ISSN 1392-3110
Socialiniai tyrimai / Social Research. 2011. Nr. 1 (22), 52–64

Translation Quality of Survey Instruments: Assessment of Questionnaire on Motivation and Academic Studies Quality

Zemyna Pauliukaite[1], Ausra Turcinskaite-Balciuniene[1], Jonas Balciunas[2]
[1]Kaunas University of Technology, K. Donelaicio str. 20, LT-44239 Kaunas, Lithuania,
e-mail: zemyna.pauliukaite@ktu.lt, ausra.turcinskaite-andujar@ktu.lt
[2]Vilnius Pedagogical University, T. Sevcenkos str. 31, LT-03111 Vilnius, Lithuania,
e-mail: jonas.balciunas@vpu.lt

Abstract

This article seizes the problematic of survey instruments' translation quality assessment. Transfer of a survey instrument from one sociocultural environment to another is a very complex task therefore translation quality of the instrument is one of the key elements to be assessed before making any further data analysis. In this paper, according to the contemporary scientists' recommendations the assessment of translation quality of questionnaire on Motivation and Academic Studies Quality (QUISS II[1]) is presented.

Keywords: translation quality of survey instruments, qualitative methods of translation quality assessment, quantitative methods of translation quality assessment, expert panel review, psychometric characteristics, questionnaire on Motivation and Academic Studies Quality (QUISS II).

Introduction

For several decades the demand for cross-national and cross-cultural survey researches has been growing apace. Different research organizations around the world face the necessity for data collection from respondents of different cultural and linguistic backgrounds. Growing scale of intercultural research and interest in how best to do it is observed in Lithuania, too. Various sociological, psychological, health care and other questionnaires developed abroad could be considered a sort of imported technology, a product. Consistent attempts to adapt such questionnaires are found almost in every field of social sciences: Bagdonas (2007), Mazeikiene (2001), Zydziunaite (2003), Zvirdauskas (2006) adapted questionnaires in educology; Kalinauskaite (2007), Velickaite et al. (2009) – in sociology; Dromantas (2008) – in management; Matonyte and Morkevicius (2009) – in political science, etc. There are many questionnaire translation and adaptation practice examples in the field of medical sciences and health care as well (Furmonavicius et al., 2004; Miltiniene et al., 2008; Riklikiene, 2009b; Rugiene et al., 2005; Vanagas et al., 2006). Translated and adapted questionnaires are also used in Lithuania in the context of different world-wide and European-wide surveys: World Values Survey, Eurobarometer, European Social Survey[2], International Research of Educational Achievement[3], Quality of Life Research, etc (Merkys and Pauliukaite, 2010).

Increasing number of inter-cultural researches and frequent practices of translation and adaptation of foreign questionnaires suggest the relevance of methodological issues about survey instruments' translation. The last few decades have seen great changes in the methodology of adaptation of instruments. New specific methods were created, recommendations were formulated aiming to improve the quality of cross-cultural comparative research, and many methods of translation of survey instruments and evaluation of translation quality have been developed and proposed. Reviewing a wealth of theoretical and empirical publications it appeared that researchers involved in comparative research are concerned about measurement issues, comparability, reliability and validity of their data (Behling and Law, 2000; Dean et al., 2007; Humbleton, 2005; Kalinauskaite and Merkys, 2007; Maneesriwongul and Dixon, 2004; Mazeikiene and Merkys, 2001, 2002; McGorry, 2000; Merkys, 1999). Nevertheless in Lithuanian publications presenting results obtained with transferred survey instruments a particular tendency is observed. Different authors (Bagdonas, 2009; Merkys and Pauliukaite, 2010; Pauliukaite, 2009; Riklikiene, 2007, 2009a, 2009b) have observed that in some publications the problematic of the research instrument translation and adaptation is not sufficiently developed. Transferred instruments are often applied carelessly: compulsory procedures of quality control (validity and reliability tests) which are a must in intercultural research methodology most often are respected only formally or not applied at all. Some authors do not describe the translation technique they have used nor they present the psy-

[1] Qualitätsverbesserung in Schulen und Schulsysthemen II.

[2] Performed by the Institute of Politics and Public Administration of Kaunas University of Technology (LiDA project)

[3] Administrated by the International Educational Achievement Evaluation Association (IEA), in Lithuania administrated by the National Examination Centre

chometric characteristics of the questionnaire which were obtained in Lithuanian population (Merkys and Pauliukaite, 2010), some focus on psychometric measures and evaluation while translation procedure, language, cultural and practical adaptation process are barely mentioned (Riklikiene, 2007).

In this paper we focus on issues concerning quality of translation of survey instruments as an important part of cross-cultural studies. Quality of translation of survey instruments plays a central and important role in survey projects whenever researchers apply an instrument that has been conceived in a foreign language. Due to linguistic and cultural differences a translated questionnaire may not automatically follow the same factor structure, and translated items may contain additional *emic* meaning (limited to a single culture), and be misinterpreted by respondents. Incorrectly translated questionnaires can result in failure of the whole research project due to the poor translation quality impeding collection of comparable data by researchers (Behling and Law, 2000; Maneesriwongul and Dixon, 2004; Mazeikiene and Merkys, 2001, 2002; Merkys, 1999; Pauliukaite, 2009; Riklikiene, 2007, 2009 a, b; Survey Research Center, 2010). Good quality translation of a survey instrument does not ensure the success of a survey, however, poor translation is the most important source of item bias. In conclusion, the transfer of a survey instrument from one sociocultural environment to another is a very complex task and translation quality is one of the key elements to be assessed before making any further data analysis.

Scientific problem of this article may be expressed by the following question: how to ensure and assess translation quality of a transferred survey instrument?

The aim of the article is to seize the problematic of survey instruments' translation quality assessment and according to the contemporary scientists' recommendations to analyse quality of translation of questionnaire on Motivation and Academic Studies Quality (QUISS II).

Research methodology: analysis of scientific literature on translation quality of survey instruments; QUISS II translation quality assessment by expert panel review description; psychometric characteristics (reliability and construct validity) of QUISS II scales for *Motivation* and *Academic Studies Quality* analysis: 1) Cronbach's alpha; 2) Cronbach's alpha if item deleted; 3) corrected item-total correlation; 4) factor analysis.

Translation quality of survey instruments

In cross-national/cross-cultural research the translated version of research instrument must satisfy two sets of requirements: a) it must meet the basic standards set for all measures, translated or not. That is, it must be valid, reliable, and cost-effective (i.e., possess utility); b) it must meet requirements for equivalence relative to the source language measure. Researchers must aim to establish comparability of concepts, norms and semantics (Behling and Law, 2000; Dean et al., 2007).

A literature review indicates that the main problems involved in questionnaire transfer to target language are various. First of all, standard monolingual survey instruments quality testing requirements and assessment procedures, surprisingly, are not required for translated questionnaires (Behling and Law, 2000; Harkness et al., 2004). Second, there is no commonly accepted set of standards and procedures established in the survey research community neither for translating questionnaires nor for assessing the quality of translations procedures. Without some clear criteria for translation quality assessment, it is impossible to know if the translation quality is good, and consequently whether to keep it that way or to improve it (Schiaffino, 2005; Survey Research Center, 2010). Third, particular goals of the translation are rarely specified, i.e., articulated. This means that if the goals to be met by the product are not specified in advance, the criteria of assessment cannot be set in any manner fair to translators. Without specifications, it is impossible to evaluate the usefulness of assessment procedures (Harkness et al., 2004; Survey Research Center, 2010). Fourth, more complex measurement methods (e.g., item response models and factor analysis) that appear to be useful in formally establishing the equivalence of scores obtained from instruments translated (adapted) into multiple languages are not well-known to people involved in studies using translated instruments (Geisinger, 1994).

Quality of translation and validation of the translated instrument plays a significant role in ensuring that the results obtained in cross-cultural research are not due to errors in translation, but rather are due to real differences or similarities between cultures in the phenomena being measured (Maneesriwongul and Dixon, 2004). Lots of translation techniques exist aiming to ensure quality of a translated questionnaire: Simply Direct Translation, Modified Direct Translation, Parallel Blind Technique, Committee and Modified Committee Translation, etc. (Dean et al., 2007; Hambleton et al., 2005; Martinez et al., 2006; McGorry, 2000; Pauliukaite, 2009). For many researchers, combining translation techniques is seen as the best and the most efficient way to deal with translation-related problems (Birbili, cited by Riklikiene, 2007). Time, personnel and funds available for translation quality ensuring are usually limited. Therefore, decisions on which translation procedure to adopt are influenced by the time, funding, expertise and per-

sonnel available, as well as by specific aspects of a given study. All these factors are interrelated and interfering. Once a questionnaire or a survey instrument is constructed, its quality is assessed by using statistical methods. As mentioned above, there is the relevance of survey translation quality to data quality. In spite of this, few research projects have been familiarized with some current translation assessment techniques and/or the assessment outputs (Harkness et al., 2010).

