EDUCATIONAL RESOURCE
Volume 11 Issue 3 2019
DOI: 10.21315/eimj2019.11.3.6
ARTICLE INFO
Submitted: 29-04-2019
Accepted: 28-05-2019
Online: 31-10-2019

ABC of Response Process Validation and Face Validity Index Calculation

Muhamad Saiful Bahri Yusoff

Department of Medical Education, School of Medical Sciences, Universiti Sains Malaysia, Kelantan, MALAYSIA

To cite this article: Yusoff MSB. ABC of response process validation and face validity index calculation. Education in Medicine Journal. 2019;11(3):55–61. https://doi.org/10.21315/eimj2019.11.3.6

To link to this article: https://doi.org/10.21315/eimj2019.11.3.6

ABSTRACT

Validity evidence can be supported by five sources that are content, response process, internal structure, relation to other variables, and consequences. Response process validity measures the thought processes of users of the tested inventory as they respond to the assessment tool. These are commonly evaluated in the form of clarity of instructions and language used in the assessment tool, as well as the comprehension of instruction after training or an observation session. Response process validity contributes to the overall validity of an assessment tool; therefore, it should be quantified systematically based on the evidence and best practice. This paper describes a systematic approach to quantify response process validity in the form of face validity index based on the evidence.

Keywords: *Validity, Response process, Face validity, Questionnaire, Face validity index*

CORRESPONDING AUTHOR Associate Professor Dr. Muhamad Saiful Bahri Yusoff, Department of Medical Education, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia | E-mail: msaiful_bahri@usm.my

INTRODUCTION

In 1947, Mosier analysed the various definitions of face validity concept (1). Commonly, response process validity evidence is performed after content validity has been established (2, 3) and response process validity is also known as the face validity that refers to the degree to which test respondents view the content of a test and its items as relevant to the context in which the test is being administered (4). Similarly, other researchers define face validity as the degree raters judge the items of an assessment instrument as appropriate to the targeted construct and assessment objectives (5, 6). The raters of face validity include: (a) the person who actually takes the test; (b) the nonprofessional users who work with the results of the test; and (c) the general public (6). In other words, the people who are involved with the test taking should be asked to do the rating, in which they cannot be replaced by professional, experts or psychometricians (6). The raters' understanding and interpretation about the items will determine the accuracy of an assessment tool to measure the targeted construct. People with a similar background rate test face validity similarly, and they rate the face validity of different tests differently (6). Due to so much concern about the face validity concept, Cook and Beck (2006) have avoided using the face validity term, instead, the researchers use the response process evidence of validity as the term to reflect the thought processes of users of

the tested assessment as they respond to the tool (7, 8) and it can be quantified by face validity index (FVI) (8–11). These are commonly evaluated in the form of clarity and comprehensibility of instructions and language used in the assessment tool by the raters (7, 8). According to Cook and Beck (2006), validity evidence can be supported by content, response process, internal structure, relation to other variables, and consequence of an assessment tool (7). The clarity of instructions and language refers to whether there were ambiguities or multiple ways to interpret the items, whereas the comprehensibility of instructions and language refers to whether words and sentences of the constructed items can be understood easily by raters. It is important to establish response process validity to support the overall validity of an assessment tool such as questionnaires, especially for research purposes. The response process validity can be represented by FVI and several studies (8–11) have calculated it to support the validity of an assessment tool. Based on the evidence, this paper describes the best practice to perform response process validation and calculate the FVI of an assessment tool.

Response Process Validation Procedure

The following are the six steps of response process validation:

1. Preparing response process validation form.
2. Selecting a panel of raters.
3. Conducting response process validation.
4. Reviewing items for clarity and comprehension.
5. Providing score for each item based on the clarity and comprehensibility rating scale.
6. Calculating FVI.

Each step will be elaborated in the subsequent sections.

Step 1: Preparing response process validation form

The first step of response process validation (also known as face validation) is to prepare the response process validation form to ensure that the panel of raters, who are the intended respondents, will have a clear expectation and understanding about the task. An example for the instruction and rating scale is provided in Figure 1. The rating scales of clarity and comprehension have been used for scoring individual items (12–15) (Figure 2).

