The Validity and Reliability of the Teaching Supervision Instruments in Culinary Field: An Analysis of the Rasch Measurement Model

Suriani Mohamed, Juniza Jubri, Irdayanti Nashir

Faculty of Technical and Vocational, Universiti Pendidikan Sultan Idris, Tg. Malim, Perak. Malaysia

Abstract

The implementation of teaching supervision is mandatory for all teachers. Feedback in teaching supervision is a very important aspect. Not only that, supervisors need to have sufficient skills, updated knowledge on the changes in subject syllabus, and reasonable recommendations to improve the teachers they supervised. There is a notion that an accurate teaching supervision model using specific instruments needs to be created so that the teaching and learning supervision can be carried out more effectively. Along with the issues and focus, this study aimed to test the validity and reliability of teaching supervision instruments in the culinary field. The design of the study involves the development of instruments based on the data gathered from interviews, where this study utilised a Modified Delphi Methodology approach. The researcher conducted a test on 100 respondents who were among administrators tasked to carry out teaching supervision in classrooms. The Rasch Measurement Model with Winsteps 3.73.0 software was used to obtain the validity and reliability of each item in the study of teaching supervision in culinary field. In the instrument evaluation phase, the data analysis went through five stages to fulfil the criteria and conditions for instrument usability. The results of the analysis show that of the 45 items in the teaching supervision in culinary field study, two items were dropped and 43 remained. With this, the research instrument can be used as a reference for school administrators in the actual teaching supervision.

Keywords: Teaching supervision, Culinary, Rasch Measurement Model.

Introduction

The supervision of teaching and learning is vital to ensure the effectiveness of any teaching method. According to Glickman, Gordon, and Ross-Gordon (2014), the teaching supervision is akin to the "glue of a successful school", where it becomes an essential factor in guaranteeing school's efficacy. In short, when teachers and students have the same objective in class, the teaching and learning process is considered complete. This aspect is also in the

same vein with teachers' supervision by the school administration, where if done consistently, drives the attainment of school's objective. To this end, a teaching supervisor must have the knowledge, interpersonal skills and technical skills to accomplish the teaching observation successfully (Anizam & Farzeeha, 2014).

At school level, teaching supervision is compulsory in all classrooms. This exercise has been mandated by the Ministry of Education Malaysia through the Professional Circular No. 3/1987: The Supervision of Teaching and Learning in Classrooms by School Principal/Headmaster. The circular explains that headmasters and principals in schools are required to monitor and supervise the teaching and learning that happens in classrooms. The circular also includes a clear guideline to implement the supervision process.

To make sure that the teaching and learning supervision runs smoothly, headmasters and principals must select the teaching supervisor among the school administrator body, such as the Senior Assistant of Administration and Curriculum, Senior Assistant for Students Affairs, Senior Assistance for Co-Curriculum, Afternoon Session Supervisor, Heads of Subjects, and Head of Subjects Committee. However, the headmaster or principal assumes the final responsibility in the supervision of teaching and learning.

Jemaah Nazir Sekolah (JNS), a professional body authorised in supervising teaching and learning at schools, reported in the Malaysian Education Blueprint (2013 - 2025) the different understanding that schools and Jemaah Nazir had regarding the meaning of quality teaching and learning. The outcomes from the School Self-Evaluation Report done by schools' administrators in Malaysia showed 63% schools determined they achieved 'Good' and 'Excellent' ratings for their teaching and learning process. However, the evaluation done by JNS revealed only 13% schools in all of Malaysia actually reached 'Good' and 'Excellent' level. This demonstrated the dissimilar understanding mentioned earlier, even though both parties utilised the same evaluation instruments. This raises the questions as to how schools and JNS define quality teaching and learning, when both were given the same tools to determine the ratings. To present the perspective of how the current supervision of teaching and learning is done, Abdul Rahman (2011) summarised the situation by saying how a physics teacher who graduated in 2000 needs the support and advice from a principal who obtained a degree in Malay Literature in 1980. Ong'ondo and Borg (2011) confirmed this observation by detailing how supervisors' feedback mostly focused on the common pedagogical aspects of teaching that had little to no relation with the subject taught. It is clear through this revelation that the teaching supervision in schools concentrates on the methodology side of teaching but leaves the teachers unclear on how the methodology can help better deliver the subjects involved.