Different qualitative and quantitative methods are used for translation quality assessment. For example, expert panel review, cognitive interviews, interviewer and respondent debriefing, focus groups, back-translation and etc. are used for qualitative evaluation of translation quality. For quantitative assessment of translation quality in the review process different statistical tests are used (Harkness et al., 2004; Survey Research Center, 2010). Many factors can determine which assessment strategies are suitable and when. However, an overview of theoretical and empirical publications indicates that questionnaire translation must be subjected to both qualitative and quantitative assessment (Harkness et al., 2004; Maneesriwongul and Dixon, 2004; Survey Research Center, 2010). As Figure 1 demonstrates, qualitative assessment of translation quality is usually used first, that is, when the process of instrument translation is not finished yet in order to improve it. Quantitative procedures to evaluate the translated survey instrument (questionnaire) constitute a further step and are used after data collection (e.g. in a pilot study). The obvious route to take is first to make translation as good as possible by using design and translation strategies, then to produce a qualitative testing and afterwards to produce statistical analysis confirming the problematic issues detected by the qualitative procedures or discovering some new incorrectly translated items that were not detected in qualitative appraisal (Harkness et al., 2010).

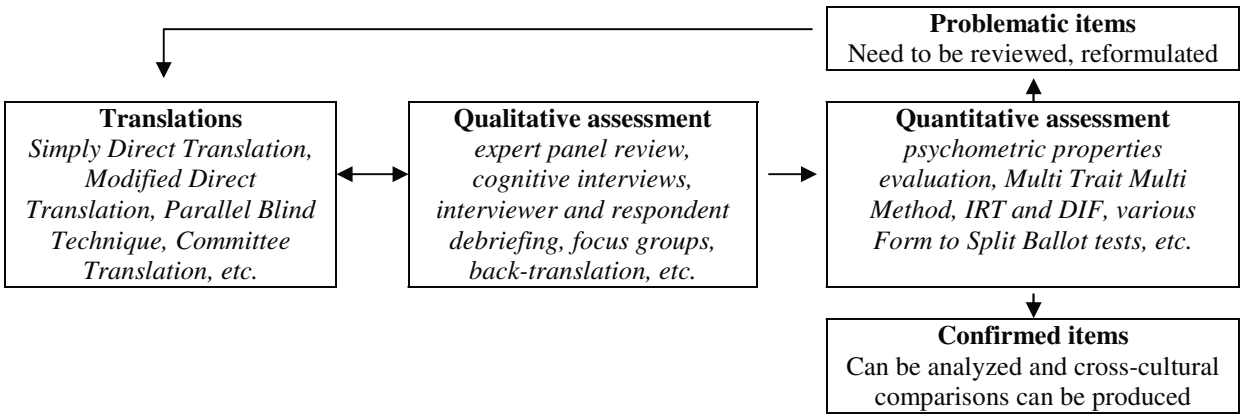

Fig. 1. Cycle of Questionnaire Translation Quality Assessment
Source: compiled by the authors

As qualitative methods of translation quality assessment are not sufficient to indicate whether items truly function as required across different cultures, quantitative procedures are used to assess whether translated instruments perform as expected (Harkness et al., 2004).

Statistical tests take various forms, depending on the characteristics of an instrument, the number of items, the sample sizes available, and the target of assessment. Some of the quantitative assessment procedures used mainly to assess whether translated instruments perform as expected include: Multi Trait Multi Method, Item Response Theory (IRT) and Differential Item Functioning (DIF), various Form to Split Ballot tests (Survey Research Center, 2010).

Statistical analyses are required to investigate the measurement characteristics of items. That is, they determine whether the item functions similarly in both the adapted and the source language versions of the text or not. This is accomplished through the use of an item bias study (often called a "Differential Item Functioning" or DIF study). Item bias may cause problems if such situations as poor wording, inaccurate translations, and inappropriateness of item content in a cultural group exist at the item level of the measurement (Ellis, 1989 cited by Yi-Hsiu et al., 2005). If there are items that function differently for each group, it is necessary to rewrite or retranslate, readministrate, and reanalyze those items to determine whether they function in the same manner for both groups. The statistical techniques developed to detect item bias are divided into two main categories: one procedure developed for dichotomously scored items is the Mantel-Haenszel procedure (Holland and Thayer, 1988 cited by Yi-Hsiu et al., 2005), another procedure developed for detecting differential item functioning in test scores with interval-scale properties was based on the analysis of variance (ANOVA) (Hambleton and Patsula, 1999; Yi-Hsiu et al., 2005).

When assessing whether the translated instrument performs as expected psychometric parameters are widely used (Falk et al., 2007; Kalinauskaite and Merkys, 2007; Riklikiene and Pukenas, 2007b; Rugiene et al., 2005; Tafas et al., 2002). Seeking to test the clarity and appropriateness of the target language version questionnaires psychometric characteristics including internal consistency, reliability and indicators of validity are used (Clifford et al., 1999; Tafas et al., 2002; Maneesriwongul and Dixon, 2004). Cronbach's alpha coefficient analysis indicates items that should be reviewed and corrected. When some high psychometric indicators are obtained by using translated questionnaire in the target population, it confirms that the transfer procedures were successful.

Our recommendation is that to make the process of questionnaire translation and quality assessment clear as well as to confirm the quality of the transferred questionnaire, authors who use translated instruments should always describe their translation technique in detail (e.g. using schemes) and present translated questionnaire psychometric characteristics obtained in target population. Thus, according to the use of qualitative and quantitative survey instruments assessment methods presented above, a) description of QUISS II translation quality assessment by expert panel review; and b) analysis of psychometric characteristics of questionnaire on Motivation and Academic Studies Quality (QUISS II) are presented next.

Research methodology

After completion of analysis of scientific publications aiming to seize survey instruments translation quality assessment issues (Bagdonas, 2009; Behling and Law, 2000; Clifford et al., 1999; Dean et al., 2007; Harkness et al., 2010; Harkness et al., 2004; Humbleton, 2005; Kalinauskaite and Merkys, 2007; Maneesriwongul and Dixon, 2004; Mazeikiene and Merkys, 2001, 2002; Merkys, 1997, 1999; Merkys and Pauliukaite, 2010; Pauliukaite, 2009; Riklikiene, 2007, 2009 a, b; Yi-Hsiu et al., 2005), psychometric characteristics of QUISS II[4] questionnaire on *Motivation* and *Academic Studies Quality* will be assessed. QUISS survey methodology was elaborated by the team of scholars directed by Professor Werner Georg at Konztanz University (Germany) in 1983, since it was used repeatedly in surveys in Germany and systematically improved by authors. In 2000-2002 one of its recent versions, QUISS I, was used on international level: respondents from 16 universities took part in the survey (2 German universities – 984 respondents, 6 Spanish universities – 1823 respondents and 8 French universities – 1230 respondents). In 2009 the latest version of the instrument, QUISS II, was used in students' survey in Lithuania. Lithuanian State Studies Foundation financed the project "*Academic Studies Quality and Social Context Survey*" by a group of scientists directed by Professor Gediminas Merkys in which QUISS II was translated and culturally adapted (Merkys et al., 2009; Turcinskaite-Balciuniene and Merkys, 2010).

Translated paper-pencil type questionnaire dominated by psychometric Likert-type scales of different levels (from 3 to 9), most often with central categories, was filled in by 964 respondents from Kaunas, Klaipeda, Siauliai and Vilnius higher education institutions constituting a data basis for quantitative questionnaire translation quality assessment procedures.

In this paper, psychometric characteristics of QUISS II scales for *Motivation* and *Academic Studies Quality* are assessed: 34 scales consisting of 297 items.

A. Motivation scales (5 scales – 42 items):

1. Choice of Higher Education Institution (8 items);
2. Choice of Studies (7 items);
3. Meaningfulness of Studies (10 items);
4. Motivation for Doctoral Studies (11 items);
5. Intentions to Study Abroad (6 items).

