Journal of Contemporary Educational Research, 2022, Volume 6, Issue 3
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

An Assessment Scale for Evaluating the Experimental Design Ability of Elementary Science Teachers Based on Primary Trait Analysis

Saiqi Tian[1]*, Jinxing Zhu[1], Chuyin Chen[1,2], Zhiqing Wang[1]

[1]College of Education, Wenzhou University, Wenzhou 325035, Zhejiang Province, China

[2]Tiantai Primary School, Taizhou 317299, Zhejiang Province, China

******Corresponding author:*** Saiqi Tian, tiansaiqi@wzu.edu.cn

Abstract: In this study, an assessment scale for evaluating the experimental design ability of elementary science teachers was constructed based on primary trait analysis. This assessment scale contains three first-level indexes and eleven second-level indexes. The corresponding weights of indexes were determined by the objective weighting method. The scores of all the descriptions of the indexes were also assigned. After a trial test, this assessment scale was verified to be reliable and valid for evaluating the experimental design ability of elementary science teachers.

Keywords: Assessment scale; Experimental design ability; Elementary science teachers; Primary trait analysis

Online publication: March 15, 2022

1. Introduction

Science education plays a critical role in preparing a sufficient supply of competent workforce and citizens for the twenty-first century, which is featured by a globalized science and technology-driven society [1-3]. As a core subject in elementary school, science is an integrated discipline based on experimental activities, with the purpose of promoting science literacy among elementary school students [4]. It has been emphasized that experimental teaching must be much accounted for, and students should be encouraged to actively involve themselves in laboratory courses [5,6]. Experimental design is one of the most important processes in an inquiry-based experiment [7]. Prior to class, a reasonable and rigorous experimental design scheme can greatly improve the teaching efficiency.

The experimental design ability of elementary science teachers directly influences the improvement of students' science literacy [8]. It is therefore of practical significance to evaluate the experimental design ability of elementary science teachers. The assessment of elementary science teachers' experimental design ability is a strong basis for making policies for improving the skills and abilities of teachers. This has led to much research on an assessment scale that can objectively evaluate the experimental design ability of elementary science teachers.

Among the tools for educational assessment, primary trait analysis (PTA) [9,10], which was first introduced by Walvoord and other academicians in the 1990s, has been regarded as an objective, explicit, and criterion-based assessment tool. It has been adopted by many schools, colleges, and universities with diverse disciplines [11,12]. A common scale is employed to assess one specific course assignment but not to evaluate a primary trait ability directly. In contrast, the PTA scale can measure discipline or programspecific abilities in a curriculum once the evaluation indexes have been identified.

In light of these issues, an assessment scale was designed in this study to evaluate the experimental design ability of elementary science teachers based on PTA. Above all, the first-level and second-level indexes were identified. The weights of the indexes were determined by the objective weighting method. Subsequently, the scores of all the descriptions of the indexes were assigned. Finally, the reliability and validity of the assessment scale were verified through a trial test. The research scheme and structure of this paper are shown in **Figure 1**.

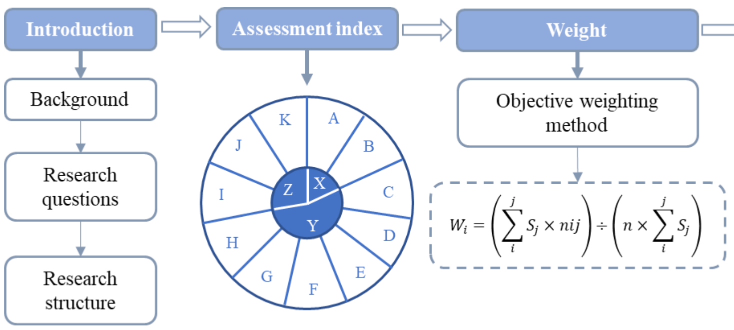

Figure 1. Research scheme and structure of the paper

2. Designing the assessment scale

2.1. Identification of first-level indexes

Experimental design ability refers to the ability of one to design experiments using scientific knowledge and principles based on experimental problems, assumptions, and objectives. Several studies have defined experimental design ability. Liu and other researchers divided it into five dimensions: determining qualitative and quantitative methods, distinguishing variables, designing experimental procedures logically, selecting appropriate materials and tools to collect data, as well as using tables and figures to present results [13]. In another research, experimental design ability was evaluated based on whether the experiment scheme can test a hypothesis, whether the method is reasonable, and whether the description of the design is clear [14]. In a different study, experimental design ability was defined as having principles, plan, and reflection [15]. In recent research, the constituent elements of experimental design ability were identified as scientificity, feasibility, and quality of experimental design description [16]. Taking into consideration of all the definitions and analyzing the characteristics of elementary science education, first-level indexes have been identified and coded as the ability of making assumptions (X), the ability of planning experiments (Y), and the ability of explanations, reflections, and improvements (Z).

