INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 13, Issue 3, (2024) E-ISSN: 2226-6348

Validity and Reliability of Preschool Teacher's Creative Teaching Instrument for Children's Language Arts: A Pilot Study of The Rasch Measurement Model

Nur Atiqah Mat Saberi, Nor Mashitah Mohd Radzi & Loy Chee Luen

Faculty of Human Development, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak Darul Ridzuan, Malaysia

To Link this Article: http://dx.doi.org/10.6007/IJARPED/v13-i3/22170 DOI:10.6007/IJARPED/v13-i3/22170

Published Online: 18 July 2024

Abstract

This pilot study was conducted to examine and verify the validity and reliability of Preschool Teacher's Creative Teaching Instrument for Children's Language Arts. This questionnaire instrument consisted of 45 items and was distributed to 50 preschool teachers from the Larut Matang and Selama district in Perak. This instrument was developed to measure the nine main constructs of this study, as follows: encouraging students to learn freely, working together with students in learning, motivating students, withholding judgement on creative ideas, encouraging flexible thinking, encouraging students to conduct self-assessment of the ideas generated, taking into account all suggestions and questions put forward by students, offering students the chance to do tasks, and helping students to overcome disappointment, if they face failure upon trying something new. This approach was used to examine the validity and reliability of the items and the respondents in this study, which emanated from the Rasch Measurement Model. This approach was more valid and well-grounded compared to only focusing on the output produced by Cronbach's alpha. The Winsteps software, version 3.68.2, was used to analyse four diagnoses that were performed on functional items: (i) item reliability and separation; (ii) detect the polarity of items that measure the constructs based on the value of PTMEA CORR; (iii) the fit of constructing items; and (iv) unidimensionality to determine the dependent items based on the standard residual correlation value. At the end of the analysis, 45 items were found to meet the inspection criteria, in accordance with the Rasch Model. The final instrument recorded 45 items that can be used to measure the nine constructs of this study.

Keywords: Creative Teaching, Validity, Reliability, Rasch Measurement Model

Introduction

The instrument used in this study was based on the creative assessment criteria proposed by Soh (2015), as follows: to encourage students to learn freely, to collaborate with other students in their learning, to motivate students frequently, to hold back judgement towards a creative idea, to encourage flexible thinking, to encourage students to conduct self-

assessment towards ideas produced, to take into account all suggestions and questions put forward by students, to offer students the chance to do assignments, and to help them in handling disappointment when they fail upon trying something new.

By measuring these creative assessment criteria, this study can help preschool teachers achieve their goals and improve their teaching performance in language arts. Thus, this pilot study was conducted to ensure that the questionnaire instrument has good validity and reliability via the Rasch Model approach. This approach was chosen, as each item can be thoroughly checked and discussed rather than by just looking solely at the values of Cronbach's alpha.

Literature Review

The Communication Strand is one of the most important strands in this curriculum. The Communication Strand emphasises on oral and non-oral language skills during interaction. It contains language knowledge discipline that is compulsory to be learned by all preschoolers. Among the key objectives of this strand is to give students the chance to use language to communicate effectively. Therefore, teachers are encouraged to use their creativity to choose, organise, modify, and diversify activities guided by the Learning Standard according to preschoolers' suitability as preparation to continue their education later at the primary school level (Kementerian Pendidikan Malaysia, 2017).

Teachers are advised to conduct teaching and learning of language skills through various activities and fun language games during classes (Kementerian Pendidikan Malaysia, 2017). However, the National Preschool Curriculum Standard does not provide details on how the process of teaching and learning through play activities during class can enhance children's language skills (Ali & Mahamod, 2015). Therefore, there is a need for a valid and trusted instrument to determine the aspects of teachers' creative teaching that are subjective to be measured, but not impossible to envision.

This issue needs to be studied, as it is closely related to the development of preschool children. In their study, Abdullah et al (2021) found that some teachers preferred to adopt a teaching approach that is centred on himself or herself. However, teachers' initiatives to come up with more creative lessons are aligned with the government policy to develop creativity and innovation (Ab. Jawas & Zulkifli, 2022). This is in line with the findings of a study conducted by Mokhlis (2019), which reported that to foster creativity, preschool teachers need to understand the differing problems and needs of children. In this context, the key action that teachers should take is to ensure that their planned teaching and learning activities can satisfy the children's interests and level of ability by taking into account their individual differences, giving suitable incentives, and not making excessive comparisons.

The assessment of creative teaching in this study was based on the creative assessment criteria proposed by Soh (2015), who introduced the Creativity Fostering Teacher Index (CFTI). This instrument has nine elements, namely, to encourage students to learn freely, to collaborate with other students in their learning, to motivate students, to hold back judgement towards a creative idea, to encourage flexible thinking, to encourage students to conduct self-assessment towards ideas produced, to take into account all suggestions and questions put forward by students, to offer students the chance to do assignments, and to help them in handling disappointment when they fail upon trying something new.

Each element has its own strengths in being an important aspect of a teacher's creative teaching. The use of domains in the construction of this instrument was the preliminary step in ensuring the viability of teaching approaches that can foster creativity, and in contributing

to the efforts towards producing creative and innovative new generations, particularly in children's language arts. Although the CFTI is standardised and used by many countries around the world, its suitability in the context of preschool teachers' teaching, especially in the aspect of language arts, is still open for discussion.