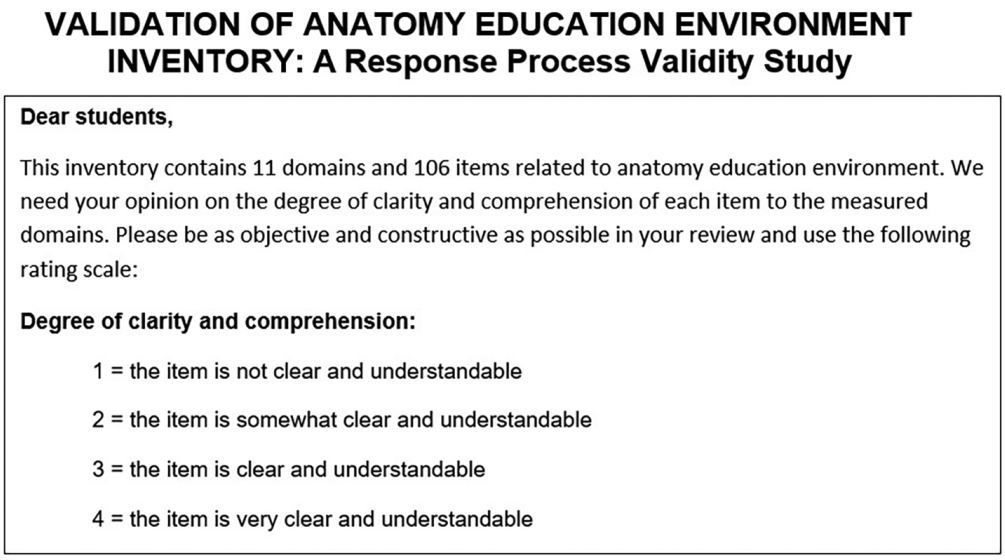
Step 2: Selecting a panel of raters

The selection of raters to review and critique an assessment tool (e.g., questionnaire) is usually based on the target user of the tool, for example, students, public, and teachers. Table 1 summarises the number of raters with its implication on the acceptable cut-off score of FVI based on previous studies (8–11, 16–18).

It can be concurred that for response process validation, the minimum acceptable number of raters is 10; however, most studies had at least 30 raters. Considering the previous studies (Table 1) and the author's experience, the number of experts for content validation should not be less than 10 raters.

Step 3: Conducting response process validation

The response process validation can be conducted through face-to-face or online survey (Table 1). For the face-to-face survey, the researcher facilitates the response process validation process by holding a meeting with the raters followed by ***Step 4*** and ***Step 5*** (as elaborated further). For the online survey, an online response process validation form is sent to the raters and clear instructions are provided (Figure 1) to facilitate the validation process. Based on the author's experience, the face-to-face approach is very efficient to increase the response rate, whereas the online survey is efficient in terms of cost and time.



VALIDATION OF ANATOMY EDUCATION ENVIRONMENT INVENTORY: A Response Process Validity Study

Dear students,

This inventory contains 11 domains and 106 items related to anatomy education environment. We need your opinion on the degree of clarity and comprehension of each item to the measured domains. Please be as objective and constructive as possible in your review and use the following rating scale:

Degree of clarity and comprehension:

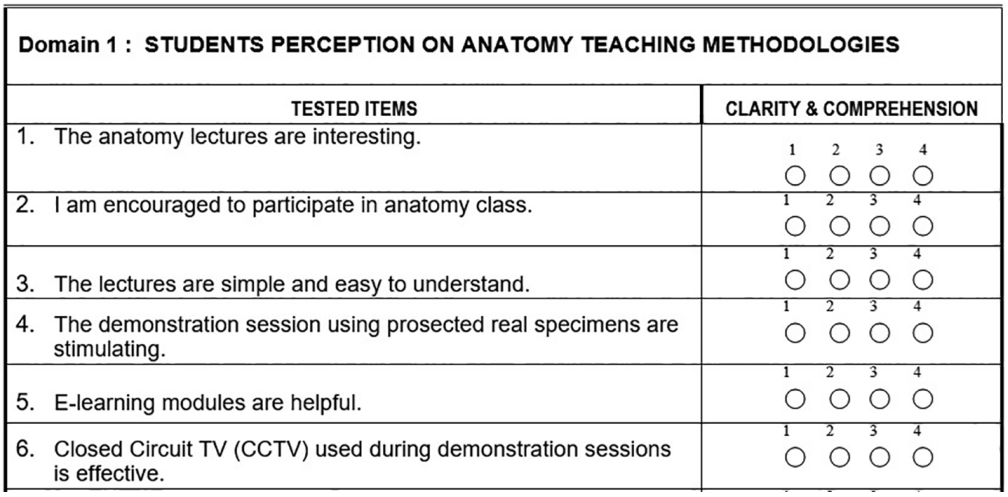
1 = the item is not clear and understandable

