

Reliability study of parallel forms in a cognitive functioning test.

Aitana García Domenech 48560453-R.

Keywords: Reliability, parallel forms, cognitive functioning, disability.

Introduction:

When we talk about reliability we refer to "the stability of the measurements when there are no theoretical or empirical reasons to suppose that the variable to be measured has been differentially modified for people, so that its stability is assumed, until proven otherwise" (Muñiz, 2018, p.25). One of the procedures used to verify reliability is the parallel tests, which results in a reliability coefficient. It is a statistical procedure consisting of the application to a sample of people on the same occasion or on different occasions of two or more versions of the test equivalent in content, difficulty, etc. The forms must be applied under the same conditions. The use of parallel tests provides a very firm basis for estimating the accuracy of a measurement. The consistency with which it maintains its position from one measurement to another reflects the reliability of the test.

The objective of this work is to perform a reliability study of parallel forms in a cognitive functioning test, which must be identical and independent, in order that the different forms can be used interchangeably in the different moments of evaluation in the clinical process to avoid that the patient's responses are biased by the possible training or memory effects.

The scale used for this analysis is the Aprendizaje y Aplicación del Conocimiento de la Batería de Evaluación de la Capacidad y el desempeño (BECAD). The scale used for this analysis is the Knowledge and Application of Battery Knowledge Assessment of Capacity and Performance (BECAD). It is a screening test to determine the cognitive status of a patient. Among the objectives of this instrument, cognitive evaluation stands out, both in people with deterioration and with severe mental disorder and the evaluation of the results of clinical interventions, both to assess the improvement of a patient after establishing a treatment, and during the sessions of follow-up, due to this it is necessary to emphasize the importance of the correct creation of parallel forms of the scale in order to minimize the effects of practice and learning.

The theoretical framework of the BECAD battery is based on the Activities and Participation component of the CIF. Also in screening tests for cognitive impairment, such as The Montreal Cognitive Assessment (MoCA), the Minimental State Examination (MMSE) or Screen for Cognitive Impairment in Psychiatry (SCIP-S), which are characterized by providing a fast assessment of cognitive deficits presented by people suffering from mental illness. It includes relevant aspects for the reduction of disability, addressing both the individual functioning and the characteristics of the context from an integrating point of view between the health professional and those close to them.

Process:

participants:

A sample of 34 subjects was used, comprised between ages 22 and 88 years old with an average of 51.5, which were divided into two groups, a group formed with patients with deterioration and a control group of healthy subjects.

The group of healthy subjects has 26 subjects, 12 men and 14 women aged between 22 and 81 years with an average of 43.15 years of age and an average of schooling years of 15.73, which obtained mean scores of 58,154 points with a standard deviation of 5,167 in Form A, 57,77 points with a standard deviation of 6,18 in Form B, 56,65 points with a standard deviation of 6,42 in Form

C and 58.61 points with a standard deviation of 4.4 in the D form, all this on a maximum score of 65 points.

The group with cognitive impairment consists of 8 subjects, 2 men and 6 women between 60 and 88 years old with an average age of 78.62 and a mean of 9.62 years of schooling, which present different types of conditions, among those that are schizophrenia, Alzheimer's, senile dementia, stroke ... These subjects obtained an average score of 36 points with a standard deviation of 13.005 in Form A, 34.5 points with a typical deviation of 10.99 in Form B, 33.88 points with a standard deviation of 13.94. in Form C and 33.63 points with a standard deviation of 10.17 in Form D.

Method:

The BECAD Learning and Application of Knowledge scale was administered individually to all subjects, in a controlled environment in which there were no distractions or sounds that could be annoying. It was taken into account that the light was adequate as well as the tone of voice of the evaluator who read the questions, because of many of the people evaluated had some visual or auditory deficit.

The subjects performed the test on a desk comfortably seated and had a pen and a sheet in which they could write anything they needed to point knowing that it would not be taken into account for the final result of the test.

Both the duration of the sessions and the number of sessions depended exclusively on the evaluated person, since there is a great diversity among the subjects and their capacities. Due to this, the time used for the complete completion of the 4 forms was approximately one and a quarter hours in the control group, reaching up to 4 hours in the case of people with serious deterioration of the experimental group. In the same way, the number of sessions also varied, the people in the control group usually did all four forms in the same session, unlike the experimental group, which needed a minimum of two sessions, reaching in some cases up to four sessions. These variations were made depending on the needs of each subject, because before starting the test they were explained that if they felt tired or had any problem they could stop as soon as they needed it without this having an impact on the final result of the test.

