

Testing the Validity and Reliability of the “Learn, Pick, Flip, Check, Reward” (LPFCR) Card Game in Homophone Comprehension

Alvin Wong & Hamidah Yamat

To Link this Article: <http://dx.doi.org/10.6007/IJARBSS/v10-i1/6803>

DOI: 10.6007/IJARBSS/v10-i1/6803

Received: 22 November 2019, Revised: 21 December 2019, Accepted: 03 January 2020

Published Online: 25 January 2020

In-Text Citation: (Wong & Yamat, 2020)

To Cite this Article: Wong, A., & Yamat, H. (2020). Testing the Validity and Reliability of the “Learn, Pick, Flip, Check, Reward” (LPFCR) Card Game in Homophone Comprehension. *International Journal of Academic Research in Business and Social Sciences*, 10(1), 22–32.

Copyright: © 2020 The Author(s)

Published by Human Resource Management Academic Research Society (www.hrmars.com)

This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen

at: <http://creativecommons.org/licenses/by/4.0/legalcode>

Vol. 10, No. 1, 2020, Pg. 22 - 32

<http://hrmars.com/index.php/pages/detail/IJARBSS>

JOURNAL HOMEPAGE

Full Terms & Conditions of access and use can be found at
<http://hrmars.com/index.php/pages/detail/publication-ethics>

Testing the Validity and Reliability of the “Learn, Pick, Flip, Check, Reward” (LPFCR) Card Game in Homophone Comprehension

Alvin Wong¹ & Hamidah Yamat²

Fakulti Pendidikan, Universiti Kebangsaan Malaysia

Email: alvinwong1893@gmail.com¹, hamidah_yamat@ukm.edu.my²

Abstract

The “Learn, Pick, Flip, Check, Reward” (LPFCR) Card Game was developed to enhance Primary 4 Pupils' comprehension of homophones. This paper reports the pilot test of the instruments used in the implementation of the game. It aims to examine the validity and reliability of the instruments. Four data collection instruments which were the pre and post tests, observation checklist and questionnaire were analysed statistically and thematically through the method of data analysis. 20 Primary 4 pupils of similar characteristics with the actual samples were involved in this pilot study. Findings of the pilot tests revealed that both pre and post tests achieved a high internal consistency and reliability of over, $\alpha < 0.8$. Meanwhile, the observation checklist was analysed thematically and was found to be valid and reliable as similarity in the given themes of observations were obtained. The questionnaire was analysed through SPSS Version 25 and a reliability coefficient of over, $\alpha < 0.8$ was obtained, which also indicated high internal consistency and reliability. These results proved that the instruments are valid and reliable to be used in evaluating the LPFCR card game.

Keywords: Validity, Reliability, Validity and Reliability Test, Data Collection Instruments.

Introduction

Testing the validity and reliability of a data collection instrument is fundamental in ensuring feasibility and consistency of the instrument in measuring an intended outcome of a research (Hazzi & Maldaon, 2015; Kinchin, Ismail, & Edwards, 2018). As the main study intends to measure the pupils' learning comprehension on homophones through the use of LPFCR Card Game, the instruments must first be developed and tested. The development and testing of the instruments are crucial for the main study as most available instruments do not specifically target on measuring learning comprehension on homophones. Although several past studies conducted by Dautrichea, Fiblad, Fievetb, and Christopheb (2018), Kenan and Hakki (2018) and Treiman, Seidenberg, and Kessler (2015) focused on the learning of homophones, the instruments used in those studies mainly gather participants' behavioural interaction and motivational level. Hence, the major purpose of this pilot study is to

determine the validity and reliability of the instruments prior to implementation in the main research. Four instruments were developed to measure the validity and reliability of the LPFCR Card Game. They were a pre and post test, observation checklist and questionnaire. This paper will first discuss literature on development of instruments and testing validity and reliability of instruments. This will be followed by a discussion on the methodological concerns in conducting the pilot study. The findings and discussion of the pilot test are then presented to determine the validity and reliability of the instruments.

