

Procedures for Successful Data Collection through Psychological Tests in the Elderly

Procedimentos para uma Recolha de Dados Bem-Sucedida junto de Idosos

António Moreira Diniz* & Nuno Amado
Universidade de Évora, Évora, Distrito de Évora, Portugal

Abstract

Drawing on the experience of sampling and assessing thousands of elders, who dwell both in rural and urban areas, from all socio-economic strata, we present data collection procedures recommended to assess elderly populations. Based on International Test Commission's Guidelines (2010) for test translation and adaptation, we propose the administration of psychological tests to older adults in an interview format arguing that it is better suited for such population. We provide and discuss practical examples and common problems of data collection regarding: the need of adapting the translation to make cultural and psychological sense; the lack of familiarity of most elders with the psychological assessment setting and process; and, ethical concerns for this specific population.

Keywords: Elderly, psychological assessment, protocol adaptation, International Test Commission, illiteracy.

Resumo

Baseados na experiência de seleção e avaliação de milhares de idosos residentes em meio rural e urbano, pertencentes aos diversos estratos socioeconómicos, apresentamos procedimentos de recolha de dados recomendáveis para esta população. Considerando as diretrizes de tradução e adaptação de testes da *International Test Commission* (2010), propomos a administração de testes psicológicos a idosos sob a forma de entrevista, argumentando que esta é a forma mais adequada. Apresentamos e discutimos exemplos práticos e problemas comuns relativos à recolha de dados, considerando: a necessidade de nas traduções ser atendido o significado cultural e psicológico; a falta de familiaridade da maioria dos idosos com o cenário e o processo de avaliação psicológica; e, aspetos éticos vis-à-vis esta população.

Palavras-chave: Idosos, avaliação psicológica, adaptação de protocolos, *International Test Commission*, analfabetismo.

There is a growing consensus regarding the need for the careful adaptation of psychological tests for use in diverse linguistic and cultural contexts (Hambleton, 2001; Hambleton & Patsula, 1999; International Test Commission, 2010; Muñiz & Bartram, 2007; Peña, 2007; Tanzer & Sim, 1999; Vijver & Tanzer, 2004). Gone, one hopes, are the days when psychological tests developed in a specific language and country would be used in another culture and language with a poor translation of the items. Nowadays, it is recognized that the specific characteristics of a population must be carefully taken into account to ensure the linguistic and cultural fit of the adapted instrument. It is not just a matter of its linguistic equivalence, but also of its' functional, cultural, measurement, structural and nomological equivalence.

The International Test Commission's (2010) guidelines for translating and adapting tests (ITC Guidelines) present a framework that allows researchers of many different fields to successfully conduct the adaptation of tests. However, given its universal nature, the guidelines may be insufficient when one considers particular populations where the adaptation process, when put into practice, faces numerous particular difficulties and challenges. Such is the elderly population that due to its profile of functioning, namely in the fourth age (which can be considered individually, according to a person-based definition, or according to a population-base definition, beginning between 75-80 or 80-85 years, depending on the country; Baltes & Smith, 2003), and socio-economic and history-graded features, presents a daunting challenge for a successful adaptation of psychological instruments.

We propose some procedures for data collection through psychological tests relevant to elderly populations based on the experience obtained through the adaptation and validation of several instruments that originated more

* Mailing Address: Departamento de Psicologia, Universidade de Évora, Largo dos Colegiais, 2, Évora, Portugal 7004-516. E-mail: amdiniz@uevora.pt and nunoamado@uevora.pt

than 20 master theses, and several publications. Among these we can mention, as examples of the width of the research conducted: A doctoral dissertation on successful aging (Amado, 2008); a book chapter concerning the adaptation of the Life Satisfaction Index (Diniz, Amado, & Gouveia, 2011); a paper contrasting institutionalized and community elders (Pimentel & Diniz, 2012); and, conference proceedings on religious faith (Amado & Diniz, 2008) and on sexual attitudes (Margalha, Diniz, & Vaz, 2008). Our research has involved the sampling and assessment of thousands of seniors from all socio-economic strata, residing both in rural and urban areas.

