and syntaxes according to the different bibliographic databases (see Appendix 1 in Supplementary Material for the complete search syntax used for each electronic database).

Inclusion and exclusion criteria

The search was focused on human studies considering adults from the age of 19 years. No language restriction was used. We included only original reports which investigated the effect of at least one cultural variable on performance at CDT. Specifically, we explored the effect of: (a) language; (b) education; (c) illiteracy; (d) level of acculturation; and (e) ethnicity. Exclusion criteria were: (1) studies including minors and (2) studies that considered only the performance on the Clock Reading Test or the Copy of a Clock.

Selection of the studies

Searched results were systematically screened by three reviewers (GM, IC, and AN) for inclusion and exclusion criteria. We used a three steps screening process. Firstly, duplicates were removed both automatically and manually. Then, papers were screened for titles and abstracts. Finally, the full texts of relevant studies were searched and further assessed for eligibility criteria. In case of doubts about eligibility, the paper was reviewed by all three authors and included if two out of three were in agreement. Manual searches were extended to papers describing scoring methods of CDT cited in the selected articles to ensure that significant studies would not be missed. The study selection process is detailed in a PRISMA flow diagram (Figure 1).

Data extraction and synthesis

Data were extracted from the selected studies and reported in a dedicated database. The following information was retrieved from each study: title, authors, year of publication, geographic area of the study (defined as the one in which the patients were enrolled), patient population, sample size, scoring systems used for the CDT, and modalities of administration of CDT (free drawn, pre-drawn, or only hands setting), cultural variables potentially affecting the CDT score and type of influence of each cultural variable considered in the study (Appendix 2 in Supplementary materials).

In the papers where multiple CDT scoring systems were used, each method was considered independently as a single analysis. Therefore the total number of analyses considered is higher than the number of studies included in the review. In the case of more than one publication on the same population, the most informative paper was considered. Data were synthesized qualitatively and descriptive analyses were performed to describe the frequency of evaluation of each of the five cultural variables considered.

Results

Search results

The PRISMA flow diagram of the literature search is shown in Figure 1. Overall, 4431 papers were identified from a structured search of three databases. Seven hundred thirty-one duplicates were identified and removed. We then excluded 3411 articles by screening the titles and the abstracts. A total of 289 papers were assessed for eligibility and searched for the full-text screening, 28 of which were not available. One hundred and two papers fitted

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Table 1. Effects of cultural variables on the most frequently used clock drawing test scoring systems

Scoring system	Number of analyses	Effect of education *n/N* (%)	Effect of literacy *n/N* (%)	Effect of ethnicity (or race?) *n/N* (%)	Effect of language *n/N* (%)	Effect of at least one cultural variable *n/N* (%)
(Shulman et al., 1993)	22	18/20 (90%)	2/3 (66. 7%)	2/2 (100%)	–	21/22 (95.5%)
(Sunderland et al., 1989)	15	12/15 (80%)	–	1/2 (50%)	–	12/15 (80%)
(Royall et al., 1998)	14	12/13 (92.3%)		2/3 (66. 7%)	0/3 (0%)	14/14 (100%)
(Mendez et al., 1992)	9	6/9 (66. 7%)	–	1/2 (50%)	–	7/9 (77. 8%)
(Rouleau et al., 1992)	8	6/8 (75%)	1/1 (100%)	0/1 (0%)	–	6/8 (75%)
(Freedman et al., 1994) – Fd	8	4/7 (57.1%)	–	1/1 (100%)	–	5/8 (62.5%)
(Freedman et al., 1994) – Pd	6	3/6 (50%)	–	–	–	3/6 (50%)
(Freedman et al., 1994) – Hs	6	3/6 (50%)	–	–	–	3/6 (50%)
(Wolf-Klein et al., 1989)	6	3/6 (50%)	–	0/1 (0%)	–	3/6 (50%)
(Nasreddine et al., 2005) – MoCA	5	3/4 (75%)	2/2 (100%)	–	–	4/5 (80%)
(Manos & Wu, 1994)	4	4/4 (100%)	–	–	–	4/4 (100%)
(Morris et al., 1989) – CERAD	4	2/4 (50%)	–	1/2 (50%)	0/2 (0%)	3/4 (75%)

CERAD: consortium to establish a registry for Alzheimer's disease; Fd: free-drawn; Hs: hands setting; MoCA: Montreal cognitive assessment; *n/N* (%): number of analysis where the effects were found/Number of analysis where the effects were investigated (percentage of studies where the effects were found); Pd: pre-drawn.

Table 2. Effects of language on clock drawing test performance

Authors and year	Country	Populations	Languages	Scoring system	Effect	Type of effect	Mediators
(Alegret et al., 2012)	ESP	nHW	Spanish *versus* Catalan	(Golden, 1980) – Luria's Clock test	no		
(Borson et al., 1999)	USA	AA, nHW, HW, AS, AM	English *versus* non-English speaking*	(Morris et al., 1989) – CERAD	no		
(Borson et al., 2000)	USA	AA, nHW, HW, AS, AM	English *versus* non-English speaking*	(Morris et al., 1989) – CERAD	no		
(LaRue et al., 1999)	USA	HW, nHW	English *versus* Spanish	(Goodglass & Kaplan, 1983)	yes	English > Spanish ($p = 0.04$)	possible: education, income, rural *versus* urban origin
(Lessig et al., 2008)	USA	HW, nHW, AA, AS	English *versus* Chinese *versus* Spanish	(Lessig et al., 2008)	no		
(Menon et al., 2012)	USA	HW, nHW	English *versus* Spanish	(Royall et al., 1998) – CLOX 1	no		
(O'Bryant et al., 2018)	USA	HW	English *versus* Spanish	(Royall et al., 1998) – CLOX 1	no		
(Royall et al., 2003)	USA	HW	English *versus* Spanish	(Royall et al., 1998) – CLOX 1	no		

AA: African American; AM: American; AS: Asian; ESP: Spain; HW: Hispanic White; nHW: non-Hispanic White; USA: United States of America.
*Spanish, Korean, Chinese, and Filipino dialect.

the inclusion criteria. Three additional studies were identified from citation searches on the relevant articles. A total of 105 studies were thus included in the systematic review.

In 28 studies, more than one CDT scoring system was used; since we considered each scoring method separately, we extrapolated a total of 160 analyses from the 105 studies included. The studies that were ultimately considered in the present analysis are listed in Appendix 2 in Supplementary materials.

Relevant studies were conducted in 37 different countries worldwide, mostly in the United States of America and Brazil, followed by Italy, China, and Japan. The geographical distribution of the considered studies is shown in Figure 2.

