



Global Conference on Business and Economics Research (GCBER) 2017
14-15 August 2017, Universiti Putra Malaysia, Malaysia

Measuring International Students' Satisfaction: the Development of Survey Instrument

Pui-Yee Chong^{a*}, Siew-Yean Tham^b, Jia-Yi Andrew Kam^b

^a *Universiti Tenaga Nasional, Malaysia*

^b *Universiti Kebangsaan Malaysia, Malaysia*

Abstract

Survey questionnaire is the most common instrument used in social science research. Whether the approach is “adapt & adopt” from currently available instrument or if it is newly developed, researchers would need to test the validity and reliability of the instrument. There are advantages in the “adapt & adopt” approach, however developing one own survey questionnaire has its benefit, it helps to ensure the currency or up-to-date construct and applicability to the context of the study, for example in a particular country or location of study. Hence, this paper aims to share the process of the development of survey instrument to measure international students' satisfaction. The development of the survey instrument involved four stages; (1) Conceptualization, (2) Development, (3) Validity and (4) Reliability. The first stage which is the conceptualization stage involved the review of literatures and interviews with 20 international students. The results from the review of literatures and interviews serve as the input for the second stage, which is the writing of the items in the questionnaire. Researchers organized and wrote the questions based on constructs found from literature reviews and interviews in the form of items. Upon completion of the survey questionnaire, proof reading and pre-test were conducted to validate the survey questionnaire. Subsequently, in stage four, a pilot study was executed at a private university located in Klang Valley. Data from pilot test were used to test the reliability of the instrument using the Cronbach's Alpha test on each construct. After the removal and re-grouping of constructs, the final questionnaire has 48 constructs and 186 items. Finally, the second pre-test of the instrument was conducted at two private universities, before the instrument was ready for data collection. It is hope that the sharing of this process will help and provide novice researchers some guidelines in the development of a survey instrument.

Keywords: International students, measuring satisfaction, survey instrument, piloting

1. INTRODUCTION

Researchers in general tend to perceive quantitative research as easier to conduct and analyze as compared to qualitative research (Summers, 2001). The most common quantitative research method is through survey questionnaires. Particularly in the context of social science research, questionnaires are the most widely used method in data collection (Rowley, 2014; Hinkin, 1998; Radhakrishna, 2007). Survey questionnaire is suitable to be used as data collection method if it is used to conduct (1) profiling and descriptive research, (2) analytical and predictive research to identify relationships among variables using multiple regression and SEM and (3) developing and testing measurement scale using EFA and CPA (Rowley, 2014). It is also used to gather data from

*Corresponding author. Tel.: +0-603-89212020; Fax: +0-603-89212161
E-mail: pychong@uniten.edu.my

large sample size (Rowley, 2014). Developing a new survey questionnaire posed many questions and fear to novice researchers. Novice researchers have always been advised to “adapt and adopt” past survey instrument developed by other researchers. The decision whether to develop a new instrument or to use “adapt & adopt” approach will depend on the objective and contribution of the study. Developing one own survey questionnaire has its benefit, it helps to ensure the currency of the construct or if the construct is up-to-date, as instrument of 5 years old might be outdated (John W. Creswell, 2008) and may not be applicable to the context of the study, for example in a particular country or location of study. In the context of this paper, it is Malaysia. Some reasons why new researchers do not attempt to develop their own survey instrument include (1) the lack of knowledge on how to, (2) the lack of confidence in producing a valid and reliable instrument and (3) do not want to risk any measurement error. This paper targets the novice researchers and aims to share the process of developing survey instrument to measure international students’ satisfaction. It is hoped that the sharing of this process will help and provide young researchers some guidelines in the development of a survey instrument. The discussion of this paper will be organized in a sequential order of a survey instrument development process.