B. Academic studies quality aspects (29 scales – 255 items):

1. Availability of Information (3 scales – 29 items): a) Information Available (10 items); b) Information Needed (6 items); c) Use of Internet (13 items);
2. Major Problem Areas (3 scales – 42 items): a) Areas in Need for Improvement (16 items); b) Problematic Areas (12 items); c) Urgent Problems (14 items).
3. Studies and Teaching (7 scales – 76 items): a) Speciality Studies (12 items); b) Studies Requirements (14 items); c) Studies and Teaching Quality (9 items); d) Teaching Quality (10 items); e) Studies Experience (7 items); f) Anxiety over Possible Failures (10 items); g) Studies Results (14 items);
4. Communication and Consulting (7 scales – 30 items): a) Satisfaction with Communication Possibilities (3 items); b) Consulting Possibilities (4 items); c) Consultations (4 items); d) Attentiveness to Students' Problems (3 items); e) Consultations Provided by Student Offices (3 items); f) Benefits of Consultations by Student Offices (3 items); g) Functions of Student Representative Office (10 items);
5. Studies Results (6 scales –50 items): a) Personal Growth Factors (13 items); b) Professional Perspectives Factors (13 items); c) Seeking for Improvement (6 items);

[4] Qualitätsverbesserung in Schulen und Schulsysthemen II.

d) Interests in Improvement Possibilities (6 items); e) Autonomous Search for Improvement (5 items); f) Studies (9 items);

6. European Higher Education Area (3 scales – 28 items): a) EHEA Means (9 items); b) EHEA Studies Organization Forms (7 items); c) EHEA Objectives (12 items).

Aiming to ensure QUISS II questionnaire translation quality expert panel review was used. Evaluating QUISS II psychometric characteristics (reliability and construct validity) statistical package SPSS 13.0 for Windows was used when computing various statistics: 1) Cronbach's alpha in order to assess internal consistency of the items in the scales; 2) Cronbach's alpha if item deleted analysis aiming to determine whether some of the items have significantly different scores; 3) corrected item-total correlation analysis in order to determine how well score of one item is internally consistent with composite scores from all other items that remain; 4) factor analysis aiming to determine the dimensionality of the scales.

Research results

Qualitative assessment of translation of questionnaire was ensured by expert panel review. There were 3 versions of the questionnaire: English, German (original), and French. Each version of the questionnaire was translated separately and then the three translations were confronted aiming to elaborate the best possible translation by the panel of three experts (an expert and two translators) assessing and ensuring translation quality. Two of the three experts were researchers. Once the experts validated translation, the questionnaire was reviewed by another independent expert (trilingual) and returned to the expert panel for finalisation of questionnaire translation procedure (see Figure 2). After final expert panel translation quality review, the final version of the Lithuanian questionnaire was confirmed to be of good quality and suitable to be applied as a survey instrument. As translation was a part of the scientists' group project "*Academic Studies Quality and Social Context Survey*", according to deadlines pre-test and pilot-study were not administrated.

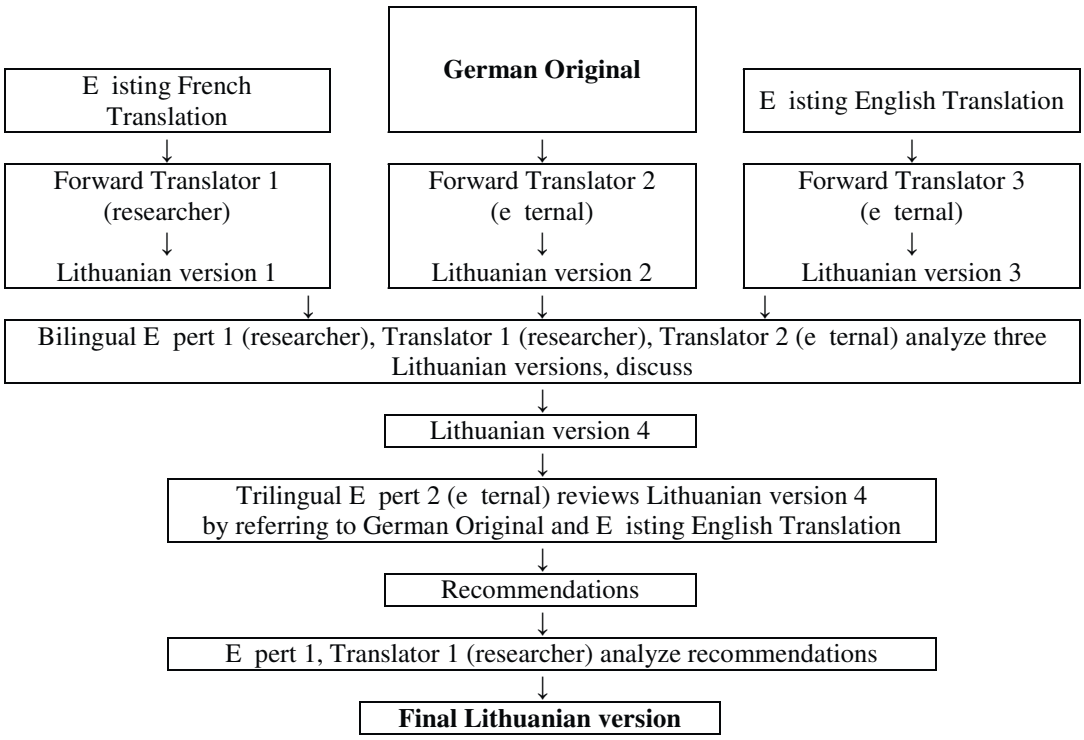

Fig. 2. QUISS II Questionnaire Translation Quality Review by Expert Panel within Translation Process
Source: authors' questionnaire development protocol

Next, psychometric properties (translated items reliability and construct validity) were assessed. When both validity and reliability analyses produce reasonably good results, then the translated questionnaire can be concluded and declared to have acceptable psychometric properties. Statistical methods evaluating internal consistency (see Table 1 and Table 4) and dimensionality (see Table 2 and Table 3) were applied for quantitative translation quality assessment of QUISS II *Motivation* and *Academic Studies Quality* scales.

Table 1

Internal Consistency Statistics – Translated Items Reliability

Scales (N=964)	Items	Cronbach's alpha		Cronbach's alpha if item deleted	Corrected item-total correlation
Motivation scales (42 items)					
Choice of Higher Education Institution	8	0.82	good	0.79 – 0.82	0.36 – 0.60
Choice of Studies	7	0.87	good	0.84 – 0.87	0.50 – 0.77
Meaningfulness of Studies	10	0.88	good	0.87 – 0.89	0.40 – 0.70
Motivation for Doctoral Studies	11	0.88	good	0.87 – 0.88	0.47 – 0.72
Intentions to Study Abroad	6	0.95	excellent	0.94 – 0.95	0.80 – 0.89
Academic Studies Quality: Availability of Information (29 items)					
Information Available	10	0.88	good	0.87 – 0.88	0.50 – 0.68
Information Needed	6	0.90	excellent	0.87 – 0.91	0.58 – 0.82
Use of Internet	13	0.96	excellent	0.96	0.77 – 0.84
Academic Studies Quality: Major Problem Areas (42 items)					
Areas in Need for Improvement	16	0.90	excellent	0.90	0.48 – 0.67
Problematic Areas	12	0.95	excellent	0.95	0.71 – 0.83
Urgent Problems	14	0.93	excellent	0.92 – 0.93	0.59 – 0.74
Academic Studies Quality: Studies and Teaching (76 items)					
Speciality Studies	12	0.81	good	0.79 – 0.81	0.35 – 0.61
Studies Requirements	14	0.94	excellent	0.93 – 0.94	0.48 – 0.79
Studies and Teaching Quality	*9*	*0.80*	*good*	*0.76 – 0.82*	*0.23 – 0.65*
Teaching Quality	10	0.94	excellent	0.93 – 0.94	0.62 – 0.84
Studies Experience	7	0.90	excellent	0.88 – 0.90	0.61 – 0.78
Anxiety over Possible Failures	10	0.83	good	0.81 – 0.84	0.36 – 0.62
Studies Results	14	0.95	excellent	0.95	0.67 – 0.78
Academic Studies Quality: Communication and Consulting (30 items)					
Satisfaction with Communication Possibilities	*3*	*0.79*	*acceptable*	*0.71 – 0.84*	*0.50 – 0.77*
Consulting Possibilities	4	0.95	excellent	0.93 – 0.94	0.86 – 0.90
Consultations	4	0.86	good	0.81 – 0.83	0.69 – 0.72
Attentiveness to Students' Problems	*3*	*0.46*	*unacceptable*	*0.25 – 0.43*	*0.23 – 0.35*
Consultations Provided by Student Offices	3	0.96	excellent	0.94 – 0.95	0.92 – 0.94
Benefits of Consultations by Student Offices	3	0.90	excellent	0.85 – 0.89	0.78 – 0.83
Functions of Student Representative Office	10	0.97	excellent	0.96 – 0.97	0.74 – 0.90
Academic Studies Quality: Studies Results (50 items)					
Personal Growth Factors	13	0.94	excellent	0.93 – 0.94	0.56 – 0.76
Professional Perspectives Factors	13	0.96	excellent	0.95 – 0.96	0.66 – 0.82
Seeking for Improvement	6	0.97	excellent	0.96	0.88 – 0.91
Interests in Improvement Possibilities	6	0.92	excellent	0.90 – 0.91	0.75 – 0.82
Autonomous Search for Improvement	5	0.91	excellent	0.88 – 0.90	0.33 – 0.58
Studies	*9*	*0.76*	*acceptable*	*0.72 – 0.76*	*0.71 – 0.82*
Academic Studies Quality: European Higher Education Area (28 items)					
EHEA Means	9	0.90	excellent	0.88 – 0.90	0.62 – 0.75
EHEA Studies Organization Forms	7	0.93	excellent	0.92 – 0.93	0.74 – 0.81
EHEA Objectives	12	0.95	excellent	0.94 – 0.95	0.65 – 0.81

