Kuram ve Uygulamada Eğitim Bilimleri • Educational Sciences: Theory & Practice - 11(1) • Winter • 330-334

Instruments Used in Doctoral Dissertations in Educational Sciences in Turkey: Quality of Research and Analytical Errors

Engin KARADAĞ [a]
Eskişehir Osmangazi University

Abstract

The aim of this study was to define the level of quality and types of analytical errors for measurement instruments used [i.e., interview forms, achievement tests and scales] in doctoral dissertations produced in educational sciences in Turkey. The study was designed to determine the levels of factors concerning quality in research methods and the case study model was used. Theoretical universe for the study was 324 doctoral dissertations in educational sciences in Turkey from 2003 to 2007. Sampling group was consisted of 211 doctoral dissertations accessed through online in the National Thesis Center. In order to collect the data, an evaluation form was developed by the researcher and the data analysis method was epistemological document analysis. In the analysis process, frequencies, descriptive statistics, and typology analysis techniques were used. The findings indicate that the properties of measurement tools used in dissertations in educational sciences were absent and that the most common analytical mistake was the absence of validity.

Key Words

Research in Educational Sciences, Measurement Instrument, and Design Errors.

The number of educational research concerning the educational system has gained an important role in the decade. While some of the published studies have created a basis for educational reforms, another portion has tested the results of previous research and the reliability via literature review (Balcı & Apaydın, 2009; Odom et al., 2005; Onwuegbuzie & Daniel, 2003). In parallel to this growth in quantity, inquiry process of educational research, the results reached through the conduct of research and the availability of high quality works are quite important. When these studies were examined, some authors' findings included misconceptions, contrary to the reality, opposite findings were expressed. This study provides an important and necessary synthesis of studies (Dunkin, 1996).

a *Correspondence:* Assoc. Prof. Engin KARADAĞ, Eskisehir Osmangazi University, College of Education, Department of Elementary Education, 26480 Eskişehir / Turkey. E-mail: enginkaradag@ogu.edu.tr & engin.karadag@hotmail.com. Phone: +90 222 339 37 50 / 1644 Fax: +90 222 229 31 24.

Kieffer, Reese, and Thompson (2001) determined that variance and covariance analyses, regression analyses, and correlation analyses had been frequently used in 756 articles issued in Journal of Counseling Psychology (JCP) and American Educational Research Journal (AERJ) between 1988 and 1997. In addition to American oriented journal analysis, in Onwuegbuzie's (2002) study published in British Journal of Education Psychology (BJEP) in 1998, the most frequently used analyses procedures were variance, covariance, and factor analysis.

Purpose

The purpose of this study was to examine the quality of research, measurement instruments and the types of analytical errors in doctoral dissertations in educational sciences produced in the Turkish universities.

Method

Research Design

While conducting the research, which aimed to determine the measurement instruments used and the analytical errors in the doctoral dissertations in educational sciences between the years of 2003 and 2007, the case study design had been used in addition to the other qualitative research designs (Creswell, 1998; Merriam, 1998, Yıldırım & Şimşek, 2005).

Universe and Sampling

The theoretical universe of this research was the doctoral dissertations produced in Turkey in education sciences. Yet, the theoretical universe to study, which was identified by taking into consideration the improvements in methodology and whether they were up to date, included 324 doctoral dissertations education sciences between the years of 2003 and 2007 (Yüksek Öğretim Kurulu [YÖK], 2008). The distribution of doctoral dissertations when the year was taken as the unit of analysis was as follows: 2003 (*n*=43, 13.2%), 2004 (*n*=50, 15.4%), 2005 (*n*=50, 15.4%), 2006 (*n*=84, 25. 9%), and 2007 (*n*=97, 29.4%). In the research, a sampling was not used since the researcher was able to reach all the dissertations except 211 dissertations due to the restrictions of usage. As a result, the examined distribution of doctoral dissertations based on years was as follows: 2003 (*n*=6, 2.8%), 2004 (*n*=7, 3.3%), 2005 (*n*=30, 14.2%), 2006 (*n*=79, 37.4%), and 2007 (*n*=89, 42.1%).

