Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 174 (2015) 2849 – 2855

Procedia Social and Behavioral Sciences

INTE 2014

Psychometric properties of data gathering tools used in thesis

Zekeriya NARTGÜN[a], Merve ŞAHİN[b]

[a]Assoc.Prof.Dr.,Abant Izzet Baysal University, Bolu, 14280, Turkey

[b]Research Assistant, Abant Izzet Baysal University, Bolu, 14280, Turkey

Abstract

Reliability and validity of data gathering tools used in postgraduate theses are crucial to obtain error-free evaluation results and to ensure that the data gathered through these means serve their purpose. This study investigated the presented evidence regarding the reliability and validity of data gathering tools (measurement tools) used in master's and PhD theses. The population of the study undertaken via document review method was composed of a total of 111 theses obtained between the years of 2011-2014 from the Institute of Educational Sciences of a university situated in the Western Black Sea region of Turkey. 93 of these theses were master's theses whereas 18 were PhD theses. Sample of the study was composed of a total of 46 theses open to publication (39 master's theses, 7 PhD theses). Data were analyzed descriptively (%, f). Findings point to the important problems faced during proving process of reliability and validity of data gathering tools used in theses. Suggestions were provided in the light of the findings.

Peer-review under responsibility of the Sakarya University

Keywords: Data gathering tools, validity, reliability

1.Introduction

Measurement is the process of observing specific properties of individuals or objects through the use of appropriate instruments and expressing the observation results with the help of numbers or symbols (Büyüköztürk et.al., 2013; Tan, 2012; Turgut and Baykul, 2012). Each measurement process includes a characteristic to measure along with a measurement tool to assess the specified feature (Özçelik, 2011). It is necessary for measurement tools to possess specific psychometric properties in order to have close to reality, objective and pertinent measurement results. Reliability and validity are the most important psychometric properties that ought to be included in measurement tools and they cannot be substituted for one another (Güler, 2012; Atılgan, 2013).

Validity is defined as the degree to which the instrument serves its purpose or measures what it is supposed to measure without confusing it with other properties or variables (Baykul, 2010). Various methods have been developed to prove the validity of measurement tools. Face validity, content validity, construct validity, predictive

* Corresponding author. Tel.: +00 -000-0000-0000
E-mail address: sahinmerv@gmail.com

Peer-review under responsibility of the Sakarya University
doi:10.1016/j.sbspro.2015.01.978

validity, concurrent validity and validity based on expert views are the major methods used. Different validity evidences generate different assumptions. Therefore, more than one validity evidence can be used to validate the assumptions obtained in studies (Crocker and Algina, 1986).

The concept of reliability is the extent of error-free measurement of a specified characteristic by measurement tools (Atılgan, 2013). Stability, consistency and sensitivity levels of measurement tools are considered to be indicators of reliability. In this framework, several methods have been developed to identify reliability levels of measurement tools. Test-retest, parallel forms, split-half , KR-20, KR-21, Cronbach Alpha and Inter-observer agreement are the major methods used. Different evidences for reliability generates different assumptions regarding the stability, consistency and sensitivity levels of the measurement tool. Therefore, more than one reliability evidence can be used to validate the assumptions obtained in studies (Crocker and Algina, 1986).

1.1.Purpose of the Study

The main purpose of this study was the investigation of reliability and validity of data gathering tools (measurement tools) used in postgraduate theses. In this framework, the main research areas included whether reliability and validity evidences were presented and which methods were used if evidence was presented. Distribution of data gathering tools based on type and development-adaptation was also investigated.

Answers to the questions provided below were sought in line with the specified purpose:

1. What is the distribution of tools used in theses for data gathering purposes according to type?
2. What is the distribution of tools used in theses for data gathering purposes according to their development-adaptation?
3. What is the distribution of tools used in theses for data gathering purposes according to presentation of reliability and validity evidences?
4. What is the distribution of tools used in theses for data gathering purposes according to methods used to prove validity?
5. What is the distribution of tools used in theses for data gathering purposes according to methods used to prove reliability?