Use of Likert-type scales is imperative to calculate Cronbach's alpha, which is also a good reliability indicator for the translated scales. According to George and Mallery (2003), Cronbach's alpha scores lower than 0.50 are unacceptable; scores from 0.50 to 0.60 are poor; scores from 0.60 to 0.70 are questionable; scores from 0.70 to 0.80 are acceptable; scores from 0.80 to 0.90 are good; and those higher than 0.90 are excellent. An alpha of 0.80 is considered to be a reasonable goal.

According to the results, in general strong internal consistency among items is observed. 91% of the scales demonstrate good or excellent overall Cronbach's alpha scores: 22 scales demonstrate excellent scores and 9 scales display good Cronbach's alpha scores. *Satisfaction with Communication* scale of 3 items and *Studies* scale of 9 items show acceptable scores and *Attentiveness to Students' Problems* scale shows unacceptable Cronbach's alpha scores, indicating some lack of internal consistency among items.

According to Cronbach's alpha if item deleted analysis, it appears that in some cases internal consistency of the scale is so strong that after deleting any of items overall Cronbach's alpha would drop, for example in the scales of *Consultations* or *Benefits of Consultations by Student Offices*. In general Cronbach's alpha if item deleted varies around the overall Cronbach's alpha score within four points, nevertheless, in the case of *Studies and Teaching Quality* scale, *Satisfaction with Communication Possibilities* scale and *Attentiveness to Students' Problems* scale it varies by 6-18 points reflecting unequal contributions of different items to total alpha and pointing at the lack of internal consistency.

Finally, corrected item-total correlation analysis for each of the item of the scale was done. It is the correlation between a given item and the total score of the other items of the scale. The last column in Table 1 displays for each scale the range of corrected item-total correlations of its items. This is a way to assess how well the score of one item is internally consistent with composite scores of all other items that remain. This correlation is considered weak when the score is lower than 0.30 (De Vaus, 2004). There are 2 scales with weak corrected item-total correlation: *Studies and Teaching Quality* (9 items) and *Attentiveness to Students' Problems* (3 items). In the case of *Studies and Teaching Quality* scale, only 2 items have weak corrected item-total correlation of 0.23 and 0.25 respectively, the remaining 7 items range from 0.48 to 0.65, meanwhile all the data indicate that *Attentiveness to Students' Problems* scale lacks internal consistency: 1) overall Cronbach's alpha is unacceptable, 2) Cronbach's alpha if alpha deleted varies by up to 18 points for only 3 items in the scale, and 3) corrected item-total correlation is between 0.23 and 0.35.

Internal consistency analysis of different QUISS II *Motivation* and *Academic Studies Quality* scales allows us to conclude that only one scale of 34 analysed obviously lacks internal consistency by all analysed criteria. This result may be related to translation quality issues, or it may be also due to different original methodology conception aspects. Other two scales *Satisfaction with Communication Possibilities* and *Studies* show acceptable internal consistency score and 2 items of *Studies and Teaching Quality* scale demonstrate weak corrected item-total correlation score. Based on internal consistency analysis, 280 items seem to be reliable and consistent, some 17 items lack internal consistency and need to be reviewed and corrected.

High value of Cronbach's alpha indicates good internal consistency of the items on the scale, but it does not mean that the scale is one-dimensional. Exploratory factor analysis involving the principle component analysis extraction and varimax rotation is commonly used to assess the construct validity (see Table 2).

Table 2

Factor Analysis Statistics – Construct Validity

Scales (N=964)	Items	Percent of the total variance that factor accounts for		Initial eigenvalues		Kaiser-Meyer-Olkin measure	
		1	2	1	2		
Motivation scales (42 items)							
Choice of Higher Education Institution	8	45	-	3.60	0.91	0.86	good
Choice of Studies	7	57	18	3.97	1.29	0.84	good
Meaningfulness of Studies	10	51	11	5.09	1.06	0.87	good
Motivation for Doctoral Studies	11	47	13	5.15	1.39	0.88	good
Intentions to Study Abroad	6	81	-	4.86	0.34	0.93	excellent
Academic Studies Quality: Availability of Information (29 items)							
Information Available	10	49	10	4.86	1.01	0.91	excellent
Information Needed	6	69	-	4.12	0.66	0.90	good
Use of Internet	13	70	-	9.08	0.70	0.97	excellent
Academic Studies Quality: Major Problem Areas (42 items)							
Areas in Need for Improvement	16	43	8	6.88	1.34	0.92	excellent
Problematic Areas	12	67	-	8.00	0.61	0.97	excellent
Urgent Problems	14	78	-	7.81	0.44	0.98	excellent
Academic Studies Quality: Studies and Teaching (76 items)							
Speciality Studies	12	35	14	4.14	1.64	0.87	good
Studies Requirements	14	56	-	7.91	0.85	0.96	excellent
Studies and Teaching Quality	9	46	17	4.16	1.51	0.87	good
Teaching Quality	10	66	-	6.58	0.65	0.96	excellent

Continued to Table 2

Studies Experience	7	64	-	4.47	0.83	0.89	good
Anxiety over Possible Failures	10	43	10	4.3	1.06	0.88	good
Studies Results	14	61	-	8.53	0.90	0.96	excellent
Academic Studies Quality: Communication and Consulting (30 items)							
Satisfaction with Communication Possibilities	3	71	-	2.13	0.63	0.61	questionable
Consulting Possibilities	4	87	-	3.48	0.21	0.87	good
Consultations	4	70	-	2.81	0.47	0.82	good
Attentiveness to Students' Problems	3	49	-	1.47	0.85	0.58	poor
Consultations Provided by Student Offices	3	93	-	2.80	0.11	0.78	acceptable
Benefits of Consultations by Student Offices	3	84	-	2.52	0.28	0.75	acceptable
Functions of Student Representative Office	10	78	-	7.81	0.44	0.98	excellent
Academic Studies Quality: Studies Results (50 items)							
Personal Growth Factors	13	58	-	7.54	0.87	0.95	excellent
Professional Perspectives Factors	13	67	-	8.76	0.63	0.96	excellent
Seeking for Improvement	6	86	-	5.14	0.23	0.94	excellent
Interests in Improvement Possibilities	6	72	-	4.34	0.45	0.92	excellent
Autonomous Search for Improvement	5	73	-	3.67	0.45	0.87	good
Studies	9	36	20	3.20	1.75	0.72	acceptable
Academic Studies Quality: European Higher Education Area (28 items)							
EHEA Means	9	58	-	5.18	0.88	0.93	excellent
EHEA Studies Organization Forms	7	71	-	4.98	0.52	0.93	excellent
EHEA Objectives	12	65	-	7.76	0.89	0.96	excellent

Factor analysis is based on the correlation matrix of the variables involved, and correlations usually need a large sample size for them to stabilize. Sample size is expected to be at least 300, better 500, and 1000 or more is excellent. As a rule, a bare minimum of 10 observations per variable is necessary to avoid computation difficulties. As our sample size is 964, factor analysis can be applied for dimensionality analysis of the scales.

Before conducting a factor analysis for each scale, Kaiser-Meyer-Olkin measure of sampling adequacy needs to be run: a value lower than 0.50 is unacceptable; from 0.50 to 0.60 is poor; from 0.60 to 0.70 is questionable; from 0.70 to 0.80 is acceptable; from 0.80 to 0.90 is good; and higher than 0.90 is excellent (Cekanavicius and Murauskas, 2004).