2 = the item is somewhat clear and understandable

3 = the item is clear and understandable

4 = the item is very clear and understandable

Figure 1: An example of instruction and rating scale in the response process validation form to the raters (students)
Source: Author



Domain 1 : STUDENTS PERCEPTION ON ANATOMY TEACHING METHODOLOGIES

TESTED ITEMS	CLARITY & COMPREHENSION
1. The anatomy lectures are interesting.	1 2 3 4
2. I am encouraged to participate in anatomy class.	1 2 3 4
3. The lectures are simple and easy to understand.	1 2 3 4
4. The demonstration session using prosected real specimens are stimulating.	1 2 3 4
5. E-learning modules are helpful.	1 2 3 4
6. Closed Circuit TV (CCTV) used during demonstration sessions is effective.	1 2 3 4

Figure 2: An example of layout for response process validation form with domain and items
Source: Author

Table 1: The number of raters and its implication on the acceptable cut-off score of FVI

Source	Number of raters	Acceptable FVI value	Method
Hadie et al. (2017) (8)	30 medical students	At least 0.80	Face-to-face survey
Ozair et al. (2017) (9)	30 paramedics	At least 0.83	Face-to-face survey
Lau et al. (2017) (16)	30 parents of pre-school children	At least 0.80	Face-to-face survey
Lau et al. (2018) (10)	30 parents of pre-school children	At least 0.80	Face-to-face survey
Marzuki et al. (2018) (11)	10 users of medical apps	At least 0.83	Online survey
Chin et al. (2018) (17)	32 medical students	At least 0.80	Online survey
Mahadi et al. (2018) (18)	32 medical students	At least 0.80	Online survey

Step 4: Reviewing items for clarity and comprehension

In the response process validation form, the domain and its items are provided to the raters as shown in Figure 2. The raters are requested to review all items before providing score for each item. The raters are encouraged to provide verbal comment or written comment to improve the clarity and comprehension of the items. All comments are taken into consideration to refine items.

Step 5: Providing score for each item based on the clarity and comprehensibility rating scale

Upon completion of reviewing all items, the raters are requested to provide score for each item independently based on the clarity and comprehension scale (Figures 1 and 2). The raters are required to submit their responses to the researcher once they have provided a score for all the items.

Step 6: Calculating FVI

There are two forms of FVI, i.e. FVI for item (I-FVI) and FVI for scale (S-FVI). Two methods for calculating S-FVI, in which the average of the I-FVI scores for all the items on the scale (S-FVI/Ave) and the proportion of items on the scale that achieve a clarity and comprehension scale of 3 or 4 by all raters (S-FVI/UA) (Table 2). The definition and formula of the FVI indices are summarised in Table 2.

Prior to the calculation of FVI, the clarity and comprehension rating must be recoded as 1 (the scale of 3 or 4) or 0 (the scale of 1 or 2) as shown in Table 3. To illustrate the calculation of different FVI indices, the clarity and comprehension ratings on item scale by 10 raters are provided in Table 3.

To illustrate the calculation for the FVI indices (please refer to Table 2), the following are examples of calculation based on the data provided in Table 3:

i. ***Raters in agreement***: just sum up the relevant rating provided by all raters for each item, for example, the raters in agreement for Q2 (1 + 0 + 1 +1 + 1 + 1 + 1+ 1 + 1 + 1) = 9.

ii. ***Universal agreement***: score '1' is assigned to the item that achieved 100% raters in agreement, for example, Q1 obtained 1 because all the raters provided rating of 1, while Q2 obtained 0 because not all raters provided rating of 1.

iii. ***I-FVI***: the raters in agreement divided by the number of raters, for example I-FVI of Q2 is 9 divided by 10 raters that is equal to 0.9.