Most studies investigated the performance of the CDT in healthy subjects. Some studies enrolled outpatients referred to memory clinics whereas others compared patients with dementia and healthy controls. The remaining papers examined subjects with other neurological disorders. The minimum sample size of the relevant articles was 40 subjects and the maximum was 1873 subjects. Only 17 studies involved less than 100 subjects, whereas 31 studies involved more than 400 subjects. In 77 out of 109 analyses (70.6%), age showed a negative correlation with performance at CDT.

Due to the considerable number of scoring systems used, we performed an additional evaluation of the most used and culturally influenced scoring systems.

Scoring systems for CDT

Three studies did not specify the adopted CDT scoring method. Twenty-eight studies used more than one scoring method. Overall, 46 different CDT scoring systems were used in the selected papers; free-drawn clocks and quantitative systems were used in most of the cases. The most used scoring system was that of Shulman et al. (Shulman et al., 1993), followed by Sunderland

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Table 3. Accuracy of studies reporting an effect of education on CDT performance

Authors and year	Country	Scoring system	Definition of low education	LO ED		HI ED	
				SE (%)	SP (%)	*SE* (%)	SP (%)
(Ainslie & Murden, 1993)	USA	(Shulman et al., 1986)	< 9	90.1	**41.8**	78.3	83.9
(Ainslie & Murden, 1993)	USA	(Shulman et al., 1993)	< 9	**61.1**	70.9	69.6	96.8
(Ainslie & Murden, 1993)	USA	(Sunderland et al., 1989)	< 9	74.1	**44.3**	78.3	83.9
(Ainslie & Murden, 1993)	USA	(Wolf-Klein et al., 1989)	< 9	**48.2**	89.9	**47.8**	100
(Aprahamian et al., 2010)	BRA	(Mendez et al., 1992)	< 5	80.5	87.5	74.4	**61.0**
(Aprahamian et al., 2010)	BRA	(Shulman et al., 1993)	< 5	81.8	92.5	72.1	74.6
(Aprahamian et al., 2010)	BRA	(Sunderland et al., 1989)	< 5	80.5	87.5	**51.2**	94.9
(Borson et al., 1999)	USA	(Morris et al., 1989)	< 9	94	85	70	100
(Kim & Chey, 2010)	KOR	(Todd et al., 1995)	< 7	**53**	88	**60**	92
(Lessig et al., 2008)	USA	(Lessig et al., 2008)	< 5	–	–	71	88
(Lolekha et al., 2021)	THA	(Lolekha et al., 2021)	< 6	89.7	74.4	80	87.9
(Lourenço et al., 2008)	BRA	(Manos & Wu, 1994)	< 5	**63**	**59**	–	–
(Lourenço et al., 2008)	BRA	(Shulman et al., 1993)	< 5	65	**57**	–	–
(Lourenço et al., 2008)	BRA	(Wolf-Klein et al., 1989)	< 5	**62**	71	–	–
(Lourenço et al., 2008)	BRA	(Sunderland et al., 1989)	< 5	**59**	**64**	–	–
(Yap et al., 2007)	SGP	(Royall et al., 1998)	< 7	79	**65**	88	68

Note. BRA: Brazil; HI HE: higher education; KOR: Korea; LO ED: lower education; *SE*: sensitivity; SGP: Singapore; SP: specificity; THA: Thailand; USA: United States of America.
In bold are sensitivity and specificity ≤ 65.

et al. (Sunderland et al., 1989) and Royall et al. colleagues (Royall et al., 1998) (Table 1). As shown in Table 2, none of the most frequently used scoring methods is free from the influence of the level of education.

The influence of culture

An influence of at least one cultural variable on performance at CDT was found in 127 of the analyses (79.4%). Considering only the 18 studies that were conducted recruiting a multi-cultural sample, all but two found the effect of at least one cultural variable.

Language of administration

The language used for the administration of the CDT was examined in 8 studies (Table 2), mostly comparing performance between English-speaking and Spanish-speaking subjects. In three studies other languages were considered: Chinese, Korean, and Filipino dialects. The language used for test administration significantly affected the test score in only one study (LaRue et al., 1999) and the authors suggested that it can be mediated by differences in educational and income level or rural/urban origin.

Education

Quality of education

The effect of the quality of education on performance at CDT was investigated in two studies (Hubbard et al., 2008; Johnson et al., 2006). In both, the Wide-Range Achievement Test-3 (WRAT-3) Reading subtest (Wilkinson, 1993) was administered along with CDT. WRAT-3 is a test of word familiarity and reading ability, considered a measure of estimated premorbid intelligence and a marker of quality of education (Manly et al., 2002). The authors found conflicting results. Hubbard et al. identified age and WRAT-3 reading scores as the only predictors of CDT scores assessed by Freund's, Mendez's, and Cahn's global score scoring methods (Cahn, 1996; Freund et al., 2005; Mendez et al., 1992). The authors also showed that including the WRAT-3 reading scores as covariates reduces the effect of education and race on performance at the CDT; therefore, they suggested that normative scores for CDT could be based on scores at the WRAT-3 instead of on subjects' education and race. On the contrary, Johnson and colleagues found a significant effect of WRAT-3 reading scores on several executive function tests, but not on CDT performance.

Level of education

The influence of the level of education was investigated in 154 analyses from 100 studies, documenting an influence on test performance in 118 analyses (76.6%), all revealing a positive correlation between educational level and CDT. Some authors found an effect of level of education on performance at the CDT only between subjects with a very poor education (differently defined between the studies) when compared with all the others (Lessig et al., 2008; Ravaglia et al., 2003; Senger et al., 2019; Shao et al., 2020; Wolf-Klein et al., 1989). On the contrary, Cooke et al. (Cooke et al., 2009) found that only completion of a tertiary educational level had a significant correlation with CDT performance. Sixteen analyses from 8 studies examined CDT suitability for low- or high-educated subjects (Table 3) finding hardly comparable results since authors arbitrarily chose different cut-offs (between 5 and 9 years) to distinguish between low and high education. A limited specificity or sensitivity of CDT in the assessment of low-educated and high-educated subjects was found in 10 and 4 analyses respectively. In addition, Cecato and colleagues (Cecato et al., 2012) investigated the ability of CDT, assessed with different scoring methods, to differentiate patients with different levels of education and scores on the Clinical Dementia Rating (CDR) Scale (Hughes et al., 1982). However, in all subjects with a high educational level (> 11 years of education), CDT scores were not able to differentiate patients with very different scores at the CDR (0 *vs.* 2). Using modified Shulman and Sunderland scoring methods all the subjects with more than 11 years of education obtained CDT scores above the cut-off, regardless of the score obtained at CDR. Only in subjects with less than 4 years of education, CDT was sufficiently accurate in identifying each level of CDR. Conversely, Scarabelot et al. (Scarabelot et al., 2019) showed that the use of CDT in subjects with less than 4 years of education could be impaired by the high rate of refusals to perform the test.