Although largely applicable in the context of creative teaching, some aspects of CFTI are too general and less suitable for measuring preschool teachers' creative teaching of children's language arts. The approach used to examine the validity and reliability of the items and respondents in this study emanated from the Rasch Measurement Model, which provides more valid and well-grounded output compared to the output produced by Cronbach's alpha.

Rasch Measurement Model

The Rasch model approach is often used to analyse the validity and reliability of a study instrument more deeply through several diagnostic methods (Huei et al., 2020). This approach is intended to test and examine the validity and reliability of a constructed questionnaire. Some of the diagnoses that re performed in this approach are as follows: to test the reliability and the separation index of items and respondents; to identify the polarity item that measures the constructs; to examine the suitability of the item instrument (item fit); to determine the item difficulty level and the ability of the respondents; to determine the structure functionality of the measurement scale category; and to distinguish a unidimensional construct.

The Rasch model incorporates a method for ordering persons according to their ability and ordering items according to their difficulty (Bond & Fox, 2015). The criteria listed in the following Table 1 have been used as benchmarks for determining the validity of the instrument.

Table 1
Summary of item validity and reliability using the Rasch Model

Criteria Statistical	**Value**	**Reference**
Person Reliability Value	> 0.8	Bond & Fox (2015) Sumintono & Widhiarso (2015)
Item Reliability Value	> 0.8	Bond & Fox (2015) Sumintono & Widhiarso (2015)
Separation (SE) All items show	≥ 2.0	Linacre (2006)
Item Fit Total Mean-square infit and outfit (Item Misfit)	0.5–1.5 logits	Sumintono & Widhiarso (2015)
Item Polarity PTMEA CORR	Positive (+)	Bond & Fox (2015)
Unidimensionality – Value Principal Component Analysis of Residual (PCA)	Minimum 20%	Sumintono & Widhiarso (2015) Aziz et al. (2015)
Unidimensionality – Unexplained variance in 1st contrast	Maximum 15%	Sumintono & Widhiarso (2015)

		Aziz et al. (2015)
Unidimensionality – Standardised Residual Correlation	< .70	Aziz et al. (2015)

For this pilot study, the Rasch Model was used to examine the validity and reliability of the questionnaire instrument that has been adapted through the quantitative data collection. The validity and reliability of the respective items in a questionnaire will usually be discerned through the overall value of Cronbach's alpha. On the other hand, the Rasch approach is used to determine the validity and reliability of a questionnaire more deeply through some diagnoses. Only four diagnoses were performed in this study to check functional items in terms of the following: (i) item reliability and separation; (ii) detect the polarity of items that measure the constructs based on the value of PTMEA CORR; (iii) the fit of constructing items; and (iv) determine the dependent items based on the standard residual correlation value.

Objective

The objective of this pilot study was to test the reliability of the instrument that has been developed and to detect its weakness. In this pilot study, several inspections have been performed on item functionality in terms of its reliability, and the separation of item-respondent, polarity items, the suitability of the item, and distinguish a unidimensional construct.

Methodology

This pilot study was conducted using a quantitative approach by distributing the developed questionnaire to 50 selected respondents (preschool teachers). This was considered adequate based on report of Linacre (1994) that sample size of as low as 30 to 50 was adequate to run Rasch analysis. According to Cooper and Schindler (2011), the number of respondents considered as adequate for a pilot study is between 25 and 100 people. The findings generated from this pilot study were then analysed using the Winsteps software, version 3.68.2, alongside the Rasch Measurement Model.

Results and Findings

The description and explanation for each item tested on the functionality are presented in the following subsection.

Reliability and Item Separation

In the Rasch Measurement Approach, reliability is based on the value of Cronbach's alpha (α) that can be accepted between 0.71 and 0.99. Thus, this value is at its best at 71% to 99%), as described in Table 2 (Sumintono & Widhiarso, 2015).

Table 2
The Interpretation of Cronbach's Alpha Score

The Score of Cronbach's Alpha	Reliability
0.8–1.0	Very good and effective with a high level of consistency
0.7–0.8	Good and is acceptable
0.6–0.7	Acceptable
0.5–0.6	The item needs refinement
< 0.5	The item needs to be discarded

Rasch analysis was also conducted with reference to the reliability value of the respective items and the value of item separation. The reliability value based on Cronbach's alpha (α) is 0.92, as shown in Table 3 below. This value showed that this instrument was in a good condition and was acceptable. Thus, it can be used in the real research.

Table 3
The Reliability Score (Cronbach's alpha)

Cronbach's alpha (Kr-20) Person Raw Score Test Reliability	0.93

The reliability value and the separation values of the items and respondents for this entire instrument were also analysed. Table 4 shows that the reliability value of the items is 0.87, which indicates that they are in a good condition and are acceptable (Sumintono & Widhiarso, 2015). Meanwhile, the value of item separation was 2.54. As suggested by Linacre (2005), a good index separation value should be greater or more than 2.0.