1.2.Significance of the Study

This study is regarded to be significant since it will guide those concerned (researchers, advisors and members of the jury) about the points to be taken into consideration in the context of reliability and validity of data gathering tools in line with the findings obtained in this study.

1.3.Limitations of the Study

This study is limited to the analysis of evidence presented in the context of basic psychometric properties (reliability and validity) of data gathering tools used in thesis writing process.

2.Method

2.1.Design of the Study

The data were obtained via document review in this study which utilized the survey model that identifies the qualitative and quantitative cases regarding the reliability and validity of data gathering tools used in master's and PhD theses. Document review includes the analysis of written materials that consist of information regarding the topics targeted for research. Documents are effective data collection instruments used in qualitative studies (Yıldırım and Şimşek, 2011).

2.2.Population and Sample

The population of the study was composed of a total of 111 theses obtained between the years of 2011-2014 from the Institute of Educational Sciences of a university situated in the Western Black Sea region of Turkey. 93 of these theses were master's theses whereas 18 were PhD theses. Sample of the study was composed of a total of

46 theses open to publication (39 master's theses, 7 doctorate theses). Random sampling method was used in sample selection. Department and disciplines related to the theses used in the sample are provided below:

- Department of Educational Sciences: Curriculum and Instruction, Educational Management and Supervision, Measurement and Evaluation, Psychological Counseling and Guidance
- Department of Elementary Education: Mathematics Education, Classroom Teaching, Social Studies Education, Pre-School Education
- Department of Fine and Arts: Music Education
- Department of Special Education: Teaching People with Mental Disabilities
- Department of Turkish Language: Turkish Education
- Department of Foreign Languages: English Education

2.3.Data Gathering Tool, Data Collection and Analysis

A *survey form* was developed by the researchers to collect data regarding the psychometric qualities of reliability and validity for data gathering tools used in the theses included in the sample. Criteria included in the survey form were identified by the researchers after reviewing the literature and taking expert views. Data were collected with this form. In order to ensure reliability of the collected data, 8 master's and 3 PhD theses were examined twice with a 20-day interval and consistency between results were examined. Frequencies and percentages were used in data analyses.

3.Findings and Interpretations

What is the distribution of tools used in theses for data gathering purposes according to type?

Table1. Distribution of Data Gathering Tools According to Type

Type of Tool	f	%
Scale	44	48.35
Test	14	15.38
Questionnaire	12	13.19
Inventory	7	7.70
Interview Form	7	7.70
Form	3	3.30
Observation Form	2	2.19
Scenario	2	2.19
Total	91	100

Examination of data presented in Table 1 shows that the majority of tools used for data gathering purposes in postgraduate theses were scales (48.35%), tests (15.38%) and questionnaires (13.19%). They were followed by inventories (7.70%), interview forms (7.70%), forms (3.30%), observation forms (2.19%) and scenarios (2.19%) respectively.

Total number of data gathering tools used in the 46 theses that were examined was 91. This number shows that more than one data gathering tools were used in some theses. The theses that utilized more than one data gathering tool were the theses that used interview forms, observation forms, scenarios and forms. Findings show that half of the theses that utilized other data gathering tools used only one data gathering tool whereas the other half of the theses used more than one data gathering tools together. Examination of the data presents that questionnaire, scale and inventory terms were used interchangeably in these theses and researchers had difficulty in naming data gathering tools. It was also identified that the same data gathering tool was termed inventory, questionnaire or scale in the same thesis. These findings are important since they show that both the students who wrote the theses and their advisors and also the members of the thesis defense jury had lack of information about the types of data gathering tools or that they did not show the required sensitivity in the process of giving feedback.

What is the distribution of tools used in theses for data gathering purposes according to their development-adaptation?