The detailed account of how the guidelines fit a specific population illustrates the theoretical and practical constraints present when adapting psychological assessment protocols. We hope that this alerts researchers of different fields to carefully consider how the guidelines fit their target populations.

Specificity of the Adaptation Process Regarding Target Populations

When adapting a test, one should be acquainted with the manifold characteristics of the population for which it is being adapted. We do not mean to provide a complete description of aging, but it's important to mention that one cannot adapt tests for the elderly without considering issues of health and well-being (Isaacowitz & Smith, 2003; Kunzmann, Little, & Smith, 2000), namely the impact of functional health and of chronic illness, depression or anxiety in the administration process and how they may influence the validity of the measures obtained.

It is known that the historical conditions influence the individual development (Baltes, 1987; Elder, 1991). So, one may expect that the elderly, being born in an undoubtedly different socio-economic and cultural environment from the actual one, in domains such as educational system, labor issues, mass media and public opinion, income, family structure, values, habits and daily practices, show significant differences *vis-à-vis* other age groups. The way these issues are reflected on elder's current life circumstances needs to be taken into account.

The following Fernández-Ballesteros, Hambleton, and Vijver's (1999) statement highlights the spirit of this article: "Psychological measures developed for one age group may not be suitable for other age groups without some modifications" (p. 169). Beyond changes in instruments, the administration procedure for a specific population must also be pondered. As Brazier, Walters, Nicholl, and Kohler (1996) suggest, the use of self-administrated measures in a test-based interview to assess elderly populations can overcome problems of format of administration, consistency of response, and floor effects. Moreover, the administration of psychological tests to elders through an oral presentation when visual or motor impairment exists (Ricker & Axelrod, 1994), or an interview when the

socio-economic context demands (Ganguli et al., 1999; Tajvar, Arab, & Montazeri, 2008), has been producing good results. Based on ITC Guidelines (cf. Appendix), the main scope of the article is dedicated to providing a framework for the administration-through-interview format of psychological tests in elderly populations. We also highlight aspects of the cultural adaptation of tests and consider the use of this format to group comparisons.

Procedures Regarding Cultural Adaptation

The ITC Guidelines D.1 and D.2 (cf. Appendix) stress the need for taking into account not only the linguistic but also the cultural differences between the population for which the test was originally developed and the population for which it is being adapted.

Functional Equivalence

There are methodological sources of bias when conducting the processes regarding the translation of tests (Fernández-Ballesteros et al., 1999; Hambleton, 2001; Peña, 2007; Tanzer & Sim, 1999; Vijver & Tanzer, 2004). The combination of translation, back-translation (from the target language to the original language), and the use of a panel of experts to evaluate the result of the back-translation, to consider issues such as item difficulty and cultural equivalence, enhances the quality of the translation. The expert panel enables us to overcome the bias caused by translators. Our practice has shown that, when not specifically warned, translators do not use simple and clear language (Hambleton, 2001) but, instead, create items that are too erudite to be applied to populations with poor schooling.

Also, when translating tests, not only linguistic equivalence must be attained but functional equivalence too (Rogler, 1999). This means that not only must the translated items be transposed with full meaning from the original language but the cultural meaning must remain intact. We present an example from the translation of the Webster's (2003) Self-Assessed Wisdom Scale. An item contained the English idiomatic expression "do not judge a book by its cover", which is a popular saying in many English speaking countries, was literally translated to a Latin language. Had an expert panel not been consulted, the item would have remained this way in the final version. However, the expert panel concluded that the expression does not resonate in the same way amongst the elderly population of interest. So, to achieve not only linguistic but functional equivalence we re-wrote the item as "do not make judgments based on appearances" which is easier to understand, but still keeps the original meaning. Items should be written in the vernacular, bearing in mind, in a user-centred and user-friendly approach, the culture and the use of the language by the population of interest (Skevington, Sartorius, Amir, & The WHOQOL-Group, 2004).

Beside the expert panel, we recommend that the first version of the instrument in the new language be applied in a small group of participants, using a thinking-aloud procedure wherein the reasoning behind the answers is expressed. These should represent the main characteristics of the population for which it is being adapted, namely, when regarding the elderly, members of high as well as low socio-economic strata, and women and men. This assessment can be conducted in individual face-to-face setting or in group setting.