2. SIGNIFICANCE OF THE STUDY

There are limited papers that discuss the detail development of survey instrument in measuring international students’ satisfaction. Majority discussed the broad view of methodology, such as sampling, sample size and data administration (Hishamuddin Fitri Abu Hasan, Azleen Ilias, Rahida Rahman, 2008; Ong Chee Hui, 2014). Some researchers would adopt and adapt from available instrument for example SERVQUAL in measuring service quality (Arambewela, 2003) and HEdPERF (Abdullah, 2006). To fill this gap, researchers reckon the need to document and report the detail of instrument development in measuring international students’ satisfaction in the context of Malaysia.

3. METHODOLOGY

Researchers used mixture of methods and approaches in the instrument development process. These include literature review, interviews with international students, engagement with experts, survey during piloting and finally statistical tools used to test reliability of instrument. The detail elaboration on each methods used is found in section 4 of this paper.

4. INSTRUMENT DEVELOPMENT PROCESS

The development of a valid and reliable instrument is imperative to reduce measurement error (Uma Sekaran, 2003; Hinkin,1998). Hence, this instrument development process aims to reduce measurement error. With that, researcher proposed four stages in the development of survey questionnaires, Stage 1-Conceptualization; Stage 2–Development; Stage 3–Validity Test and finally Stage 4-Reliability Test. Within each stage, there are several steps in the development of a survey questionnaire in measuring international students’ satisfaction in Malaysia. Fig. 1 shows the instrument development process.

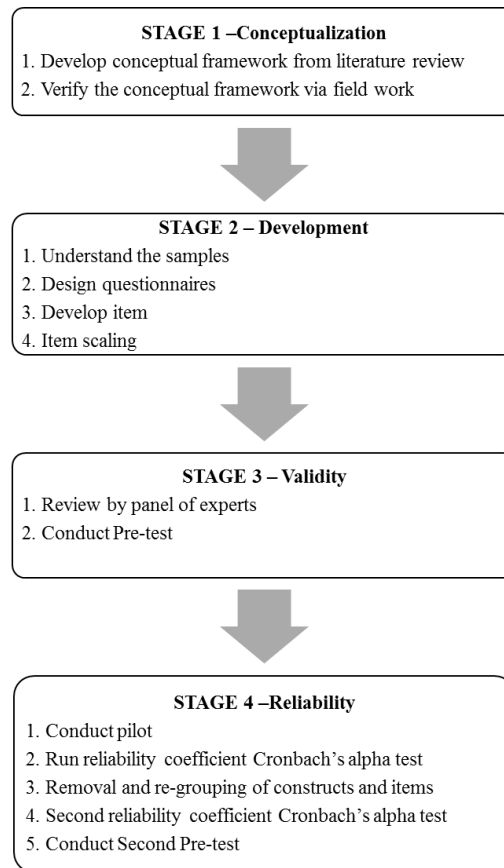


Fig. 1 : Instrument Development Process

5. STAGE 1 – CONCEPTUALIZATION

5.1 Developing Conceptual Framework

Researchers have identified the conceptual framework in measuring international students' satisfaction through rigorous review of forty two primary literatures and thirty two secondary literatures. A framework was developed in measuring factors that influence international students' satisfaction (Chong, 2015). In the framework, researchers have identified five independent domains which include (1) Academic Service on Campus, (2) Non-Academic Service on Campus, (3) External Living Environment out of Campus, (4) Image and Brand of University and Country and (5) Perceived Value or Value-for-Money. Each of the domain consists of variables, a total of 36 variables were identified. In addition to the literature on international students, literatures on measuring customer satisfaction were studied. With these two theories, formed the initial proposed model on factors influencing international student's satisfaction, for detail literature review on the development of this model, refer to Chong (Chong, 2015).

5.2 Verifying the Conceptual Framework

Subsequently, interviews were conducted to verify the variables and to identify new variables that are relevant to the time and context of Malaysia. A total of 20 international students were interviewed from 11 private universities located in Klang Valley. International students who participated in the one-to-one semi-structured interview came from 14 different nationalities. The outcome from the field work has confirmed the 5 domains and additional of 19 variable, thus making it a total of 58 variables. The detail of the improved model is found in Fig. 2.