According to Kaiser-Meyer-Olkin measure of sampling adequacy, two scales are not appropriate for factor analysis: *Satisfaction with Communication Possibilities* and *Attentiveness to Students' Problems* (see Table 2). To all other scales factor analysis can be applied – none of the correlation matrixes is an identity matrix (Bartlett's tests of sphericity are significant with probabilities lower than 0.05).

Factor analysis has shown that 23 scales are one-dimensional and demonstrate strong internal consistency. Other 7 scales are two-dimensional and 2 scales are three-dimensional. Further analysis of items constituting different factors is needed (see Table 3) in order to determine whether it may be the result of translation or it is due to the original scales conception particularities.

Two-dimensional *Motivation* scales: a) The first factor of the *Choice of Studies* scale consists of 4 items and expresses socio-professional aspects of studies motivation, the second factor of 3 items expresses personal growth aspects of studies motivation, hence, the two factors are obviously in original scale and not necessarily the result of translation; b) The first factor of the *Meaningfulness of Studies* scale consists of 8 items, the second factor consists of 2 items with quite different ideas: 1) willingness to spend time studying and to postpone as far as possible involvement in professional activities; 2) higher education enabling to be helpful to other people. For this scale the two factors are difficult to label and two-dimensional character of the scale may be the case of inadequate translation; c) The first factor of the *Motivation for Doctoral Studies* scale consists of 7 items and expresses positive aspects of choosing doctoral studies, the second factor consists of 4 items and expresses negative aspects of choosing doctoral studies. These two factors are obviously in original scale rather than resulting from inadequate translation.

Two-dimensional *Availability of Information* scale *Information Available* consists of 2 factors each containing 5 items and respectively expressing 1) information about profession and institution's policy related aspects; 2) information about different studies aspects. In this case the two factors may also be in original scale and may not be the result of translation.

Two-dimensional *Studies and Teaching* scales: a) *Speciality Studies* scale consists of 2 factors each containing 6 items expressing 1) studies organization aspects and competition among students; 2) social aspects and studying difficulty level. Two-dimensional character of the scale may be in original scale; b) the first factor of the *Studies and Teaching Quality* scale contains 7 items and concerns teaching quality and the second factor is of 2 items and concerns studies quality. This scale in original language is definitely two-dimensional; c) each of two factors of *Anxiety over Possible Failures* scale contains 5 items: 1) anxiety related to social and personal aspects; 2) anxiety related to perspectives and studies requirements. The two factors are obviously in original scale and not necessarily the result of translation.

The first factor of *Areas in Need for Improvement* of the three-dimensional *Major Problem Areas* scale consists of 8 items covering different studies organization aspects (studies content, didactics, studies programmes, studies forms, teaching, etc.); the second factor contains 5 items and covers different aspects of requirements for the students; the third factor contains 3 items and covers such aspects as admission to the institution and institution's staff. The three factors may be conceived in the original scales and may not be related to the translation issues.

Three-dimensional *Studies Results* scale *Studies* of 9 items consists of three factors: 1) anxiety over the studies results – 4 items; 2) studying abilities – 3 items; 3) examinations and studies achievements – 2 items.

Table 3

Two-dimensional and Three-dimensional Scales

Scales (N=964)	Items	Percent of the total variance that factor accounts for		
		1	2	3
Two-dimensional Motivation scales				
Choice of Studies	7	57 (4 items)	18 (3 items)	-
Meaningfulness of Studies	10	51 (8 items)	11 (2 items)	-
Motivation for Doctoral Studies	11	47 (7 items)	13 (4 items)	-
Two-dimensional Availability of Information scale				
Information Available	10	49 (5 items)	10 (5 items)	-
Two-dimensional Studies and Teaching scales				
Speciality Studies	12	35 (6 items)	14 (6 items)	-
Studies and Teaching Quality	9	46 (7 items)	17 (2 items)	-
Anxiety over Possible Failures	10	43 (5 items)	10 (5 items)	-
Three-dimensional Major Problem Areas scale				
Areas in Need for Improvement	16	43 (8 items)	8 (5 items)	7 (3 items)
Three-dimensional Studies Results scale				
Studies	9	36 (4 items)	20 (3 items)	12 (2 items)

Producing translation quality verification on a set of QUISS II *Motivation* and *Academic Studies Quality* scales was started with internal consistency tests considering that scales are supposedly one-dimensional. As it appeared, it is not the case for 9 of 34 scales. Therefore internal consistency statistics for all the factors were computed next (see Table 4).

Analysing internal consistency statistics for each factor separately, it appears that 53 items regrouped in 10 factors demonstrate excellent or good overall Cronbach's alpha scores, and 39 items regrouped in 10 factors show acceptable or poor overall Cronbach's alpha scores. None of the items display weak corrected item-total correlation score (less than 0.30), therefore scores of the items are more or less internally consistent with composite scores from all other items constituting separate factors. In general Cronbach's alpha if item deleted varies around the overall Cronbach's alpha score within five points, nevertheless, *Choice of Studies: Socio-Professional Dimension* Cronbach's alpha if item deleted varies within 8 points; *Choice of Studies: Personal Growth Dimension* – varies within 14 points; *Anxiety over Possible Failures: Perspectives and Studies Requirements Dimension* – varies within 13 points; *Areas in Need for Improvement: Admission to the Institution and Institution's Staff Dimension* – varies within 25 points; *Studies: Anxiety over the Studies Results Dimension* – varies within 11 points; and *Studies: Studying Abilities Dimension* – varies within 11 points reflecting unequal contributions of different items to total alpha.

Table 4

Internal Consistency of Two-dimensional and Three-dimensional Scales

Scales (N=964)	Items	Cronbach's alpha		Cronbach's alpha if item deleted	Corrected item-total correlation
Two-dimensional Motivation scales (28 items)					
Choice of Studies: Socio-Professional Dimension	4	0.90	excellent	0.82 – 0.90	0.68 – 0.83
Choice of Studies: Personal Growth Dimension	3	0.80	good	0.66 – 0.80	0.59 – 0.70
Meaningfulness of Studies, Dimension I	8	0.89	good	0.87 – 0.88	0.60 – 0.71
Meaningfulness of Studies, Dimension II	2	0.53	poor	–	0.37
Motivation for Doctoral Studies: Positive Dimension	7	0.83	good	0.80 – 0.84	0.38 – 0.71
Motivation for Doctoral Studies: Negative Dimension	4	0.78	acceptable	0.70 – 0.75	0.51 – 0.64
Academic Studies Quality: Two-dimensional Availability of Information scale (10 items)					
Information Available: Professional and Institution's Policy Dimension	5	0.77	acceptable	0.72 – 0.75	0.51 – 0.60
Information Available: Different Studies Aspects Dimension	5	0.85	good	0.81 – 0.84	0.60 – 0.71
Academic Studies Quality: Two-dimensional Studies and Teaching scales (31 items)					
Speciality Studies: Studies Organization and Competition among Students Dimension	6	0.81	good	0.76 – 0.80	0.48 – 0.67
Speciality Studies: Social Aspects and Studies Difficulty Level Dimension	6	0.71	acceptable	0.66 – 0.70	0.38 – 0.52
Studies and Teaching Quality : Teaching Quality Dimension	7	0.85	good	0.84 – 0.89	0.50 – 0.71
Studies and Teaching Quality: Studies Quality Dimension	2	0.70	acceptable	–	0.54
Anxiety over Possible Failures: Social and Personal Dimension	5	0.79	acceptable	0.73 – 0.76	0.55 – 0.63
Anxiety over Possible Failures: Perspectives and Studies Requirements Dimension	5	0.70	acceptable	0.61 – 0.74	0.34 – 0.59
Academic Studies Quality: Three-dimensional Major Problem Areas scale (16 items)					
Areas in Need for Improvement: Studies Organization Dimension	8	0.88	good	0.86 – 0.88	0.53 – 0.70
Areas in Need for Improvement: Requirements for the Students Dimension	3	0.80	good	0.75 – 0.78	0.54 – 0.64
Areas in Need for Improvement: Admission to the Institution and Institution's Staff Dimension	3	0.72	acceptable	0.51 – 0.76	0.43 – 0.64
Academic Studies Quality: Three-dimensional Studies Results scale (9 items)					
Studies: Anxiety over the Studies Results Dimension	4	0.72	acceptable	0.61 – 0.72	0.38 – 0.58
Studies: Studying Abilities Dimension	3	0.75	acceptable	0.59 – 0.70	0.54 – 0.63
Studies: Examinations and Achievements Dimension	2	0.84	good	–	0.73

Analysing internal consistency statistics for separate factors some observations may be made

- If *Choice of Studies* scale is originally constituted of two separate dimensions: *Socio-Professional Dimension* and *Personal Growth Dimension*, its translation may be considered of good quality.
- Translation and conception issues about the two items constituting the second dimension of *Meaningfulness of Studies* scale and of the two items constituting *Studies Quality Dimension* of *Studies and Teaching Quality* scale should be reviewed, other items of those two scales may be considered of good translation quality.
- Translation quality of half (6 items) of *Speciality Studies* scale (*Studies Organization and Competition among Students Dimension*) may be considered of good quality while its *Social Aspects and Studies Difficulty Level Dimension* (6 items) lacks internal consistency. Translation and conception issues about this scale are to be reviewed.
- Translation and conception issues about two-dimensional scales *Motivation for Doctoral Studies*, *Information Available*, and *Anxiety over Possible Failures* should be reviewed.
- Translation and conception issues about three-dimensional scales should be reviewed. Especially problematic is *Studies* scale.