Table 2. Distribution of Data Gathering Tools according to their Development-Adaptation

Development-Adaptation	f	%
Developed by the Thesis Author	42	46.16
Developed by Other Researchers	41	45.05
Adapted	8	8.79
Total	91	100

Examination of data presented in Table 2 shows that 46.16% of the tools used in theses to gather data were developed by the theses authors, 45.05% of the tools used in these theses were developed by others and 8.79% of the tools were used through adaptation by the theses authors or others. It may be significant for the originality of the research that approximately half of the tools used for data gathering were developed by the researchers in line with their aims. On the other hand, the fact that approximately half of data gathering tools used in these theses was developed by others may be significant since it shows these theses were implemented on different sample groups in line with the same research goals. Based on these information, it can be stated that developing and adapting data gathering tools is crucial in the process of preparing a thesis. However, it was identified that the process of data gathering tool development and presentation of evidence for validity and reliability were not undertaken in a satisfactory manner as can be seen from the findings of the next research question.

What is the distribution of tools used in theses for data gathering purposes according to presentation of validity and reliability evidences?

When the data presented in Table 3 are examined as a whole, it was seen that 57.14% and 34.07% of the theses that were investigated did not provide any evidence regarding the validity and reliability of data gathering tools used in these theses respectively. As stated in the introduction part, validity and reliability are among the basic psychometric properties of data gathering tools. Validity provides information about the degree of serving the specified purpose whereas reliability gives information regarding the extent of error-free results obtained via the tool used in data collection. Therefore, proving validity and reliability of data gathering tools with appropriate methods and presenting them are crucial steps in development and adaptation process of data gathering tools. This finding is important since it displays that researchers that develop and adapt tools are inadequate in terms of knowledge, skills and sensitivity in terms of the steps they need to take in the process of data gathering tool development. This finding is crucial since it signifies the fact that research results obtained via data that cannot be proven valid or reliable should be questioned.

One of the fundamental findings regarding the data gathering tools for which reliability and validity evidences are provided is related to the tools developed especially by other researchers. Thesis authors report the evidence provided by the researchers who developed or adapted the tools but they do not study the reliability of data obtained from the sample to whom they implement the tools. This finding is also important to show the need for validation of research results obtained in these types of theses.

Table 3. Evidence for Validity and Reliability Presented for Data Gathering Tools

		Evidence for Validity (f)	Evidence for Reliability (f)
Scale (44)	Specified	19	32
	Unspecified	25	12
Test (14)	Specified	8	14
	Unspecified	6	0
Questionnaire (12)	Specified	5	6
	Unspecified	7	6
Inventory (7)	Specified	2	4
	Unspecified	5	3
Interview Form (7)	Specified	3	2
	Unspecified	4	5
Observation Form (3)	Specified	2	2
	Unspecified	1	1
Scenario (2)	Specified	0	0
	Unspecified	2	2
Form (2)	Specified	0	0
	Unspecified	2	2
Total (91)	Specified	39 (% 42.86)	60 (%65.93)
	Unspecified	52 (%57.14)	31(%34.07)

What is the distribution of tools used in theses for data gathering purposes according to methods used to prove validity?

As stated in Table 3, out of 46 theses, the number of tools for which evidence of validity was provided was 39 (42.86%) and the number of tools for which no evidence of validity was provided was 52 (57.14%). Examination of data presented in Table 4 shows that a total of 41 evidences for validity were presented for 39 tools. Distribution of methods used for evidencing validity is as follows: 14 (34.1%) content validity, 9 (21.95%) exploratory factor analysis, 9 (21.95%) validity based on expert views, 8 (19.5%) confirmatory factor analysis and 1 (2.45%) validity based on a criterion.