Thirty-six analyses failed to find a significant effect of education on CDT performance. However, these studies involved smaller populations, and most of them considered only specific ranges of education, primarily mid range (i.e. > 10 years) (Bruce-Keller

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Table 4. Effect of ethnicity on CDT performance

Authors and year	Country	Definition of ethnicity	Populations	Scoring system	Effect	Type of effect	Mediators
(Amini et al., 2021)	USA	Race	AA, HW, nHV	Unspecified	Yes	HW < nHW ($p < 0.01$); AA < nHW ($p < 0.001$)	
(Garrett et al., 2019)	USA	Race	AA, nHW	(Royall et al., 1998)	Yes	AA < nHW ($p < 0.001$)	possible: age, education, income, marriage
(Grober et al., 2008)	USA	Race	AA, nHW	(Freedman et al., 1994)	Yes	in patients with CDR = 0.5: AA < nHW ($p < 0.05$)	
(Hubbard et al., 2008)	USA	Race	AA, nHW	(Freund et al., 2005)	No		
(Hubbard et al., 2008)	USA	Race	AA, nHW	(Mendez et al., 1992)	No		
(Hubbard et al., 2008)	USA	Race	AA, nHW	(Cahn, 1996) – quantitative	No		
(Hubbard et al., 2008)	USA	Race	AA, nHW	(Cahn, 1996) – qualitative	Yes	AA < nHW ($p = 0.008$)	proved: quality of education measured with WRAT-3 Reading raw score
(Hubbard et al., 2008)	USA	Race	AA, nHW	(Cahn, 1996) – global	Yes	AA < nHW ($p = 0.023$)	proved: quality of education measured with WRAT-3 Reading raw score
(Hubbard et al., 2008)	USA	Race	AA, nHW	(Freedman et al., 1994) – center	No		
(LaRue et al., 1999)	USA	Race	HW, nHW	(Goodglass & Kaplan, 1983)	Yes	HW < nHW ($p < 0.05$)	possible: education, income, rural *versus* urban origin
(Lessig et al., 2008)	USA	Race	Ch, nHW, AA, nChAs, HW	(Lessig et al., 2008)	No		
(Marcopulos et al., 1997)	USA	Race	AA, nHW	(Sunderland et al., 1989)	No		
(Menon et al., 2012)	USA	Race	HW, nHW	(Royall et al., 1998)	No		proved: - In HW: age - In nHW: age, education and gender
(Nielsen et al., 2018)	DEU, BEL, DNK, SWE, NOR, GRC		WE ethnic majority, migrant minorities	(Shulman et al., 1993)	Yes	Migrant minorities < WE majorities ($p < 0.001$)	
(Schillerstrom et al., 2007)	USA	Race	AA, nHW	(Royall et al., 1998)	Yes	AA < nHW ($p < 0.001$)	proved: age, education, acculturation

AA: African American; BEL: Belgium; Ch: Chinese; DEU: Germany; DNK; Denmark; GRC: Greece; HW: Hispanic white; nChAs: non-Chinese Asian; nHW: non-Hispanic white; NOR: Norway; SWE: Sweden; USA: United States of America; WRAT-3: wide-range achievement test-3; WE: Western Europeans.

Table 5. Suitability of CDT' scoring system for a multicultural population on performance at CDT

Authors and year	Country	Populations	Scoring system	Suitability*	Sensitivity (%)	Specificity (%)	Mediators
(Borson et al., 1999)	USA	AA, nHW, HW, AS, AM	(Morris et al., 1989)	Yes	82	92	proved: education
(Storey et al., 2002)	AUS	EU, AS, SA, AF	(Morris et al., 1989)	No	90	28	
(Storey et al., 2002)	AUS	EU, AS, SA, AF	(Mendez et al., 1992)	No	98	16	
(Storey et al., 2002)	AUS	EU AS, SA, AF	(Shulman et al., 1993)	No	90	28	
(Storey et al., 2002)	AUS	EU, AS, SA, AF	(Sunderland et al., 1989)	No	86	35	proved: education
(Storey et al., 2002)	AUS	EU, AS, SA, AF	(Watson et al., 1993)	No	82	30	
(Storey et al., 2002)	AUS	EU, AS, SA, AF	(Wolf-Klein et al., 1989)	Yes	78	58	

AA: African American; AF: African; AM: American; As: Asian; AUS: Australia; EU: European; HW: Hispanic white; nHW: non-Hispanic white; SA: South American; USA: United States of America.
*When a scoring system is defined as "suitable" it means that the author of the relative paper identified it as sufficiently accurate for the investigated multicultural population.

et al., 2012; Caffarra et al., 2011; Gruber et al., 1997; Hill et al., 1995; Lowery et al., 2003; Royall et al., 1999; Yamamoto et al., 2004) or low range (i.e. < 6 years) (Alegret et al., 2012; Chan et al., 2005; Marcopulos et al., 1999; Storey et al., 2002).

Illiteracy

In 10 analyses from 8 studies, illiterate subjects were involved. Three studies defined illiterate as those subjects who never attended school or attended school for less than 1 year; in the other four studies, illiterate subjects were the ones who considered themselves unable to read and/or write, for example using the Literacy Questionnaire interview (Moon & Chey, 2004), or unable to respond to the "close your eyes" and "write a sentence" items of Mini-Mental State Examination (Folstein et al., 1975). In three studies, to be considered illiterate, subjects had to adhere to both the previous definitions. All studies but one showed a negative influence of illiteracy on CDT performance, however, the latter (Cassimiro et al., 2016) involved only subjects with less than 4 years of education and found better performance in subjects with 3–4 years of education compared to subjects with less than 3 years.