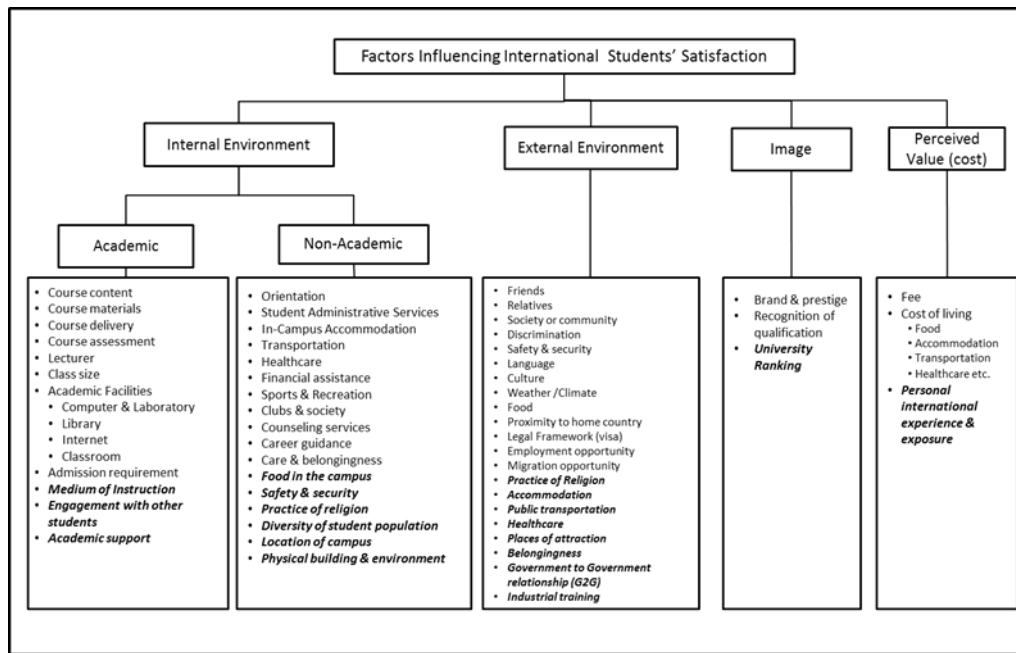


Fig. 2: Improved Model of Factors Influencing International Students' Satisfaction

6. STAGE 2 – DEVELOPMENT

6.1 Understanding the Samples

Before the attempt to develop the survey instrument, researcher needs to understand the background of their sample for the study (Radhakrishna, 2007; Rowley, 2014). As for the discussion of this paper, the targeted samples are the international students studying at private universities located in the Klang Valley in Malaysia. International students come from diverse backgrounds, some may come from countries where English is commonly spoken while others may not. However, in the case of Malaysia, among the top sending countries are China, Indonesia, Middle East & Northern Africa (MENA) (Verbik, L. & Lasanowski, 2007) where English is not their first or second language. Researcher decided to use English language in the survey because it is the common language that all international students will understand. This is because the medium of instruction at private universities in Malaysia is in English language. In catering to the diverse levels of English proficiency, the questions in the survey were crafted in the simplest English to ensure that it is easy for all respondents to understand and therefore reap a more valid and reliable answers with higher response rate.

6.2 Designing the Questionnaire

Before writing the questions or items, researcher must first plan the layout and organization of the survey questionnaire. Typically, there are two broad sections; (1) basic demographic or profile of respondents, and (2) items that measure variables or constructs. Some basic guidelines in developing item which include statement should be simple, as short as possible and language used should be familiar to target respondents (Hinkin, 1998; Rowley, 2014). It is advisable not to use negatively worded statement to reduce comprehension error and reading time (Buttle, 1996). The survey measuring international students' satisfaction and factors that influence satisfaction has two parts. Part 1 are questions to obtain respondents' profile such as gender, country of origin, years in Malaysia and institution of learning and field of study and others. In Part 2, it has 7 Sections to capture six independent domains (1) Academic Service on Campus, (2) Non-Academic Service on Campus, (3) External Living Environment out of Campus, (4) Image and Brand of University and Country, (5) Perceived Value or Value-for-Money, (6) Word-of-Mouth and (7) Rank Factors that Influence Satisfaction.