Table 4. Methods Used to Prove Validity of Data Gathering Tools

Validity Methods	f	%
Content Validity	14	34.15
Exploratory Factor Analysis	9	21.95
Validity based on Expert Views	9	21.95
Confirmatory Factor Analysis	8	19.5
Validity based on a Criterion	1	2.45
Total	41	100

Based on literature reviews, content validity is an often used method in proving the validity of tests. Exploratory factor analysis is mostly used in the identification of construct validity for psychological properties aimed to be generally measured through scales. It was seen that confirmatory factor analysis was the most often used validity evidence method during the adaptation process of data collection tools. Researchers mostly preferred observation and interview forms as methods for validity evidence based on expert views.

The finding that evidence for the validity of one data gathering tool was presented whereas no evidence was submitted for the others when more than one data gathering tools were used in the same thesis is one of the findings of the study that can be deemed important.

Presenting evidence for validity based on more than one method is especially desired during the process of data gathering tool development. Findings obtained in this study display that the methods used to prove validity of tools are mostly limited to only one method. In addition to the abundance of data gathering tools for which no evidence of validity was submitted (57.14%), the limitation of validity evidences with only one method may be important to show the inadequacies of the involved parties in terms of knowledge and sensitivity.

What is the distribution of tools used in theses for data gathering purposes according to methods used to prove reliability?

As stated in Table 3, out of 46 theses, the number of tools for which evidence of reliability was provided was 60 (65.93%) and the number of tools for which no evidence of reliability was provided was 31 (34,07%). Examination of data presented in Table 5 shows that a total of 64 evidences for reliability were presented for 60 tools. Distribution of methods used for evidencing reliability is as follows: 39 (60.94%) Cronbach Alpha, 10 (15.63%) KR-20, 8 (12.5%) inter-rater reliability, 6 (9.37%) test-retest and 1 (1.56%) split halves.

Investigation shows that Cronbach Alpha was used to prove the reliability of scales and KR-20 was used to prove the reliability of tests. Inter-rater reliability was the preferred method to identify the reliability of data obtained from observation and interview forms. Test-retest method was used to validate reliability of data obtained from questionnaires. The finding that evidence for reliability of one data gathering tool was presented whereas no evidence was submitted for the others when more than one data gathering tools were used in the same thesis is frequently seen.

Table 5. Methods Used to Prove Reliability of Data Gathering Tools

Reliability Methods	f	%
Cronbach Alpha	39	60.94
KR-20	10	15.63
Inter-Rater Reliability	8	12.5
Test-Retest	6	9.37
Split Halves	1	1.56
Total	64	100

When thesis authors used a tool developed by others, they reported the evidence of the researcher who developed the original tool. It is also a frequent case that thesis authors did not study the reliability of data obtained from the groups on whom the study implemented.

Proving the reliability of the questionnaires that resemble Likert type scales in format but prepared in a rating system with independent scoring for items with Cronbach Alpha method is one of the problems frequently faced.

All these three findings identified here are among the findings of the study that are deemed significant.

Presenting reliability evidence based on more than one method is a desired cased during the development of data gathering tools. The obtained findings show that methods used to prove the reliability of tools are mostly limited to one method. In addition to the abundance of tools for which no evidence of reliability was submitted (34.07%), the limitation of reliability evidences with only one method may be important to show the inadequacies of the involved parties in terms of knowledge and sensitivity.

4.Results and Discussion

Led by scales, main data gathering tools used in theses are tests, questionnaires, inventories, interview forms, observation forms, general forms and scenarios. More than one data gathering tools were used together in some of the theses utilized in the study. Therefore, the number of theses that were investigated was 46 whereas the number of tools that were studied was 91. Approximately half of the tools used in these theses was developed by the thesis authors. The other half was composed of tools developed by other researchers. Although comparatively lower in number, the number of tools adapted for use cannot be underestimated.

Thesis authors use the concepts of questionnaires, scales and inventories interchangeably as if they all meant the same thing.

No evidence for validity was observed in more than half of the theses that were investigated and no evidence for reliability was seen in approximately one third of these theses.