Instrument

The Education Research Evaluation Form: The form developed within the scope of this research was prepared with the intention to determine the methodological quality of the studies conducted in education (Martuza, 1977). The form was also assessed for content validity by 10 members in educational sciences (Lawshe, 1975). Lawshe (Vickery, 1998) computed a content validity ratio. The results of content validity were evaluated by the experts who work in the field of educational sciences and it was found between 9.50 and 10.00. [X=9.94; Median=10.00, SD=0.14]. The content validation rates were .80 and 1.00 for the 50 item-form. The expressions used in the form were evaluated on a horizontal line ranging from *completely* (*10*) to *none* (*0)* with the help of 11 point Likert-type grading scale. A high point received on the basis of sub-scales of the form indicates the efficiency level of the variable that the sub-scale stands for.

Process

In this study, the epistemological document analysis was used as the data collection method (*i*) accessing the documents which were the first stage of document analysis; the doctoral dissertations in the pre-identified sampling group had been obtained from the YOK (the Turkish Higher Education Council) Documentation Unit. The identified dissertations had been downloaded from the web-site of the YOK Documentation Unit to the computer and coded. In the next stage, (*ii)* the downloaded dissertations were analyzed according to the *educational research evaluation scale* which was organized in the *Likert type scaling* (Forster, 1995; Rowlinson, 2004). In this context, for the objectives related to the sampling method dimensions and quality levels of the research, *frequency analysis* among the qualitative data analysis types and content analysis types (Bilgin, 2006; Köhler & Stemmler, 1997; Lienert & Oeveste, 1985; Martinmäki & Rusko, 2008) was used. In the analysis of the analytical errors, *descriptive analysis* (Kümbetoğlu, 2005) and *typological analysis* (Dey, 2007; Mayring, 2000) among the qualitative data analysis were used. In the research, the mean of the findings obtained from frequency analysis (*X*) and standard deviation (*SD*) values were presented. The descriptive analysis was used in the research consisted of four stages.

Findings

The measurement instruments used in doctoral theses in educational sciences and their percentages were as follows: (1) *scale* (*n*=163, %63.1) and (2) *interview form* (*n*=54, %20.9), (3) *achievement test* (*n*=32, %12.4) and (4) *observation form* (*n*=9, %12.4).

The general quality levels of measurement instruments vary from 0.68 to 5.57 based on the item. The total average point of measuring instruments was calculated as 2.18 [SD=1.46, Median=1.86]. The quality levels of interview forms vary from 2.51 to 5.34 based on the item. The total average point of interview forms was calculated as 3.85 [SD=2.08, Median=3.67]. The quality levels of achievement tests vary from 1.19 to 4.97 based on the item. The total average point of achievement tests was calculated as 2.39 [SD=1.28, Median=2.20].

The quality levels of measurement instruments in doctoral theses were insufficient and vary from 0.05 to 7.56 based on the item. The total average point of measurement instruments in doctoral theses was insufficient and it was calculated as 3.63

[SD=1.27, Median=3.55]. The quality levels of validity property which scales must have was insufficient and varied from 0.37 to 4.74 based on the item. The total average point of validity property which scales must have was insufficient and it was calculated as 2.39 [SD=2.13, Median=2.00]. The quality levels of reliability property which scales must have was insufficient and varied from 1.14 to 6.82 based on the item. The total average point of reliability property which scales must have was insufficient and it was calculated as 4.21 [SD=1.86, Median=4.02]

The quality levels of scale adaptations were also insufficient and varied from 0.30 to 1.34 based on the item. The total average point of scale adaptations was also insufficient and it was calculated as 0.76 [SD=1.82, Median=0.00].

Discussion

The level of measurement instruments used in doctoral theses in educational sciences by means of general properties were normally insufficient while the info was not provided regarding the properties of the measurement instruments in the qualitative studies which was considered as an error (Punch, 2005; Neuman, 2007). These finding were similar to the research findings of Onwuegbuzie (2002), Stevenson (2000), and West, Carmody and Stallings (1983).

The quality level of *interview forms* used as measurement instrument in doctoral theses was insufficient. The most important error in this dimension was the lack of pilot studies or lack of explanations of pilot test results in the theses where pilot study has been conducted (Mason, 1996; Patton, 2002).

The quality level of *achievement tests* used as other measurement instruments in doctoral theses was insufficient. This result of the research was in correspondence with some of the earlier research findings [*see*: Hall, 1986; Kırcaali-İftar, 1999; Onwuegbuzie, 2002].