2.2. Identification and description of first-level indexes

2.2.1. Ability of making assumptions

Rather than subjectively conjecturing, making assumptions are raising solutions based on the perception of problems. According to the 2017 *Primary School Science Curriculum Standards for Compulsory Education* [17], making assumptions are raising targeted hypotheses from the structure, function, change, and interrelation of things, on the strength of knowledge acquired, in addition to explaining the reasons for these assumptions. In the 2011 *Biology Curriculum Standards for Compulsory Education* [18], making assumptions is defined as constructing new knowledge on the basis of previous knowledge, raising hypothetical answers as much as possible, thereby forecasting the inspectability of assumptions.

To sum up, making assumptions should have the following characteristics:

(1) the assumptions are well-founded, instead of guesswork or subjective conjecture;

(2) although well-founded, hypothesis and speculation are also involved;
(3) the assumptions are verifiable;
(4) the assumptions are diverse.

In the light of the aforementioned characteristics, the ability of making assumptions has been identified and coded as diversity of assumptions (A) and rationality of assumptions (B) in this study. The corresponding second-level indexes and description of making assumptions are shown in **Table 1**.

Table 1. Second-level indexes and description of making assumptions

Second-level index	Description
Diversity of assumptions (A)	Two or more assumptions (A_1) Only one assumption (A_2) Unable to make assumptions (A_3)
Rationality of assumptions (B)	Assumptions are well-founded (B_1) Assumptions are unfounded (B_2)

2.2.2. Ability of planning experiments

Planning experiments refer to developing feasible proposals according to problems and assumptions, combined with the provided materials and equipment. Experimental proposals include experimental objectives, experimental principles, methods, materials, equipment, procedures, and possible results. Owing to the contents of planning experiments, the ability of planning experiments has been identified and coded as clarity of experimental objectives (C), scientificity of experimental principles (D), scientificity of methods (E), rationality of materials and equipment (F), feasibility of procedures (G), and prediction of results (H). The corresponding second-level indexes and description of planning experiments are presented in **Table 2**.

Table 2. Second-level indexes and description of planning experiments

Second-level index	Description
Clarity of experimental objectives (C)	The experimental objectives can be well achieved (C_1) The experimental objectives can be achieved (C_2) The experimental objectives cannot be achieved (C_3)
Scientificity of experimental principles (D)	The experimental principles are scientific (D_1) The experimental principles are non-scientific (D_2)
Scientificity of methods (E)	The methods are reasonable and the concept of controlling variable is involved (E_1) The methods are reasonable but the concept of controlling variable is not involved (E_2) The methods are unreasonable (E_3)
Rationality of materials and equipment (F)	The utilized materials and equipment are reasonable, safe, and easy to use (F_1) The utilized materials and equipment are reasonable and safe but too complicated to use (F_2) The utilized materials and equipment are unreasonable (F_3)
Feasibility of procedures (G)	The procedures are practical and well-stated (G_1) The procedures are practical but poorly stated (G_2) The procedures are complete but not easy to operate (G_3) The procedures are incomplete (G_4)

(Continued on next page)

(Continued from previous page)

Second-level index	Description
Prediction of results (H)	More than one results predicted (H_1) Only one result predicted (H_2) No results predicted (H_3)

2.2.3. Ability of explanations, reflections, and improvements

A complete experimental design should contain explanations of predicted results, reflections, and improvements. The ability of explanations, reflections, and improvements has been identified and coded as rationality of explanations (I), reflection of experimental design (J), and improvement of experimental design (K). The corresponding second-level indexes and description of explanations, reflections, and improvements are presented in **Table 3**.

Table 3. Second-level indexes and description of explanations, reflections, and improvements

Second-level index	Description
Rationality of explanations (I)	The explanations of experimental results are reasonable (I_1) The explanations of experimental results are unreasonable (I_2)
Reflection of experimental design (J)	The shortcomings of experimental design and various interference factors can be reflected (J_1) The shortcomings of experimental design and only one interference factor can be reflected (J_2) The shortcomings of experimental design can be reflected, but the interference factors cannot be reflected (J_3) Neither the shortcomings of experimental design nor the interference factors can be reflected (J_4)
Improvement of experimental design (K)	The experimental design can be improved with creativity (K_1) The experimental design can be improved but without creativity (K_2) The experimental design can be improved but without reasonability (K_3) The experimental design cannot be improved (K_4)

