EasyChair Preprint

№ 4535

The Application of Fuzzy Delphi Method in Content Validity Analysis

Mohd Azli Jailani

November 11, 2020

The Application of Fuzzy Delphi Method in Content Validity Analysis

By Mohd Azli bin Jailani

UNITAR International University

Abstract

The purpose of this study was to explore about an alternative approach of testing the content validity of an instrument. For this reason, a newly developed counseling needs assessment instrument was used to quantify the agreement that was made by 16 of the panel experts counselors. To compute the feedbacks quantitively, this study has chosen the Fuzzy Delphi Method (FDM). The final results have shown that out of 126 items, 5 items namely item 62, item 63, item 76, item 77 and item 115 were discarded from the list as they were scored 0.5 and below of the defuzzification values and the items were also ranked 122 (item 77), 123 (item 63), 124 (item 76), 125 (item 62) and 126 (item 115).

Introduction

Testing has been regarded as a fundamental component in counseling services. This is because counselors constitutes a major group of test users (Anastasi, 1992; Leppma & Jones, 2018). The nature of counseling work that need to deal with different types of client's populations would require counselors to use various types of instruments, which include measurements of cognitive and affective behavior, self-administered inventories, computerized testing and several other approaches or methods.

To ensure the procedure can truly benefit the clients, the instrument itself need to meet certain types of requirements. One of the essential requirements is validity. According to Taherdoost (2016), validity is simply can be defined as measure what it supposed to measure. Fraenkel and Wallen (2008, p.147) define validity as "the appropriateness, meaningful, correctness and usefulness of the inferences a researcher makes." Without validity, a phenomenon such as intelligence can't be explained and described by an intelligent test because the test didn't measure what it supposed to measure.

In general, validity is divided into four different types of categories (Taherdoost, 2016); a) face validity, b) content validity, c) construct validity and d) criterion validity. Yaghmale (2003) delineates that measuring content validity of instruments are important because it gives confidence to the readers and researchers about instruments. Bahri Yusoff (2019) defined content validity as the degree to which elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose.

Heale and Twycross (2015) mention that content validity is a subset of face validity. Therefore, Drost (2011), Yaghmale (2003) and Fraenkel and Wallen (2008) describe content validity is a qualitative type of validity rather than quantitative. The

procedure of measuring content validity normally involves a group of panel experts will be given an evaluation form which contains several open-ended questions for them to highlight their comments or responses about the items of an instrument. But this procedure depicted some limitations as explained by Rubio, Berg-Weger, Tebb, Lee and Rauch (2003) and Mustafa and Darusalam (2018). Among of the limitations are experts' feedback is subjective, thus it is subject to be bias, especially when a quite senior expert deliberates his or her comment, the tendency to accept his or her comment would be obvious as compared to a comment which is made by a junior lecturer, even though sometimes his or her comment makes sense.

The other limitation of measuring content validity via conventional method is the feedbacks that were produced can't be calculated since it is qualitative in nature. Therefore, it is difficult to reduce errors or inconsistencies in the comments made by the expert panels (Mustafa & Darusalam, 2018).

Problem statement

To overcome the problems mentioned above, a quantitative approach is seen as the best approach to conduct content validity studies as an alternative to the conventional approach.

One of the quantitative methods is the Fuzzy Delphi Method (FDM). Abdul Rahman, Md Nor, Ahmad Nadzim, Mohd Radzi and Moktar (2016) deliberate that FDM is not a new technique but an innovated technique from the traditional Delphi method. Delphi method is an expert opinion survey that was constructed with three different features; anonymous response, iteration and controlled feedback and finally statistical group response. However, this procedure requires researchers to conduct the survey for several times, which could delay the research progress and increase the overall costs. FDM as explained by Yu, Cheng and Kreng (2010) was introduced to quickly get consensus of experts without going through many rounds of survey exercises. This also eliminates some ambiguity caused by differences in meaning and interpretation of expert opinions.

