Copyright © 2015 by Academic Publishing House *Researcher*

Published in the Russian Federation
European Journal of Physical Education and Sport
Has been issued since 2013.
ISSN: 2310-0133
Vol. 8, Is. 2, pp. 109-117, 2015

DOI: 10.13187/ejpe.2015.8.109
www.ejournal7.com

UDC 79

Reliability of the DESK 3-6 for 3-Years Old Children

[1] Miltiadis Proios
[2] Vasilios Tsimaras
[3] Maria Sidiropoulou
[4] Despina Arzoglou
[5] Kosmas Christoulas
[6] Theofilos Pillianidis

[1] Aristotle University of Thessaloniki, Greece
Department of Physical Education and Sport Science
E-mail: mproios@phed.auth.gr
[2] Aristotle University of Thessaloniki, Greece
Department of Physical Education and Sport Science
E-mail: tsimaras@phed.auth.gr
[3] Aristotle University of Thessaloniki, Greece
Department of Physical Education and Sport Science
E-mail: sidiropo@phed.auth.gr
[4] Aristotle University of Thessaloniki, Greece
Department of Physical Education and Sport Science
E-mail: desarz@yahoo.gr
[5] Aristotle University of Thessaloniki, Greece
Department of Physical Education and Sport Science
E-mail: kchristo@phed.auth.gr
[6] Democritus University of Thrace, Greece
Department of Physical Education and Sport Science
E-mail: thpillian@phyed.duth.gr

Abstract

An early detection of possible disorders mainly in preschool children could help the prediction of disorders relating to learning disabilities and problems in school performance. For this reason in the present study investigated the reliability of DESK 3-6 for children 3-years old of age through the use of methods such are internal consistency reliability, test-retest and Cronbach's alpha, as well as check its suitability on a different sample. Participants were 383 preschool children (207 boys and 176 girls) with ages ranging from 36 to 47 months. The Dortmund Developmental Screening for Preschools was used. Results revealed that the original version of DESK 3-6 for children 3-years olds of age shows problems in internal consistency, adequate Cronbach alpha and test-retest coefficients.

Keywords: DESK 3-6, Reliability, Internal consistency reliability, Cronbach alpha, Test-retest reliability.

Introduction

The appearance of developmental disorders in children is an often phenomenon that continuously rise throughout the years (Fombonne, Zakarian, Bennett, Meng, & McLean-Heywood, 2006; Chakrabarti & Fombonne, 2001). An early detection of possible disorders mainly in preschool children could help the prediction of disorders relating to learning disabilities and problems in school performance (Esser, 1993). For this the instrument Dortmund Developmental Screening for Preschools (DESK 3-6; Tröster, Flender, & Reineke, 2004) was developed.

DESK 3-6 is an instrument developed to identify 3, 4, and 5-6 years old children with developmental disorders (Tröster et al., 2004). It is partly composed of monitoring tasks completed by the teacher based on his/her daily observations and performance tasks also applied by the teacher. DESK 3-6 includes developmental fields such are: fine and gross motor skills, linguistic and cognitive skills as well as social skills. The above are divided in four groups of tasks (factors). Structural validity of the four factors was checked by the subjectivity and reliability of the DESK 3-6 measurements (Troster et al., 2004). In order to check subjectivity for DESK 3-6 the criterion of matching assessments among teachers, while for reliability coefficient alpha Cronbach's and test-retest were used.

To assess reliability there are different ways such as internal consistency methods, test-retest, and Cronbach alpha (Anastasi & Urbina, 1997; Cortina, 1993).

Internal consistency reliability: This method contributes to the check of the reliability of the instrument by estimating how well the items that reflect the same construct yield similar results. There is a wide variety of internal consistency measures that can be used, such are for example the intercorrelations of items within a scale (average inter-item correlation) and the correlations with item-to-scale (average item-total correlation) (DeVellis, 2003). The average inter-item correlation compares correlations between all pairs of questions that test the same construct by calculating the mean of all paired correlations, while average item-total correlation takes the average inter-item correlations and calculates a total score for each item, then averages these (Priest, McColl, Thomas, & Bond, 1995).