This thinking-aloud procedure can also permit the manifestation of important cultural information regarding the items. For example, when adapting an instrument to measure positive and negative affect in old age, derived from the PANAS (Watson, Clark, & Tellegen, 1988), it allowed the detection of a problematic item. This item was “proud”, which had been correctly translated to a Latin language. However, the pre-test of the instrument, with an administration-through-interview procedure, showed that some elderly persons viewed pride not as a positive affect but as a negative one, due to religious connotation. Catholic upbringing stresses the mortal sins, of which pride (the opposite of the virtue of modesty) is one. To avoid this connotation, we kept the correctly translated word but added a caveat to imply that it was to be understood in its positive way.

We also suggest that the adaptation process develops versions of items able to be understood by the least educated members of this population. Doing this, one follows the principle of social inclusion: If the item can be answered by the least educated it can be answered by the rest of the population. For example, when adapting the Selection Optimization and Compensation Questionnaire (Short-Form; Freund & Baltes, 1998), the previous mentioned thinking-aloud procedure with a heterogeneous small group of elderly showed the necessity to create paraphrases for some of the translated items, rewriting expressions like “focusing on important goals” in a simpler form such as “achieving what is important to me”. Making tests able to be applied to all the elderly allows for their use in representative samples.

Taken as a whole, the proposed sequential combination of procedures to achieve functional equivalence of adapted tests ensures more effectively its cultural and linguistic fit to the population of interest.

Procedures Regarding Administration Adaptation

The ITC Guidelines A.1, A.2, A.3 and A.4 (cf. Appendix) mention the need to anticipate and minimize, as far as possible, the types of problems that can be expected, preparing the materials and administration procedures and response modes to avoid threats to validity.

Administration-Through-Interview

When considering the elderly population, and while taking into account the previous guidelines, we suggest

the administration of psychological tests by an interviewer trained for this purpose. Although ITC Guideline A.6 (cf. Appendix) recommends the minimization of the interaction between administrator-examinee, there are instrumental and ethical-deontological reasons for the administration-through-interview.

The first instrumental reason is the lack of familiarity and ease of the elderly with an assessment setting or procedure. The ubiquity of written tests can make researchers unaware of how it can feel strange and menacing to sit in silence writing answers to unfamiliar questions, often demanding ones. To have these same questions made by another person in the course of a conversation, although it doesn't eliminate all of the menacing aspects of being assessed, can make the procedure closer to everyday life. A second instrumental reason to use an interview regards issues of functional health. Poor eyesight, deficient fine motor control and other problems make it harder for the elderly to answer written tests. The third instrumental reason to use an interview, when working with elders, is that some of them are illiterate (Carr-Hill, 2008).

To the instrumental reasons for using an interview we add the ethical-deontological arguments. The content of many psychological tests may elicit strong responses in the elderly living in social isolation, suffering from an array of health and/or economic problems and/or facing death. The administration of psychological tests by trained interviewers attends to these emotional responses and provides a safer context for the elderly to answer truthfully. Also, to require the elderly to answer by writing is to exclude the participation of many and, consequently, to paint an incorrect portrait of aging.

The administration of psychological tests in an interview replaces the impersonal data collection context by a context of interpersonal relationship between the elderly person and the interviewer. This personal connection is of special utility while developing the adapted version of the test. Using the interview, researchers can gather data of how the elderly responds to its administration and also identify possible threats to validity, such as problematic items or typical and socially desirable answers.

Framework for the Administration

When working with the elderly, we recommend some measures to ensure that the administration is as unobtrusive as possible, and that it fits the culture (ITC Guidelines A.5 and A.6, cf. Appendix).

Provide a Safe Context and Ensure Confidentiality.

The environment surrounding the administration should feel safe for the elderly person. It must be a location where privacy is assured, where the lighting and sound conditions are favorable and where both the interviewer and the interviewee feel comfortable. Following ethical principles applied to psychological assessment, confidentiality of the answers must be assured and, most importantly, understood by the elderly.