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Acculturation

Only two studies investigated the effect of acculturation on performance at the CDT (Nielsen & Jørgensen, 2013; Royall et al., 2003). Authors defined acculturation as a multidimensional process whereby members of one cultural group adopt the attitudes, values, and behaviors of another (Gordon, 1964). Acculturation was assessed with the Hazuda scale (Hazuda et al., 1988), investigating both English proficiency and the pattern of English *versus* Spanish usage, or the Turkish adaptation of the Short Acculturation Scale for Hispanics (SASH) (Marin et al., 1987; Nielsen et al., 2012). Authors obtained conflicting results: Royall et al. (Royall et al., 2003) identified a significant but small effect of acculturation on performance at CDT (all the sociodemographic variables combined explained the 8% of the CDT variance; $p < 0.001$). On the contrary, Nielsen and Jørgensen did not find a significant correlation between the CDT performance of Turkish migrants with years of residence in Denmark and SASH score in both literate and illiterate subjects.

Ethnicity

Ten studies investigated the effect of ethnicity on CDT scores, and most of them were conducted in the USA. In all but two studies, authors operationalized ethnicity as different races, comparing Caucasians with other races. Better performance of Caucasians was found in half of the cases (Table 4). Nielsen et al. considered ethnicity as the country of origin and found that migrant minorities (Polish, Yugoslavian, Turkish, and Moroccan) display lower scores than Western European majorities (Belgian, Danish, German, Greek, Norwegian, and Swedish) (Nielsen et al., 2018). Some possible factors underlying the association between ethnicity and CDT score were identified in different papers, mainly age, level and quality of education, and degree of acculturation, nevertheless ethnicity maintained an influence on test scores even when controlling for these variables.

Two studies examined the accuracy of different scoring methods in detecting dementia in a multicultural population (Borson et al., 1999; Storey et al., 2002). Only two scoring systems turned out to be sufficiently accurate in the target population, even if with conflicting results (Table 5).

Discussion

The present study represents the first attempt to systematically present and discuss the available evidence on the influence of culture on the performance at the CDT. An influence of the considered cultural variables was found in most studies, in particular in three-quarters of the studies regarding the level of education (and almost all those regarding literacy) and in half of the studies regarding ethnicity and acculturation. Conversely, the language of administration of CDT seemed to have a negligible effect.

Most of the studies included in this review have been conducted in America and Europe, few studies have been conducted in Asia (mainly China and Japan) and Africa. Few studies recruited a multicultural sample. Sample sizes were very heterogeneous and several systems have been used to score CDT, thus limiting the validity of comparisons. None of the most used scoring systems showed to be free from the influence of cultural variables.

Only a few studies investigated the influence of language in which CDT is administered, most of which found no significant effect. It is not surprising since the CDT requires limited linguistic competence.

When investigating the effect of education on CDT performance, it is necessary to consider the quality of education as a possible confounding variable. Attending the same number of years of school does not mean having the same education in qualitative terms. The Reading Recognition subtest from the WRAT–3 (Wilkinson, 1993) can be used as a measure of reading ability and quality of education (Manly et al., 2002). However, the relationship between quality of education and scores at CDT is still poorly investigated, and the results are mixed. Many studies focused on the effect of educational level on CDT performance, identifying a positive correlation: as the level of education increases, performance on the test improves. We hypothesize, following Ardila (Ardila, 2020), that two factors could mediate the above relationship. First, the concept of familiarity: the subjects with a higher educational level may be more familiar not only with the material administered but also with the drawing activity and paper and pencil assignments; they may also be more accustomed to assessment contexts, including the one-to-one relationship, the background authority of the examiner and the concept of best performance. Secondly, the relationship between the level of education and CDT performance might be mediated by cognitive reserve, defined as the "adaptability (i.e., efficiency, capacity, flexibility) of cognitive processes that help to explain differential susceptibility of cognitive abilities or day-to-day function to brain aging, pathology, or insult" (Stern et al., 2020). The authors suggest that differences in cognitive results are determined by processes influenced by not only innate differences but also lifetime exposure, including education, occupation, and social engagement. Different studies showed that higher levels of education are associated with a lower risk of dementia (Evans et al., 1997; Karp, 2004; Stern et al., 1994). Therefore, it is possible that a higher level of education, contributing to increasing the level of cognitive reserve, induces an improvement in CDT performance. Results are mixed when investigating the effect of education on the test in specific education cohorts. Some studies have shown low accuracy of CDT in subjects with low levels of education, while others have identified a low specificity or sensitivity of the test in subjects with high levels of education. CDT may not be suitable for detecting cognitive impairment in low educated and illiterate subjects because they may be excessively disadvantaged by the limited training of the skills useful to perform the test; in addition, they may suffer from unfamiliarity with the task and the assessment setting. On the contrary, in highly educated subjects the greater cognitive reserve and the increased familiarity with the task could make the test too easy, leading to overestimating their cognitive abilities. The high heterogeneity of enrolled cohorts and different cut-offs used to identify different levels of education contributed to the difficulty in interpreting these results. A small percentage of studies identified no correlation between educational level and performance on the CDT, but most of these studies recruited smaller cohorts.

All the analyses except one found a significant effect of illiteracy on CDT performance, both when illiteracy is defined as the absence of formal education and when defined as the inability to read and write. The authors explained the effect of illiteracy as a consequence of poor development of constructional skills and planning, organization, simultaneous processing, and self-monitoring, all directly or indirectly trained in school (Kim & Chey, 2010; Mokri et al., 2012; Nielsen & Jørgensen, 2013). Kim et al. showed that illiterate older people made errors similar to those of the Alzheimer's dementia patients, specifically conceptual errors. It is noteworthy that no one of the studies investigating the

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

effect of illiteracy on the performance at the CDT was conducted in the USA (where most of the studies included in the present review were conducted), suggesting the possibility of an underestimation of this issue.

The influence of acculturation on performance at the CDT has been investigated in few studies, identifying at most a small effect. One possible explanation can be found in the characteristics of questionnaires designed to assess the level of acculturation. Both of them investigated acculturation levels comparing in different contexts the frequency of speaking in the native language *versus* in the language spoken in the host nation. Given the absence of the effect of language on test performance, it is not surprising that acculturation as measured in this way also shows no significant effect. However, acculturation is a multidimensional process, before ruling out an influence on the test further studies would be desirable that would investigate the construct from a different perspective.