6.3 Developing Items

Questions developed were anchored based upon the conceptual framework derived from theories available in literatures (Chua Yan Piaw, 2006; Hinkin, 1998; Mackenzie, Podsakoff, & Podsakoff, 2011; Rowley, 2014) and fieldworks. A total of 59 variables with 204 items were developed. All these items were crafted based on the

theoretical construct from Stage 1-Conceptualization and reference to items developed by other researchers. There is no hard-and-fast rules on the number of items (Hinkin, 1998). There should be at least four items to test each construct, however minimum of three items are adequate to achieve internal consistency reliability (Hinkin, 1998). For example in the survey to test the construct “A1-Course/Syllabus/Subject”. There are 4 items and these include “Item A1.1-The program that I register for has syllabus/subjects relevant to my future job and career prospect”; “Item A1.2- The program that I register for has syllabus/subject that increase my knowledge and interest”; “Item A1.3-The program that I register for has syllabus/subject that is of international standard”; “Item A1.4-The program that I register for has syllabus/subject that is up-to-date”. All these items formed the different aspects or dimensions are correlated to measure the construct “Course/Syllabus/Subject”. Researcher should ensure that the time taken to complete a survey should not be too long to reduce response bias caused by boredom or fatigue (Hinkin, 1998). Trying to keep about 3-4 pages of questionnaire for the general population. However, questionnaires of 15 pages may be possible for highly educated respondents and a salient topic (W. Lawrence Neuman, 2006).

6.4 Item Scaling

The type of scaling techniques used will determine the type of analysis that researcher planned to conduct (Rowley, 2014). For example, for nominal or ordinal data, researcher could use Chi-Square or Spearman rho test. However if the type of data is ratio or interval, then researcher may need to use T-Test or ANOVA (Chua Yan Piaw, 2006). Therefore, before deciding the scale when developing the questionnaire, researcher needs to be clear on what type of analysis technique should be conducted once data is collected to answer its research question.

Although there is a variety of different scaling techniques available such as dichotomous scale, category scale, Likert scale, numerical scale, itemized rating scale and many more (Uma Sekaran, 2003), Likert-type scales is the most frequently used in survey questionnaire research (Eutsler & Lang, 2015; Hinkin, 1998). It is suitable to be used when conducting behavioral research and factor analysis (Hinkin, 1998). On top of that, it is easy for respondents to fill out and hence increase response rate. There are many different point scales in the Likert-type scale, these include 3-point, 4-point, 5-point and 7-point (Vagias, 2006). It is recommended to use either 5- or 7-point scale, as there found no significant difference in means, variance and measures of normality. Variance appears maximized with 7-point scales, additional points do not increase variance (Eutsler & Lang, 2015). Researchers decided to use 5-point Likert scale in this research as it would reduce the frustration and confusion level of respondents and increase response rate (Buttle, 1996; W. Lawrence Neuman, 2006) moreover there is no significant difference statistically.