Watch Out for False Volunteers. When working in institutions, such as day care centers, even though researchers stress that participation in the research is optional and voluntary, some elderly persons may feel forced, by other or by the institution staff, to participate. The interviewers should always, at the beginning of the interview inform the elderly person being assessed that participation is voluntary. Usually, from our experience, false volunteers show themselves suspicious towards the interviewer and his questions, and answer with unusual speed or start to answer even before the end of the question has been asked: This behavior translates a wish to end the interview as fast as possible.

Watch Out for Social Desirability Bias. The interview format of test administration has its downside on the will of the person being interviewed to please the interviewer (Dillman & Christian, 2005; Rosenthal, 1976). The interviewer should try to control/minimize answer bias (Cronbach, 1946; Rorer, 1965). One strong source of bias is related to item content, when it implies a socially desirable answer. Another source of bias can be unrelated to item content, where answering styles provide bias (acquiescence, negativity, random answer and central tendency answer). For certain specific psychological tests (for example, regarding sexual attitudes and behaviors) the need for gender correspondence between interviewer and the elderly person should be made a priori.

Consider the Time/Season when to Conduct the Interview. Taking into account an array of factors, such as circadian rhythms (Li, Hasher, Jonas, Rahhal, & May, 1998) and the schedules of senior care institutions, we recommend the administration of tests, when possible, at mid-morning. Also, well-being related tests should not be applied during holiday seasons or special dates, such as Christmas and anniversaries (for example, death of significant persons). From our experience we have found a halo effect that artificially and temporarily boosts well-being assessments, even when these regard trait-like affect.

Conducting the Interview

Prior to the Administration. When collecting data in institutions, the one-on-one process will feel atypical to many elderly persons, whose daily routine in institutions many times involves schedules, activities and spaces designed for groups, and where individuality can be easily forgotten. Stress the individual contribution of the participants for the research being conducted, which reminds the elderly person that his personal opinions and thoughts matter and that he is still of importance to society.

Also, before starting, explain that the following interview is part of a large study of seniors. As such, it was developed to be as inclusive and clear as possible, to allow for its application among seniors of varied backgrounds and characteristics. This preamble prevents that the more literate and functional elderly feel patronized.

Throughout the whole process, exhibit a disposition to conduct the interview according to the particular rhythm of

the elderly person. Many show a need to interact and talk with the interviewer; usually the administration-through-interview of the tests consists in a welcomed break from routine and a chance to be heard by someone intent on listening.

When the interviewer is also a psychologist, diminish the belief that this must give the interview either a clinical or a performance assessment bent. This can be done by explaining that the research is not focused on individual results and that the researcher is not interested in rating or assessing the performance of the respondent, but simply in honest, genuine answers; so as to diminish the anxiety of being assessed – the elderly want to perform well and are afraid of making mistakes. Also, the elderly should be informed that any concerns or matters that they wish to discuss regarding any items or questions can be fully addressed at the end of the interview.

When working with ordered-categorical measures, interviewers should also have some practice items to assure that the elderly person understands the response format. For instance, some elderly show difficulties when asked to answer items using more than three rating categories. With a Likert scale response format, interviewers should first ask them if the answer is positioned in the agreement or disagreement side of the scale and after ask for the accurate answer.

During the Administration. The interviewer should read the items and the answer formats aloud, in a moderately loud tone of voice. When necessary, items or answer formats should be re-read. The interviewer should also avoid the induction of answers, by minimizing interactions. Often the elderly will comment, reminisce or even pose questions to try to understand the expectations of the interviewer regarding the answer. The interviewer should politely postpone discussion for after the administration and ask for the answer.

It is also important to take into account non-verbal behavior of the elderly, such as body language, the use of silence or tone of voice that may signal grief or other negative emotions triggered by the questions. If necessary, the interviewer should make a note of this and address it after the administration. Manifest positive emotions triggered by the questions should also be noted. They can be used, at the end of the procedure, to frame the experience of the administration in a positive way.