Table 4
Reliability and Item Separation Values for The Entire Instrument

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	200.5	50.0	.00	.29	.98	-.1	1.01	.0
S.D.	8.5	.0	.84	.05	.25	1.2	.25	1.2
MAX.	213.0	50.0	2.04	.43	1.43	2.2	1.44	1.8
MIN.	180.0	50.0	-1.32	.20	.52	-2.8	.54	-2.8
REAL RMSE	.31	ADJ.SD	.79	SEPARATION	2.54	ITEM	RELIABILITY	.87
MODEL RMSE	.30	ADJ.SD	.79	SEPARATION	2.66	ITEM	RELIABILITY	.88
S.E. OF ITEM MEAN = .13								

Based on the following Table 5, the reliability value of the respondents is 0.92 and the respondent separation value is 3.30. These results showed that the reliability of the respondents was very good and effective, with a high level of consistency. Bond and Fox (2015) have explained that reliability value that exceeds 0.71 is good and acceptable. The value of the items and respondent separation index, which were higher than 2.0, were considered as good (Fox & Jones, 1998; Linacre, 2005; Bond & Fox, 2007).

Table 5
Reliability and Respondent Separation Value for the Entire Instrumen

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	180.4	45.0	.20	.29	.98	-.2	1.01	-.2
S.D.	13.2	.0	1.11	.01	.60	2.4	.68	2.5
MAX.	215.0	45.0	3.36	.37	2.99	8.1	3.03	7.7
MIN.	152.0	45.0	-2.08	.28	.27	-3.8	.28	-3.7
REAL RMSE	.32	ADJ.SD	1.06	SEPARATION	3.30	PERSON RELIABILITY	.92	
MODEL RMSE	.29	ADJ.SD	1.07	SEPARATION	3.66	PERSON RELIABILITY	.93	
S.E. OF PERSON MEAN = .16								

Polarity Item by PTMEA CORR Value

The Point Measure Correlation (PTMEA CORR) value was calculated to identify the polarity items in the study, which would test the extent of which the established constructs can achieve their goals. If the PTMEA CORR value is positive (+), it indicates that the respective item can achieve its goals of measuring the construct that needs to be measured (Bond & Fox, 2015). In contrast, if the value is negative (-), then, the established item does not measure the construct that needs to be measured. Therefore, this item needs to be revised or discarded because it does not address the question, or it is too difficult for the respondents to answer. Based on the following Table 6, no items have recorded negative PRMEA CORR values.

Item Fit in Measuring the Constructs

Item fit measures the constructs through the infit and outfit Mean-square (MNSQ). According to Sumintono and Widhiarso (2015), the outfit and infit MNSQ should be in the range of 0.50 to 1.50 to ensure that the items are suitable for measuring the constructs.

If the infit or outfit MNSQ value is higher than 1.50 logit, it is referring to a confusing item. If the MNSQ value is less than 0.50 logit, it shows that the item is easily anticipated by the respondents (Linacre, 2006). Additionally, the outfit and infit ZSTD values should be within the -2.00 to +2.00 range (Sumintono & Widhiarso, 2015). However, if the outfit and infit MNSQ values are accepted, then, the ZSTD index can be ignored (Linacre, 2006). Therefore, if this condition is not met, the item should either be removed or revised. The following Table 7 shows the list of items that are suitable for measuring the constructs in this study.

The infit and outfit results for these items showed that all items were within the infit and outfit range of 0.5 to 1.5, as follows: D30 (infit 1.38 and outfit 1.44), D20 (infit 1.43 and outfit 1.38), D32 (infit 1.11 and outfit 1.43), D5 (infit 1.39 and outfit 1.33), D6 (infit 1.39 and outfit 1.37), D28 (infit 1.24 and outfit 1.37), D27 (infit 1.35 and outfit 1.21), D40 (infit 1.29 and outfit 1.27), D2 (infit 1.25 and outfit 1.03), D9 (infit 1.08 and outfit 1.21), D22 (infit 1.19 and outfit 119), D41 (infit 1.1 5 and outfit 1.19), D7 (infit 1.10 and outfit 1.18), D17 (infit 1.18 and outfit 1.18), D21 (infit 0.97 and outfit 1.18), D3 (infit 1.08 and outfit 1.17), D34 (infit 1.10 and outfit 1.15), D1 (infit 1.06 and outfit 0.14), D29 (infit 1.11 and outfit 1.12), D31 (infit 1.10 and outfit 1.08), D39 (infit 1.09 and outfit 1.07), D16 (infit 1.05 and outfit 1.06), D15 (infit 1.05 and outfit 1.05), D36 (infit 1.04 and outfit 1.04), D11 (infit 0.86 and outfit 1.02), D42 (infit 0.95 and outfit 1.01), D12 (infit 0.98 and outfit 1.00), D8 (infit 0.92 and outfit 0.98), D4 (infit 0.93 and outfit 0.94), D45 (infit 0.91 and outfit 0.93), D37 (infit 0.89 and outfit 0.85), D43 (infit 0.84 and

outfit 0.88), D13 (infit 0.76 and outfit 0.87), D33 (infit 0.75 and outfit 0.87), D35 (infit 0.85 and outfit 0.75), D38 (infit 0.80 and outfit 0.73), D19 (infit 0.73 and outfit 0.75), D26 (infit 0.73 and outfit 0.71), D14 (infit 0.70 and outfit 0.59), D23 (infit 0.66 and outfit 0.66), D18 (infit 0.52 and outfit 0.65), D24 (infit 0.60 and outfit 0.62), D25 (infit 0.59 and outfit 0.57), D44 (infit 0.56 and outfit 0.56), D10 (infit 0.54 and outfit 0.54), D8 (infit 0.92 and outfit 0.98), and D4 (infit 0.93 and outfit 0.94). The items where the reading approached or was 1.00 were the highest-quality items, where 100% of the items measured the said construct.