Another widely used measurement instrument in doctoral theses was scales. The quality level of scales used as measurement instruments in doctoral theses was insufficient. In the qualitative studies, 5 different mistakes were detected. (*i*) article numbers were not being presented, (*ii*) psychometric properties of the scale were not being presented (Dawson-Saunders & Trapp-Robert, 1994; Gronlund & Linn, 1990), (*iii*) presence of mistakes in determination of limit values in behavior scales (Morris, 2002; Turgut & Baykul, 1992), (*iv*) number of articles were insufficient (Bryman & Cramer, 1997, Büyüköztürk, 2002; Kline, 1994; Mertens, 1998; Tosun & Karadağ, 2008) and (*v*) negative and positive items quantities were not equal in behavior scales.

The quality level of validity property which scales must have was insufficient; according to the qualitative resolution, 8 different mistakes were detected. (*i*) Dividing scale to factors without making factor analysis [see: Balcı, 2007; Bryman & Cramer, 1997; Deniz, 2007; Dennis & Winston, 2003; Kangwa & Olubodun, 2003; Wang & Ahmed, 2004], (*ii*) not giving sufficient info about factor analysis results, (*iii*) presence of mistakes regarding article factor loadings (Şencan, 2005), (*iv*) Elimination of too much articles (items) in content validity (McMillan & Schumacher, 2006), (*v*) only content validity carried out in the development of scale (Young, 1996), (*vi*) Presence of unnecessary info in the content validity study (Şencan, 2005; Vickery, 1998), (*vii*) making factor analysis separately for scale dimensions. (*viii*) The findings of the research on the same subject 20 years ago showing the weakness of scale validity levels proved that the problem has still continued today in Turkey [*see:* Baykul, 2000; Chapman, 1988; Emmons, 1988; Hersom, 1980; Uysal, 1971; Vockell & Asher, 1974; Ward, Hall, & Schramm, 1975].

The quality level of reliability properties which was another property of the scales had been also insufficient; according to the qualitative resolution, 4 types of different mistakes were detected. (*i*) Presence of mistakes in excuses of chosen confidence tests. (*ii*) Presence of mistakes in the detection of security confidence levels. (*iii*) Lack of information provided confidence parameters (Cohen, Manion, & Morrison, 2005). (*iv*) Deleting item after reliability operations of the scale had been made (Appleton, 1995; Carmines & Zeller, 1982; Chen, 2003; Henson, 2001).

Although new scales were developed in most of the theses, it was detected that 26 scales in English were adapted to Turkish. The quality level of scale adaptations was also insufficient; based on the qualitative resolution, 3 different mistakes were detected (*i*) lack of information regarding adaptation. (*ii*) Adaptation being only limited by translation (Baloğlu & Karadağ, 2008). (*iii*) Adaptation being made in groups which had different properties than the research subject. The main problem in scale adaptation was that many verbs do not have Turkish translations. For example, *like, enjoy,* or *love are translated into Turkish as "sevmek"* (Gülgöz, 2005). Besides, the fact that a word not

being present in a language does not prove that the concept was not present in that culture had always been a matter of discussion (Hambleton & De Jong, 2003; Whorf, 1956). The results of Büyüköztürk and Kutlu (2006) on the same subject support the results of this research and the comments that have been made.

References/Kaynakça

Appleton, J. V. (1995). Analysing qualitative interview data: Adressing issues of validity and reliability. *Journal of Advanced Nursing*, *22*, 993-997.

Balcı, A. (2007). *Sosyal bilimlerde araştırma: Yöntem, teknik ve ilkeler.* Ankara: Pegema Yayıncılık.

Balcı, A. ve Apaydın, Ç. (2009). Türkiye'de eğitim yönetimi araştırmalarının durumu: Kuram ve Uygulamada Eğitim Yönetimi Dergisi örneği. *Kuram ve Uygulamada Eğitim Yönetimi, 15* (59), 325-343.

Baloğlu, N. ve Karadağ, E. (2008). Öğretmen yetkinliğinin tarihsel gelişimi ve Ohio öğretmen yetkinlik ölçeği: Türk kültürüne uyarlama, dil geçerliği ve faktör yapısının incelenmesi. *Kuram ve Uygulamada Eğitim Yönetimi Dergisi*, *56*, 571-606.

Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klâsik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.

Bilgin, N. (2006). *Sosyal bilimlerde içerik analizi -Teknikler ve örnek çalışmalar-*. Ankara: Siyasal Kitapevi.

Bryman, A., & Cramer, D. (1997). *Quantitative daha analysis with SPSS for Windows: A guide for social scientists*. New York: Routledge.