In the present review, we decided to include all papers investigating the effect of "ethnicity" intended as a broad category. This would have allowed us to consider all the studies that subdivided subjects based on any ethnic characteristic such as race, culture of reference, country of origin, and language spoken; nonetheless, most studies defined ethnicity as race. Since most of these studies were conducted in the USA, the reason for this categorization is related to the prevalence of a long-standing migration in the USA. Given that many migrants are second or third-generation migrants, it should be difficult to categorize them based on their country of origin. We found that the majority of the scoring systems developed to evaluate CDT performance seem to be inaccurate in detecting dementia in a multicultural population. Moreover, half of the studies found a better performance of the Caucasian population when compared with other races. These results can be explained partly by the mediating effect of the quality of education. Manly et al. (Manly et al., 2002) suggest that in the USA there is a great deal of discordance in the quality of education between Caucasian and African American subjects, and Avila et al. (Avila et al., 2021), comparing the contribution of the level of education to cognitive reserve in Whites, Blacks, and Hispanics, found that educational attainment does not contribute to cognitive reserve similarly across different racial groups. Through the effect on cognitive reserve, the differences in the quality of education could explain the residual effect of ethnicity on performance at the CDT even when controlling for the level of education. In addition, it is well known that most of the neuropsychological tools are designed in an occidental context explicitly for a WEIRD Population (Henrich et al., 2010) (that is Western, Educated, Industrialized, Rich, and Democratic), so it should not be surprising that patients belonging to the same cultural group of the test developer usually obtain better results.

This study has some limitations. Firstly, we did not use specific scales to assess the quality of the included studies. Secondly, we included all the studies in which the effect of cultural variables on CDT performance was considered, regardless of their sample size and their sensitivity and specificity. Thirdly, the heterogeneity of definitions used (such as the definition of "low" and "high" education, ethnicity, etc.) and the several different CDT scoring methods used across studies do not allow us to perform a meta-analysis. Finally, we were not able to find 21 papers. However, we would highlight the strengths of our study too. We have considered several cultural variables which can affect CDT performance and we have taken into account also different scoring systems. Moreover, we were able to include all the studies with no limitation of languages of publication.

In future studies, we suggest better investigating the role of quality of education and the level of acculturation on CDT performance, especially as mediating factors of ethnicity and level of education. Also, the suitability of the test for illiterate or low- high-education level subjects should be better studied to avoid an overestimation or an underestimation of cognitive impairment in these populations.

Conclusion

Based on these findings, CDT does not seem to provide a culturally unbiased assessment of global cognition. These results suggest caution when using neuropsychological tests in a multicultural context, even when limited linguistic competence is required.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355617722000662

Acknowledgements. Alessia Nicotra, Giorgia Maestri, and Marco Canevelli are supported by a research grant from the Italian Ministry of Health for the project "Dementia in immigrants and ethnic minorities living in Italy: clinical-epidemiological aspects and public health perspectives" (ImmiDem) (GR-2016-02364975). Author contributions Conceived and designed the study: SP, GM, AN. Collected the data: IC, GM, AN. Contributed data or analysis tools: GM, AN, SP, IC Performed the analysis: GM, AN, SP, IC. Writing—original draft: IC, GM, AN, MC. Writing—review and editing: all authors discussed the results and contributed to the final manuscript. Conflict of interest The authors declare no competing interests. Funding The authors received no specifc funding for this work.

References

Ainslie, N. K., & Murden, R. A. (1993). Effect of education on the clock-drawing dementia screen in non-demented elderly persons. *Journal of the American Geriatrics Society*, *41*, 249–252. https://doi.org/10.1111/j.1532-5415.1993.tb06701.x

Alegret, M., Espinosa, A., Vinyes-Junqué, G., Valero, S., Hernández, I., Tárraga, L., Becker, J. T., & Boada, M. (2012). Normative data of a brief neuropsychological battery for Spanish individuals older than 49. *Journal of Clinical and Experimental Neuropsychology*, *34*, 209–219. https://doi.org/10.1080/13803395.2011.630652

Amini, R., Sahli, M., & Ganai, S. (2021). Cigarette smoking and cognitive function among older adults living in the community. *Aging, Neuropsychology, and Cognition*, *28*, 616–631. https://doi.org/10.1080/13825585.2020.1806199

Aprahamian, I., Martinelli, J. E., Neri, A. L., & Yassuda, M. S. (2010). The accuracy of the clock drawing test compared to that of standard screening tests for Alzheimer's disease: Results from a study of Brazilian elderly with heterogeneous educational backgrounds. *International Psychogeriatrics*, *22*, 64–71. https://doi.org/10.1017/S1041610209991141

Ardila, A. (2020). Cross-cultural neuropsychology: History and prospects. *RUDN Journal of Psychology and Pedagogics*, *17*, 64–78. https://doi.org/10.22363/2313-1683-2020-17-1-64-78

Avila, J. F., Rentería, M. A., Jones, R. N., Vonk, J. M. J., Turney, I., Sol, K., Seblova, D., Arias, F., Hill-Jarrett, T., Levy, S., Meyer, O., Racine, A. M., Tom, S. E., Melrose, R. J., Deters, K., Medina, L. D., Carriòn, C. D., Diaz-Santos, M., Byrd, D. R., Chesebro, A., Colon, J., Igwe, K. C., Maas, B., Brickman, A. M., Schupf, N., Mayeux, R., & Manly, J. J. (2021). Education differentially contributes to cognitive reserve across racial/ethnic groups. *Alzheimer's & Dementia*, *17*, 70–80. https://doi.org/10.1002/alz.12176

Borson, S., Brush, M., Gil, E., Scanlan, J., Vitaliano, P., Chen, J., Cashman, J., Sta Maria, M. M., Barnhart, R., & Roques, J. (1999). The clock drawing test: Utility for dementia detection in multiethnic elders. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *54*, M534–M540. https://doi.org/10.1093/gerona/54.11.M534

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The Mini-Cog: A cognitive ? Vital signs? Measure for dementia screening in multilingual elderly. *International Journal of Geriatric Psychiatry*, *15*, 1021–1027.

https://doi.org/10.1002/1099–1166(200011)15:11<1021:AID-GPS234>3.0.CO;2–

Bruce-Keller, A. J., Brouillette, R. M., Tudor-Locke, C., Foil, H. C., Gahan, W. P., Nye, D. M., Guillory, L., & Keller, J. N. (2012). Relationship between cognitive domains, physical performance, and gait in elderly and demented subjects. *Journal of Alzheimer's Disease*, *30*, 899–908. https://doi.org/10.3233/JAD-2012-120025

Caffarra, P., Gardini, S., Zonato, F., Concari, L., Dieci, F., Copelli, S., Freedman, M., Stracciari, A., & Venneri, A. (2011). Italian norms for the Freedman version of the clock drawing test. *Journal of Clinical and Experimental Neuropsychology*, *33*, 982–988. https://doi.org/10.1080/13803395.2011.589373

Cahn, D. (1996). Screening for dementia of the Alzheimer type in the community: The utility of the clock drawing test. *Archives of Clinical Neuropsychology*, *11*, 529–539. https://doi.org/10.1016/0887-6177(95)00041-0

