Journal of Affective Disorders 227 (2018) 471–476

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Does the 15-item Geriatric Depression Scale function differently in old people with different levels of cognitive functioning?

Francesca Chiesi[a,*], Caterina Primi[a], Martina Pigliautile[b], Marta Baroni[b], Sara Ercolani[b], Lucia Paolacci[b], Virginia Boccardi[b], Patrizia Mecocci[b]

[a] Department of Neuroscience, Psychology, Drug, and Child's Health (NEUROFARBA), Section of Psychology, University of Florence, Italy
[b] Institute of Gerontology and Geriatrics, Department of Medicine, University of Perugia, Italy

ABSTRACT

Background: The 15-item version of the Geriatric Depression Scale (GDS-15) is widely employed to screen depression among elderly but little is known about the scale functioning in cognitively impaired individuals when compared to normal ones. The aim of the current study is to investigate Differential Item Functioning (DIF) across groups of older people that differ in terms of cognitive functioning applying Item Response Theory (IRT)-based analyses.

Methods: Data from an Italian multi-centric clinical-based study on cognitive impairment and dementia in old people were employed (N = 1903; Age: M = 77.33, SD = 7.05, 62% women). All the participants underwent a comprehensive evaluation (including clinical examination, laboratory screening, neuroimaging, and cognitive and behavioral assessments) and they were assigned to three different groups on the basis of their cognitive functioning (normal, mild cognitive impairment, cognitive impairment)

Results: Two items showed uniform DIF but their differential functioning does not propagate to the GDS-15 total scores in such a way that a differential interpretation is needed

Limitations: Whereas an advantage of the study is the large sample size, the relatively small size of the mild cognitive impairment group might reduce the stability of the present results

Conclusions: Since a screening tool for elderly is intended to apply to everyone in the target population, the current findings support the clinical utility of the GDS-15 as screening tool for depression.

1. Introduction

The 15-item Geriatric Depression Scale (GDS-15; Yesavage and Sheikh, 1986) is a widely employed screening test that offers an added value in the primary care detection of late-life depression. Indeed, it requires short time and effort to administer and it has good psychometric properties (for reviews, see Azulai and Walsh, 2015; Mitchell et al., 2010; Pocklington et al., 2016).

Regardless the extensive use (for a recent review, see Pocklington et al., 2016), it exists a debate on the possibility to administer the scale to people with a cognitive impairment, which is a very common and comorbid to depression disease in old age (for a review, see Wang and Blazer, 2015). On one hand, Wancata et al. (2006) affirmed that the scale should not be used with persons with marked cognitive impairment and Burke and colleagues (Burke et al., 1989) stated that only people with levels of 0 and 1 (i.e., cognitively intact and mildly demented individuals, respectively) on the Clinical Dementia Rating (CDR; Berg, 1984) were able to complete the test. On the other hand, Lach et al. (2010) built evidence for the use of the GDS-15 in populations that include people with mild to moderate dementia. Additionally, several studies (Conradsson et al., 2013; Jongenelis et al., 2005; McGivney et al., 1994; Smalbrugge et al., 2008) attested that the GDS-15 seems to be valid for people with mild cognitive impairments —assessed by the Mini-Mental State Examination (MMSE; Folstein et al., 1975)—, but it is unclear if the scale can be used with people with lower cognitive functioning (Conradsson et al., 2013).

This brief review suggests that further research is needed to understand whether cognitive impairments might produce biased responses or prevent the correct use of the GDS-15 (Luppa et al., 2012; Watson and Pignone, 2003). Specifically, it is important to ensure that other variables different from depression (i.e., the construct that the test seeks to measure) do not have an impact on the total test score. Since cognitive impairment might be a respondent's characteristics that produces biases on the GDS-15 total score, it is important to establish

* Correspondence to: Department of Neuroscience, Psychology, Drug, and Child's Health (NEUROFARBA), Section of Psychology, via di San Salvi 12- Padiglione 26, 50135 Firenze, Italy.
E-mail address: francesca.chiesi@unifi.it (F. Chiesi).

https://doi.org/10.1016/j.jad.2017.11.045
Received 28 January 2017; Received in revised form 9 September 2017; Accepted 11 November 2017
Available online 13 November 2017

empirically whether test items work in the same way across people with different cognitive functioning.

From a psychometric point of view, this issue is adequately addressed by Item Response Theory (IRT) that allows the assessment of Differential Item Functioning (DIF; Embretson and Reise, 2000; Reise and Waller, 2009). DIF analysis is used to study the performance of items in scales, and it examines whether or not the likelihood of an item endorsement is equal across subgroups (e.g., people with different levels of cognitive functioning), which are matched on the trait measured (e.g., depression). For example, a randomly selected individual with a normal cognitive functioning and a specific level of depression and a randomly selected individual with an impaired cognitive functioning but the same level of depression should have the same likelihood of endorsing an item measuring depression. This aspect is of particular relevance to screening tools because, as Zumbo (2003) pointed out, the presence of DIF might produce a systematic bias in the total test scores and, as a consequence, interpretations based on cut-off scores might be biased (Hidalgo, Galindo-Garre, and Gómez-Benito, 2015; Jones and Raju, 2000; Stark et al., 2004).