After the Administration. Any question, stories, doubts and so on, that came up during the administration should be addressed now. The interviewer should pay special attention to the emotional impact of certain tests contents (for example, death anxiety, well-being). It is possible that answering these tests elicits strong emotional responses and/or the need to share. Although it is not the role of the interviewer to act as a therapist (and it is recommended strongly that one does not act so), he must not end its task when all the information he needs is gathered, but must listen emphatically and treat the elderly with all consideration and ethical care required.

Overall, these procedures provide the background to adapt standardized contents (including contingencies) for data collection regarding psychological assessment in elderly populations. Following the protocol, administration bias (Vijver & Tanzer, 2004) is avoided as long as the trained interviewers possess a framework grounded in the specificity of the target population. Likewise, threats to validity due to instrumental reasons, bounded by the elderly physical, cognitive and emotional conditions, can be overcome.

Protocol of Adapted Administration Procedures: Issues Regarding Group Comparisons

The ITC Guideline D.6 (cf. Appendix) reminds the need for the data collection design to allow the empirical study of item equivalence between different versions of the test. The proposed changes in the test administration from self-administration to test-based interview stem from this guideline.

The protocol of adapted data collection procedures through tests in elderly can produce biased test results in comparisons with self-administrated data from the same tests in the population. Between-group differences (for example, elders in the third and in the fourth age) at measurement, structural and nomological level of a test can be produced only because a method effect jeopardizes the research's statistical conclusions.

The method effect bias due to administration can be controlled in two ways: (a) Standardizing the data collection procedures for all participants; or, (b) testing across matched samples the measurement, structural and nomological equivalence of the two administration procedures. The former option has been the one we selected to work with elders. All of them were interviewed, despite the human resources and the time consumption it evolves. The later is rarely used, maybe because authors consider it as just an empirical issue. Often both administration techniques (self-administration and test-based interview) are used interchangeably, due, for example, to a visual impairment or to an inability to write because of rheumatoid arthritis (Schuurmans et al., 2005), and also to illiteracy (Fernández-Ballesteros, 2001). In these cases, method effect bias could be studied regarding the potential of means and covariance structures analysis to assess differential item functioning and also latent means, variances and covariances differences across multiple group data (e.g., Byrne, Shavelson, & Muthén, 1989; Little, 1997; Meredith, 1993; Millsap & Yun-Tein, 2004). Although the details of these procedures exceed the scope of this article, we are currently conducting this research.

Conclusions

Our experience suggests the possibility of successfully adapting a wide range of tests to the elderly population.

However, the adaptation process is far from simple or immediate.

The process of adaptation should not be a mechanical transposition of items from a language and cultural context to another, but a carefully tailoring of the original instrument to the new target population. Linguistic equivalence is a necessary but not sufficient condition; functional equivalence between source and target languages should also be achieved. We have seen the usefulness of an expert panel and of a thinking-aloud procedure with a heterogeneous small-group of the target population to examine, respectively, the work of the translators and experts, and the instrument's first version.

Using an administration-through-interview procedure allows for a better and fairer use of tests with elderly persons, but it also implies a tremendous amount of effort and time spent in training administrators and in gathering data. Nonetheless, it provides valid and new information regarding items' functional equivalence, problematic items or typical and socially desirable answers, and allows the overcome of instrumental problems in test administration due to illness/disability, illiteracy, and lack of familiarity with tests. Also, taking into account the low socio-economic status of sectors of the elderly population, even in developed countries, such as the United States and England (Population Reference Bureau, 2008), it can be considered an ethical-deontological obligation. The administration of psychological tests by trained interviewers to the elderly living in social isolation, suffering from an array of health and economic problems and/or facing death, permits to emphatically attend emotional responses after the administration and provides a safer context for the elderly to answer truthfully.

The presentation of the procedures for data collection through psychological tests recommended to elderly populations illustrates how the ITC Guidelines (cf. Appendix) adjust to the theoretical and practical constraints present when adapting psychological assessment protocols. They stem from our research on aged populations, being relevant to elders in the third and especially in the fourth age. However, this does not mean that many of the recommended procedures cannot be used in other populations, namely illiterate and with sensory and motor disabilities populations. The essential of a successful adaptation lies in the thorough knowledge of the target population. Researchers adapting psychological tests should consider both the theoretical knowledge of their domains and the methodological knowledge of test adaptation to successfully adapt assessment protocols for their target populations.