Canevelli, M., Lacorte, E., Cova, I., Zaccaria, V., Valletta, M., Raganato, R., Bruno, G., Bargagli, A. M., Pomati, S., Pantoni, L., & Vanacore, N. (2019). Estimating dementia cases amongst migrants living in Europe. *European Journal of Neurology*, *26*, 1191–1199. https://doi.org/10.1111/ene.13964

Cassimiro, L., Fuentes, D., Nitrini, R., & Yassuda, M. S. (2016). Decision-making in cognitively unimpaired illiterate and low-educated older Women: Results on the Iowa gambling task. *Archives of Clinical Neuropsychology*, *32*, acw080v1. https://doi.org/10.1093/arclin/acw080

Cecato, J. F., Fiorese, B., Montiel, J. M., Bartholomeu, D., & Martinelli, J. E. (2012). Clock drawing test in elderly individuals with different education levels: Correlation with clinical dementia rating. *American Journal of Alzheimer's Disease & Other Dementias*, *27*, 620–624. https://doi.org/10.1177/1533317512463954

Chan, C. C., Yung, C. Y., & Pan, P. C. (2005). Screening of dementia in Chinese elderly adults by the clock drawing test and the time and change test. *Hong Kong Medical Journal = Xianggang Yi Xue Za Zhi*, *11*, 13–19.

Cooke, D. M., Gustafsson, L., & Tardiani, D. L. (2009). Clock drawing from the occupational therapy adult perceptual screening test: Its correlation with demographic and clinical factors in the stroke population: Clock drawing in the stroke population. *Australian Occupational Therapy Journal*, *57*, 183–189. https://doi.org/10.1111/j.1440-1630.2009.00795.x

Ehreke, L., Luppa, M., König, H.-H., & Riedel-Heller, S. G. (2010). Is the clock drawing test a screening tool for the diagnosis of mild cognitive impairment? A systematic review. *International Psychogeriatrics*, *22*, 56–63. https://doi.org/10.1017/S1041610209990676

Eknoyan, D., Hurley, R. A., & Taber, K. H. (2012). The clock drawing task: Common errors and functional neuroanatomy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *24*, 260–265. https://doi.org/10.1176/appi.neuropsych.12070180

Evans, D. A., Hebert, L. E., Beckett, L. A., Scherr, P. A., Albert, M. S., Chown, M. J., Pilgrim, D. M., & Taylor, J. O. (1997). Education and other measures of socioeconomic status and risk of incident Alzheimer disease in a defined population of older persons. *Archives of Neurology*, *54*, 1399–1405. https://doi.org/10.1001/archneur.1997.00550230066019

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". *Journal of Psychiatric Research*, *12*, 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Franzen, S., van den Berg, E., Goudsmit, M., Jurgens, C. K., van de Wiel, L., Kalkisim, Y., Uysal-Bozkir, O., Ayhan, Y., Nielsen, T. R., & Papma, J. M. (2020). A systematic review of neuropsychological tests for the assessment of dementia in non-western, low-educated or illiterate populations. *Journal of the International Neuropsychological Society*, *26*, 331–351. https://doi.org/10.1017/S1355617719000894

Freedman, M., Leach, L., Kaplan, E., Winocur, G., Shulman, K., & Delis, D. C. (1994). *Clock drawing: a neuropsychological analysis*. Oxford University Press.

Freund, B., Gravenstein, S., Ferris, R., Burke, B. L., & Shaheen, E. (2005). Drawing clocks and driving cars: Use of brief tests of cognition to screen driving competency in older adults. *Journal of General Internal Medicine*, *20*, 240–244. https://doi.org/10.1111/j.1525-1497.2005.40069.x

Garrett, S. L., Kennedy, R. E., Sawyer, P., Williams, C. P., Brown, C. J., & Allman, R. M. (2019). Association between executive dysfunction and instrumental activities of daily living: Racial and ethnic differences among community-dwelling older adults in the southeastern US. *Journal of the National Medical Association*, *111*, 320–327. https://doi.org/10.1016/j.jnma.2018.11.002

Golden, C. J. (1980). In reply to Adams "In search of Luria's battery: A false start". *Journal of Consulting and Clinical Psychology*, *48*, 517–521.

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Lea & Febiger.

Gordon, M. (1964). *Assimilation in American life*. Oxford University Press.

Grober, E., Hall, C., Mcginn, M., Nicholls, T., Stanford, S., Ehrlich, A., Jacobs, L. G., Kennedy, G., Sanders, A., & Lipton, R. B. (2008). Neuropsychological strategies for detecting early dementia. *Journal of the International Neuropsychological Society*, *14*, 130–142. https://doi.org/10.1017/S1355617708080156

Gruber, N. P., Varner, R. V., Chen, Y.-W., & Lesser, J. M. (1997). A comparison of the clock drawing test and the Pfeiffer short portable mental status questionnaire in a geropsychiatry clinic. *International Journal of Geriatric Psychiatry*, *12*, 526–532. –https://doi.org/10.1002/(SICI)1099–1166(199705)12:5<526:AID-GPS535>3.0.CO;2-V

Harris, M. (1983). *Culture, people, nature: an introduction to general anthropology* (3rd ed.). Harper & Row.

Hazuda, H. P., Stern, M. P., & Haffner, S. M. (1988). Acculturation and assimilation among Mexican Americans: Scales and population-based data. *Social Science Quarterly*, *69*, 687–706.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–83. https://doi.org/10.1017/S0140525X0999152X

Hill, R. D., Bäckman, L., Wahlin, Å., & Winblad, B. (1995). Visuospatial performance in very old demented persons: An individual difference analysis. *Dementia and Geriatric Cognitive Disorders*, *6*, 49–54. https://doi.org/10.1159/000106921

Hubbard, E., Santini, V., Blankevoort, C., Volkers, K., Barrup, M., Byerly, L., Chaisson, C., Jefferson, A. L., Kaplan, E., Green, R. C., & Green, R. (2008). Clock drawing performance in cognitively normal elderly. *Archives of Clinical Neuropsychology*, *23*, 295–327. https://doi.org/10.1016/j.acn.2007.12.003

Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, *140*, 566–572. https://doi.org/10.1192/bjp.140.6.566

Johnson, A. S., Flicker, L. J., & Lichtenberg, P. A. (2006). Reading ability mediates the relationship between education and executive function tasks. *Journal of the International Neuropsychological Society*, *12*, 64–71. https://doi.org/10.1017/S1355617706060073