Qualitative assessment of translation of the questionnaire was ensured by expert panel review, quantitative assessment of translation was produced by analysing psychometric characteristics of the questionnaire – reliability and construct validity. Summing up the analysis of psychometric characteristics of QUISS II *Motivation* and *Academic Studies Quality* scales, evidence is provided that based on all the statistics computed 23 scales (67.7%) are one-dimensional and exhibit strong internal consistency. Therefore translation of 199 items (67%) can be concluded to have acceptable psychometric properties (reliability and validity), hence it may be considered being of good quality. Based on statistics of internal consistency of separate factors, some additional 28 items may be considered of good translation quality as well. Nevertheless 9 scales (32.3%) that are not one-dimensional and 2 scales that are not appropriate for factor analysis need some further cross-cultural exploration by comparing the data obtained using original questionnaire within German students' population and the data obtained using translated questionnaire within Lithuanian students' population. In addition to quantitative methods of translation quality assessment it is also recommended to apply different qualitative methods while reviewing the items that psychometric characteristics analysis revealed to be problematic.

Conclusions

The transfer of a survey instrument from one sociocultural environment to another is a very complex task and translation quality is one of the key elements to be assessed before making any further data analysis.

Translation assessment includes different forms of qualitative and quantitative approaches. Many factors can determine which assessment strategies are suitable and when. However, an overview of theoretical and empirical publications indicates that questionnaire translation must be subjected to both qualitative and quantitative assessment.

To confirm the quality of the transferred questionnaire it is always necessary to describe its translation technique in detail (e.g. using schemes) and to present translated questionnaire psychometric characteristics obtained in target population.

QUISS II translation quality assessment was performed a) by expert panel review; and b) by analyzing psychometric properties of different *Motivation* and *Academic Studies Quality* scales: reliability evaluating internal consistency statistics (Cronbach's alpha, Cronbach's alpha if item deleted and corrected item-total correlation) and construct validity analysing factor analysis.

After the final expert panel translation quality review, the final Lithuanian questionnaire version was approved to be of good quality and suitable to be applied as a survey instrument, pre-test and pilot-study were not administrated.

Based on all the statistics computed 23 scales are one-dimensional and demonstrate strong internal consistency. Therefore translation of 199 items (67%) can be concluded to have acceptable psychometric properties and may be considered being of good quality.

Some further cross-cultural exploration by comparing the data obtained using original questionnaire within German students' population and the data obtained using translated questionnaire within Lithuanian students' population is recommended for the 9 scales that are not one-dimensional and 2 scales that are not appropriate for factor analysis need. Cross-cultural comparisons could provide additional and reliable data concerning some survey instrument translation quality aspects that are difficult to evaluate otherwise.

In addition to quantitative methods of translation quality assessment it is also recommended to apply different qualitative methods while reviewing the items that psychometric characteristics analysis revealed to be problematic.

References

1. Bagdonas, A. (2007). Mokyklos baimės raiška ir jos įveikimo edukaciniai veiksniai (Daktaro disertacija, Kauno technologijos universitetas).
2. Bagdonas, A. (2009). Mažos šalies vargai dėl proto. Available online at http://www.technologijos.lt/p/spausdinti?name=straipsnis-6721.
3. Behling, O., Law, K. S. (2000). *Translating Questionnaires and Other Research Instruments: Problems and Solutions.* Sage.
4. Clifford, C., Day, A., Cox, J., Werrett, J. (1999). A Cross-cultural Analysis of the Use of the Edinburgh Post-Natal Depression Scale (EPDS) in Health Visiting Practice. *Journal of Advanced Nursing*, 30, 655–664.
5. Čekanavičius, V., Murauskas, G. (2004). *Statistika ir jos taikymo metodai II.* Vilnius: TEV.
6. Dean, E., Caspar, R., McAvinchey, G., Reed, L., Quiroz, R. (2007). Developing a Low-cost Technique for Parallel Cross-cultural Instrument Development: The Question Appraisal System (QAS-04). *International Journal of Social Research Methodology,* 10 (3), 227–241.
7. De Vaus, D. (2004). *Surveys in Social Research*. Routledge.
8. Dromantas, M. (2008). Komandinis darbas Lietuvos organizacijose: diagnostinis aspektas (Daktaro disertacija, Lietuvos žemės ūkio universitetas, Vytauto Didžiojo universitetas).
9. Falk, H., Persson, L. O., Wijk, H. (2007). A Psychometric Evaluation of a Swedish Version of the Quality of Life in Late-Stage Dementia (QUALID) Scale. *International Psychogeriatrics,* 19 (6), 1040–1050.