Table 6
Point Measure Correlation (PMEA CORR) Values

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
12	209	50	2.04	.40	.98	.0	1.00	.2	.39	.39	86.0	83.8	D12	0
1	212	50	1.60	.37	1.06	.4	1.14	.5	.35	.41	76.0	78.9	D1	0
32	213	50	1.47	.36	1.11	.7	1.43	1.3	.29	.41	78.0	77.6	D32	0
27	180	50	1.35	.25	1.35	1.8	1.21	1.0	.36	.53	56.0	60.0	D27	0
19	189	50	1.20	.30	.73	-1.5	.75	-1.3	.68	.47	82.0	70.9	D19	0
23	191	50	1.09	.31	.66	-1.8	.66	-1.7	.73	.46	84.0	73.2	D23	0
45	189	50	1.08	.28	.91	-.4	.93	-.3	.55	.49	76.0	68.3	D45	0
20	186	50	.97	.24	1.43	2.2	1.38	1.8	.31	.54	56.0	56.4	D20	0
25	193	50	.86	.30	.59	-2.2	.57	-2.2	.80	.47	86.0	72.7	D25	0
36	194	50	.79	.31	1.04	.3	1.04	.3	.44	.46	68.0	73.7	D36	0
41	195	50	.72	.32	1.15	.7	1.19	.8	.34	.45	70.0	75.0	D41	0
18	181	50	.47	.22	.52	-2.4	.65	-1.4	.75	.57	74.0	68.2	D18	0
40	197	50	.47	.30	1.29	1.3	1.27	1.2	.25	.48	70.0	71.9	D40	0
43	196	50	.45	.25	.84	-.9	.88	-.6	.62	.53	68.0	63.0	D43	0
42	197	50	.37	.25	.95	-.2	1.01	.1	.55	.53	66.0	61.6	D42	0
17	198	50	.35	.27	1.18	.9	1.18	.9	.37	.50	64.0	67.5	D17	0
31	198	50	.33	.26	1.10	.6	1.08	.5	.45	.52	64.0	65.2	D31	0
37	202	50	.05	.27	.89	-.5	.85	-.7	.59	.50	74.0	67.0	D37	0
6	203	50	.00	.26	1.39	1.9	1.37	1.8	.26	.52	52.0	63.0	D6	0
34	184	50	-.01	.20	1.10	.6	1.15	.8	.52	.61	50.0	51.5	D34	0
22	204	50	-.06	.25	1.19	1.0	1.19	1.0	.40	.52	58.0	61.8	D22	0
44	204	50	-.06	.25	.56	-2.8	.56	-2.8	.82	.52	72.0	61.8	D44	0
29	189	50	-.14	.21	1.11	.6	1.12	.6	.48	.60	48.0	58.3	D29	0
5	204	50	-.17	.31	1.39	1.6	1.33	1.3	.13	.47	74.0	74.0	D5	0
3	204	50	-.22	.33	1.08	.4	1.17	.7	.35	.45	80.0	77.6	D3	0
30	204	50	-.22	.33	1.38	1.5	1.44	1.5	.11	.45	72.0	77.6	D30	0
33	196	50	-.29	.20	.75	-1.0	.87	-.5	.61	.61	58.0	59.7	D33	0
39	205	50	-.30	.32	1.09	.5	1.07	.3	.37	.45	74.0	75.9	D39	0
8	203	50	-.30	.42	.92	-.1	.98	.1	.38	.38	88.0	86.5	D8	0
38	207	50	-.44	.30	.80	-.9	.73	-1.2	.63	.47	76.0	72.7	D38	0
4	206	50	-.46	.33	.93	-.2	.94	-.1	.49	.44	80.0	77.9	D4	0
15	208	50	-.50	.30	1.05	.3	1.05	.3	.43	.47	68.0	71.2	D15	0
35	208	50	-.60	.31	.85	-.7	.75	-1.1	.57	.45	78.0	74.6	D35	0
16	209	50	-.66	.31	1.05	.3	1.06	.3	.42	.46	74.0	73.1	D16	0
28	201	50	-.72	.22	1.24	1.1	1.37	1.6	.45	.57	70.0	57.9	D28	0
13	199	50	-.78	.23	.76	-1.2	.87	-.6	.70	.56	82.0	59.5	D13	0
7	205	50	-.78	.43	1.00	.1	1.18	.6	.29	.37	90.0	86.9	D7	0
2	201	50	-.93	.28	1.25	1.0	1.03	.2	.32	.49	74.0	72.6	D2	0
26	211	50	-.95	.32	.73	-1.4	.71	-1.3	.65	.45	82.0	73.7	D26	0
9	208	50	-1.05	.37	1.08	.4	1.21	.7	.26	.41	84.0	81.9	D9	0
24	202	50	-1.09	.23	.60	-2.3	.62	-2.1	.79	.54	80.0	60.2	D24	0
10	203	50	-1.11	.24	.54	-2.7	.54	-2.6	.83	.54	82.0	61.1	D10	0
14	210	50	-1.19	.35	.70	-1.4	.59	-1.6	.65	.42	82.0	78.6	D14	0
11	211	50	-1.28	.29	.86	-.4	1.02	.2	.58	.45	84.0	75.7	D11	0
21	212	50	-1.32	.29	.97	.0	1.18	.7	.51	.46	74.0	74.5	D21	0
MEAN	200.5	50.0	.00	.29	.98	-.1	1.01	.0			73.0	70.1		
S.D.	8.5	.0	.84	.05	.25	1.2	.25	1.2			10.3	8.4		