Methods used to prove validity of tools were mostly limited to one method. Content validity in tests, exploratory factor analysis in scales, confirmatory factor analysis in adaptations and validity confirmation via expert views were methods that were often used as evidence for validity.

Methods used to prove reliability of tools were mostly limited to one method. Cronbach Alpha, a method to validate reliability, was used to confirm reliability of scales and KR-20 was used to confirm reliability of the tests. The method used to determine reliability of the data obtained from observation and interview forms was inter-rater agreement. Test-Retest was utilized to examine reliability of data obtained through questionnaires.

Findings and results are similar to findings obtained in studies that focused on data gathering tool development processes and various aspects of postgraduate theses (Erkuş, 1999; Kabaca and Erdoğan, 2007; Tavşancıl, 2008; Tavşancıl et. al., 2010; Erkuş, 2010).

Science should be established on valid and reliable data. Research and theses that based on data gathering tools with questionable reliability and validity are far from scientific foundations. Results obtained in this study are rather discouraging in this context. The facts that these theses were prepared under the guidance of advisors and approved by a jury are topics that merit separate discussions

5.Suggestions

Some suggestions based on the findings of this study are provided below:

1. The use of data gathering tools whose validity and reliability are not proven through appropriate methods can be prevented with the help of investigations through institutes,
2. Related institutes can warn advisors regarding the results mentioned in this study and can provide in-service training opportunities about data gathering tools used in theses and the process of data gathering tool development.

References

Atılgan, H.(2013). *Measurement and Evaluation in Education (6th Edition).* Ankara: Anı Publishing.

Baykul, Y. (2010). *Measurement in Education and Psychology: Classical test theory and its practice (2nd Edition).*Ankara: Pegem Akademi.

Büyüköztürk, Ş., Çakmak, E.K., Akgün, Ö.E., Karadeniz, Ş. & Demirel, F., (2013). *Scientific research methods (14th Edition).* Ankara: Pegem Academy.

Crocker, L. &Algina, J. (1986). *Introduction to classical and modern test theory*. USA: Harcourt Brace Jovanovich College Publishers.

Erkuş, A. (1999). Postgraduate theses in three universities in psychometric. *4. National educational Sciences Congress* (75-87). Eskişehir.

Erkuş, A. (2010). Discrepancy between the Turkish equivalents of psychometric terms and the procedures that are implemented. *Journal of Measurement and Evaluation in Education and Psychology, 1*(2), 72-77.

Güler, N. (2012). *Measurement and Evaluation in Education (4th Edition).* Ankara: Pegem Academy.

Kabaca, T. & Erdoğan, Y. (2007). Statistical investigation of theses in science, computers and mathematics teaching disciplines. *Pamukkale University Faculty of Education Journal, 22* (2), 54-63.

Özçelik, D.A. (2011). *Measurement and Evaluation (4th Edition).* Ankara: Pegem Academy.

Tan, Ş. (2012). *Public Personnel Selection Examination (PPSE) Booklet for measurement and evaluation in teaching (7th Edition).* Ankara: Pegem Academy.

Tavşancıl, E. (2008). Content analysis of postgraduate theses' method section .*I. National Congress of Measurement and Evaluation in Education and Psychology* (107-122). Ankara.

Tavşancıl, E.,Ömay, Ç., Çıtak, G.G., Kezer, F., Yıldırım, Y., Bilican, S., Büyükturan, E.B., Şekercioğlu, G., Yalçın, N., Erdem, D. & Özmen, D.T. (2010). *Investigation of completed graduate theses in educational sciences institutes*. Ankara University Scientific Research Project Final Report.

Turgut, M.F. & Baykul, Y. (2012). *Measurement and evaluation in education (4th Edition).* Ankara: Pegem Academy.

Yıldırım A. & Şimşek, H. (2011). *Qualitative research methods in social sciences (8th Edition).*Ankara: Seçkin Publishing.