The same problem was also detected by Hamidah, Jamal, Sharifah, and Syed (2016) who found teachers to provide unfavourable responses when inquired about the feedback of the supervision they received. This is worrying, as Weber, Chandler, and Finley (2016) highlighted the importance of feedback to guarantee a continuous improvement, where the readiness and acceptance of comments on the teacher's performance must be present. Further, Hamdan and Nurlia (2015) also agreed on the significance of supervision feedbacks. They opined that supervisors must have sufficient pedagogical knowledge, updated knowledge of subject syllabus they are assessing, and the skills to identify areas of improvements.

Therefore, it is important to have a set of valid instruments when conducting teaching supervisions. However, past research findings on the instruments used when supervising culinary teachers' pedagogical skills elucidated a lack of differentiation between general practice of teaching from those relevant for teaching culinary subjects (Feng, Su, & Yang, 2011). The same issue was also raised by Ambarwangi and Rachman (2015) and Paulsen and Martin (2013), where they put forth the need to develop an accurate model for teaching supervision for culinary subjects to warrant an effective teaching and learning supervision.

As such, it stands to reason that the instrument to supervise teaching and learning must be suitable with a particular field, which in this case, the culinary field. Any instruments developed should be put to the test to determine their efficacy to be used in all culinary schools in the country. In line with the issue and focus of this study, this research aims to test the validity and reliability of teaching supervision instruments in the culinary field.

Research Methodology

The methodology used in this study involved the instruments developed through interviews, where the Modified Delphi Methodology was utilised. Through this method, the researcher obtained data from interviews with culinary experts, literature reviews and documents analysis. Next, the researcher developed an instrument based on the data gathered. This instrument was put through the validity and reliability test by administering a survey on culinary teachers in the instrument evaluation phase.

The instrument evaluation phase began by developing the questionnaire on teaching supervision for culinary field from the JAK analysis in the second and third cycle of Modified Delphi Methodology. This was followed by two pilot tests. An assessment was conducted to determine the suitability of the instrument with the identified constructs. The validity of the constructs obtained from the actual study proved that the constructs in the instrument were appropriate. The researcher then carried out an actual test on 100 respondents, which consisted of administrators who conducted teaching supervision in secondary schools.

The Rasch Measurement Model with Winsteps 3.73.0 software was used to determine the validity and reliability of each item in the teaching supervision in the culinary field study. The Rasch analysis was utilised to measure the difficulty level for each item in the questionnaire, which aided in the instrument refining process. According to Wright and Mok (2004), the Rasch Model is useful for the measurement process in social sciences because it meets the following five criteria or assumptions: (a) produces a consistent size with the same time interval, (b) provides accurate estimation process, (c) detects misfits or outliers, (d) overcomes missing data and (e) produces non-repetitive measurements (independent of the parameters examined). The validity of the instrument can be identified by referring to key analysis such as item polarity, item-individual map, item-individual mismatch, item-segregation, unidimensional, item-matching and rating scale.

Data Analysis

In the instrument evaluation phase, the data analysis must go through several stages by fulfilling the conditions and requirement for instrument usability. They are:

STAGE 1
Testing the reliability and items isolation/respondent index.

Table 1 shows the summary of the reliability and items segregation index for the constructs in the Teaching Supervision for Culinary Field research. The items reliability index shows a high value of 0.97. Further, the items segregation index indicates the respondent's ability to isolate the difficulty of the items to several levels in the measured construct. Index for the isolation of the items in the Teaching Supervision for Culinary Field study shows a good value of 5.66.