Table 7
Item Fit Based on MNSQ Values

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
30	204	50	-.22	.33	1.38	1.5	1.44	1.5	A .11	.45	72.0	77.6	D30	0
20	186	50	.97	.24	1.43	2.2	1.38	1.8	B .31	.54	56.0	56.4	D20	0
32	213	50	1.47	.36	1.11	.7	1.43	1.3	C .29	.41	78.0	77.6	D32	0
5	204	50	-.17	.31	1.39	1.6	1.33	1.3	D .13	.47	74.0	74.0	D5	0
6	203	50	.00	.26	1.39	1.9	1.37	1.8	E .26	.52	52.0	63.0	D6	0
28	201	50	-.72	.22	1.24	1.1	1.37	1.6	F .45	.57	70.0	57.9	D28	0
27	180	50	1.35	.25	1.35	1.8	1.21	1.0	G .36	.53	56.0	60.0	D27	0
40	197	50	.47	.30	1.29	1.3	1.27	1.2	H .25	.48	70.0	71.9	D40	0
2	201	50	-.93	.28	1.25	1.0	1.03	.2	I .32	.49	74.0	72.6	D2	0
9	208	50	-1.05	.37	1.08	.4	1.21	.7	J .26	.41	84.0	81.9	D9	0
22	204	50	-.06	.25	1.19	1.0	1.19	1.0	K .40	.52	58.0	61.8	D22	0
41	195	50	.72	.32	1.15	.7	1.19	.8	L .34	.45	70.0	75.0	D41	0
7	205	50	-.78	.43	1.00	.1	1.18	.6	M .29	.37	90.0	86.9	D7	0
17	198	50	.35	.27	1.18	.9	1.18	.9	N .37	.50	64.0	67.5	D17	0
21	212	50	-1.32	.29	.97	.0	1.18	.7	O .51	.46	74.0	74.5	D21	0
3	204	50	-.22	.33	1.08	.4	1.17	.7	P .35	.45	80.0	77.6	D3	0
34	184	50	-.01	.20	1.10	.6	1.15	.8	Q .52	.61	50.0	51.5	D34	0
1	212	50	1.60	.37	1.06	.4	1.14	.5	R .35	.41	76.0	78.9	D1	0
29	189	50	-.14	.21	1.11	.6	1.12	.6	S .48	.60	48.0	58.3	D29	0
31	198	50	.33	.26	1.10	.6	1.08	.5	T .45	.52	64.0	65.2	D31	0
39	205	50	-.30	.32	1.09	.5	1.07	.3	U .37	.45	74.0	75.9	D39	0
16	209	50	-.66	.31	1.05	.3	1.06	.3	V .42	.46	74.0	73.1	D16	0
15	208	50	-.50	.30	1.05	.3	1.05	.3	W .43	.47	68.0	71.2	D15	0
36	194	50	.79	.31	1.04	.3	1.04	.3	v .44	.46	68.0	73.7	D36	0
11	211	50	-1.28	.29	.86	-.4	1.02	.2	u .58	.45	84.0	75.7	D11	0
42	197	50	.37	.25	.95	-.2	1.01	.1	t .55	.53	66.0	61.6	D42	0
12	209	50	2.04	.40	.98	.0	1.00	.2	s .39	.39	86.0	83.8	D12	0
8	203	50	-.30	.42	.92	-.1	.98	.1	r .38	.38	88.0	86.5	D8	0
4	206	50	-.46	.33	.93	-.2	.94	-.1	q .49	.44	80.0	77.9	D4	0
45	189	50	1.08	.28	.91	-.4	.93	-.3	p .55	.49	76.0	68.3	D45	0
37	202	50	.05	.27	.89	-.5	.85	-.7	o .59	.50	74.0	67.0	D37	0
43	196	50	.45	.25	.84	-.9	.88	-.6	n .62	.53	68.0	63.0	D43	0
13	199	50	-.78	.23	.76	-1.2	.87	-.6	m .70	.56	82.0	59.5	D13	0
33	196	50	-.29	.20	.75	-1.0	.87	-.5	l .61	.61	58.0	59.7	D33	0
35	208	50	-.60	.31	.85	-.7	.75	-1.1	k .57	.45	78.0	74.6	D35	0
38	207	50	-.44	.30	.80	-.9	.73	-1.2	j .63	.47	76.0	72.7	D38	0
19	189	50	1.20	.30	.73	-1.5	.75	-1.3	i .68	.47	82.0	70.9	D19	0
26	211	50	-.95	.32	.73	-1.4	.71	-1.3	h .65	.45	82.0	73.7	D26	0
14	210	50	-1.19	.35	.70	-1.4	.59	-1.6	g .65	.42	82.0	78.6	D14	0
23	191	50	1.09	.31	.66	-1.8	.66	-1.7	f .73	.46	84.0	73.2	D23	0
18	181	50	.47	.22	.52	-2.4	.65	-1.4	e .75	.57	74.0	68.2	D18	0
24	202	50	-1.09	.23	.60	-2.3	.62	-2.1	d .79	.54	80.0	60.2	D24	0
25	193	50	.86	.30	.59	-2.2	.57	-2.2	c .80	.47	86.0	72.7	D25	0
44	204	50	-.06	.25	.56	-2.8	.56	-2.8	b .82	.52	72.0	61.8	D44	0
10	203	50	-1.11	.24	.54	-2.7	.54	-2.6	a .83	.54	82.0	61.1	D10	0
MEAN	200.5	50.0	.00	.29	.98	-.1	1.01	.0			73.0	70.1		
S.D.	8.5	.0	.84	.05	.25	1.2	.25	1.2			10.3	8.4		