Literature Review

Development and Selection of Research Instruments

In order to select and develop research instruments, the researchers must first identify the contextual and psychometric aspects of the intended study that will be conducted (Mimi, Nor, Lai, & Kahiol, 2015). The contextual aspect of a research refers to the setting of the study which includes the objective and purpose of the study, characteristics of the participants and limitations of the research (Omar, 2015; Watson & de Wit, 2018). Meanwhile, the psychometric aspect of a study touches on the construction and the properties of the instrument in functioning and gauging the objectives of the research and also the feasibility of administration (Lin & Tsai, 2017; Zamanzadeh et al., 2015). Both contextual and psychometric aspects of a research have prominent impact on the type of instruments to be implemented in a study. Thus, researchers must develop and select proper instruments to go in line with the study.

Another important aspect to be taken into consideration is the research design of the study. Chu and Ke (2017) and Klenke, Martin and Wallace (2016) agreed that qualitative and quantitative research designs employ different data collection instruments in accordance to the aims of the research. Instruments that are thematically analyzed such as interviews, field notes and observation are mostly applied in qualitative-based research while statistically analyzed instruments such as test and assessments, surveys and questionnaires with Likert scales are employed in quantitative-based research (Charman, 2017; Creswell & Guetterman, 2019; Dick, 2015). On another perspective, associational and interventional researches such as correlational and action research, often implement both qualitative and quantitative instruments (Dick, 2015; Klenke et al., 2016). Depending on the types of research design, researchers should select and develop appropriate instruments to cater to the needs of the study. As the research discussed in this paper employs data analysis methodology, the instruments involved were a pre and post test, observation checklist and questionnaire, in accordance to the characteristics of the participants, limitations of the research and the aim of determining the pupils' comprehension in homophones.

Validity and Reliability of Research Instruments

Validity is defined as the accuracy of an instrument in measuring the anticipated construct within a research (Klenke et al., 2016; Noble & Smith, 2015). In other words, when an instrument measures what it is supposed to measure, then the instrument is valid. Validity includes several forms, namely face validity, content validity, and construct validity and criterion validity. For the purpose of this pilot study, face validity and content validity were emphasized in relation to the four data collection instruments. Face validity on one hand, measures the degree of an instrument at a surface level in

the aspects of appropriateness and suitability in relation to the purpose of a study (Heale & Twycross, 2015). This validity determines whether an instrument has a “face value” that engages the needs of a research. Content validity on the other hand, focuses on the capacity of data collection items to gather, reflect and portray the variables that are measured (Appelman & Sundar, 2016; Mimi et al., 2015). This type of validity addresses the magnitude of each items in an instrument in gauging the context and construct of a research. On the whole, a valid data collection instrument is appropriate in collecting and measuring the outcome based on the context and psychometric aspects of a research.

Meanwhile, reliability is defined as the stability and consistency of scores from an instrument (Braun, Clarke, Hayfield, & Terry, 2019; Sharifah, Jamal, & Hamidah, 2017). This means that instruments that are reliable can be used several times in different timeline and produce explicit and consistent result. There are several measures of testing reliability which are, test-retest, equivalent form, internal consistency and reliability statistics. In this pilot study, the internal consistency and reliability statistics were used. On the aspect of internal consistency, an instrument is reliable when the items are consistent in measuring the same construct and produces similar scores or results (Heale & Twycross, 2015; Taber, 2018). An instrument that achieves internal consistency yield similar scores on the same aspect of skills or traits that are being evaluated. Meanwhile reliability statistics is related to Cronbach’s Alpha, which is a measure of consistency within a scale in an instrument (Noble & Smith, 2015; Vaske, Beaman, & Sponarski, 2017). For most research purposes, a minimum reliability of 0.8 is necessitated (Taber, 2018). This minimum reliability is to ensure that the set questions or items in an instrument is consistent and has high relatedness in gauging the targeted aims of a study. Hence, the pilot study discussed in this paper involved testing the validity and reliability of the instruments that will then be used in gauging and measuring pupils’ comprehension of homophones.