Table 1
Reliability and Index of Item Isolation for Teaching Supervision Instruments in the Culinary Field: Actual Study

Summary of 45 Measured Item

```
--------------------------------------------------------------------------------
|     TOTAL                  MODEL      INFIT      OUTFIT    |
|     SCORE    COUNT    MEASURE  ERROR    MNSQ  ZSTD  MNSQ  ZSTD |
|------------------------------------------------------------------------------|
| MEAN   439.5   100.0      .00    .21    1.00   -.3  1.01   -.3 |
| S.D.    26.5     .0     1.28    .02     .20   1.5   .24   1.5 |
| MAX.   486.0   100.0     3.41    .30    1.78   3.7  1.87   3.8 |
| MIN.   369.0   100.0    -2.45    .20     .77  -2.6   .74  -2.7 |
|------------------------------------------------------------------------------|
| REAL RMSE   .22 TRUE SD   1.26 SEPARATION 5.66 ITEM  RELIABILITY .97 |
|MODEL RMSE   .21 TRUE SD   1.26 SEPARATION 5.88 ITEM  RELIABILITY .97 |
| S.E. OF ITEM MEAN = .19                                    |
--------------------------------------------------------------------------------
```

Table 2 shows the summary of the reliability and isolation index of the respondents measuring the constructs in the Teaching Supervision for Culinary Field study. The reliability index of the respondents recorded a high value of 0.81. The items segregation index indicates that the quality of the items is capable of indicating respondents to several abilities. Further, the isolation index of respondents showed a good value of 2.06.

Table 2
Reliability and Respondent Isolation Index for Teaching Supervision Instruments in the Culinary Field: Actual Study

```
TABLE 3.1 ACTUAL STUDY.sav            ZOU952WS.TXT  Sep 22 15:48 2019
INPUT: 100 PERSON  45 ITEM  REPORTED: 100 PERSON  45 ITEM  4 CATS  WINSTEPS 3.73
--------------------------------------------------------------------------------

SUMMARY OF 100 MEASURED PERSON
--------------------------------------------------------------------------------
|    TOTAL                MODEL     INFIT     OUTFIT  |
|    SCORE    COUNT   MEASURE ERROR   MNSQ ZSTD MNSQ ZSTD |
|------------------------------------------------------------------------|
| MEAN   197.8    45.0     3.80   .32    .99   .0 1.01   .0 |
| S.D.    7.6     .0      .76   .02    .23  1.4  .28  1.4 |
| MAX.   214.0    45.0     5.63   .41   1.66  3.7 1.96  3.5 |
| MIN.   179.0    45.0     1.88   .31    .30 -3.5  .27 -3.5 |
|------------------------------------------------------------------------|
| REAL RMSE   .33 TRUE SD    .69  SEPARATION  2.06  PERSON RELIABILITY .81 |
|MODEL RMSE   .32 TRUE SD    .69  SEPARATION  2.17  PERSON RELIABILITY .82 |
| S.E. OF PERSON MEAN = .08                              |
--------------------------------------------------------------------------------
PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .82
```

STAGE 2
Determining and Identifying Items Polarity

Table 3 shows the summary for the items polarity that measured identical constructs. The value for items polarity, The Point Measure Correlation (PTMEA), shows positive value with 0.3 logits. Items that complied with the set PTMEA values confirmed that they were measured in only the same direction. Items A10 and B11 were dropped as their minimum PTMEA values were less than 0.3 logits.

Table 3
Item Polarity for Teaching Supervision in the Culinary Field: Actual Study

No	Construct	PTMEA CORR	Item	Total
1.	Planning and Preparation	0.23	A10	1
2.	Execution and Presentation	0.23	B11	1
3.	Evaluation Method	0.56	C5	1
4.	Objectives Attainment and Reflection	0.45	D8	1
5.	Teacher's Professional Responsibility	0.33	E2	1

STAGE 3
Identifying items' suitability

Table 4 shows the summary for the mismatched constructs in the Teaching Supervision in the Culinary Field study. Mean-square (MNSQ) and z-standardised (ZSTD) was used to detect data that contradicted the Rasch measurement model. The MNSQ value for Likert scale items is between 0.60 to 1.40 logits. Items outside the MNSQ value rate need to be dropped or repaired. The ZSTD values indicate whether the data matches the Rasch measurement model perfectly. It also represents the probability of the data being significant

if the data corresponds to the Rasch measurement model. The ZSTD values range from -2.00 to +2.00. However, the researcher decided to ignore the value for ZSTD if the MNSQ value was accepted with reference to Linacre (2011).