Karp, A. (2004). Relation of education and occupation-based socioeconomic status to incident Alzheimer's disease. *American Journal of Epidemiology*, *159*, 175–183. https://doi.org/10.1093/aje/kwh018

Kim, H., & Chey, J. (2010). Effects of education, literacy, and dementia on the clock drawing test performance. *Journal of the International Neuropsychological Society*, *16*, 1138–1146. https://doi.org/10.1017/S1355617710000731

LaRue, A., Romero, L. J., Ortiz, I. E., Chi Lang, H., & Lindeman, R. D. (1999). Neuropsychological performance of Hispanic and Non-Hispanic older adults: An epidemiologic survey. *The Clinical Neuropsychologist*, *13*, 474–486. https://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT474

Lessig, M. C., Scanlan, J. M., Nazemi, H., & Borson, S. (2008). Time that tells: Critical clock-drawing errors for dementia screening. *International Psychogeriatrics*, *20*, 459–470. https://doi.org/10.1017/S1041610207006035

Lolekha, P., Tangkanakul, C., Saengchatri, T., & Kulkeartprasert, P. (2021). The six-item clock-drawing scoring system: A rapid screening for cognitive impairment in Parkinson's disease. *Psychogeriatrics*, *21*, 24–31. https://doi.org/10.1111/psyg.12605

Lourenço, R. A., Ribeiro-Filho, S. T., Moreira, I. D. F. H., Paradela, E. M. P., & Miranda, A. S. D. (2008). The clock drawing test: Performance among elderly with low educational level. *Revista Brasileira de Psiquiatria*, *30*, 309–315. https://doi.org/10.1590/S1516-44462008000400002

Lowery, N., Giovanni, L., Mozley, L. H., Arnold, S. E., Bilker, W. B., Gur, R. E., & Moberg, P. J. (2003). Relationship between clock-drawing and neuropsychological and functional status in elderly institutionalized patients with

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Schizophrenia. *American Journal of Geriatric Psychiatry*, *11*, 621–628. https://doi.org/10.1097/00019442-200311000-00006

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, *8*, 341–348. https://doi.org/10.1017/S1355617702813157

Manos, P. J., & Wu, R. (1994). The ten point clock test: A quick screen and grading method for cognitive impairment in medical and surgical patients. *The International Journal of Psychiatry in Medicine*, *24*, 229–244. https://doi.org/10.2190/5A0F-936P-VG8N-0F5R

Marcopulos, B. A., Gripshover, D. L., Broshek, D. K., McLain, C. A., & McLain, R. H. R. (1999). Neuropsychological assessment of psychogeriatric patients with limited education. *The Clinical Neuropsychologist*, *13*, 147–156. https://doi.org/10.1076/clin.13.2.147.1968

Marcopulos, B. A., McLain, C. A., & Giuliano, A. J. (1997). Cognitive impairment or inadequate norms? A study of healthy, rural, older adults with limited education. *The Clinical Neuropsychologist*, *11*, 111–131. https://doi.org/10.1080/13854049708407040

Marin, G., Sabogal, F., Marin, B. V., Otero-Sabogal, R., & Perez-Stable, E. J. (1987). Development of a short acculturation scale for Hispanics. *Hispanic Journal of Behavioral Sciences*, *9*, 183–205. https://doi.org/10.1177/07399863870092005

Mendez, M. F., Ala, T., & Underwood, K. L. (1992). Development of scoring criteria for the clock drawing task in Alzheimer's disease. *Journal of the American Geriatrics Society*, *40*, 1095–1099. https://doi.org/10.1111/j.1532-5415.1992.tb01796.x

Menon, C., Hall, J., Hobson, V., Johnson, L., & O'Bryant, S. E. (2012). Normative performance on the executive clock drawing task in a multi-ethnic bilingual cohort: A project FRONTIER study: An executive clock drawing task: Normative performance. *International Journal of Geriatric Psychiatry*, *27*, 959–966. https://doi.org/10.1002/gps.2810

Mokri, H., Ávila-Funes, J. A., Le Goff, M., Ruiz-Arregui, L., Gutierrez Robledo, L. M., & Amieva, H. (2012). Self-reported reading and writing skills in elderly who never attended school influence cognitive performances: Results from the coyoacan cohort study. *The Journal of Nutrition, Health & Aging*, *16*, 621–624. https://doi.org/10.1007/s12603-012-0070-8

Moon, H. S., & Chey, J. (2004). Literacy and neuropsychological functions in the older Korean adults. *Journal of Korean Geriatric Psychiatry*, *8*, 113–120.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellitis, E. D., & Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*, 1159–1165. https://doi.org/10.1212/wnl.39.9.1159

Nasreddine, Z. S., Phillips, N. A., BÃcdirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment: MOCA: A brief screening tool for MCI. *Journal of the American Geriatrics Society*, *53*, 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nielsen, T. R., & Jørgensen, K. (2013). Visuoconstructional abilities in cognitively healthy illiterate Turkish immigrants: A quantitative and qualitative investigation. *The Clinical Neuropsychologist*, *27*, 681–692. https://doi.org/10.1080/13854046.2013.767379

Nielsen, T. R., Segers, K., Vanderaspoilden, V., Bekkhus-Wetterberg, P., Minthon, L., Pissiota, A., Bjørkløof, G. H., Beinhoff, U., Tsolaki, M., Gkioka, M., & Waldemar, G. (2018). Performance of middle-aged and elderly European minority and majority populations on a cross-cultural neuropsychological test battery (CNTB). *The Clinical Neuropsychologist*, *32*, 1411–1430. https://doi.org/10.1080/13854046.2018.1430256

Nielsen, T. R., Vogel, A., Gade, A., & Waldemar, G. (2012). Cognitive testing in non-demented Turkish immigrants – comparison of the RUDAS and the MMSE: Cognitive testing in Turkish immigrants. *Scandinavian Journal of Psychology*, *53*, 455–460. https://doi.org/10.1111/sjop.12018

O'Bryant, S. E., Edwards, M., Johnson, L., Hall, J., Gamboa, A., & O'jile, J. (2018). Texas Mexican American adult normative studies: Normative data for commonly used clinical neuropsychological measures for English- and Spanish-speakers. *Developmental Neuropsychology*, *43*, 1–26. https://doi.org/10.1080/87565641.2017.1401628

Parker, C., & Philp, I. (2004). Screening for cognitive impairment among older people in black and minority ethnic groups. *Age and Ageing*, *33*, 447–452. https://doi.org/10.1093/ageing/afh135