Table 2: The definition and formula of I-FVI, S-FVI/Ave and S-FVI/UA

The CVI indices	Definition	Formula
I-FVI (item-level face validity index)	The proportion of rater giving an item a clarity and comprehension rating of 3 or 4	**I-FVI** = (agreed item)/ (number of rater)
S-FVI/Ave (scale-level face validity index based on the average method)	The average of the I-FVI scores for all the items on the scale or the average of proportion clarity and comprehension judged by all raters. The proportion clarity and comprehension is the average of rating by individual rater.	**S-FVI/Ave** = (sum of I-FVI scores)/(number of item) **S-FVI/Ave** = (sum of proportion clarity and comprehension rating)/ (number of rater)
S-FVI/UA (scale-level face validity index based on the universal agreement method)	The proportion of items on the scale that achieve a clarity and comprehension scale of 3 or 4 by all raters. Universal agreement (UA) score is given as 1 when the item achieved 100% raters in agreement, otherwise the UA score is given as 0.	**S-FVI/UA** = (sum of UA scores)/(number of item)

Note: The definition and formula was based on Ozair et al. (2017) (9) and the content validity index formula was reported in Yusoff (2019) (3)

Table 3: The clarity and comprehension ratings on the item scale by 10 raters

	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Rater 9	Rater 10		Raters in Agreement	I-FVI	UA
Item														
Q1	1	1	1	1	1	1	1	1	1	1		10	1	1
Q2	1	0	1	1	1	1	1	1	1	1		9	0.9	0
Q3	0	0	0	0	0	0	0	0	0	0		0	0	0
Q4	1	1	1	1	1	1	1	1	1	1		10	1	1
Q5	1	1	1	1	1	1	1	1	1	1		10	1	1
Q6	1	1	1	1	1	1	1	1	1	1		10	1	1
Q7	1	1	1	1	1	1	1	1	1	1		10	1	1
Q8	1	1	1	1	1	1	1	1	1	1		10	1	1
Q9	1	1	1	1	1	1	1	1	1	1		10	1	1
Q10	1	1	1	1	1	1	1	1	1	1		10	1	1
Q11	1	1	1	1	1	1	1	1	1	1		10	1	1
Q12	1	1	1	1	1	1	1	1	1	1		10	1	1
												S-FVI/Ave	0.91	
Proportion Clarity & Comprehension	0.92	0.83	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92		S-FVI/UA		0.83
				Average proportion of items judged clarity & comprehension across the 10 raters							0.91			

iv. ***S-FVI/Ave*** *(based on I-FVI)*: the average of I-FVI scores across all items, for example, the S-FVI/Ave [(10 + 9 + 0 + 10 + 10 + 10 + 10 + 10 + 10 + 10 + 10 + 10)/12] is equal to 0.91.

v. ***S-FVI/Ave*** *(based on proportion clarity and comprehension)*: the average of proportion clarity and comprehension scores across all raters, for example, the S-FVI/Ave [(0.92 + 0.83 + 0.92 + 0.92 + 0.92 + 0.92 + 0.92 + 0.92 + 0.92 + 0.92)/10] is equal to 0.91.

vi. ***S-FVI/UA***: the average of UA scores across all items, for example, the S-FVI/UA [(1 + 0 + 0 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1)/12] is equal to 0.83.

Based on the above calculations, it can be concluded that I-FVI, S-FVI/Ave, and S-FVI/UA meet satisfactory level, and thus the scale of questionnaire has achieved satisfactory level of response process validity. For more examples on how to report the response process validity index, please refer to papers written by Hadie et al. (2017) (8), Ozair et al. (2017) (9), Lau et al. (2017) (16), Lau et al. (2018) (10), Marzuki et al. (2018) (11), Chin et al. (2018) (17), and Mahadi et al. (2018) (18).