Table 4
Summary of items mismatches: Actual Study

No	Constructs	Item	Infit		Outfit		PTMEA
			MNSQ	ZSTD	MNSQ	ZSTD	CORR
1	Planning and Preparation	A10	1.43	2.9	1.42	1.3	0.23
		A9	0.66	-3.1	0.57	-2.7	0.74
		A8	0.65	-3.5	0.52	-3.4	0.75
2	Execution and Presentation	B5	1.46	3.0	1.41	1.9	0.41
		B11	1.43	2.2	1.45	0.8	0.23
		B10	0.65	-2.8	0.52	-2.2	0.76
		B4	0.62	-3.2	0.45	-3.2	0.80
3	Evaluation Method	C4	0.65	-2.4	0.41	-2.6	0.82
		C2	0.58	-2.7	0.39	-2.7	0.87
4	Objectives Attainment and Reflection	D8	1.29	2.1	1.57	1.7	0.45
		D5	1.50	3.2	1.37	1.2	0.46
		D6	0.72	-2.3	0.51	-1.9	0.69
		D10	0.56	-3.6	0.37	-2.5	0.80
5	Teacher's Professional Responsibility	E5	1.38	2.3	1.51	1.3	0.35
		E2	1.36	2.8	1.47	2.7	0.33

A total of 15 misfit items were identified out of 45 from the Teaching Supervision in Culinary Field study based on the value of the MNSQ infit/outfit index. They were three items that were taken from the Planning and Preparation constructs, four items from the Execution and Presentation construct, two items from the Evaluation Method construct, four items from the Objectives Attainment and Reflection construct, and two from Teacher's Professional Responsibility construct.

STAGE 4
Determining and Detecting Items Measuring Single or Unidimensional Constructs (Standard Residual Variance)

Dimensional uniformity ensures that the instrument can be firmly measured and achieved based on the Residual Principal Component Analysis technique. This technique is able to detect the ability of an instrument to measure in a uniform dimension with an acceptable level of interference. Table 5 shows a summary of the Principal Component Analysis (PCA) based on the variance explained by measuring the competency constructs. The PCA value is accepted because it exceeds 20%, which is 31.2%. For unexplained variance by 1st contrast (size), the degree of interference of the item in one contrast is accepted if it is worth less than 15%. The value of unexplained variance by 1st contrast (size) for the constructs in the Teaching Supervision in Culinary Field research is 8.6%, which is within the acceptable value.

Table 5
Unidimensional: Standardised Residual Variance for the Five Constructs in the Teaching Supervision in the Culinary Field: Actual Study

TABLE 23.0 ACTUAL STUDY.sav
INPUT: 100 PERSON 45 ITEM REPORTED: 100 PERSON 45 ITEM 4 CATS WINSTEPS 3.73

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
-- Empirical --   Modeled
Total raw variance in observations =   65.4 100.0%   100.0%
Raw variance explained by measures =   20.4 31.2%   30.9%
Raw variance explained by persons =   4.3   6.6%   6.6%
Raw Variance explained by items   =   16.1 24.6%   24.4%
Raw unexplained variance (total)   =   45.0 68.8% 100.0%   69.1%
Unexplned variance in 1st contrast =   5.6   8.6% 12.5%
Unexplned variance in 2nd contrast =   4.4   6.8%   9.9%
Unexplned variance in 3rd contrast =   4.2   6.3% 9.2%
Unexplned variance in 4th contrast =   3.4   5.1% 7.5%
Unexplned variance in 5th contrast =   2.8   4.3% 6.2%

Looking at Table 6, six items have residual correlation values that are more than 0.7 logits. This situation was caused by respondents being confused by items that shared overlapping qualities. These items need to be repaired or removed by looking at their MNSQ and ZSTD values. The item with MNSQ infit value approaching 1.00 and ZSTD value approaching 0.00 should be retained, and the other should be removed. Table 4.32 shows the items that were dropped after going through the screening process by looking at the MNSQ value approaching the value of 1.00 and the ZSTD approaching the value of 0.00. After going through the screening process, no items were dropped.