Büyüköztürk, Ş. (2002). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegema Yayıncılık.

Büyüköztürk, Ş. ve Kutlu, Ö. (2006). Sosyal bilim araştırmalarında yöntem sorunu. K. Karakütük (Ed.), *Sosyal Bilimlerde Süreli Yayıncılık I. Ulusal Kurultay Bildirileri Kitabı* içinde (s.113-122). Ankara: TÜBİTAK Yayınları.

Carmines, E. G., & Zeller, R. A. (1982). *Reliability and validity assessment*. Beverly Hills: Sage Publications Inc.

Chapman, J. (1988). Research and reflection: Thirty years of the "Australian Journal of Education". *Australian Journal of Education*, *32* (3), 259-273.

Chen, L. (2003). Examination of reliability and validity of scale of coaching performance (SCP) using NCAA sample. *Journal of Measurement in Physical Education and Exercise Science, 7* (3), 175-197.

Cohen, L., Manion, L., & Morrison, K. (2005). *Research methods in education*. New York: RoutledgeFalmer.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Dawson-Saunders, B., & Trapp-Robert, G. (1994). *Basic & clinical biostatistics*. London: Appleton and Lange.

Deniz, K. Z. (2007). Psikolojik ölçme aracı uyarlama. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 40* (1), 1-16.

Dennis, R., & Winston, B. (2003). A factor analysis of Page and Wong's servant leadership instrument. *Leadership & Organization Development Journal*, *24* (8), 455-459.

Dey, I. (2007). *Qualitative data analysis: A user-friendly guide for social scientists*. London: Routledge Taylor & Francis Group.

Dunkin, M. J. (1996). Types of errors in synthesizing research in education. *Review of Educational Research, 66* (2), 87-97.

Emmons, N. J. (1988). *Statistical methods used in "American Educational Research Journal", "Journal of Educational Psychology", and "Sociology of Education" from 1972 through 1987* (ERIC Document Reproduction Service No. ED319797).

Forster, N. (1995). The analysis of company documentation. C. Cassell & G. Symon (Eds.), *Qualitative methods in organizational research: A practical guide* (pp. 83-106). London: Sage.

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching.* New York: Mac Millian Publishing.

Gülgöz, S. (2005). Five factor theory and NEO-PI-R in Turkey. In J. Allik & R. R. McCrae (Eds.), *The five-factor model of personality across cultures* (pp. 175-196). Dordrecht, Netherlands: Kluwer Academic Publishers.

Hall, B. W. (1986). Validity, reliability, and norms of popular versus less popular published educational achievement tests. *Journal of Educational Research*, *79* (3), 145-150

Hambleton, R. K., & De Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing*, *20* (2), 127-134.

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on Coefficient Alpha. *Measurement and Evaluation in Counseling and Development*, *34* (3), 177-189.

Hersom, N. (1980). Perspectives: Twenty-five years of research in education: "The Alberta Journal of Educational Research" 1955 to 1979. *Alberta Journal of Educational Research*, *26* (4), 262-275.

Kangwa, J., & Olubodun, J. (2003). A factor approach to analysis of home maintenance outcomes and attributes of management successes in the owner-occupied sector. *Structural Survey*, *21* (4), 158-171.

Kırcaali-İftar, G. (1999). Ölçme. A. A. Bir (Ed.), *Sosyal bilimlerde araştırma yöntemleri* içinde (s.12-22). Eskişehir: Anadolu Üniversitesi Yayınları.

Kieffer, K. M., Reese, R. J., & Thompson, B. (2001). Statistical techniques employed in AERJ and JCP articles from 1988 to 1997: A methodological review. *Journal of Experimental Education, 69* (3), 280-309.

Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge.

Köhler, T., & Stemmler, M. (1997). Normative versus impassive configure frequency analysis in personality research -their use discussed in a reanalysis of data on situation- bound anxiety. *European Journal of Personality*, *11* (1), 69-79.

Kümbetoğlu, B. (2005). *Sosyolojide ve antropolojide niteliksel yöntem ve araştırma.* İstanbul: Bağlam Yayıncılık.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563-575.

Lienert, G. A., & Oeveste, H. Z. (1985). Configural frequency analysis as a statistical tool for developmental research. *Educational and Psychological Measurement*, *45* (2), 301-307.

Martinmäki, K., & Rusko, H. (2008). Time-frequency analysis of heart rate variability during immediate recovery from low and high intensity exercise. *European Journal of Applied Physiology*, *102* (3), 353-360.