At the beginning of the administration of the test, the evaluator reads the items out loud and records the answers that the subject gives, except for the items in which the own elaboration is necessary.

At the end of the evaluation phase of the subjects, four templates of variables were created in the SPSS program to be able to perform the statistical calculations, one for each form of the scale (form A, B, C and D), in which the variables were each item of the test and the subjects were differentiated between control subjects and subjects with deterioration, then they were manually filled with the scores of the individual booklets of each subject.

First, the descriptive analyzes were carried out, including the average score and the standard deviation of the groups based on age, sex and schooling years. Then the parallel forms analysis between the four forms of BECAD and another analysis of internal consistency of the forms was made. Also an Anova.

Results:

The results obtained from the descriptive statistical analyzes indicate the mean scores in the different forms of the scale of the total group, the healthy group and the group with deterioration. There is a significant difference between the two groups, this difference is 21.61 points in Form A, 23.27 in Form B, 22.78 in Form C and 21.99 in Form D, giving an average difference between the groups of 22.41 points. It was shown that these differences were significant through an Anova.

The reliability analysis of parallel tests was also performed based on the total group, healthy subjects and those with deterioration. The results show Pearson correlations for a 95% confidence interval. In the total group, the lowest correlation is in the AB forms with a correlation of 0.9462, and the highest is 0.9597 for the B C forms. In the control group, the lowest correlation is the CD forms with 0.8157, and the highest is 0.9193 in the A C forms. Finally, in the group with deterioration, the lowest correlation belongs to the AD forms with 0.8423, and the highest is 0.9779 in the B forms D.

With respect to internal consistency, the alpha coefficient of each form reflects the degree to which the items that constitute each form of the scale would co-vary. In the total sample, Form A has a score of 0.95, B has 0.95, C is 0.96 and D is 0.94. In the control group the score of form A is 0.81, the score of 0.85, C of 0.85 and D of 0.63. In the group with deterioration the scores are 0.95 for the A form, 0.94 for the B, 0.96 for the C and 0.92 for the D.

Conclusions:

The statistical results in the analysis of parallel forms show that the four forms of the BECAD are parallel to each other, with rather high reliability coefficients, which shows that they are equivalent, and therefore any of the four forms could be used in different moments of the evaluation to avoid training effects and memory of the subjects. In addition, the Anova analysis shows that there are significant differences between the scores of the group of healthy subjects and that of deterioration, and that there are no differences between the four forms in the two groups. The internal consistency analyzes indicate that the four forms have a high internal consistency.

Among some of the limitations that have been found at the time of conducting the study, it can be highlighted that the order of administration of the tests was always the same, ABCD and the process could have been randomized so that the statistical analyzes were more exhaustive. In addition, some subjects, especially those in the deterioration group, needed more than one session to complete the scale, which can produce learning effects. It has also been found that among the subjects of the sample there is a great difference in the schooling years, and that these correlate negatively with age, so it would be interesting to return to the study with more elderly people among those who have more variety in the schooling years.

Bibliography:

- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699.
- Moro Ipola, M., & Mezquita Guillamón, L. (2013). Desarrollo y validación de la BECAD: un instrumento de evaluación basado en la Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud. *Revista Española de Salud Pública*, 87(1), 11-24.
- Muñiz, J. (2018). *Introducción a la Psicometría. Teoría clásica y TRI*. Madrid, Pirámide.
- OMS, Organización Mundial de la Salud; Organización Panamericana de la, & OMS. (2001). *Clasificación Intenacional del Funcionamiento, de la Discapacidad y de la Salud. Versión abreviada*. Madrid: IMSERSO.
- Piera, P. J. F., Lorenzo-Seva, U., & González, R. P. (2009). Implementación de procedimientos gráficos y analíticos para la construcción de formas paralelas. *Psicothema*, 21(2), 321-325.
- Pino, O., Guilera, G., Rojo, J. E., Gómez-Benito, J., & Purdon, S. E. (2014). SCIP-S, screening del deterioro cognitivo en psiquiatría. TEA Ediciones, Madrid.
- Prieto, G., & Delgado, A. R. (2010). Fiabilidad y validez. *Papeles del psicólogo*, 31(1), 67-74.
- Poggi, M., BERTONI, A., & Teobaldo, M. (1996). Evaluación. Nuevos significados para una práctica compleja. Bs. As. Editorial: Kapelusz.
- Thorndike, R. L. (1989). *Psicometría aplicada*. Limusa.