Table 6
Dropped Items with Residual Correlation Values exceeding 0.7 Logits.

```
TABLE 23.99 ACTUAL STUDY.sav
INPUT: 100 PERSON  45 ITEM  REPORTED: 100 PERSON  45 ITEM  4 CATS  WINSTEPS 3.73
--------------------------------------------------------------------------------
LARGEST STANDARDIZED RESIDUAL CORRELATIONS
USED TO IDENTIFY DEPENDENT ITEM
---------------------------------
|CORREL-| ENTRY    | ENTRY    |
|  ATION|NUMBER ITE |NUMBER ITE |
|-------+-----------+-----------|
|  .86 |   14 B04 |   22 B12 |
|  .84 |   25 C02 |   29 C06 |
|  .83 |   32 D03 |   36 D07 |
|  .81 |   13 B03 |   17 B07 |
|  .78 |   30 D01 |   35 D06 |
|  .73 |   14 B04 |   19 B09 |
|  .72 |   18 B08 |   23 B13 |
|  .71 |   11 B01 |   16 B06 |
|  .69 |   24 C01 |   25 C02 |
|  .68 |   33 D04 |   38 D09 |
---------------------------------
```

STAGE 5
Determining and confirming the level of difficulty of items and the abilities of respondents.

Table 7 shows that Teaching Supervision in Culinary Field research had respondents who were able to answer items with the highest level of difficulty. The measurement value of the 11th and 46th respondents is +5.63 logits, above the value of item C3, which is +3.41. Figure 1 shows the Wright Map item distribution and respondents' abilities for the items in the Teaching Supervision in Culinary Field study. There is also an easy item that can measure weak respondents, which is B11 at -2.45.

Table 7

Items Difficulty and Respondents' Ability for Five Constructs in the Teaching Supervision in Culinary Field Study

Constructs	Item Measurement	Item	Respondents' Ability	Respondents
Teaching Supervision in Culinary Field	-2.45/+3.41	B11/C3	1.88/5.63	17/11,46