Test-retest reliability: The Standards for Educational and Psychological Testing state that test-retest reliability is 'a reliability coefficient obtained by administering the same test a second time to the same group after a time interval and correlating the two sets of scores' (AERA, 1999). In order to measure the test-retest reliability, we have to give the same test to the same test respondents on two separate occasions. We can refer to the first time the test is given as T1 and the second time that the test is given as T2. The scores on the two occasions are then correlated. This correlation is known as the test-retest-reliability coefficient or the coefficient of stability. Stability is an aspect of reliability and many researchers report that a highly reliable test indicates that the test is stable over time (AERA, 1999).

Cronbach alpha: Cronbach's alpha measures the internal consistency of a group of items by measuring the homogeneity of the group of items—"it is an indication of how well the different items complement each other in their measurement of different aspects of the same variable or quality" (Litwin, 2003, p. 22). Cronbach's alpha ranges in value between zero and one. Cronbach's statements (1947, 1951) about reliability, suggest that the reliability of a multidimensional measure can only be estimated by correlating scores on parallel forms of a test that each represent the same factor structure.

As already mentioned for the estimation of structural validity of DESK 3-6 Cronbach alpha was used to check reliability of measurements. Cronbach alpha though, is not the only estimate of reliability (Cortina, 1993). The particular estimate of reliability that one may use depends on the particular error-producing factor that one seeks to identify (Cronbach, Gieser, Nanda, & Rajaratnam, 1972). The importance of the present study is found within the use of other methods of testing the reliability of measures of the DESK 3-6.

The purpose of the present study is to further investigate the reliability of DESK 3-6 and moreover the edition for 3-years old children, through the use of methods such are internal consistency reliability, test-retest and Cronbach's alpha. Additionally, the present study will make an effort to examine the reliability of this instrument at a different sample than that initially used.

Method

Participants

Participants were 383 preschool children (207 boys and 176 girls) ages ranging from 36 to 47 months (M = 42.93, SD = 3.17). These children were recruited from 22 private and public preschool classes. Initially classes were selected and the relevant license was acquired by the Ministry of Education for the conducting of measurements. The next step was to ask for the parents' consent for the participation of their children at the study.

Instrument

DESK 3-6 for 3 years old children (Tröster et al., 2004) was used. Standardized back-translation procedures were used to develop a Greek version of the DESK using three independent bilingual translators (Brislin, 1986). The back-translation procedure was repeated iteratively until the original and back-translated German versions of the questionnaires were identical. DESK includes four developmental fields: *fine motor skills* including screening tests (10 items) for fine hand motor skills that identify the coordination of eyes-hands and hand skills, *gross motor skills* including screening tests (10 items) for body-balance coordination, *linguistic and cognitive skills* including screening tests (15 items) for the development level of speech and cognition and *social skills* including tests (10 items) for the ability of child to deal with everyday issues with no help as well as the kid follows social rules. Screening should be conducted by school teachers. In the present study school teachers were trained to perform developmental screening.

Data analysis

For the test of internal consistency reliability of the DESK 3-6 for 3 years old children, correlations (average inter-item and item-total correlation) were used. Correlation coefficients whose magnitude was between .9 and 1.0 indicate variables that can be considered as very highly correlated. Correlation coefficients whose magnitude was between .7 and .9 indicate variables that can be considered as highly correlated. Correlation coefficients whose magnitude was between .5 and .7 indicate variables that can be considered as moderately correlated. Correlation coefficients whose magnitude was between .3 and .5 indicate variables that have a low correlation. Correlation coefficients whose magnitude was less than .3 have little if any (linear) correlation. We can readily see that $.9 < |r| < 1.0$ corresponds with $.81 < r^2 < 1.00$; $.7 < |r| < .9$ corresponds with $.49 < r^2 < .81$; $.5 < |r| < .7$ corresponds with $.25 < r^2 < .49$; $.3 < |r| < .5$ corresponds with $.09 < r^2 < .25$; and $.0 < |r| < .3$ corresponds with $.0 < r^2 < .09$. The r^2 provides a measure of how well observed outcomes are replicated by the model, as the proportion of total variation of outcomes explained by the model (Draper & Smith, 1998; Glantz & Slinker, 1990; Steel & Torrie, 1960). Regarding homogeneity of each group of items Cronbach alpha coefficient was used, while for the control of stability of measure over time the test-retest-reliability coefficient was used.