Martuza, V. R. (1977). *Applying norm-referenced and criterion-referenced measurement in education*. Boston: Allyn & Bacom.

Mason, J. (1996). *Qualtitative researching*. London: Sage.

Mayring, R. (2000). *Nitel sosyal araştırmaya giriş* (çev. A. Gümüş ve M. S. Durgun). Adana: Baki Kitapevi.

McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence based inquiry.* Boston: Brown and Company.

Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Mertens, D. M. (1998). *Research methods in education and psychology: Integrating diversity with quantitative and quantities approaches*. London: Sage.

Morris, C. (2002). *Psikolojiyi anlamak* (çev. A. Erkuş, A. D. Batıgün, B. Ayvaşık) Ankara: TPD Yayınları.

Neuman, L. W. (2007). *Toplumsal araştırma yöntemleri: Nitel ve nicel yaklaşımlar* (çev. S. Özge). İstanbul: Yayın Odası Yayıncılık.

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Council for Exiceptional Children*, *71* (2), 137-148.

Onwuegbuzie, A. J. (2002). Common analytical and interpretational errors in educational research: an analysis of the 1998 volume of the British Journal of Educational Psychology. *Educational Research Quarterly, 26*, 11-22.

Onwuegbuzie, A. J., & Daniel L.G. (2003). Typology of analytical and interpretational errors in quantitative and qualitative educational research [Electronic Version]. *Current Issues in Education*, *6* (2). Retrieved December 21, 2007 from http://cie.ed.asu.edu/ volume6/number2/.

Patton, M. Q. (2002). *Qualitative evaluation and research methods*. Thousand, CA: Sage.

Punch, K. F. (2005). *Sosyal araştırmalara giriş* (çev. D. Bayrak, H. B. Arslan ve Z. Akyüz). Ankara: Siyasal Kitabevi.

Rowlinson, M. (2004). Historical analysis of company documents. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organizational research* (pp. 301-312). London: Sage.

Stevenson, J. (2000). Seven years of vocational education research: a review of material published in the "Australian and new Zealand journal of vocational education research." *Australian and New Zealand Journal of Vocational Education Research*, *8* (1), 93-133.

Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.

Tosun, Ü. ve Karadağ, E. (2008). Yapılandırmacı düşünme envanterinin Türkçeye uyarlanması dil geçerliği ve psikometrik incelemesi. *Kuram ve Uygulamada Eğitim Bilimleri, 8*, 225-264.

Turgut, M. F. ve Baykul, Y. (1992). *Ölçme teknikleri*. Ankara: ÖSYM Yayınları.

Uysal, Ş. (1971). Metodoloji açısından Türkiye'de yapılan sosyolojik araştırmalar ve bir örnek köy araştırması. N. H. Fişek (Ed.), *Türkiye'de sosyal araştırmaların gelişimi sempozyumu bildirileri kitabı* içinde (s.139-151). Ankara: Hacettepe Üniversitesi Yayınları.

Vickery, S. (1998). Let's not overlook content validity. *Decision Line*, *29* (4), 10-13.

Vockell, E. L., & Asher, W. (1974). Perceptions of document quality and use by educational decision makers and researchers. *American Educational Research Journal, 11*, 249-258.

Wang, L. C., & Ahmed, P. K. (2004). The development and validation of the organizational innovativeness construct using confirmatory factor analysis. *European Journal of Innovation Management*, *7* (4), 303-313.

Ward, A. W., Hall, B. W., & Schramm, C. E. (1975). Evaluation of published educational research: A national survey. *American Educational Research Journal, 12,* 109-128.

West, C. K., Carmody, C., & Stallings, W. M. (1983). The quality of research articles in the journal of education research, 1970 and 1980. *Journal of Educational Research*, *77* (2), 28-36.

Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge: MIT Press.

Yıldırım, A. ve Şimşek, H. (2005). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.

Young, C. A. (1996). *Validity issues in measuring psychological constructs: The case of emotional intelligence*. Retrived October 07, 2008 from http://www.socialresearchmethods.net/tutorial/Young/eiweb2.htm#validity.

Yüksek Öğretim Kurulu [YÖK] (2008). *Yüksek öğretim kurulu ulusal tez merkezi*. http://tez2.yok.gov.tr adresinden 07 Temmuz 2008 tarihinde edinilmiştir.