7. STAGE 3 – VALIDITY TEST

7.1 Panel of Experts

Once the questionnaire has been developed, the next stage would be to validate the instrument. Validity is defined as “the degree to which a measure accurately represents what it is supposed to” (Joseph F. Hair, William C. Black, 2010). There are a few validity tests, these include three broad categories, content, criterion-related and construct validity (Hinkin, 1998; John W. Creswell, 2008; Radhakrishna, 2007; Uma Sekaran, 2003). Content validity ensure that items created are adequate and representative to measure the construct (Uma Sekaran, 2003). A panel of experts can attest to the content of the instrument. The experts examine the information about (1) the objective of the instrument, where the questions are comprehensive enough to collect all the information needed to answer the purpose and goal of the study, (2) the content areas where it measures what it is intended to measure, (3) the level of difficulty of the questions that is appropriate for the sample and finally and (4) if the instrument looks like a questionnaire (John W. Creswell, 2008; Radhakrishna, 2007). In the case of the international students’ satisfaction instrument, a total of 10 academicians were invited to review and judge the content of the questionnaire. Feedbacks obtained served as improvement to the questionnaire. For researchers who are interested, detailed explanation on criterion-related validity and construct validity can be found from Creswell (John W. Creswell, 2008).

7.2 Pre-Test

A field test or pre-test is used to examine how well the questionnaire works (Rowley, 2014; Shelby D. Hunt, Richard D. Sparkman Jr., 1982). It is an important step to gain feedback from respondents of the overall structure, organization and content of the survey. During pre-test, it helps to ascertain any questions that seem ambiguous to respondents to test their understanding and interpretation of each question. Besides these aims, it helps researcher to estimate time taken to answer the survey.

When conducting the pre-test, personal interview is highly recommended (Shelby D. Hunt, Richard D. Sparkman Jr., 1982), as researcher may be able to observe cues such as reaction, hesitation from body language and facial expression and therefore able to obtain immediate feedback from respondents. Besides making observation, researchers may ask two important questions (1) what are other ways to interpret the questions and (2) why they answer in such a way (Summers, 2001).

For this research, ten international students from a private university in Klang Valley were invited to pre-test the questionnaire. They were asked to fill out the questionnaire while researcher made observations. It was conducted on a one-to-one session with each respondent. Feedbacks and ambiguities were identified and rephrasing of the items was done. For instance, the item "Item A3.2-There is a two-way communication between lecturers and students in the class". The word "two-way communication" is not clear to them. Hence to rephrase "There is a two-way communication between lecturers and students in the class. (e.g. I can ask questions in the class)". Another example "Item B6.2-The healthcare providers have competent medical officers". Respondents have difficulty understanding the word "competent". Hence, researcher has added descriptions to clear the ambiguity to "The clinic on my campus has competent (qualified & good) doctors or pharmacist".

8. STAGE 4 – RELIABILITY TEST

"Reliability means that scores from an instrument are stable and consistent" (John W. Creswell, 2008). In other words, when researcher administers the survey, the results or score would be consistent when you conduct the same survey at another time, the questionnaire consistently measures whatever it measures (Radhakrishna, 2007; Uma Sekaran, 2003). There are a few types of reliability tests, researcher can use one or more types of procedures to test reliability. For the purpose of this research, the Cronbach's coefficient alpha was used to test internal consistency. It is to test if the degree of items are independent measures of the same construct and if these items are correlated to each other (Uma Sekaran, 2003). If item score is in a form of continuous variable, e.g. strongly agree, to strongly disagree, alpha provides a coefficient to estimate consistency of scores on an instrument (John W. Creswell, 2008).

8.1 Pilot Study

Reliability test can only be achieved through a pilot study. A pilot study is a trial run by administering questionnaire for about 30 respondents that are not included in the sample (Chua Yan Piau, 2006; Radhakrishna, 2007; Summers, 2001). Data collected from the pilot are then analyzed using Statistical Package for Social Sciences (SPSS) or other similar software packages. One of the valuable functions provided by SPSS is to test reliability of items to measure a particular construct. The reliability coefficient Cronbach's alpha score can range from 0 to 1, with 0 representing an instrument with full of error and 1 representing total absence of error (Radhakrishna, 2007). Cronbach's alpha score is deemed reliable when score is 0.70 and above (John W. Creswell, 2008; Nancy L. Leech, Karen C. Barrett, 2005). A pilot study was conducted on 45 international students at a private university located in Klang Valley.