Discussion

In the present study the reliability of a version of DESK 3-6 for 3-years old children was examined. For this reason a series of methods such are internal consistency reliability, Cronbach's alpha and test-retest were used.

The findings of the present study on the internal consistency reliability revealed that the inter-item and corrected item-total correlations were low to moderate. A large number of items in all four scales showed lower scores than those considered as adequate (Kline, 1986). This means that while internal consistency is certainly necessary-but not sufficient- for homogeneity (Schmitt, 1996), the specific items with low correlation do not reflect the same construct to other items of the corresponding scale (Kline, 1979). Beavers, Lounsbury, Richards, Huck, Skolits, and Esquivel (2013) report that if an item is not significantly correlated to any of the factors (generally considered to be less than .30) and does not provide a conceptually vital dimension to the measure, the item should be removed. However, Hayes, Nelson, and Jarrett (1987, p. 972) supported that "a measure could readily have treatment utility without internal consistency… high internal consistency should not necessarily be expected."

In contrast, the results of the analyses in this study showed very good internal reliability (high coefficients alpha and test-retest). Similar results for Cronbach alpha are reported in the manual for DESK 3-6 (Tröster et al., 2004). Nevertheless, the finding related to high reliability in the present study was linked to low items' homogeneity. This is a finding enhanced by findings of other researchers that have shown either high or low item homogeneity that can be associated to

either high or low reliability despite classical itemetric opinion (Allen & Potkay, 1983; Lachar & Wirt, 1981; McDonald, 1981). High coefficients alpha does not reflect the degree that a scale is homogeneous (e.g., Cortina, 1993; Green, Lissitz, & Mulaik, 1977; Miller, 1995; Schmitt, 1996). Even though coefficient alpha is sensitive to the internal consistency of a scale, the revealing of high alpha coefficients in the present study contrary to internal consistency coefficients, is probably due to the affect of the large number of items in the each scale (Cortina, 1993; Yang & Green, 2011).

The findings of the present study show that the original version of DESK 3-6 for 3-years old children appears to be problematic in internal consistency, that is homogeneity among items of each scale. Cattell (1973, 1978, 1982) has argued that generally there is an optimally *low* level of item homogeneity. He also provided a conceptual demonstration of high item validity in the context of zero item homogeneity. Another conclusion is the sufficient degree of instrument reliability of DESK 3-6 for 3-years old coming from the high Cronbach alpha and test-retest coefficients.

Results

Fine motor skill

Fine motor skill was assessed by eleven items (FM1 to FM10). The results of the inter-item correlation indicated low to medium relation among items (Table 1). An average inter-item correlation of .30 or higher indicates acceptable reliability (Robinson, Shaver, & Wrightsman, 1991). Tabachnick and Fidell (2001) suggest that correlations exceeding .30 provide enough evidence to indicate that there is enough commonality to justify comprising factors. Low to medium were also the values of the squared multiple correlation on almost all items (Table 2).

Table 1: Average Inter-Item Correlation for Fine Motor Skill Children 3-year

Variable	FM1	FM2	FM3	FM4	FM5	FM6	FM7	FM8	FM9
FM1	-	-	-	-	-	-	-	-	-
FM2	.58	-	-	-	-	-	-	-	-
FM3	.49	.53	-	-	-	-	-	-	-
FM4	.12	.07	.15	-	-	-	-	-	-
FM5	.36	.28	.23	.09	-	-	-	-	-
FM6	.22	.18	.21	.11	.22	-	-	-	-
FM7	.21	.23	.27	.07	.28	.29	-	-	-
FM8	.35	.29	.35	.19	.20	.15	.25	-	-
FM9	.33	.32	.34	.10	.17	.17	.18	.31	-
FM10	.32	.33	.37	.01	.18	.15	.25	.25	.19

Note: $p < .01$

Table 2: Item-Total Statistics for Fine Motor Skill Children 3-year

Variable	Item-Total Correlations	Squared Multiple Correlations (r^2)	Cronbach's Alpha* if Item Deleted
FM1	.590	.426	.725
FM2	.576	.431	.729
FM3	.601	.392	.724
FM4	.167	.056	.777
FM5	.370	.159	.758
FM6	.315	.131	.763
FM7	.382	.191	.756
FM8	.463	.231	.746
FM9	.419	.193	.752
FM10	.406	.198	.753

Note: * Cronbach's alpha = .769; Standardised Cronbach's alpha = .760

Moreover, the item-total correlation coefficient values were examined and found to alter from .167 to .601. Values for an item-total correlation (point-biserial) between 0 and .19 may indicate that the question is not discriminating well, values between .2 and .39 indicate good discrimination, and values .4 and above indicate very good discrimination. In item analysis, in order to protect the summability aspect of the scale, it has to be higher than .30 (Kline, 1993), or at least .40 (Gliem & Gliem, 2003).