Methodology

Development and Purpose of Research Instruments

The development of a research instrument is a necessary step to determine the content of the instrument and the scope of analysis based on the respondents for the research (Hult & Johnson, 2015). Hence, this section of the pilot study discusses the development of the instruments that will be administered for the main research.

The first and second instruments are the pre-test. It consists of three sections, which are Sections A, B and C, with a total of 20 questions. Section A comprises of eight “fill-in the blanks” questions, Section B contains six “circle the correct answer” questions and Section C covers six “underline the correct answer” questions. Meanwhile, the post-test contains four sections - Sections A, B, C and D, with a total of 24 questions. Section A consists of nine “underline the correct answer” questions, Section B comprises of five “true or false” questions, Section C covers six “multiple choice” questions and Section D contains four “build sentences” questions. Both of the pre and post tests include ten homophone pairs, selected from the High Frequency Words in the Standard Curriculum Document (DSKP) Primary 4 (Ministry of Education, 2015).

Another aspect of the test involves the thinking taxonomy. One of the modules by Hoffman, Kennedy, LoPilato, Monahan and Lance (2015) indicate that, an assessment must cover several levels

of thinking taxonomy or classification to challenge examinee based on the intended topics of learning. Since the participants were Primary 4 pupils, both the pre and post tests were set at the lower level of remembering, understanding and applying. For example, in the pre-test, Section A, Section B and Section C were set based on the remembering and understanding level where questions such as “fill-in the blanks”, “circle the correct answer” and “underline the correct answer” were involved (Hambleton & Li, 2014; Hoffman et al., 2015). Conversely, the level of thinking taxonomy for the post-test was also similar to that of pre-test with addition of applying level for Section D where pupils are required to “build sentences” based on the learnt homophones.

After the administration of the pre and post tests, the scores were then weighted to 100%. After identifying the pupils' scores, the scores are analysed and placed with reference to the Criterion Referenced Assessment (CRA) based on Malaysian Primary School Achievement Test (UPSR) as it is a demand in determining the grade of a learner against a set of qualities or criteria, without reference to other individuals' achievement (Hambleton & Li, 2014; Lok, McNaught, & Young, 2016). CRA allows an examiner to grade, categorize and rank students based on a given set of achievement scores. With the categorical level indicated in the CRA, the researcher will be able to determine and verify the participants' level of homophones comprehension through the scores obtained from the pre and post tests in the main research. Table 1 shows the percentage (%) of scores to indicate the pupils' level of homophone comprehension.

Table 1. Percentage Score Formula CRA based on UPSR

Score	Grade	Descriptor	
80 – 100	A	Excellent	PASS
65 – 79	B	Good	PASS
50 – 64	C	Satisfactory	PASS
40 – 49	D	Achieve Minimum Level (Adequate)	PASS
0 – 39	E	Below Minimum Level (Poor)	FAIL

The third and fourth instruments are the observation checklist and questionnaire adapted from the Model of Academic Success by York, Gibson III and Rankin (2015). The model outlined that learning is effective when learners are able to achieve comprehension and success in five main elements: (1) attainment of learning outcome, (2) satisfaction in learning, (3) persistence in learning, (4) acquired learning skills and (5) performance or academic achievement. Hence, the observation checklist consisted of 15 observation statements which revolve around the five elements of academic success with three statements for each element. The questionnaire too consists of 15 close-ended items, gauging on the comprehension aspects in the learning of homophones. Each of the items are rated on a Likert scale of 1 to 5, to indicate the degree of which the participants agree or disagree to any of the given statements based on their comprehension level as they learned through the application of LPFCR Card Game.