TABLE 12.2 ACTUAL STUDY.sav
INPUT: 100 PERSON 45 ITEM REPORTED: 100 PERSON 45 ITEM 4 CATS WINSTEPS 3.73

```
    PERSON - MAP - ITEM
       <more>|<rare>
 6        +    ┌──────────┐ ┌──────────┐ ┌────────┐ ┌──────────┐ ┌──────────┐
       #  |    │Planning and│ │Execution and│ │Evaluatio│ │Objectives │ │ Teacher's │
          |    │ Preparation│ │Presentation│ │  n      │ │Attainment │ │Professional│
       # T|    │          │ │          │ │ Method  │ │          │ │Responsibility│
 5       # +   │          │ │          │ │        │ │  and     │ │          │
     .####  |  │          │ │          │ │        │ │          │ │          │
    .###### S| │          │ │          │ │        │ │ Reflection│ │          │
       ##  |   └──────────┘ └──────────┘ └────────┘ └──────────┘ └──────────┘
 4    ###### +
     ##### M|
    .######## |                          C03
    .##### |────────────────────────────C01─C02────────────────────────────
 3    .## S+                             C04
     ###  |                             C06
     .## |T
        T|
 2     # +
        |                    C05
        |
        |S
 1       +
        | A02
        |          B02            D02
     ───| A07─────────────────────────────────────────────────────────────
 0      +M A06      B06  B08           D06 D08      E03 E04 E06
        | A01 A03 A05 A08   B01 B05 B07 B13
                                  D01 D05 D07
        | A04      B03  B09         D03 D09    E02
        |          B04              D04      E01
-1      + A09      B10 B12          D10
        |S
        |
```

```
            |  A10
   -2       +                                                        E05
            |
            |T            B11
         ───┼─────────────────────────────────────────────────────────
   -3       +
          <less>|<frequ>
EACH "#" IS 2. EACH "." IS 1.
```
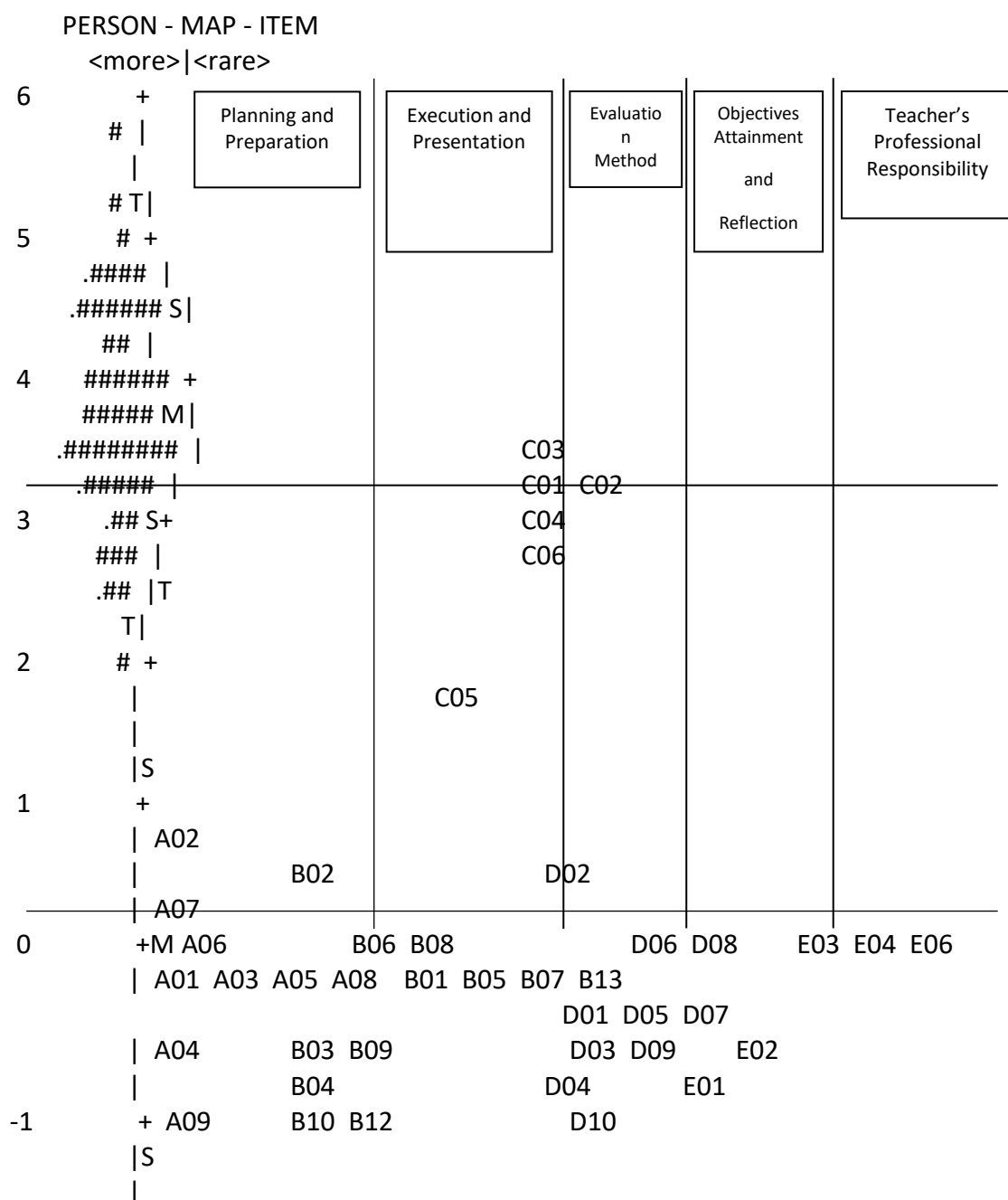
Figure 1. Wright Map Item Distribution and Respondents' Ability in the Teaching Supervision in Culinary Field Study

The final findings of the study at the validation stage for the five constructs of food preparation and serving instruments showed item reliability index of 0.97 and respondent reliability of 0.81 then this study instrument is acceptable, high and excellent in parallel as suggested by Bond and Fox (2015). This indicates that in general the Food Preparation and Presentation instruments are consistent and stable when administered on other samples that have similar and almost identical characteristics.