Cronbach's alpha and standardized Cronbach's alpha were .769 and .760 respectively showing acceptable alpha coefficient ($\alpha > .70$; Kline, 2005). Details of reliability statistics are shown in table 2. Test-retest reliability for the stability of measured scores over time was used. For this, 30 children were used who completed the DESK 4 and were then retested 2 weeks later. The subscale demonstrated adequate stability with test–retest coefficient of .91 score.

Gross motor skills

Gross motor skill was assessed by eleven items (GM1 to GM10). The results of the inter-item correlation, just as in fine motor skills, indicated low correlation among items (Table 3). Low were also the values of the squared multiple correlation on almost all items (Table 4). The item-total correlation coefficient values were found to alter from .307 to .598. Cronbach's alpha and standardized Cronbach's alpha were .786 and .801 respectively showing acceptable alpha coefficient. Details of the reliability statistics are shown in table 4. Test–retest coefficient was .90 proving adequate stability.

Table 3: Average Inter-Item Correlation for Gross Motor Skill Children 3-year

Variable	GM1	GM2	GM3	GM4	GM5	GM6	GM7	GM8	GM9
GM1	-	-	-	-	-	-	-	-	-
GM2	.31	-	-	-	-	-	-	-	-
GM3	.42	.42	-	-	-	-	-	-	-
GM4	.47	.27	.41	-	-	-	-	-	-
GM5	.42	.27	.30	.36	-	-	-	-	-
GM6	.22	.25	.22	.12	.10	-	-	-	-
GM7	.41	.27	.33	.42	.38	.20	-	-	-
GM8	.34	.18	.20	.24	.29	.09	.29	-	-
GM9	.24	.34	.30	.20	.10	.27	.17	.11	-
GM10	.43	.34	.28	.44	.38	.21	.34	.13	.36

Note: $p < .01$

Table 4: Item-Total Statistics for Gross Motor Skill Children 3-year

Variable	Item-Total Correlations	Squared Multiple Correlations (r^2)	Cronbach's Alpha* if Item Deleted
GM1	.598	.400	.764
GM2	.492	.271	.765
GM3	.557	.334	.755
GM4	.525	.358	.760
GM5	.453	.296	.769
GM6	.307	.130	.789
GM7	.510	.305	.761
GM8	.318	.170	.783
GM9	.387	.223	.778
GM10	.564	.375	.755

Note: * Cronbach's alpha = .786; Standardised Cronbach's alpha = .801

Cognitive skills - linguistic

Cognitive and linguistic skills were measured by twelve items SK1 to SK15. As in previous scales the inter-item correlation among items of this scale was low to medium (Table 5). Low to

medium were also the values of the squared multiple correlation on almost all items (Table 6). The item-total correlation coefficient values were found to alter from .380 to .675. Cronbach's alpha and standardized Cronbach's alpha were .873 and .877 respectively showing acceptable alpha coefficient. Details of the reliability statistics are shown in table 6. Test–retest coefficient was .91 showing adequate stability.

Table 5: Average Inter-Item Correlation for Cognitive Skills-Linguistic Children 3-year