Unidimensionality

The Principal Component Analysis of Residuals (PCA) is used in Rasch analysis to ensure the consistency of the dimensions of the instrument (Aziz et al., 2015). This study referred to two criteria for testing the unidimensionality of the developed instrument, namely, the value

of the PCA and the level of item distortion or unexplained variance found in the first contrast (Aziz et al., 2015). According to Sumintono and Widhiarso (2013), a good PCA value is at least 20% and more than 40%, and the unexplained variance in the first contrast is 15% maximum (Aziz et al., 2015).

Table 8 below presents the findings of the PCA based on variances that can be explained by measure for content knowledge. The PCA value of 40% for content knowledge was accepted, as it exceeded the 20% minimum. The value of unexplained variance found in the first contrast (size) that can be in the desired specification was the content knowledge at 10.6%.

Table 8
Findings of the Principal Component Analysis (PCA)

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --		Modeled
Total raw variance in observations	=	75.0	100.0%	100.0%
Raw variance explained by measures	=	30.0	40.0%	37.1%
Raw variance explained by persons	=	13.0	17.4%	16.1%
Raw Variance explained by items	=	17.0	22.6%	21.0%
Raw unexplained variance (total)	=	45.0	60.0% 100.0%	62.9%
Unexplned variance in 1st contrast	=	8.0	10.6% 17.7%	
Unexplned variance in 2nd contrast	=	5.1	6.8% 11.3%	
Unexplned variance in 3rd contrast	=	4.4	5.8% 9.7%	
Unexplned variance in 4th contrast	=	3.6	4.8% 8.0%	
Unexplned variance in 5th contrast	=	2.5	3.4% 5.6%	

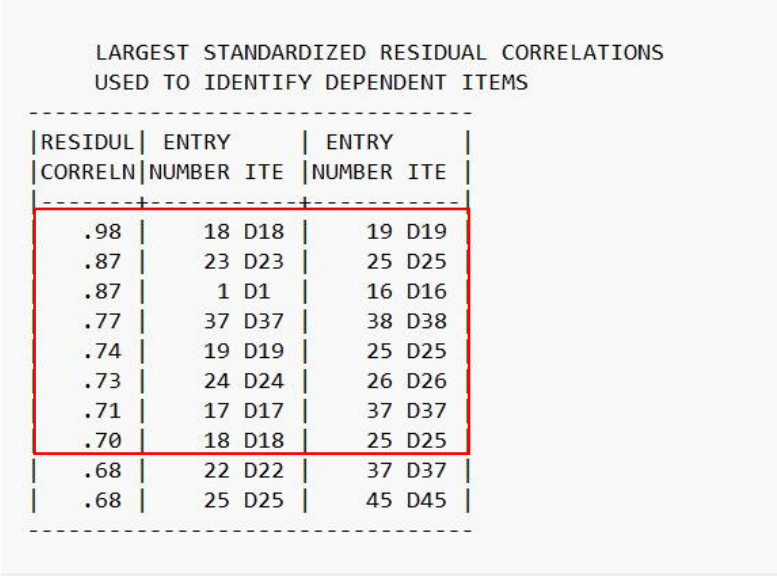
In order to identify whether there were confusing or overlapping items, the Standardised Residual Correlation test was conducted to ensure that the instrument was free from any confusion and missed objectives. If there are two items that equals or exceeds 0.70, this shows a high correlation value, given the same characteristics between them and combining a few other dimensions that are shared together.