To identify whether FDM could deliberate information about content validity, a newly developed counseling needs assessment instrument was used to quantify the consensus of the panel experts over the instrument. The counseling needs assessment instrument consists 126 items was developed to assist school counselors in determining priorities of their guidance and counseling programs or activities. FDM was deployed to verify whether the 126 items could really assess students counseling needs under the six components of needs namely Academic needs, Emotional needs, Personal Development needs, Career needs, Peer Relationship and finally Family needs,

Research objective

To validate the content validity of the 126-tems counseling needs assessment instrument through the Fuzzy Delphi method (FDM).

Research question

What do the panel experts say about the 126-items counseling needs assessment instrument through the Fuzzy Delphi method (FDM) of analysis?

Methodology

The process of FDM is illustrated as follows:

Step 1 - selection of panel experts: 16 panel experts were selected based on Hsu and Sanford (2007) model of expert panel selection. The panels were school counselors from schools situated in two States; Selangor and Perak. The selection process was taking place based on two criteria; a) the deliberate practice and b) the length of occupational background, minimum 4 to 5 years of experience as deliberated by Turner (1995).

Step 2 – determining linguistic scale: Each of the panel expert was given a questionnaire/survey which contained 126 items for them to verify based on the following triangular fuzzy and linguistic scales (Table 1.0):

Linguistic scale (5 points)	Fuzzy scale
1 = Strongly disagree	(0.0, 0.1, 0.2)
2 = Disagree	(0.1, 0.2, 0.4)
3 = Not sure	(0.2, 0.4, 0.6)
4 - Agree	(0.4, 0.6, 0.8)
5 - Strongly agree	(0.6, 0.8, 1.0)

Table 1.0: 5 points linguistic scale and the fuzzy scales

The panel experts were asked to verify whether the items could assess students' counseling needs by circling the items with 5 different linguistic scales as mentioned above (Strongly disagree, Disagree, Not sure, Agree and Strongly agreed). Figure 1.0 illustrates the sample of the survey form.

Item	Sebagai seorang pelajar…	Skor	Keadaan saya ketika ini…	Skor
1.	Sebagai seorang pelajar, kita perlu memahami pelbagai jenis perasaan yang kita alami seperti gembira, sedih dan marah.	1 2 3 4 5 6	Saya mampu untuk memahami pelbagai jenis perasaan yang saya alami seperti gembira, sedih dan marah.	1 2 3 4 5 6

Skor untuk kesesuaian item:				
Sangat sesuai	**Sesuai**	**Sederhana sesuai**	**Tidak sesuai**	**Sangat tidak sesuai**
5	**4**	**3**	**2**	**1**

Catatan:
……
……
……

Figure 1.0: Sample of the survey form

Step 3 - Step 3 - determining the Threshold "d" value: The Threshold *"d"* value is important in determining the levels of agreement among the panel experts upon the items. The data from this exercise were entered onto a Microsoft Excel worksheet and analyzed with the following formula:

$$d(m,n) = \sqrt{\frac{1}{3}[(m1 + n1)2 + (m2 - n2)2 + (m3 - n3)2]}$$

Mustafa and Darusalam (2018) explain if the threshold *"d"* value is lower or equivalent to 0.2 (≤ 0.2) it can be said the panel experts have achieved their consensus on the items.

Step 4 - determining group consensus percentage: the data from the Microsoft Excel worksheet file than were analyzed to quantify the value of group consensus percentage.

As deliberated by Ahmad, Patahol Wasli, Mohd Mohd Fauzi, Mohd Jamil and Siraj (2014) the group consensus percentage must exceed 75% (> 75%) for the next process to be taken in place. If the group consensus percentage is lower than 75% it means the items need to be removed or the process need to be redone again.

Step 5 - determining the ∝-cut value (defuzzification value): to determine the ∝-cut value, the data were analyzed with the following formula:

$$\text{Amax} = 1/3 * [(\text{m1} + \text{m2} + \text{m3})]$$

The formula was used to rank and discard the items which have ∝-cut value lower or greater than 0.5. The items with lower value (< 0.5) have been removed from the list as it has not achieved the levels of agreement or consensus from all panel of experts (Mustapha & Darusalam, 2018).