8.2 Run Reliability Coefficient Cronbach's Alpha Test

Data collected were analyzed using SPSS. Each construct was tested for its reliability. For example, construct "B15-Diversity of Student Population", researchers selected all the 3 items, "Item B15.1-My university has many international students"; "Item B15.2-My university has international students from different countries"; and "Item B15.3-My university provides opportunities for interaction between local and international students" to run Cronbach's alpha test. This test provided three important outputs (1) Cronbach's Alpha score to the construct, (2) Corrected Item-Total Correlation and (3) Cronbach's Alpha if Item Deleted. The Cronbach's alpha score for construct "B15-Diversity of Student Population" is .730. While the score for each item for Corrected Item-Total Correlation is above .40, this indicates that all the three items are moderately correlated with most of the items and will make a good component of this summated rating scale. Finally, the score for Cronbach's Alpha if Item Deleted for each item is .760, .643 and .482. As the Cronbach's Alpha for the construct is .730, researchers decided to keep all the three items without deleting any of the item. Even by deleting item B15.1 to increase alpha score to .760 would not make significant improvement.

8.3 Removal and Re-grouping of Constructs and Items

It is important to take note that the degree of reliability increases as the items are added to the instrument (John W. Creswell, 2008). This is because with multiple items, it tends to be more stable, reliable and reflective than one item measure (Summers, 2001; W. Lawrence Neuman, 2006). For example, there are 9 items in the "A8-

Lecturer” construct and it yielded Cronbach alpha score of .902 as compared to three items for construct A12-Internet Access & Connection with Cronbach Alpha score of .891. Reliability test could not be run if there is only 1 item to measure a construct. In this research, the construct “B9-Counseling Service” having only 1 item “My university provides counseling services for students when they have problems” and the construct “B13-Practice of Religion on the Campus” with only 1 item “My university has the environment for me to practice my religion on the campus (e.g. place of worship, fellowship, halal food, vegetarian food, wearing of hijab etc.)”. In treating this issue, researchers have moved both of this single-items to another related construct “B14-Care & Belongingness”. Thereafter, the second time of reliability test was conducted with Cronbach’s alpha score of .797, which met the acceptable level of internal consistency reliability.

8.4 Second Reliability Coefficient Cronbach’s Alpha Test

The first reliability test was run in March 2016 on 59 constructs and 204 items. Based on the results of the first test, items that are redundant were removed and some items were moved to another construct. After this treatment of removal and movement of items, the questionnaire has 48 constructs and 186 items. Second reliability test was run in April 2016 after the removal and movement of items. The results showed that all the constructs obtained Cronbach’s alpha score of 0.7 and above except the construct “A5-Admission Requirement” with score of .650. This score is consider acceptable as it is above medium correlation score of .50.

8.5 Second Pre-test

The questionnaire was updated based on the removal and movement of variables and items. Subsequently, a second pre-test was carried out with 4 international students from two private universities to increase the confidence level of the questionnaire and the administration of the questionnaire. No additional feedback was recorded from the second pre-test. However, it was observed that the duration to complete the questionnaire has reduced from 40 minutes to 28 minutes. Upon completion of this stage, the questionnaire is ready for empirical data collection.

9. CONCLUSION

It is hoped that this paper has provided some guidelines on the procedure and systematic approach in developing survey instrument for novice researchers. A valid and reliable instrument is imperative to reduce measurement error and overall, improve the quality of research. The four stages of survey instrument development include (1) Conceptualization; (2) Development; (3) Validity and (4) Reliability are the basic rigor a researcher must anticipate in order to develop a valid and reliable instrument.

ACKNOWLEDGEMENTS

The authors acknowledge and thank the Ministry of Higher Education in providing FRGS fund (FRGS/1/2016/SS01/UKM/02/1) for this research, and the support from Universiti Tenaga Nasional (UNITEN) and Universiti Kebangsaan Malaysia (UKM). Appreciation also goes to Ms Evelyn Yeap Ewe Lin who has helped in the proofreading of the article.