Conclusion

In conclusion, the Food Preparation and Presentation instruments are of good quality based on the reliability index of items and respondents as well as the isolation index of items and respondents is achieving standards based on Rasch measurement model. The research instrument can be used as a reference for school administrators and to be adapted in the actual teaching supervision. The development of this teaching supervision instrument can also assist school administrators who do not have the background in culinary arts to still carry out effective teaching supervision.

Researcher suggest several further studies that can be done so that further studies can provide information in the preparation of instruments in the field of food preparation and serving. The suggestions are (i) build instruments related to theoretical teaching supervision in the field of food preparation and serving, (ii) conduct research on the construction of other vocational field supervision instruments, and (iii) Add construct and item instruments in the field of food preparation and serving to further strengthen the instrument.

Correspondent author
Suriani Mohamed
Faculty of Technical and Vocational,
Universiti Pendidikan Sultan Idris
Email: Suriani.mohamed@ftv.upsi.edu.my

References
Abdul Rahman, A. M. K. (2011). Penyeliaan Klinikal dari Perspektif Pengajaran di Sekolah (Pertama). Kuala Lumpur: *PTS Publications.*
Bond, T. G., & Fox, C. M. (2015). Applying the Rasch Model: Fundamental Measurement in the Human Sciences (3rd Edition). New York: *Routledge.*
Feng, L. Y., Su, Y. H., & Yang, C. C. (2011). The Construction of Professional Evaluation

Instruments for Hospitality Teachers in Taiwan's Vocational Schools. *Journal of Teaching in Travel and Tourism, 11(3),* 229–252.

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2014). Supervision and Instructional Leardership A Developmental Approach. (9th Edition). USA: *Pearson Education.*

Hamdan, S., & Nurlia, M. (2015). *Penyeliaan Pengajaran Dan Pembelajaran Di Sekolah-Sekolah Kebangsaan Di Daerah Kota Tinggi, Johor.*

Hamidah, Y., Jamal @ Nordin, Y., & Sharifah, H., & Syed, D. (2016). Aplikasi Analisis Modifikasi Kepentingan-Prestasi Amalan Penyeliaan Pengajaran Ketua Panitia seperti sedia ada dan yang diperlukan oleh guru. *Malaysian Journal of Society and Space, 8(8),* 89–99.

Linacre, J. M. (2011). Winsteps® rasch measurement computer program user's guide. Beaverton, Oregon: Winsteps.com

Mohd Anizam, M. D., & Dayana Farzeeha, A. (2014). Fungsi Penyeliaan Pengajaran dalam Pendidikan Teknik dan Vokasional. *International Seminar on Technical and Vocational Education, TVEIS,* 608–617.

Ong'ondo, C. O., & Borg, S. (2011). "We Teach Plastic Lessons to Please Them": The Influence of Supervision on the Practice of English Language Student Teachers in Kenya. *Language Teaching Research, 15(4),* 509–528.

Paulsen, T. H., & Martin, R. A. (2013). Instructional Supervision of Agricultural Education Teachers: Perceptions Regarding Selected Beliefs. Journal of Agricultural Education, 99(2), 99–113.

Weber, M. R., Chandler, J. A., & Finley, D. A. (2016). Assessing Teaching Effectiveness in a Basic Food Laboratory Setting: Pilot Testing the Instrument. *Journal of Teaching in Travel and Tourism, 11(2),* 211–225.

Wright, B. D., & Mok, M. M. C. (2004). An overview of the Family of Rasch Measurement Models. *Introduction to Rasch Measurement.*