Results and discussion

The feedbacks from the panel experts were entered onto the Microsoft Excel worksheet file (figure 2.0).

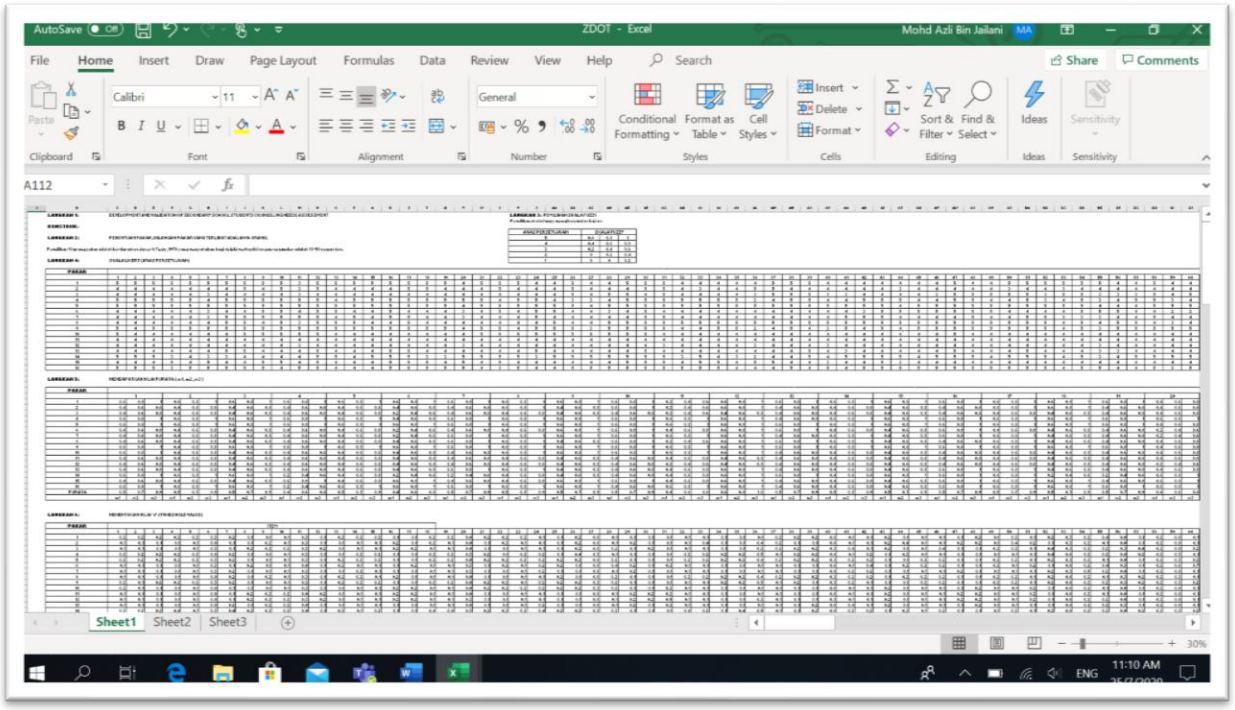

Figure 2.0: Fuzzy Delphi Method Worksheet

The feedbacks were analyzed to quantify the Threshold “d” value. From the data sheet it was found that the overall scores of the “d” value for 126-items is 0.2, which means it surpasses the requirement (≤ 0.2) to proceed to the next procedure.

Step 4 is determining group consensus percentage. As deliberated in earlier chapter, the consensus the panel experts must exceed the minimum requirement of 75 percent before it can be proceeded to the next process. From the results, it has shown that the overall percentage score for this process is 91.35 percent, which means the group agreement has exceeded the minimum percentage value of 75 percent. The decision whether to retain or discard some items that have scored low percentage values was determined by the final process, that is determining or identifying the ∝-cut value or defuzzification value.

For step 5 (defuzzification value), a formula (i – A = 1/3 * [m1 + m2 + m3]) was used to find the ∝-cut value of each item. If the score value is 0.5 and below (< 0.5), the item should need to be removed from the list. The ∝-cut value also would help the researcher to rank the items according to its priority and importance as per reviewed by the panel experts.