References

- Amado, N. (2008). *Sucesso no envelhecimento e histórias de vida em idosos sócio-culturalmente muito e pouco diferenciados* (Dissertação de doutorado, Universidade Nova de Lisboa, Instituto Superior de Psicologia Aplicada, Lisboa, Portugal).

- Amado, N., & Diniz, A. M. (2008). Estrutura factorial e poder discriminativo do Questionário da Força da Fé Religiosa de Santa Clara (versão reduzida) numa amostra de adultos idosos [CD-ROM]. In A. P. Noronha, C. Machado, L. S. Almeida, M. Gonçalves, S. Martins, & V. Ramalho (Eds.), *Actas da XIII Conferência Internacional de Avaliação Psicológica: Formas e Contextos*. Braga, Portugal: Psiquilíbrios.
- Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23*, 611-626.
- Baltes, P. B., & Smith, J. (2003). New frontiers in the future of aging: From successful aging of the young old to the dilemmas of fourth age. *Gerontology, 49*, 123-135.
- Brazier, J. E., Walters, S. J., Nicholl, J. P., & Kohler, B. (1996). Using the SF-36 and EuroQol on an elderly population. *Quality of Life Research, 5*, 195-204.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.
- Carr-Hill, R. (2008). *International literacy statistics: Review of concepts, methodology and current data* (Report). Montreal, Canada: Institute for Statistics, United Nations Educational, Scientific and Cultural Organization. Retrieved May 11, 2012, from <http://unesdoc.unesco.org/images/0016/001628/162808e.pdf>
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement, 6*, 475-494.
- Dillman, D. A., & Christian, L. M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods, 17*, 30-52.
- Diniz, A. M., Amado, N., & Gouveia, J. (2011). Índice de Satisfação com a Vida (ISV). In C. Machado, M. Gonçalves, L. Almeida, & M. R. Simões (Eds.), *Instrumentos e contextos de avaliação psicológica* (Vol. 1, pp. 45-65). Coimbra, Portugal: Almedina.
- Elder, G. H. (1991). Family transitions, cycles, and social change. In P. A. Cowan & M. Hetherington (Eds.), *Family transitions* (pp. 31-57). Hillsdale, NJ: Erlbaum.
- Fernández-Ballesteros, R. (2001). Environmental conditions, health and satisfaction among the elderly: Some empirical results. *Psicothema, 13*, 40-49.
- Fernández-Ballesteros, R., Hambleton, R., & Vijver, F. J. R. van de (1999). Protocol adaptation procedures. In J. J. F. Schroots, R. Fernández-Ballesteros, & G. Rudinger (Eds.), *Aging in Europe* (pp. 169-184). Amsterdam, Netherlands: IOS Press.
- Freund, A. M., & Baltes, P. B. (1998). Selection, optimization and compensation as strategies of life management: Correlations with subjective indicators of successful aging. *Psychology and Aging, 13*, 531-543.
- Ganguli, M., Dube, S., Johnston, J. M., Pandav, R., Chandra, V., & Dodge, H. H. (1999). Depressive symptoms, cognitive impairment, and functional impairment in a rural elderly population in India: A Hindi version of the Geriatric Depression Scale (GDS-H). *International Journal of Geriatric Psychiatry, 14*, 807-820.
- Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17*, 164-172.
- Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Test Technology, 1*, 1-30.
- International Test Commission. (2010). *International Test Commission Guidelines for translating and adapting tests*. Retrieved March 30, 2011, from <http://www.intestcom.org>
- Isaacowitz, D. M., & Smith, J. (2003). Positive and negative affect in very old age. *Journal of Gerontology: Psychological Sciences, 58B*, 143-152.
- Kunzmann, U., Little, T. D., & Smith, J. (2000). Is age-related stability of subjective well-being a paradox? Cross-sectional and longitudinal evidence from the Berlin Aging Study. *Psychology and Aging, 15*, 511-526.
- Li, K. Z. H., Hasher, L., Jonas, D., Rahhal, T., & May, C. (1998). Distractibility, circadian arousal, and aging: A boundary condition. *Psychology and Aging, 13*, 574-583.
- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53-76.
- Margalha, N., Diniz, A. M. & Vaz, I. (2008). Modelos alternativos da Escala de Atitudes Sexuais (EAS): Teste numa amostra de adultos idosos [CD-ROM]. In A. P. Noronha, C. Machado, L. S. Almeida, M. Gonçalves, S. Martins, & V. Ramalho (Eds.), *Actas da XIII Conferência Internacional de Avaliação Psicológica: Formas e Contextos*. Braga, Portugal: Psiquilíbrios.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.
- Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*, 479-515.
- Muñiz, J., & Bartram, D. (2007). Improving international tests and testing. *European Psychologist, 12*, 206-219.
- Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development, 78*, 1255-1264.
- Pimentel, F., & Diniz, A. M. (2012). Perceived motor ability and selection, optimization, and compensation: Effects of age and institutionalization. *Psico-USF, 17*, 357-368.
- Population Reference Bureau. (2008). Socioeconomic status and health disparities in old age. *Today's Research on Aging, 11*, 1-4.
- Ricker, J. H., & Axelrod, B. N. (1994). Analysis of an oral paradigm for the Trail Making Test. *Assessment, 1*, 47-52.
- Rogler, L. H. (1999). Methodological sources of cultural insensitivity in mental health research. *American Psychologist, 54*, 424-433.
- Rorer, L. G. (1965). The great response style myth. *Psychological Bulletin, 63*, 129-156.
- Rosenthal, R. (1976). *Judgement studies: Design, analysis, and meta-analysis*. Cambridge, MA: Cambridge University Press.
- Skevington, S. M., Sartorius, N., Amir, M., & The WHOQOL-Group. (2004). Developing methods for assessing quality of life in different cultural settings: The history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology, 39*, 1-8.
- Schuermans, H., Steverink, N., Frieswijk, N., Buunk, B. P., Slaets, J. P. J., & Lindenberg, S. (2005). How to measure self-management abilities in older people by self-report: The development of the SMAS-30. *Quality of Life Research, 14*, 2215-2228.
- Tajvar, M., Arab, M., & Montazeri, A. (2008). Determinants of health-related quality of life in elderly in Tehran, Iran. *BMC Public Health, 8*, 323. Retrieved June 15, 2010, from <http://www.biomedcentral.com/content/pdf/1471-2458-8-323.pdf>