Starting from this premise, the current study aimed at further investigating potential DIF in the GDS-15 applying IRT. To the best of our knowledge, only one study addressed this issue investigating differential functioning of the GDS-15 items applying Rash analyses (Tang et al., 2005). Comparing three different cognitive functioning groups (defined using the MMSE), Tang et al. (2005) investigated DIF looking at item location parameters, i.e., if one group is consistently more likely than another to endorse an item, and they reported evidence of no DIF. In the current study, to ascertain if there are biases in the measurement process among individuals with different levels of cognitive function, we studied also item discrimination parameters that can be viewed as a significant group by trait interaction, i.e., if one group is more likely to endorse an item at certain levels of the trait, while the other group is more likely to endorse the item at other levels. Additionally, we focused on the assignment of individual to each studied group. As stated above, it is difficult to derive clear indications from the literature about the use of the GSD-15 in case of cognitive impairment because cognitive impairment was measured referring mainly to a single test (e.g., the CDR or the MMSE) and a variety of inclusion criteria for cognitive functioning levels and subgroup definition have been used. Thus, in order to ensure a valid classification for the variable of interest (i.e., cognitive functioning) and go beyond the limitations of previous studies, all the participants of the current study were classified after a comprehensive evaluation, including clinical examination, laboratory screening, neuroimaging, and cognitive and behavioral assessments.

2. Methods

2.1. Participants and measure

Data were gathered from those collected at the Institute of Gerontology and Geriatrics University of Perugia according to the methodology of the ReGAl project (Rete Geriatrica Alzheimer—Geriatric Network on Alzheimer's disease), a large longitudinal Italian multicentre clinical-based study, promoted by the Italian Society of Gerontology and Geriatrics (SIGG), and focused on cognitive impairment and dementia in old people, as described elsewhere (Boccardi et al., 2016; Mariani et al., 2008). All experimental procedures were conducted in accordance with the guidelines in the Declaration of Helsinki and approved by the Ethics Committee of the Hospital-University of Perugia.

The study enrolled 1903 people (Age: $M = 77.33$, $SD = 7.05$, range 45–96; 62% women) from September 2011 to June 2014. The neuropsychological battery included the MMSE, as test of general cognition, and specific tests evaluating episodic memory (Babcock Story Recall test and the immediate and delayed recall of the Rey's Auditory Verbal Learning Test), language (Token test for verbal comprehension and the Category Fluency test for language production according to semantic cues), attention and executive functions (Visual Search test and the Letter Fluency test) and praxis (Copy Drawing test). For each test, details on administration procedures and Italian normative data for score adjustment for age and education as well as normality cut-off scores (95% of the lower tolerance limit of the normal population distribution) are available (Carlesimo et al., 1996; Spinnler and Tognoni, 1987). To avoid the underestimation of the level of functional capacity, informant based rating of functional status were carried out (Tabert et al., 2002) using the BADL (Katz et al., 1963) and the IADL scales (Lawton and Brody, 1969). In most of the cases, informants were spouses or relatives, who lived in the same household. BADL includes six activities: bathing, dressing, toileting, transferring, continence, and feeding. IADL includes eight activities: using telephone, shopping, meal preparation, housekeeping, laundry, use of transportation, self-administration of drugs, and handling finances. Because IADL items are often gender-specific, we considered not only the current ability to perform each item but also the potential capability in case of need. For BADL score ranges from 0 (total independence) to 6 (total dependence), and for IADL from 0 (total independence) to 8 (total dependence). The CDR was used to score dementia severity. Finally, the battery included the GDS-15.

General exclusion criteria were the presence of clinically severe psychiatric or systemic disease, severe sensory impairment (blindness, deafness), neurological conditions associated with severe cognitive impairment (i.e., severe dementia or advanced stages of Alzheimer disease), a history of alcohol or substance abuse or dependence, and head injury with loss of consciousness. Therefore, 598 cases were excluded from the initial pool of data.

The remaining 1305 cases were classified as no cognitive impairment (NCI), mild cognitive impairment (MCI), and cognitive impairment (CI) through the multidimensional assessment derived from the ReGAl protocol. Specifically, inclusion criteria for NCI were: (a) age- and education-adjusted MMSE score higher or equal to 24 indicating good general cognitive functions, (b) normal performance in standardized memory tests and cognitive tasks, (c) scores higher than 4 on ADL and IADL indicative of no impaired functional capacity, and (d) CDR equal to 0. This group consisted of 531 cases aged from 60 to 94 years (56.7% women). Inclusion criteria for MCI were: (a) MMSE score higher or equal to 21 indicating preserved general cognitive functions, (b) objective memory deficit, defined as a pathological score (below the normality cut-off) in at least one standardized memory test, with normal performance in the other cognitive tasks, (c) scores of 4 or higher on ADL and IADL indicative of an adequate functional capacity, (d) CDR lower or equal to 1.00, and (e) no dementia (APA, 1994). This group consisted of 182 cases aged from 61 to 91 years (52.7% women). Inclusion criteria for CI subjects were in line with the assessment criteria for dementia (APA, 1994) referring to mild or moderate levels. Specifically, (a) MMSE score higher or equal to 17 indicating unpreserved general cognitive functions, (b) objective memory deficit, defined as a pathological score (below the normality cut-off) in standardized memory test, (c) low scores on ADL and IADL indicative of functional impairment, and (d) CDR higher or equal to 1.00. This group consisted of 529 cases aged from from 61–91 years (60.7% women). Finally, some cases ($N = 63$) were not clearly classifiable into these three categories due to inconsistencies among scores or missing information. Thus, to avoid incorrect classifications, they were excluded from the analyses.

3. Results

IRT analyses were conducted employing the IRTPRO software (Cai et al., 2011). We applied the unidimensional two-parameter (2PL) logistic model, which is the most commonly used IRT model in clinical assessment (for a review see Thomas, 2011).

Preliminarily, the dimensionality of the GDS-15 was tested in each