REFERENCES

- Abdullah, F. (2006). Measuring service quality in higher education: HEDPERF versus SERVPERF. *Marketing Intelligence & Planning*, 24(1), 31–47. <https://doi.org/10.1108/02634500610641543>
- Arambewela, R. A. (2003). *Post-choice satisfaction of international postgraduate students from Asia studying in Victorian universities*. PhD Thesis, Victoria University
- Buttle, F. (1996). SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, 30(1), 8–32. <https://doi.org/10.1108/03090569610105762>
- Chong, P.-Y. (2015). Internationalisation of Higher Education : Proposed Framework on International Students eTM Satisfaction . *Pertanika Journal of Social Sciences and Humanities*, 23(Special Issue), 73–90.
- Chua Yan Piaw. (2006). *Kaedah Penyelidikan : Kaedah dan Statistik Penyelidikan Buku 1*. Kuala Lumpur: McGraw-Hill Education.
- Eutsler, J., & Lang, B. (2015). Rating scales in accounting research: The impact of scale points and labels. *Behavioral Research in Accounting*, 27(2), 35–51. <https://doi.org/10.2308/bria-51219>
- Hinkin, T. R. (1998). A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods*, 1(1), 104–121.
- Hishamuddin Fitri Abu Hasan, Azleen Ilias, Rahida Rahman, M. Z. A. R. (2008). Service Quality and Student Satisfaction : A Case Study at Private Higher Education Institutions. *International Business Research*, 1(3), 163–175. <https://doi.org/10.5539/ibr.v1n3p163>
- John W. Creswell. (2008). *Educational Research : Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (3rd Editio). New Jersey: Pearson Education.
- Joseph F. Hair, William C. Black, B. J. B. R. E. A. (2010). *Multivariate Data Analysis* (7th Editio). USA: Pearson Prentice Hall. <https://doi.org/10.1016/j.ijpharm.2011.02.019>

- Lim, E. K. S. (2013). A Qualitative Study of Factors Contributing to International Students' Satisfaction of Institutional Quality. *Asian Social Science*, 9(13), 126–131. <https://doi.org/10.5539/ass.v9n13p126>
- Mackenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct Measurement and Validation Procedures in MIS and Behavioral Research : Integrating New and Existing Techniques. *MIS Quarterly*, 35(2), 293–334.
- Nancy L. Leech, Karen C. Barrett, G. A. M. (2005). *SPSS for Intermediate Statistics : Use and Interpretation* (2nd Editio). New Jersey: Lawrence Erlbaum Associates.
- Ong Chee Hui. (2014). Determinants of International Student's Satisfaction in UUM 2014. Master Dissertation Universiti Utara Malaysia
- Radhakrishna, R. B. (2007). Tips for Developing and Testing Questionnaires/Instruments. *Journal of Extension*, 45(1).
- Rowley, J. (2014). Designing and using research questionnaires. *Management Research Review*, 37(3), 308–330. <https://doi.org/10.1108/MRR-02-2013-0027>
- Shelby D. Hunt, Richard D. Sparkman Jr., J. B. W. (1982). The Pretest in Survey Research : Issues and Preliminary Findings. *Journal of Marketing Research*, 15(May), 269–273.
- Summers, J. O. (2001). Guidelines for Conducting Research and Publishing in Marketing : From Conceptualization Through the Review Process. *Journal of the Academy of Marketing Science*, 29(4), 405–415.
- Uma Sekaran. (2003). *Research Methods for Business : A skill Building Approach*. John Wiley & Sons Inc.
- Vagias, Wade M. (2006). *Likert-Type Scale Response Anchors*. Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management.
- Verbik, L. & Lasanowski, V. (2007). International Student Mobility : Patterns and Trends. The Observatory on Borderless Higher Education (OBHE) Report, September. <https://doi.org/10.1080/15393000701488888> September 2013 from <http://www.obhe.ac.uk>
- W. Lawrence Neuman. (2006). *Social Research Methods : Qualitative and Quantitative Approaches* (6th Editio). Boston: Pearson Education Inc.