Therefore, only one of the two items would be required for measurement. Based on the following Table 9, 8 items have a large correlation value, or value of higher than .70. The following items have shown correlation values of .70 (D18 and D25), .71 (D17 and D37), .73 (D24 and D26), .74 (D19 and D25), .77 (D37 and D38), .87 (D1 and D16), 0.87 (D23 and D25), and .98 (D18 and D19). The following Table 10 lists the items with a large correlation value or exceeds the value of .70.

These results showed that the respondents viewed the coupled items as the same, which was confusing to them. Therefore, two approaches can be taken to address this issue, which is either to maintain the items or to refine them so that the purpose of the question becomes clearer, or the items are dropped based on the distillation method in Item Fit (Aziz et al., 2015).

Upon scrutiny, items with MNSQ values that were lower than 1.5 have been refined and maintained. According to discussions with the experts, these items were deemed as relevant under the construct of creative teaching.

Table 9
Standardised Residual Correlation



LARGEST STANDARDIZED RESIDUAL CORRELATIONS
USED TO IDENTIFY DEPENDENT ITEMS

RESIDUL CORRELN	ENTRY NUMBER ITE	ENTRY NUMBER ITE
.98	18 D18	19 D19
.87	23 D23	25 D25
.87	1 D1	16 D16
.77	37 D37	38 D38
.74	19 D19	25 D25
.73	24 D24	26 D26
.71	17 D17	37 D37
.70	18 D18	25 D25
.68	22 D22	37 D37
.68	25 D25	45 D45

Table 10
List of items that have a large correlation value or exceeds the value of .70.

Item	Statement	Item	Statement
D18	I consider the questions from the students, even though the questions are not very appropriate.	D19	I take follow-up action on questions that are asked by students so that they know that I take their questions seriously.
D23	I will seek students' further opinion before making any decision about the creative ideas explained by the students.	D25	I consider all of the students' ideas, and then, I will choose ideas that are really suitable for the students when performing language art activities.
D1	I encourage every student to show me the language art aspect that they have learnt previously.	D16	I listen carefully when students ask questions.
D37	I give students the chance to evaluate their own work.	D38	I give students the chance to evaluate the work of other students.
D19	I take follow-up action on questions that are asked by students so that they know that I take their questions seriously.	D25	I consider all of the students' ideas, and then, I will choose ideas that are really suitable for the students when performing language art activities.
D24	I give comments on the students' ideas after their creative ideas are thoroughly explored.	D26	I carry out language art activities according to the suitability of students' level of thinking.
D17	I listen patiently, even when students ask inappropriate questions.	D37	I give students the chance to evaluate their own work.
D18	I consider the questions from the students, even though the	D25	I consider all of the students' ideas, and then, I will choose ideas that are

	questions are not very appropriate.		really suitable for the students when performing language art activities.

Discussion

Preschool Teacher's Creative Teaching Instrument for Children's Language Arts was developed for the purpose of measuring preschool teachers' creative teaching for language arts. Apart from its importance in determining teachers' preparedness, this instrument was also a necessity to ensure the success of the aspirations of the national education system.

This study has developed nine constructs as the core of this instrument: to encourage students to learn freely; to collaborate with other students in their learning; to motivate students; to hold back judgement towards a creative idea; to encourage flexible thinking; to encourage students to conduct self-assessment towards ideas produced; to take into account all suggestions and questions put forward by the students; to offer students the chance to do assignments; and to help them in handling disappointment when they fail upon trying something new.

The Rasch model was used as the basis for measuring five assumptions that needed to be met to ensure construct validity. These assumptions were reliability, item fit, item polarity, unidimensionality, and separation index. The Rasch analysis results demonstrates that the developed instrument has good psychometric qualities. The following Table 11 shows the results of the analysed instrument.

Table 11
Analysis summary

Criteria Statistical	Result	Quality	References
Person Reliability Value	0.92	Excellent	Bond & Fox (2015) Sumintono & Widhiarso (2015)
Item Reliability Value	0.87	Good	Bond & Fox (2015) Sumintono & Widhiarso (2015)
Item Fit Total infit and outfit mean-square (Item Misfit)	0.5–1.5 logits	Good	Sumintono & Widhiarso (2015)
Polarity item	positive (+)	The positive values indicated that the items were functioning towards the same direction and in line with the measured construct.	Linacre (2002)
Unidimensionality – Value of the Principal	40.0	Exceeded min of 40%	Sumintono & Widhiarso (2015) Aziz et al. (2015)

Component Analysis of Residuals (PCA)			
Unidimensionality – Unexplained variance found in the first contrast	10.6	< 15%	Aziz et al. (2015)
Standardised Residual Correlation	< .70	There were eight pairs of items that have a large correlation value or exceeded the value of .70. These items have been refined and maintained, because according to discussions with experts, these items were relevant under the construct of creative teaching.	Aziz et al. (2015)

Based on this pilot study, it can be concluded that validity and reliability are important aspects to consider when developing a new instrument for a study. The validity and reliability tests that were conducted in this study have indicated that this instrument was fit to be used by companies or other researchers in future studies. The implications of this study can help other researchers in developing a good instrument for creative teaching by preschool teachers.