Ravaglia, G., Forti, P., Maioli, F., Arnone, G., Pantieri, G., Cocci, C., Nativio, V., Muscari, A., Pedone, V., & Mariani, E. (2003). The clock-drawing test in elderly Italian community dwellers: Associations with sociodemographic status and risk factors for vascular cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, *16*, 287–295. https://doi.org/10.1159/000072815

Rouleau, I., Salmon, D. P., Butters, N., Kennedy, C., & McGuire, K. (1992). Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. *Brain and Cognition*, *18*, 70–87. https://doi.org/10.1016/0278-2626(92)90112-Y

Royall, D. R, Cordes, J. A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology, Neurosurgery & Psychiatry*, *64*, 588–594. https://doi.org/10.1136/jnnp.64.5.588

Royall, D. R., Espino, D. V., Polk, M. J., Verdeja, R., Vale, S., Gonzales, H., Palmer, R. R., & Markides, K. P. (2003). Validation of a Spanish translation of the CLOX for use in Hispanic samples: The Hispanic EPESE study. *International Journal of Geriatric Psychiatry*, *18*, 135–141. https://doi.org/10.1002/gps.804

Royall, D. R., Mulroy, A. R., Chiodo, L. K., & Polk, M. J. (1999). Clock drawing is sensitive to executive control: A comparison of six methods. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *54B*, P328–P333. https://doi.org/10.1093/geronb/54B.5.P328

Scarabelot, L. F., Monteiro, M. D. M., Rubert, M. C. D. S., & Zetola, V. D. H. F. (2019). Is the mini-mental state examination the best cognitive screening test for less educated people? *Arquivos de Neuro-Psiquiatria*, *77*, 330–334. https://doi.org/10.1590/0004-282x20190043

Schillerstrom, J. E., Sawyer Baker, P., Allman, R. M., Rungruang, B., Zamrini, E., & Royall, D. R. (2007). Clock drawing phenotypes in community-dwelling African Americans and Caucasians: Results from the University of Alabama at Birmingham study of aging. *Neuroepidemiology*, *28*, 175–180. https://doi.org/10.1159/000104095

Senger, J., Bruscato, N. M., Werle, B., Moriguchi, E. H., & Pattussi, M. P. (2019). Nutritional status and cognitive impairment among the very old in a community sample from Southern Brazil. *The Journal of Nutrition, Health & Aging*, *23*, 923–929. https://doi.org/10.1007/s12603-019-1230-x

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & the PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*, g7647. https://doi.org/10.1136/bmj.g7647

Shao, K., Dong, F., Guo, S., Wang, W., Zhao, Z., Yang, Y., Wang, P., & Wang, J. (2020). Clock-drawing test: Normative data of three quantitative scoring methods for Chinese-speaking adults in Shijiazhuang City and clinical utility in patients with acute ischemic stroke. *Brain and Behavior*, *10*, e01806. https://doi.org/10.1002/brb3.1806

Shulman, K. I. (2000). Clock-drawing: Is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*, *15*, 548–561. –https://doi.org/10.1002/1099–1166(200006)15:6<548:AID-GPS242>3.0.CO;2-U

Shulman, K. I., Pushkar Gold, D., Cohen, C. A., & Zucchero, C. A. (1993). Clock-drawing and dementia in the community: A longitudinal study. *International Journal of Geriatric Psychiatry*, *8*, 487–496. https://doi.org/10.1002/gps.930080606

Shulman, K. I., Shedletsky, R., & Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, *1*, 135–140. https://doi.org/10.1002/gps.930010209

Spenciere, B., Alves, H., & Charchat-Fichman, H. (2017). Scoring systems for the clock drawing test: A historical review. *Dementia & Neuropsychologia*, *11*, 6–14. https://doi.org/10.1590/1980-57642016dn11-010003

Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Venezuela, M.,

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press

Vemuri, P., Vuoksimaa, E., & the Reserve, Resilience and Protective Factors PIA Empirical Definitions and Conceptual Frameworks Workgroup. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*, *16*, 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219

Stern, Y., Gurland, B., Tatemichi, T. K., Tang, M. X., Wilder, D., & Mayeux, R. (1994). Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA*, *271*, 1004–1010.

Storey, J. E., Rowland, J. T. J., Basic, D., & Conforti, D. A. (2002). Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients. *International Psychogeriatrics*, *14*, 259–271. https://doi.org/10.1017/S1041610202008463

Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., & Grafman, J. H. (1989). Clock drawing in Alzheimer's disease: A novel measure of dementia severity. *Journal of the American Geriatrics Society*, *37*, 725–729. https://doi.org/10.1111/j.1532-5415.1989.tb02233.x

Supasitthumrong, T., Herrmann, N., Tunvirachaisakul, C., & Shulman, K. (2019). Clock drawing and neuroanatomical correlates: A systematic review. *International Journal of Geriatric Psychiatry*, *34*, 223–232. https://doi.org/10.1002/gps.5013

Todd, M. E., Dammers, P. M., Adams, S. G., Todd, H. M., & Morrison, M. (1995). An examination of a proposed scoring procedure for the clock drawing test: Reliability and predictive validity of the clock scoring system (CSS). *American Journal of Alzheimer's Disease*, *10*, 22–26. https://doi.org/10.1177/153331759501000406

Watson, Y. I., Arfken, C. L., & Birge, S. J. (1993). Clock completion: An objective screening test for dementia. *Journal of the American Geriatrics Society*, *41*, 1235–1240. https://doi.org/10.1111/j.1532-5415.1993.tb07308.x

Wilkinson, G. S. (1993). *WRAT-3 administration manual*. Wide Range.

Wolf-Klein, G. P., Silverstone, F. A., Levy, A. P., Brod, M. S., & Breuer, J. (1989). Screening for Alzheimer's disease by clock drawing. *Journal of the American Geriatrics Society*, *37*, 730–734. https://doi.org/10.1111/j.1532-5415.1989.tb02234.x

Yamamoto, S., Mogi, N., Umegaki, H., Suzuki, Y., Ando, F., Shimokata, H., & Iguchi, A. (2004). The clock drawing test as a valid screening method for mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, *18*, 172–179. https://doi.org/10.1159/000079198

Yap, P. L.-K., Ng, T.-P., Niti, M., Yeo, D., & Henderson, L. (2007). Diagnostic performance of clock drawing test by CLOX in an Asian Chinese population. *Dementia and Geriatric Cognitive Disorders*, *24*, 193–200. https://doi.org/10.1159/000107080

https://doi.org/10.1017/S1355617722000662 Published online by Cambridge University Press