Process of Piloting the Research Instruments

After the development of the instruments selected for the main research, each of the instrument was pilot tested. This is an important step to determine the validity and reliability of the instruments (Noble & Smith, 2015). As mentioned by Mikuska (2017), by implementing the instruments to at least 12 to 50 people before the actual research will allow researchers to identify the strengths and

weakness of the instrument, ready for any necessary modifications. Hence, 20 pupils from the Primary 4 were selected for the pilot test. These pupils have similar proficiency as the main participants with “Average Language Proficiency” to “Low Language Proficiency”. These level of proficiency was also determined by their English language performance recorded in the School Based Assessment throughout year 2018 and 2019.

The pilot testing of the instruments took place throughout the month of May 2019 in the following order.

Table 2. Sequence of the Pilot Test of Instruments

Phase	Instrument	Action
1 st May to 3 rd May 2019		Pilot participants were introduced with the first five pairs of homophones using the LPFCR Card.
6 th May to 8 th May 2019		Pilot participants were introduced with the last five pairs of homophones using the LPFCR Card.
9 th May to 10 th May 2019	Observation Checklist	Pilot participants were given all 10 pairs of homophones using the LPFCR Card Game. Researcher and another expert teacher observe the pilot participants.
13 th May 2019	Pre-Test	Pilot participants were given the pre-test instrument as assessment.
14 th to 15 th May 2019	Observation Checklist	Pilot participants reviewed all 10 pairs of homophones using the LPFCR Card Game. Researcher and another expert teacher observe the pilot participants.
16 th May 2019	Post-Test and Observation Checklist	Pilot participants were given the post-test instrument as assessment. Researcher and another expert teacher observe the pilot participants.
17 th May 2019	Questionnaire	Pilot participants were given the questionnaire to collect their opinions.
20 th May to 21 st May 2019	Pre and Post Tests	Researcher tabulates the data from the Pre and Post Tests. Researcher refers to expert teacher and assessment development module to check for validity and reliability.
22 nd May 2019	Observation Checklist	Researcher cross-checked the data collected from the observation checklist with expert teacher to determine the similarity in phrases or notes taken. Researcher refers to expert teacher and check for face and content validity.
23 rd May 2019	Questionnaire	Researcher tabulates the data from questionnaire into SPSS Version 25 to obtain mean score, percentages and Cronbach’s Alpha reliability coefficient.

Findings and Discussion

Validity and Reliability of the Pre and Post Tests

For face validity, the pre and post tests were analyzed by an expert teacher to determine the feasibility and practicality of the questions for Primary 4 level. The expert teacher is the head of English Language Examination Resource in Daro, Sarawak. The expert teacher reported that the pre and post tests were clear and not confusing. This designates the face value of the pre and posttests;

indicating face validity. Another criteria of validity is content validity. The expert teacher examined that the pre and posttests consist of all the 10 homophones and the levels of thinking taxonomy (remembering, understanding and applying) that were assessed. Based on the expert teacher's evaluation of the face value and content of the pre and posttests, the tests were found to be valid.

Meanwhile, to determine the internal consistency reliability, the pre and posttests, were marked by the researcher and expert teacher. The overall result of the pre and posttests that were administered were as follows.

Table 3. *Percentage Score of Pre and Post Tests*

Pseudonym	Pre-Test		Post-Test		Descriptor
	Result (%)	Grade	Result (%)	Grade	
Minic	91	A	88	A	Excellent (Pass)
Stancia	90	A	89	A	Excellent (Pass)
Nessa	89	A	85	A	Excellent (Pass)
Tia	89	A	86	A	Excellent (Pass)
Morrison	88	A	83	A	Excellent (Pass)
Rai	88	A	82	A	Excellent (Pass)
Daniel	88	A	85	A	Excellent (Pass)
Aveena	87	A	88	A	Excellent (Pass)
Jerry	87	A	83	A	Excellent (Pass)
Dexter	87	A	86	A	Excellent (Pass)
Kong	86	A	83	A	Excellent (Pass)
Ciara	85	A	87	A	Excellent (Pass)
Ruth	85	A	82	A	Excellent (Pass)
Nick	84	A	81	A	Excellent (Pass)
Anna	84	A	82	A	Excellent (Pass)
Jacquel	83	A	80	A	Excellent (Pass)
Dell	82	A	80	A	Excellent (Pass)
Rica	81	A	84	A	Excellent (Pass)
Manda	80	A	82	A	Excellent (Pass)
Andy	80	A	83	A	Excellent (Pass)