Conclusion

This study offers several implications from a practical and methodological viewpoint. The clearest practical implication is the adaptation of a creative teaching instrument that combines the original items from the CFTI instrument with language art items. The incompatibility of CFTI and language arts has led to the amendments and addition of items to suit the context of this study. The implication is that this study has contributed an instrument to measure teachers' creative teaching, which is suitable to be assessed from the perspective of children's language art.

From a methodological viewpoint, the Rasch measurement model was useful for checking the validity and reliability of this creative teaching instrument. The Rasch model is capable of measuring the reliability of respondents and items in more depth and is a stronger measurement compared to only reviewing Cronbach's alpha. The Winsteps software, version 3.68.2 was used to check the functionality of items comprising reliability, separation of item-respondent, polarity, and suitability of items to measure the construct, as well as the standardised residual correlation value. The Rasch model also allowed for the refining of items that did not meet the inspection range. Therefore, the methodological implication of a strict analysis of validity and reliability using the Rasch measurement model offers a more comprehensive strategy for analysis, which can produce a reliable instrument.

Overall, this study has succeeded in developing an instrument that can measure teachers' creative teaching for language arts by following rigorous steps during the

development and validation processes. The findings of this instrument are expected to add more knowledge in the language arts field, as well as in teachers' creative teaching, especially in Malaysia.

Acknowledgement
Dr Soh Kay Cheng, Research Consultant, Singapore Centre for Chinese Language Limited, a Centre of Nanyang Technological University

Corresponding Author
Nor Mashitah Binti Mohd Radzi, Department of Early Childhood Education, Faculty of Human Development, Sultan Idris Education University.

References

Ab. Jawas, S., & Zulkifli, H. (2022). Amalan pengajaran kreatif guru Pendidikan Islam bagi murid berkeperluan khas masalah pembelajaran (Creative teaching practices by Islamic Education teachers for students with learning disabilities). *International Journal of Islamic and Humanities Research*, *2*(1), 1–9. https://nunjournal.com/index.php/qalam/article/view/50/49

Abdullah, M., Nor, M. M., & Hutagalung, F. D. (2021). Pendekatan pengajaran bermain di bilik darjah dalam kalangan guru prasekolah (Play-based teaching approach in the classroom among preschool teachers). *Jurnal Penyelidikan Pendidikan*, *39*(1), 64–74. https://ejournal.um.edu.my/index.php/JER/article/view/30732

Ali, A., & Mahamod, Z. (2015). Development of play-based instruction module for teaching preschoolers' language skills. *Australian Journal of Basic and Applied Sciences*, *9*(34), 110–118.

Aziz, A. A., Masodi, M. S., & Zaharim, A. (2015). *Asas model pengukuran Rasch: Pembentukan skala and struktur pengukuran*. Bangi: Penerbit UKM.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Cooper, D. R., & Schindler, P. S. (2011). *Business research methods* (11th ed.). Boston: McGraw-Hill.

Fox, C. M., & Jones, J. A. (1998). Uses of Rasch modelling in counselling psychology research. *Journal of Counselling Psychology*, *45*(1), 30–45. http://dx.doi.org/10.1037/0022-0167.45.1.30

Huei, O. K., Che Rus, R., & Kamis, A. (2020). Construct validity and reliability in content knowledge of design and technology subject: A Rasch measurement model approaches for pilot study. *International Journal of Academic Research in Business and Social Sciences*, *10*(3). 10.6007/IJARBSS/v10-i3/7066.

Kementerian Pendidikan Malaysia. (2017). Buku Penerangan Kurikulum Standard Prasekolah Kebangsaan [KSPK, Semakan 2017] (National Preschool Curriculum Standard Handbook [KSPK]). Putrajaya, Malaysia: Kementerian Pendidikan Malaysia.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean. *Rasch Measurement Transactions*, *16*(2), 878.

Linacre J. M. (1994). Sample size and item calibration stability. Rasch Measure. Transact. 7, (4), 328. Google Scholar

Linacre, J. M. (2005). Test validity and Rasch measurement: Construct, content, etc. *Rasch Measurement Transactions*, *18*, 970–971.

Linacre, J. M. (2006). A user's guide to WINSTEPS® MINISTEP Rasch-Model computer programs - Program Manual 3.68.0. Retrieved from www.winsteps.com.

Mokhlis, S. (2019). Pemupukan kreativiti kanak-kanak: Kajian kes amalan pengajaran kreativiti di sebuah tadika Islam (Nurturing creativity in young children: A case study of the practice of teaching for creativity in an Islamic kindergarten). *Attarbawiy: Malaysian Online Journal of Education*, *3*(1), 34–48. https://doi.org/10.53840/attarbawiy.v3i1.65

Soh, K. (2015). Creativity fostering teacher behaviour around the world: Annotations of studies using the CFTIndex. *Cogent Education*, *2*(1), 1034494. doi: 10.1080/2331186X.2015.1034494

Sumintono, B., & Widhiarso, W. (2013). *Aplikasi Model Rasch untuk penelitian ilmu-ilmu sosial* (1st ed.). Cimahi: TrimKom Publishing House.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan RASCH pada assessment pendidikan* (1st ed.). Cimahi: Trim Komunikata.