From the data, it has shown that all items have scored the ∝-cut values more than 0.5 (< 0.5), except for item 62, item 63, item 76, item 77 and item 115, which the score points were 0.5 and below. The items were ranked according to the agreement made by the panel experts. Table 2.0 explains the details.

	NILAI SKOR		
ITEM	FUZZY EVALUATION	AVERAGE OF FUZZY NUMBER	RANK
1	11	0.7	25
2	10.6	0.7	66
3	10.8	0.7	60
4	10.2	0.6	97
5	10.6	0.7	66
6	9.8	0.6	110
7	11	0.7	25
8	11.6	0.7	6
9	11.4	0.7	15
10	11.6	0.7	6
11	9.8	0.6	105
12	12.2	0.8	1
13	10.8	0.7	44
14	10.8	0.7	44
15	10.6	0.7	66
16	11.6	0.7	13
17	11.8	0.7	3
18	10.8	0.7	60
19	10.4	0.7	86
20	9.8	0.6	110
21	10.8	0.7	44
22	11	0.7	25
23	10.8	0.7	60
24	10.8	0.7	60
25	10.8	0.7	44
26	11	0.7	25
27	9.8	0.6	110
28	10.8	0.7	44
29	11.6	0.7	13
30	12.2	0.8	1
31	11.6	0.7	6
32	10.2	0.6	94
33	10.6	0.7	66

34	10.8	0.7	44
35	10.2	0.6	97
36	8.8	0.6	120
37	10.8	0.7	44
38	11	0.7	25
39	11	0.7	25
40	9.4	0.6	113
41	10.2	0.6	94
42	10.6	0.7	66
43	11.4	0.7	15
44	10.2	0.6	94
45	10.6	0.7	66
46	10.6	0.7	66
47	11.8	0.7	3
48	11.4	0.7	15
49	10.8	0.7	44
50	10.8	0.7	44
51	11.6	0.7	6
52	10.6	0.7	66
53	9.8	0.6	105
54	11.2	0.7	19
55	8.8	0.6	120
56	9.8	0.6	105
57	10.2	0.6	93
58	11.2	0.7	19
59	9.2	0.6	116
60	10.2	0.6	97
61	10.4	0.7	86
62	**6.9**	**0.4**	**126****
63	**8.6**	**0.5**	**123****
64	10	0.6	102
65	10.6	0.7	66
66	10.8	0.7	44
67	11.2	0.7	23
68	10.6	0.7	66
69	11	0.7	25
70	9.2	0.6	116
71	9.4	0.6	113
72	10.8	0.7	60
73	10.8	0.7	60
74	10.6	0.7	66
75	10.2	0.6	97
76	**7.9**	**0.5**	**124****
77	**8.6**	**0.5**	**122****
78	10	0.6	103
79	9.4	0.6	118
80	9.2	0.6	116
81	11	0.7	25

82	10.4	0.7	86
83	9.8	0.6	105
84	11	0.7	25
85	11.2	0.7	23
86	10	0.6	103
87	10.6	0.7	66
88	10.8	0.7	44
89	10.8	0.7	44
90	10.8	0.7	44
91	9.8	0.6	105
92	10.6	0.7	66
93	10.6	0.7	66
94	10.8	0.7	44
95	10.6	0.7	66
96	11.2	0.7	19
97	11	0.7	25
98	9.2	0.6	116
99	10.4	0.7	86
100	10.8	0.7	44
101	11	0.7	25
102	10.2	0.6	97
103	10.6	0.7	66
104	10.4	0.7	86
105	11	0.7	25
106	10.6	0.7	66
107	10.6	0.7	66
108	10.4	0.7	86
109	10.8	0.7	44
110	11	0.7	25
111	11	0.7	25
112	10.4	0.7	86
113	10.6	0.7	66
114	11	0.7	25
115	**7.1**	**0.4**	**125****
116	11.8	0.7	3
117	11.6	0.7	6
118	11.6	0.7	6
119	11.6	0.7	6
120	11.4	0.7	15
121	11	0.7	25
122	11.2	0.7	19
123	11	0.7	25
124	11	0.7	25
125	10.6	0.7	66
126	11	0.7	25

Table 2.0: the $\propto$-cut values and the items rankings

Finally, item 62, item 63, item 76, item 77 and item 115 were removed from the list as the items have scored the $\propto$-cut of 0.5 and below, and the items were ranked 122 (item 77), 123 (item 63), 124 (item 76), 125 (item 62) and 126 (item 115) by the panel of experts.