From the overall score, all of the pupils were able to achieve excellent result with percentage ranging from 80% to 91% for both papers. In order to determine the reliability of the pre and posttests, the Cronbach's Alpha reliability coefficient was employed based on the given formula.

$$\hat{\alpha} = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k p_i(1-p_i)}{\hat{\sigma}_X^2} \right)$$

Notes:

k = number of items in the test paper

p_i = percentage of correct responses for each item, i

$\hat{\sigma}_X^2$ = sample variance of total score

Figure 1. Cronbach's Alpha Reliability Coefficient Formula

The pre-test produced a Cronbach's Alpha of 0.854, while the post-test had a Cronbach's Alpha of 0.843. These results proved that both tests had achieved high internal consistency and are reliable. Apart from the results of the Cronbach's Alpha, assessments that produces similar results are considered to have good internal consistency, which is an indication of a reliable and valid assessment (Hoffman et al., 2015). This statement is also supported by Hambleton and Li (2014) and Lok et al. (2016) where assessment with good internal consistency correlates with the intended outcome of a lesson and produces almost equal scores among the examinees. In this case, the pilot participants were able to answer the pre and post tests with a good level of homophone comprehension after the implementation of the LPFCR Card Game. Hence, both pre and post tests showed good internal consistency which contributed to the high validity and reliability of the instruments.

Validity and Reliability of the Observation Checklist and Questionnaire

The observation checklist and questionnaire were developed with the intent to observe and measure the pupils' comprehension on the learning of homophones. According to York, Gibson and Rankin Revised Conceptual Model of Academic Success, a learner is successful in the learning process when they are able to attain five important elements of "... academic or learning achievement, satisfaction in learning, persistence in completing a given task, acquisition of skills and attainment of learning objectives" (York et al., 2015). Hence, based on these five elements, the questions and statements of the observation checklist and questionnaire were constructed to gauge on the pupils' comprehension throughout the implementation of LPFCR Card Game.

Throughout the implementation of the pilot test on the selected participants, the researcher, accompanied by the expert teacher conducted the observations. For this purpose, the observation checklist was employed to collect relevant data on the pilot participants in relation to specific patterns of behaviours. After data was collected, both the checklists by the researcher and the expert teacher were triangulated to identify similarities and differences in the statements and noted details. Table 4 shows the data triangulated from several extracted items from the observation checklists.

Table 4. Data Collected from the Triangulated Observation Checklist

Elements	Checklist Items		Written Remarks
Attainment of Learning Outcome	Pupils are able to describe the different homophone pairs when asked verbally LPFCR Card Game.	Researcher	Able to give clear description on the learnt homophones in sentences.
		Expert Teacher	Can put homophones in a sentence to describe .
Satisfaction In Learning	Pupils show excitement when receiving reward tokens from the “R – Reward” section after completing the card game.	Researcher	Pupils smiled and laughed receiving tokens.
		Expert Teacher	Giggled, smiled, enjoy getting tokens.

From the triangulated data collected, there were several similar key points and remarks that were noted. Within all five elements that were highlighted in the observation checklists, the remarks from the researchers were consistent with the remarks by the expert teacher. For example, in the first element which is the “Attainment of Learning Outcome”, both the researcher and expert teacher remarked similar keywords which are “Able to give clear description on the learnt homophones in sentences” and “Can put homophones in a sentence to describe”. The consistency in the remarks are indication of a solid and reliable instrument that can be implemented in line with the purpose of the study (Mimi et al., 2015). Nardi (2018) also supported that the similarity in observations and comments in a qualitative measure is a good indication of valid and dependable instrument. Therefore, with the indications of similarity in the observations that were conducted, the observation checklist was found to be valid and reliable.