CONCLUSION

Response process validity is vital to ensure the overall validity of an assessment, therefore, a systematic approach for validating response process should be done based on the best evidence. This paper has provided a systematic and evidence-based approach to conduct a proper response process validation through face validity index.

REFERENCES

1. Mosier CI. A critical examination of the concepts of face validity. Educ Psychol Meas. 1947;7(2):191–205. https://doi.org/10.1177/001316444700700201

2. Anthony R, Artino J, Rochelle JSL, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach. 2010;36(6):463–74. https://doi.org/10.3109/0142159X.2014.889814

3. Yusoff MSB. ABC of content validation and content validity index calculation. Education in Medicine Journal. 2019;11(2):49–54. https://doi.org/10.21315/eimj2019.11.2.6

4. Holden RR. Face validity. The corsini encyclopedia of psychology. 2010:1–2. https://doi.org/10.1002/9780470479216.corpsy0341

5. Hardesty DM, Bearden WO. The use of expert judges in scale development: implications for improving face validity of measures of unobservable constructs. Journal of Business Research. 2004;57(2):98–107. https://doi.org/10.1016/S0148-2963(01)00295-8

6. Nevo B. Face validity revisited. J Educ Meas. 1985;22(4):287–93. https://doi.org/10.1111/j.1745-3984.1985.tb01065.x

7. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med. 2006;119(2):166.e7–16. https://doi.org/10.1016/j.amjmed.2005.10.036

8. Hadie SNH, Hassan A, Ismail ZIM, Asari MA, Khan AA, Kasim F, et al. Anatomy education environment measurement inventory: a valid tool to measure the anatomy learning environment. Anat Sci Educ. 2017;10(5):423–32. https://doi.org/10.1002/ase.1683

9. Ozair MM, Baharuddin KA, Mohamed SA, Esa W, Yusoff MSB. Development and validation of the Knowledge and Clinical Reasoning of Acute Asthma Management in Emergency Department (K-CRAMED). Education in Medicine Journal. 2017;9(2):1–17. https://doi.org/10.21315/eimj2017.9.2.1

10. Lau AS, Yusoff MS, Lee Y-Y, Choi S-B, Xiao J-Z, Liong M-T. Development and validation of a Chinese translated questionnaire: a single simultaneous tool for assessing gastrointestinal and upper respiratory tract related illnesses in pre-school children. Journal of Taibah University Medical Sciences. 2018;13(2):135–41. https://doi.org/10.1016/j.jtumed.2017.11.003

11. Marzuki MFM, Yaacob NA, Yaacob NM. Translation, cross-cultural adaptation, and validation of the Malay Version of the System Usability Scale Questionnaire for the Assessment of Mobile Apps. JMIR Human Factors. 2018;5(2):e10308. https://doi.org/10.2196/10308

12. Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):381–5. https://doi.org/10.1097/00006199-198611000-00017

13. Davis LL. Instrument review: getting the most from a panel of experts. Appl Nurs Res. 1992;5(4):194–7. https://doi.org/10.1016/S0897-1897(05)80008-4

14. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health. 2006;29(5):489–97. https://doi.org/10.1002/nur.20147

15. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health. 2007;30(4):459–67. https://doi.org/10.1002/nur.20199

16. Lau AS-Y, Yusoff MSB, Lee YY, Choi S-B, Rashid F, Wahid N, et al. Development, translation and validation of questionnaires for diarrhea and respiratory-related illnesses during probiotic administration in children. Education in Medicine Journal. 2017;9(4):19–30. https://doi.org/10.21315/eimj2017.9.4.3.

17. Chin RWA, Chua YY, Chu MN, Mahadi NF, Wong MS, Yusoff MS, et al. Investigating validity evidence of the Malay translation of the Copenhagen Burnout Inventory. Journal of Taibah University Medical Sciences. 2018;13(1):1–9. https://doi.org/10.1016/j.jtumed.2017.06.003

18. Mahadi NF, Chin RWA, Chua YY, Chu MN, Wong MS, Yusoff MSB, et al. Malay language translation and validation of the Oldenburg Burnout Inventory Measuring Burnout. Education in Medicine Journal. 2018;10(2):27–40. https://doi.org/10.21315/eimj2018.10.2.4.