10. Furmonavičius, T. (2004). Su sveikata susijusios gyvenimo kokybės tyrimų metodologiniai aspektai. *Medicina,* 40 (6), 509–516.
11. Geisinger, K. F. (1994). Cross-cultural Normative Assessment: Translation and Adaptation Issues Influencing the Normative Interpretation of Assessment Instruments. *Psychological Assessment,* 6, 304–312.
12. Hambleton, R. K., Merenda, P. F., Spielberger, C. D. (2005). *Adapting Educational and Psychological Tests for Cross-cultural Assessment*. Mahwah, NJ: Lawrence Erlbaum.
13. Hambleton, R. K., Patsula, L. (1999). Increasing the Validity of Adapted Tests: Myths to Be Avoided and Guidelines for Improving Test Adaptation Practices. *JATT,* 1 (1), 1–30.
14. Harkness, J., Pennell, B. E., Schoua-Glusberg, A. (2004). Survey Questionnaire Translation and Assessment. In Presser, S. et al. (Eds.), *Questionnaire Development Evaluation and Testing Methods* (453–475). New Jersey: John Wiley and Sons Inc.
15. Harkness, J. A.,Villar, A.,Edwards, B.(2010).Translation, Adaptation, and Design. In Harkness, J. A. et al. (Eds.), *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (117–140). New Jersey: John Wiley and Sons Inc.
16. Kalinauskaitė, R. (2007). Organizacinis klimatas ir tarpgeneraciniai santykiai: diagnostinis aspektas (Daktaro disertacija, Kauno technologijos universitetas).
17. Kalinauskaitė, R., Merkys, G. (2007). A Cross-cultural Validation of the Intergenerational Tension Questionnaire: the Case of Lithuania. *Socialiniai mokslai*, 4 (58), 83–93.
18. Maneesriwongul, W, Dixon, J. K. (2004). Instrument Translation Process: a Methods Review. *Journal of Advanced Nursing*, 48 (2), 175–186.
19. Martinez, G., Marin, B. V., Schoua-Glusberg, A. (2006). Translating from English to Spanish: The 2002 National Survey of Family Growth. *Hispanic Journal of Behavioral Sciences*, 28 (4), 531–545.
20. Matonytė, I., Morkevičius, V., Best, H. (2009). Elitų Europa: ES valstybių narių politinio ir ekonominio elito europietiškumo mastai (Europe of Elites: Europeanness of Political and Economic Elites of the EU Member States), *Politologija*, 2 (54), 3–32.
21. Mažeikienė, N. (2001). Socializacijos ir ugdymo diagnostikos instrumentų kultūrinis perkeliamumas: makiavelizmo ir autoritarizmo matavimo aspektas (Daktaro disertacija, Šiaulių universitetas).
22. Mažeikienė, N., Merkys, G. (2001). Makiavelizmo skalės, kaip socializacijos diagnostikos instrumento, transkultūrinio tinkamumo kontrolė. *Socialiniai mokslai,* 3 (29), 7–16.
23. Mažeikienė, N., Merkys, G. (2002). Makiavelizmo skalės, kaip ugdymo diagnostikos instrumento, transkultūrinės validacijos patirtis. *Tiltai,* 1 (18), 1–10.
24. McGorry, S. Y. (2000). Measurement in a Cross-cultural Environment: Survey Translation Issues. *Qualitative Market Research*, 3 (2), 74–81.
25. Merkys, G. (1999). Sprachliche Adaptation und Transkulturelle Validierung der Erherbungsinstrumente: einige methodologische Erfahrubgen aus der kulturvergleichenden paedagogisch – psychologischen Forschung. Magist. *An International Journal for Psychology and Education*, 6, 22–35.
26. Merkys, G., Lekavičienė, R., Šaparnienė, D., Mikutavičienė, I., Turčinskaitė-Andujar, A. (2009). Akademinių studijų kokybė ir socialinis kontekstas. *Lietuvos studentų tarpkultūrinis palyginamasis tyrimas (ataskaita)*. Kaunas.
27. Merkys, G., Pauliukaitė, Ž. (2010). Apklausos instrumentų tarpkultūrinio perkėlimo metodologinės patirtys: Lietuvos kontekstas. *Filosofija. Sociologija*, 21 (4), 340–356.
28. Miltinienė, D., Rugienė, R., Dadonienė, J., Venalis, A. (2008). Klausimyno, skirto reumatoidiniu artritu ir spondiloartropatijomis sergančių pacientų atrankai, lietuviškos versijos adaptacija ir validizacija. *Gerontologija*, 9 (3), 162–170.
29. Pauliukaitė, Ž. (2009). Apklausos instrumentų kalbinė adaptacija tarpkultūriniuose tyrimuose. *Tarpdisciplininis diskursas socialiniuose moksluose – 2: Socialinių mokslų doktorantų ir jaunųjų mokslininkų konferencijos straipsnių rinkinys*, 85–91. Kaunas: Kauno technologijos universitetas.
30. Riklikienė, O. (2007). Tradicinių metodologinių reikalavimų pritaikymas slaugos rezultatų klasifikacijos anketos adaptavimui lietuvių kalbai, kultūrai ir slaugos praktikai. *Sveikatos mokslai*, 17 (5), 30–36.
31. Riklikienė, O. (2009a). Klausimynų pritaikymo naujoje kultūroje problematika ir išmoktos pamokos. *Tarpdisciplininis diskursas socialiniuose moksluose – 2. Socialinių mokslų doktorantų ir jaunųjų mokslininkų konferencijos straipsnių rinkinys*, 92–98. Kaunas: Kauno technologijos universitetas.
32. Riklikienė, O. (2009b). Siektini slaugos rezultatai ir jų praktinis pritaikymas palaikomojo gydymo ir slaugos ligoninėje (Daktaro disertacija, Kauno medicinos universitetas).
33. Rugienė, R., Dadonienė, J., Venalis, A. (2005). Gyvenimo kokybės klausimyno adaptavimas, jo tinkamumo kontrolinei grupei ir reumatoidiniu artritu sergantiems ligoniams įvertinimas. *Medicina*, 41 (3), 232–239.
34. Shiaffino, R., Zearo, F. (2005). Translation Quality Measurement in Practice. *Proceedings of the 46th Annual Conference of the American Translation Association*, Seattle. Available online at: http://www.translationquality.com/files/TQM_in_Practice.pdf.
35. Survey Research Center. (2010). Guidelines for Best Practice in Cross-Cultural Surveys. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. Available online at: http://www.ccsg.isr.umich.edu.
36. Tafas, C. A., Patiraki, E., McDonald, D. D., Lemonidou, C. (2002). Testing an Instrument Measuring Greek Nurses' Knowledge and Attitudes Regarding Pain. *Cancer Nursing*, 25, 8–14.
37. Turčinskaitė-Balčiūnienė, A., Merkys, G. (2010). Assessing Higher Education Quality in the Context of Conceptual Confusion and Methodological Eclecticism. *Social Research*, 4 (21), 70–81.

38. Vanagas, T., Gulbinas, A., Johnson, C., Pundzius, J., Barauskas, G. (2006). Translation and Validation of the EORTC QLQ-PAN26 Quality of the Life Questionnaire for Patients with Pancreatic Cancer. *Acta Medica Lituanica*, 13 (4), 272–275.
39. Veličkaitė, K., Matonytė, I., Morkevičius, V. (2009). Europinės ir tautinės tapatybės kriterijai elito narių požiūriu. *Sociologija. Mintis ir veiksmas*, 2 (25), 92–104.
40. Yi-Hsiu, L., Chen-Yueh, Ch., Ping-Kun, Ch. (2005). Cross-Cultural Research and Back-Translation. Available online at http://www.thesportjournal.org/article/cross-cultural-research-and-back-translation.
41. Žvirdauskas, D. (2006). Mokyklos vadovo lyderystės ypatumai švietimo subjektų interesų ir mokyklos efektyvumo raiškos aspektu (Daktaro disertacija, Kauno technologijos universitetas).
42. Žydžiūnaitė, V. (2003). Komandinio darbo kompetencijų edukacinė diagnostika ir jų vystymo, rengiant slaugytojus, pagrindimas (Daktaro disertacija, Kauno technologijos universitetas).

Pauliukaitė Ž., Turčinskaitė-Balčiūnienė A., Balčiūnas J.

Apklausos instrumentų vertimo kokybė: Motyvacijos ir Akademinių studijų kokybės klausimyno įvertinimas

Santrauka

Teoriniai tyrimo pagrindai. Visuotini žmonijos modernėjimo ir globalizacijos procesai socialinių mokslų tyrėjus vis sparčiau įtraukia į tarptautinio ir tarpkultūrinio palyginimo tyrimus. Lietuva ne išimtis – nuoseklūs bandymai adaptuoti užsienyje sukurtus klausimynus šiandien aptinkami bemaž visose Lietuvos socialinių mokslų srityse: dalyvaujama tokiuose plataus masto tyrimuose kaip Pasaulio vertybių tyrimas, Eurobarometras, Europos socialinis tyrimas, Tarptautinės švietimo pasiekimų vertinimo asociacijos (IEA) vykdomuose tarptautiniuose mokymosi pasiekimų tyrimuose, medikų atliekamuose gyvenimo kokybės, slaugos tyrimuose ir pan. Vis tik Lietuvoje pastebima tendencija perkeltą apklausos instrumentą taikyti ne visada metodologiškai korektiškai. Kai kurie autoriai nenurodo nei naudotos vertimo technikos, nei klausimyno psichometrinių charakteristikų, gautų Lietuvos populiacijose, t. y. faktorinės validacijos rodiklių, pakartotinių matavimų rodiklių (Merkys, Pauliukaitė, 2010). Kai kurie mokslininkai, pristatydami tyrimo instrumentų parengimą, daugiausia dėmesio skiria psichometriniams matavimams ir įvertinimams, kai instrumento vertimo procedūra ir kalbinės, kultūrinės ir praktinės adaptacijos procesas aprašomas labai glaustai (Riklikienė 2007). Siekiant išvengti šių nekorektiškumų, tyrėjams, naudojantiems verstinį instrumentą, siūlytina visuomet detaliai aprašyti taikytą vertimo techniką ir pateikti šio klausimyno tikslinėje populiacijoje gautas psichometrines charakteristikas.

Straipsnyje analizuojama mokslinė problema formuluojama tokiu klausimu: kaip užtikrini ir įvertinti perkelto apklausos instrumento vertimo kokybę? Vertimo kokybės įvertinimas apima įvairius kokybinius (interviu, fokus grupės, atgalinis vertimas, etc.) ir kiekybinius (statistinė analizė) metodus. Kada ir kokie metodai taikytini, lemia daugybė veiksnių, tačiau teorinių ir empirinių mokslinių straipsnių analizė atskleidžia, jog klausimyno vertimas turėtų būti įvertintas tiek kokybiniais, tiek kiekybiniais metodais. Kokybinė prieiga dažniausia pasitelkiama apklausos instrumento vertimo „darbinėje“ stadijoje, siekiant užtikrinti galutinio klausimyno kokybę, o kiekybine prieiga naudojamasi analizuojant surinktus duomenis.

Vertinant išversto ir adaptuoto klausimyno aiškumą ir tinkamumą, tikslinėje populiacijoje dažnai atsižvelgiama į psichometrines skalių charakteristikas, t. y. analizuojamas skalių vidinis nuoseklumas, patikimumas, validumo rodikliai. Kronbacho alfa koeficientas gali atskleisti, kuriuos testo žingsnius reikėtų iš naujo peržiūrėti, koreguoti vertimą. Jei adaptavus klausimyną tikslinėje populiacijoje gaunami labai aukšti psichometriniai rodikliai, toks rezultatas patvirtina, kad perkėlimo (kartu ir vertimo) procedūra pavyko.