Moving on to the aspect of testing the reliability of the questionnaire, after employing the LPFCR Card Game and post-test, the pilot participants were given a questionnaire to collect their opinions on the degree of which they agree or disagree to any of the given statements. The statements were rated on a Likert scale of 1 to 5 where “1 – Strongly Disagree” to “5 – Strongly Agree”. The data from the questionnaire were collected and tabulated into SPSS Version 25 to obtain the Cronbach’s Alpha reliability coefficient. As asserted by Taber (2018), Cronbach’s Alpha is a measure of scale reliability of which to determine the consistency in relatedness of a set of items within a group. Table 5 shows the reliability statistics of the data collected from the questionnaire.

Table 5. Reliability Statistics of Questionnaire

Cronbach’s Alpha	Cronbach’s Alpha Based on Standardized Items	N of Items
.873	.875	15

Vaske et al. (2017) stated that Cronbach’s Alpha that is more than or equal to 0.8 (≥ 0.8) is an excellent indication of a good internal consistency and a reliable set of items. As shown in the table, the sets of items of the questionnaire have a good internal consistency with a Cronbach’s Alpha of 0.873. This indicated that the questionnaire is a reliable source of instrument to gather the participants’ comprehension on the learning of homophone through the application of LPFCR Card Game.

Conclusion

The main purpose of this pilot study was to determine the validity and reliability of the instruments that will be implemented for the main research. From this pilot study, the pre-test, post-test, observation checklist and questionnaire were proven valid and reliable to be employed for data collection in the main study of LPFCR Card Game in enhancing the Year 4 pupils' comprehension of homophones. With the established validity and reliability, these instruments will be applicable for future studies in gathering data and information in relation to learners' comprehension on the learning of homophones.

References

- Appelman, A., & Sundar, S. S. (2016). Measuring Message Credibility: Construction and Validation of an Exclusive Scale. *Journalism & Mass Communication Quarterly*, 93(1), 59–79. <https://doi.org/10.1177/1077699015606057>
- Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. In P. Liamputtong (Ed.), *Handbook of Research Methods in Health Social Sciences* (pp. 843–860). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-5251-4_103
- Charman, S. (2017). The Research: Aims and Methods. In *Police Socialisation, Identity and Culture: Becoming Blue* (pp. 171–187). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-63070-0_7
- Chu, H., & Ke, Q. (2017). Research methods: What's in the name? *Library & Information Science Research*, 39(4), 284–294. <https://doi.org/https://doi.org/10.1016/j.lisr.2017.11.001>
- Creswell, J. W., & Guetterman, T. C. (2019). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (6th Editio). New York, NY: Pearson.
- Dautrichea, I., Fiblad, L., Fievetb, A.-C., & Christopheb, A. (2018). Learning Homophones In Context: Easy Cases Are Favored In The Lexicon Of Natural Languages. *Cognitive Psychology*, 104(August), 83–105. <https://doi.org/https://doi.org/10.1016/j.cogpsych.2018.04.001>
- Dick, B. (2015). Reflections on the SAGE Encyclopedia of Action Research and What It Says About Action Research and Its Methodologies. *Action Research*, 13(4), 431–444. <https://doi.org/10.1177/1476750315573593>
- Hambleton, R. K., & Li, S. (2014). Criterion-Referenced Assessment. *Wiley StatsRef: Statistics Reference Online*, 1–8.
- Hazzi, O., & Maldaon, I. (2015). A Pilot Study: Vital Methodological Issues. *Verslas: Teorija Ir Praktika*, 16(1), 53–62. <https://doi.org/10.3846/btp.2015.437>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>
- Hult, F. M., & Johnson, D. C. (2015). *Research Methods in Language Policy And Planning: A Practical Guide*. Hoboken, New Jersey: John Wiley & Sons.
- Kenan, D., & Hakki, E. İ. (2018). Teaching Homographs, Homonyms, Homophones, Synonyms, and Antonyms. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). American Cancer Society. <https://doi.org/10.1002/9781118784235.eelt0753>
- Kinchin, G., Ismail, N., & Edwards, J.-A. (2018). Pilot study, Does it really matter? Learning lessons from conducting a pilot study for a qualitative PhD thesis. *International Journal of Social*