- Tanzer, N. K., & Sim, C. Q. E. (1999). Adapting instruments for use in multiple languages and cultures: A review of the ITC guidelines for test adaptations. *European Journal of Psychological Assessment, 15*, 258-269.
- Vijver, F. J. R. van de, & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology, 54*, 119-135.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.
- Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development, 10*, 13-22.

Recebido: 08/03/2013
1ª revisão: 03/06/2013
Aceite final: 12/06/2013

Appendix

Cited ITC Guidelines (International Test Commission, 2010)

- D.1 Test developers/publishers should insure that the adaptation process takes full account of linguistic and cultural differences among the populations for whom adapted versions of the test or instrument are intended.
- D.2 Test developers/publishers should provide evidence that the language use in the directions, rubrics, and items themselves as well as in the handbook are appropriate for all cultural and language populations for whom the test or instrument is intended.
- D.6 Test developers/publishers should ensure that the data collection design permits the use of appropriate statistical techniques to establish item equivalence between the different language versions of the test or instrument.
- A.1 Test developers and administrators should try to anticipate the types of problems that can be expected, and take appropriate actions to remedy these problems through the preparation of appropriate materials and instructions.
- A.2 Test administrators should be sensitive to a number of factors related to the stimulus materials, administration procedures, and response modes that can moderate the validity of the inferences drawn from the scores.
- A.3 Those aspects of the environment that influence the administration of a test or instrument should be made as similar as possible across populations of interest.
- A.4 Test administration instructions should be in the source and target languages to minimize the influence of unwanted sources of variation across populations.
- A.5 The test manual should specify all aspects of the administration that require scrutiny in a new cultural context.
- A.6 The administrator should be unobtrusive and the administrator-examinee interaction should be minimized. Explicit rules that are described in the manual for administration should be followed.