Šio straipsnio tikslas – aptarti apklausos instrumentų vertimo kokybės įvertinimo problematiką ir, remiantis mokslo publikacijose teikiamomis rekomendacijomis bei pastebėjimais, atlikti QUISS II[5] metodologijos *Motyvacijos* ir *Akademinių studijų kokybės* klausimyno vertimo kokybės analizę.

Tyrimo metodika. QUISS apklausos instrumentą sudarė profesoriaus Werner Georg vadovaujama mokslininkų grupė iš Konstancos universiteto (Vokietija). Metodika sudaryta 1983 m., pakartotinai naudota atliekant Vokietijos studentų nuomonės apklausas ir tobulinta. 2000–2002 m. viena paskutinių jos versijų QUISS I naudota tarptautinėje studentų apklausoje, kurioje dalyvavo studentai iš šešiolikos universitetų (4 037 respondentai iš Vokietijos, Ispanijos ir Prancūzijos). 2009 m. paskutinė QUISS II versija išversta į lietuvių kalbą ir naudota Lietuvos studentų apklausoje, įgyvendinant mokslininkų grupės projektą: „Akademinių studijų kokybės ir socialinio konteksto tyrimas“. Projektą rėmė Lietuvos valstybinis mokslo ir studijų fondas, projekto vadovas – profesorius Gediminas Merkys (Merkys et al., 2009; Turčinskaitė-Balčiūnienė and Merkys, 2010). Popieriaus-pieštuko klausimynas sudarytas iš psichometrinių Laikerto tipo skalių, turinčių nuo trijų iki devynių padalų. Klausimyną užpildė 964 respondentai iš Kauno, Klaipėdos, Šiaulių ir Vilniaus aukštojo mokslo institucijų (universitetų ir kolegijų).

Šiame straipsnyje aptartos dvi apklausos instrumentų vertimo kokybės įvertinimo prieigos: kokybinė – ekspertinis vertinimas ir kiekybinė – psichometrinių charakteristikų analizė (patikimumas ir konstrukto validumas) 34 *Motyvacijos* ir *Akademinių studijų kokybės* skalių, sudarytų iš 297 testo žingsnių. 42 testo žingsniai sudaro 5 *Motyvacijos* skales: aukštosios mokyklos pasirinkimo

[5] Qualitätsverbesserung in Schulen und Schulsysthemen II.

veiksniai – 8 testo žingsniai; studijų pasirinkimo veiksniai – 7 testo žingsniai; studijų prasmingumas – 10 testo žingsnių; motyvacija daktaro laipsniui siekti – 11 testo žingsnių; studijavimo užsienyje planai – 6 testo žingsniai. 255 testo žingsniai sudaro 29 *Akademinių studijų kokybės* skales, kurios apima tokias pagrindines temas: „Informacijos pasiekiamumas“ – 3 skalės (29 testo žingsniai); „Probleminės sritys“ – 3 skalės (42 testo žingsniai); „Studijos ir dėstymas“ – 7 skalės (76 testo žingsniai); „Bendravimas ir konsultavimas“ – 7 skalės (30 testo žingsniai); „Tobulėjimas“ – 6 skalės (50 testo žingsnių); „Europos aukštojo mokslo erdvės kūrimo aspektai“ – 3 skalės (28 testo žingsniai).

Klausimyno psichometriniai rodikliai (patikimumas ir konstrukto validumas) analizuoti naudojant statistinį paketą „SPSS 13.0“. Analizuoti šie rodikliai: 1) skalių vidinis nuoseklumas apskaičiuojant Kronbacho alfa bendrus koeficientus; 2) skalių vidinis nuoseklumas, pašalinus iš skalės atskirus testo žingsnius, apskaičiuotos dalinės Kronbacho alfa kiekvieno testo žingsnio atveju; 3) skalės testo žingsnių pastovumas įvertinant skalės ir jos testo žingsnio koreliaciją; 4) skaliu komponentiškumas atliekant faktorinę analizę.

Tyrimo rezultatai. Klausimynas verstas remiantis kokybine vertimo kokybės prieiga – ekspertiniu kokybės vertinimu. Klausimyno vertimas ir kultūrinė adaptacija atlikta per seminarus, kuriuose trijų ekspertų komanda ieškojo geriausio vertimo varianto, analizuodama trijų vertėjų, nepriklausomai vienas nuo kito atliktus klausimyno vertimus iš anglų, prancūzų ir vokiečių kalbų. Klausimyno vertimo kokybę užtikrinus ekspertiniu kokybės vertinimu, atlikti kiekybiniai vertimo kokybės matavimai. Išversto klausimyno skalių vidinis nuoseklumas (patikimumas) įvertintas apskaičiavus Kronbacho alfa koeficientą: kai Kronbacho alfa reikšmė buvo 0,80 ir aukštesnė, vidinis nuoseklumas laikytas pakankamu. Atsižvelgta į Kronbacho alfa reikšmės kitimą, pašalinant atskirus testo žingsnius iš skalės: kai Kronbacho alfa reikšmės svyravo ties bendrąja Kronbacho alfa reikšme iki 0,5 punktų, vidinis nuoseklumas laikytas pakankamu. Skalės-testo žingsnio pastovumas įvertintas apskaičiavus skalės ir testo žingsnių koreliacijos koeficientus: kai ši reikšmė buvo aukštesnė už 0,30, laikyta, jog testo žingsniai pakankamai atspindi matuojamą skalės konstruktą. Atlikus faktorinę analizę (konstrukto validumo patikrinimą), paaiškėjo, kad 9 skalės yra daugiakomponentės, šių skalių rodikliai skaičiuoti ir analizuoti atskirai kiekvienam faktoriui.

Išvados. Vienoje sociokultūrinėje terpėje sukurtų apklausos instrumentų perkėlimas į kitą sociokultūrinę terpę yra sudėtinga procedūra. Vienas problemiškiausių, daug diskusijų socialinių tyrimų metodologinėje literatūroje keliančių etapų – kaip užtikrini ir patikrinti perkeliamo apklausos instrumento vertimo kokybę.

Vertimo kokybės įvertinimas apima įvairius kokybinius ir kiekybinius metodus. Kada ir kokie metodai taikytini, lemia daugybė veiksnių, tačiau teorinių ir empirinių mokslinių straipsnių analizė atskleidžia, jog klausimyno vertimas turėtų būti įvertintas tiek kokybiniais, tiek kiekybiniais metodais.

Norint patvirtinti perkelto klausimyno metodologinę kokybę, būtina visuomet šsamiiai aprašyti taikytą vertimo techniką ir pateikti šio klausimyno tikslinėje populiacijoje gautas psichometrines charakteristikas.

QUISS II klausimyno vertimo kokybė įvertinta ekspertiniu vertinimu (kokybinis metodas) ir kiekybiniais metodais, analizuojančiais psichometrines skalių charakteristikas.

Ekspertų komanda įvertino klausimyno vertimo kokybę ir nusprendė, jog klausimynas tinkamas apklausai atlikti, bandomasis tyrimas neatliktas.

Pagal visus analizuotus statistinius rodiklius 23 skalių vertimo kokybę galima pripažinti gera: 199 testo žingsniams (67 proc.) gauti aukšti psichometriniai rodikliai (patikimumas ir konstrukto validumas).

Faktorinė duomenų analizė atskleidė 9 skalių daugiakomponentiškumą, dar 2 skalių duomenims taikyti faktorinės analizės nerekomenduojama, nes jų Kaizerio-Mejerio-Olkino matai santykinai žemi. Pastarųjų 11 skalių vertimo kokybės įvertinimui būtina lyginti duomenis, surinktus originaliu klausimynu Vokietijos studentų populiacijoje, su duomenimis, gautais išverstu į lietuvių kalbą klausimynu apklausus Lietuvos studentus.

Rekomenduojama greta tolesnių kiekybinių klausimyno vertimo kokybės įvertinimo metodų, analizuojančių psichometrinius rodiklius, peržiūrėti iš naujo ir performuluoti netiksliai išverstus testo žingsnius, naudojant kokybinius vertimo įvertinimo metodus.

Pagrindiniai žodžiai: apklausos instrumentų vertimo kokybė, vertimo kokybės įvertinimo kokybiniai metodai, vertimo kokybės įvertinimo kiekybiniai metodai, ekspertinis įvertinimas, psichometriniai rodikliai, QUISS II *Motyvacijos* ir *Akademinių studijų kokybės* klausimynas.

The article has been reviewed.
Received in January, 2011; accepted in April, 2011.