- Science Research*, 6(1). Retrieved from <https://eprints.soton.ac.uk/416716/>
- Klenke, K., Martin, S., & Wallace, J. R. (2016). *Qualitative Research in the Study of Leadership*. Bingley, UK: Emerald Group Publishing Limited. <https://doi.org/doi:10.1108/9781785606502>
- Lin, T.-J., & Tsai, C.-C. (2017). Developing instruments concerning scientific epistemic beliefs and goal orientations in learning science: a validation study. *International Journal of Science Education*, 39(17), 2382–2401. <https://doi.org/10.1080/09500693.2017.1384593>
- Lok, B., McNaught, C., & Young, K. (2016). Criterion-referenced and norm-referenced assessments: compatibility and complementarity. *Assessment & Evaluation in Higher Education*, 41(3), 450–465. <https://doi.org/10.1080/02602938.2015.1022136>
- Mikuska, E. (2017). The Importance of Piloting or Pre-Testing Semi-Structured Interviews and Narratives. *SAGE Research Methods Cases*. <https://doi.org/10.4135/9781473977754>
- Mimi, M. M., Nor, L. S., Lai, C. S., & Kahirol, M. S. (2015). Measuring the Validity and Reliability of Research Instruments. *Procedia - Social and Behavioral Sciences*, 204(August), 164–171. <https://doi.org/https://doi.org/10.1016/j.sbspro.2015.08.129>
- Ministry of Education. (2015). *Standard Curriculum Document (DSKP) Year 4, 5 and 6*.
- Nardi, P. M. (2018). *Doing Survey Research: A Guide to Quantitative Method* (4th ed.). New York, NY: Routledge.
- Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-Based Nursing*, 18(2), 34–35. <https://doi.org/10.1136/eb-2015-102054>
- Omar, A. (2015). Selecting the appropriate study design for your research: Descriptive study designs. *Journal of Health Specialties*, 3(3), 153. <https://doi.org/10.4103/1658-600X.159892>
- Sharifah, H. S. D., Jamal, N. Y., & Hamidah, Y. (2017). Kesahan Dan Kebolehppercayaan Instrumen Kajian Penyeliaan Berkesan. *JuKu: Jurnal Kurikulum & Pengajaran Asia Pasifik*, 5(3), 50–61. Retrieved from <http://mjs.um.edu.my/index.php/JUKU/article/view/8227>
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Treiman, R., Seidenberg, M. S., & Kessler, B. (2015). Influences on spelling: evidence from homophones. *Language, Cognition and Neuroscience*, 30(5), 544–554. <https://doi.org/10.1080/23273798.2014.952315>
- Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking Internal Consistency in Cronbach's Alpha. *Leisure Sciences*, 39(2), 163–173. <https://doi.org/10.1080/01490400.2015.1127189>
- Watson, P., & de Wit, S. (2018). Current limits of experimental research into habits and future directions. *Current Opinion in Behavioral Sciences*, 20, 33–39. <https://doi.org/https://doi.org/10.1016/j.cobeha.2017.09.012>
- York, T., Gibson III, C., & Rankin, S. (2015). Defining and Measuring Academic Success. *Practice Assessment, Research & Evaluation*, 20(5), 1–20. Retrieved from <https://pareonline.net/getvn.asp?v=20&n=5>
- Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. *Journal of Caring Sciences*, 4(2), 165–178. <https://doi.org/10.15171/jcs.2015.017>