Variable	SK 1	SK 2	SK 3	SK 4	SK 5	SK 6	SK 7	SK 8	SK 9	SK10	SK11	SK12	SK13	SK14
SK1	-	-	-	-	-	-	-	-	-	-	-	-	-	
SK2	.39	-	-	-	-	-	-	-	-	-	-	-	-	
SK3	.39	.56	-	-	-	-	-	-	-	-	-	-	-	
SK4	.26	.28	.26	-	-	-	-	-	-	-	-	-	-	
SK5	.36	.37	.36	.38	-	-	-	-	-	-	-	-	-	
SK6	.38	.50	.50	.26	.31	-	-	-	-	-	-	-	-	
SK7	.37	.57	.55	.22	.30	.54	-	-	-	-	-	-	-	
SK8	.19	.21	.20	.17	.26	.18	.16	-	-	-	-	-	-	
SK9	.44	.54	.64	.19	.33	.48	.57	.18	-	-	-	-	-	
SK10	.30	.39	.36	.39	.38	.22	.38	.25	.35	-	-	-	-	
SK11	.35	.28	.36	.34	,39	.30	.34	.27	.32	.33	-	-	-	
SK12	.37	.40	.30	.19	.18	.26	.35	.16	.35	.24	.23	-	-	
SK13	.34	.50	.51	.18	.29	.55	.60	.19	.46	.34	.33	.31	-	
SK14	.25	.29	.27	.13	.22	.18	.27	.77	.21	.27	.30	.19	.27	-
SK15	.29	.30	.27	.26	.31	.28	.22	.16	.30	.26	.31	.21	.24	.51

Note: $p < .01$

Table 6: Item-Total Statistics for Cognitive Skills-Linguistic Children 3-year

Variable	Item-Total Correlations	Squared Multiple Correlations (r^2)	Cronbach's Alpha* if Item Deleted
SK1	.547	.334	.865
SK2	.675	.511	.858
SK3	.667	.535	.859
SK4	.380	.244	.872
SK5	.502	.344	.867
SK6	.591	.460	.862
SK7	.667	.545	.858
SK8	.387	.627	.871
SK9	.647	.535	.859
SK10	.502	.303	.866
SK11	.503	.316	.867
SK12	.435	.241	.873
SK13	.624	.478	.860
SK14	.431	.646	.870
SK15	.407	.207	.871

Note: * Cronbach's alpha = .873; Standardised Cronbach's alpha = .877

Social skills

Social skills were measured by eleven items SE1 to SE10. Similar to the results of all other scales were for this one as well. Specifically, the inter-item correlation among items was low to medium (Table 7). Low to medium were also the values of the squared multiple correlation on almost all items (Table 8). The item-total correlation coefficient values were found to alter from .111 to .555. Cronbach's alpha and standardized Cronbach's alpha were .734 and .757 respectively,

showing acceptable alpha coefficient. Details of the reliability statistics are shown in table 8. Test–retest coefficient was .97 showing adequate stability.

Table 7: Average Inter-Item Correlation for Social Skills 3-year

Variable	SE1	SE2	SE3	SE4	SE5	SE6	SE7	SE8	SE9
SE1	-	-	-	-	-	-	-	-	-
SE2	.31	-	-	-	-	-	-	-	-
SE3	.33	.33	-	-	-	-	-	-	-
SE4	.31	.07	.36	-	-	-	-	-	-
SE5	.09	.39	.21	.18	-	-	-	-	-
SE6	.28	.12	.32	.39	.20	-	-	-	-
SE7	.21	.29	.34	.32	.29	.43	-	-	-
SE8	.06	.02	.08	.04	.06	.12	.11	-	-
SE9	.57	.34	.36	.25	.20	.31	.26	.05	-
SE10	.27	.20	.41	.26	.12	.24	.22	.05	.46

Note: $p < .01$

Table 8: Item-Total Statistics for Social Skills 3-year

Variable	Item-Total Correlations	Squared Multiple Correlations (r^2)	Cronbach's Alpha* if Item Deleted
SE1	.443	.371	.709
SE2	.412	.301	.714
SE3	.533	.337	.715
SE4	.399	.253	.720
SE5	.342	.207	.732
SE6	.464	.308	.706
SE7	.498	.297	.699
SE8	.111	.022	.751
SE9	.555	.457	.692
SE10	.409	.286	.716

Note: * Cronbach's alpha = .734; Standardised Cronbach's alpha = .757

Conclusion

In conclusion it can be stated that DESK 3-6 for 3-years old children appears to carry adequate data to justify the existence of its possible use to another sample of population than the one initially used, such is the Greek population. According to Kline (1986, p. 3) "maximum validity…is obtained where test items do not all correlate with each other, but where each correlates positively with the criterion. Such a test would have only low internal-consistency reliability."

Finally, it is suggested for future studies to further examine item homogeneity of DESK 3-6 for 3-years old children. Factor analysis is one statistical technique that can be used to determine the constructs or domains within the developing measure and contribute to establish construct validity (Boyle, 1991; Tate, 2003).