Conclusion

In summary, the findings have shown that this study managed to answer the research question: “What do the panel experts say about the 126-items counseling needs assessment instrument through the Fuzzy Delphi method (FDM) of analysis?”

Through the process, the agreement of the panel experts was quantified and summarized quantitatively. This has proven the earlier assumption that the time of completing the process of getting the panel experts consensus can be shorten via FDM. This exercise also is seen to be practical in eliminating confusions that were driven from the conventional method, which is more qualitative in nature. FDM standardizes the process especially through its linguistic scale for not to limit the feedbacks with just the normal responses such as ‘good’ or ‘very good’ but with the fuzzy scales.

References

Abdul Rahman, M.N., Mariani, M.N., Ahmad Nadzim, N., Mohd Radzi, N.M., & Moktar, N. 2016. Application of fuzzy delphi approach in designing homeschooling education for early childhood islamic education. International Journal of Academic Research in Business and Social Sciences, 6(12), 837 – 846.

Ahmad, Z., Patahol Wasli, M.M., Mohd Fauzi, M.S.H., Mohd Jamil, M.R., & Siraj, S. 2014. Fuzzy Delphi analysis for future environmental education using interactive animation. 2nd International Seminar Teaching Excellence and Innovation, Universiti Malaya.

Anastasi, A. 1992. What counselors should know about the use and interpretation of psychological tests. Journal of counseling and development, 70, 610 – 615.

Bahri Yusoff, M.S. 2019. ABC of content validation and content validity index calculation. Education in Medicine Journal, 11(2), 49 – 54.

Drost, E. A. 2011. Validity and reliability in social science research. Education Research and Perspectives, 38 (1), 105-124.

Fraenkel, Jack R, & Norman E. Wallen. 2008. How to Design and Evaluate Research in Education. Boston: McGraw-Hill.

Heale, R., & Twycross. 2015. Validity and reliability in quantitative studies. Evidence-Based Nursing, 18(3), 66 – 67.

Hsu, C., & Sandford, B. 2007. The Delphi technique: Making sense of consensus. Practical Assessment, Research and Evaluation, 12(10), 1 – 8.

Leppma, M., & Jones, K.D. 2018. Multiple Assessment Methods and Sources in Counseling: Ethical Considerations. Retrieved from http://www.counseling.org/knowledge-center/vistas

Mustafa, R., & Darusalam, G. 2018. Applikasi Kaedah Fuzzy Delphi dalam penyelidikan sains sosial. Kuala Lumpur: Penerbitan Universiti Malaya.

Rubio, D.M., Berg-Weger, M., Tebb, S.S., Lee, E.S., & Rauch, S. 2003. Objectivefying content validity: Conducting a content validity study in social work research. National Association of Social Workers, Inc. 94 – 104.

Taherdoost, H. 2016. Validity and reliability of the research instrument; How to test the validation of a questionnaire/survey in a research. International Journal of Academic Research in Management (IJARM), 5, 28-36.

Tunner. D.S. 2005. Expertise in teaching: perspectives and issues. In Johnson, K. (Ed.) Expertise in second language learning and teaching. New York: Palgrave Macmillan.

Yaghmale F. 2003. Content validity and its estimation. Journal of Medical Education, 3, 25 – 27.

Yu, L.H., Cheng, H.L., & Kreng, V.B. 2010. The application of Fuzzy Delphi method and Fuzzy AHP in lubricant regenerative technology selection. Experts Systems with Applications, 37, 419 – 425.