References:

1. Allen, B. P., & Potkay, C. R. (1983). Just as arbitrary as ever: Comments on Zuckerman's rejoinder. *Journal of Personality and Social Psychology, 44*, 1087-1089. doi: 10.1037/0022-3514.44.5.1087

2. American Education Research Association APA, and The National Council on Measurement in Education. 1999. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

3. Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). New Jersey: Prentice-Hall International.

4. Beavers, A. S., Lounsbury, J.W., Richards, J. K., Huck, S.W. Skolits, G. J., & Esquivel, S.L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation, 18*(6), 1-13. Retrieved from http://pareonline.net/getvn.asp?v=18&n=6

5. Brislin, R.W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.) *Field methods in educational research* (pp. 137-164). Newbury Park, CA: Sage.

6. Cattell, R. B. 1973. *Personality and mood by questionnaire*. San Francisco, CA: Jossey-Bass.

7. Cattell, R. B. (1978). *Scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

8. Cattell, R. B. (1982). The psychometry of objective motivation measurement: A response to the critique of Cooper and Kline. *British Journal of Educational Psychology, 52*, 234-241. doi: 10.1111/j.2044-8279.1982.tb00831.x

9. Chakrabarti, S., & Fombonne, E. (2001). Pervasive Developmental Disorders in Preschool Children. *Journal of the American Medical Association, 285*(24), 3093-3099.

10. Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104. doi: 10.1037/0021-9010.78.1.98

11. Cronbach, L. J. (1947). Test reliability: Its meaning and determination. *Psychometrika, 12*, 1-16. doi: 10.1007/BF02289289

12. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. doi: 10.1007/BF02310555

13. Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: theory of generalizability for scores and profiles*. New York: Wiley.

14. DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

15. Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis*. New York: Wiley.

16. Fombonne, E., Zakarian, R., Bennett, A., Meng, L., & McLean-Heywood, D. (2006). Pervasive developmental disorders in Montreal, Quebec, Canada: Prevalence and links with immunizations. *Petiatrics, 118*(1), 139-150. doi: 10.1542/peds.2005-2993

17. Glantz, S. A., & Slinker, B. K. (1990). *Primer of Applied Regression and Analysis of Variance*. New York: McGraw-Hill.

18. Gliem, J. A., & Gliem, R. R. (2003). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Midwest Research to Practice Conference in Adult, Continuing, and Community Education, retrieved January 10, 2008 from http://www.alumni-osu.org/midwest/midwest%20papers/Gliem%20&%20Gliem-Done.pdf.

19. Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*, 827-838. doi: 10.1177/001316447703700403

20. Hayes, S. C., Nelson, R. O., & Jarrett, J. B. (1987). The treatment utility of assessment: a functional approach to evaluating assessment quality. *American Psychologist, 42*, 963-974. doi: 10.1037/0003-066X.42.11.963

21. Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. New York: Methuen.

22. Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.

23. Litwin, M. S. (2003). *How to assess and interpret survey psychometrics* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

24. McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34*, 110-117. doi: 10.1111/j.2044-8317.1981.tb00621.x

25. Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling, 2*, 255-273. doi: 10.1080/10705519509540013

26. Priest, J., McColl, B. A., Thomas, L., & Bond, S. (1995). Developing and refining a new measurement tool. *Nurse Researcher, 2*, 69–81.

27. Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). *Measures of personality and social psychological attitudes*. San Diego, CA: Academic Press.

28. Schmitt, N. (1996). Uses and abuses of coefficient alpha". *Psychological Assessment, 8*(4), 350-353. doi: 10.1037/1040-3590.8.4.350

29. Steel, R. G. D., & Torrie, J. H. (1960). *Principles and Procedures of Statistics with Special Reference to the Biological Sciences*. New York: McGraw Hill Book Company.

30. Tabachnick, B. G., &Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

31. Tröster, H., Flender, J., & Reinke, D. (2005). Dortmunder Entwicklungsscreening für den Kindergarten (DESK 3-6) [Dortmunder Development Screening for Kindergarten (DESK 3-6)]. Bern: Göttingen.

32. Yang, Y., & Green, S. B. (2011). Coefficient Alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment, 29*(4), 377-392. doi: 10.1177/0734282911406668