2.3. Determining the weights of indexes

The value of the weight reflects the degree of importance of an index. In order to avoid subjective difference, the objective weighting method was employed to determine the weight of each index. Thirty elementary science teachers were surveyed to score the importance of each index, utilizing the following scale: 3 represents "very important"; 2 represents "important"; 1 represents normal. The weight of each index was calculated using the following equation:

$$W_i = \left(\sum_i^j S_j \times nij\right) \div \left(n \times \sum_i^j S_j\right) \quad (1)$$

S refers to the scores given by the elementary science teachers, and *nij* refers to the number of teachers who choose the corresponding degree of importance of each index. For first-level indexes, there are three indexes, so $\sum S_j = 3 + 2 + 1 = 6$. The computational process is as follows:

$$W_X = (12 \times 3 + 16 \times 2 + 2 \times 1) \div (30 \times 6) = 0.39 \quad (2)$$
$$W_Y = (17 \times 3 + 13 \times 2 + 0 \times 1) \div (30 \times 6) = 0.43 \quad (3)$$
$$W_Z = (1 \times 3 + 1 \times 2 + 28 \times 1) \div (30 \times 6) = 0.18 \quad (4)$$

The results of the degree of importance of first-level indexes were summarized, and the weight for each index was calculated, as shown in **Table 4**. W_Y is 0.43, demonstrating that much importance is attached to the ability of planning experiments. W_X is 0.39, indicating that the ability of making assumptions is at the second place. W_Z is 0.18, which means that the ability of explanations, reflections, and improvements is of less concern to elementary science teachers.

Table 4. Results of the degree of importance of first-level indexes and their calculated weights

First level indexes	Numbers of elementary science teachers			Weight (W_n)
	Very important	Important	Normal	
X	12	16	2	0.39
Y	17	13	0	0.43
Z	1	1	28	0.18

The results of the degree of importance of second-level indexes as well as their weights are shown in **Table 5**. The weights were calculated using equation (1). The final weight for each second-level index was obtained by multiplying the calculated number with the weight of the corresponding first-level index.

Table 5. Results of the degree of importance of second-level indexes and their calculated weights

First-level indexes	Second-level indexes	Numbers of elementary science teachers			Weight (W_n)
		Very important	Important	Normal	
X	A	5	22	3	0.16
	B	27	3	0	0.23
Y	C	5	18	7	0.06
	D	14	12	4	0.06
	E	28	2	0	0.08
	F	24	6	0	0.08
	G	27	3	0	0.08
	H	17	8	5	0.07
Z	I	17	8	5	0.05
	J	25	3	2	0.06
	K	27	3	0	0.06

2.4. Assigning scores

The assessment scale for evaluating the experimental design ability of elementary science teachers was acquired after the scores of descriptions were assigned (**Table 6**). The bracketed number indicates the weight of the index. The score of the ability of making assumptions (S_X) is calculated as follows:

$$S_X = S_1 \times W_A + S_2 \times W_B \quad (5)$$

The score of the ability of planning experiments (S_Y) is calculated as follows:

$$S_Y = S_3 \times W_C + S_4 \times W_D + S_5 \times W_E + S_6 \times W_F + S_7 \times W_G + S_8 \times W_H \quad (6)$$

The score of the ability of explanations, reflections, and improvements (S_Z) is calculated as follows:

$$S_Z = S_9 \times W_I + S_{10} \times W_J + S_{11} \times W_K \quad (7)$$

The score of the experimental design ability (S) is calculated as follows:

$$S = S_X + S_Y + S_Z \quad (8)$$

The total score that can be obtained is 100. A score ranging from 85 to 100 reflects excellent experimental design ability; 70 to 84 reflects good experimental design ability; 55 to 69 reflects normal experimental design ability; 45 to 54 reflects low experimental design ability; a score below 44 reflects poor experimental design ability.

Table 6. Assessment scale to evaluate the experimental design ability of elementary science teachers

First-level index	Second-level index	Description	Score	Scoring formula
X (0.39)	A (0.16)	A_1	100	$S_1 = A_n$
		A_2	60	
		A_3	0	
	B (0.23)	B_1	100	$S_2 = B_n$
		B_2	0	
Y (0.43)	C (0.60)	C_1	100	$S_3 = C_n$
		C_2	60	
		C_3	0	
	D (0.06)	D_1	100	$S_4 = D_n$
		D_2	0	
	E (0.08)	E_1	100	$S_5 = E_n$
		E_2	60	
		E_3	0	
	F (0.08)	F_1	100	$S_6 = F_n$
		F_2	60	
		F_3	0	
	G (0.08)	G_1	100	$S_7 = G_n$
		G_2	80	
		G_3	60	
		G_4	0	
	H (0.07)	H_1	100	$S_8 = H_n$
		H_2	60	
		H_3	0	

(Continued on next page)

(Continued from previous page)

First-level index	Second-level index	Description	Score	Scoring formula
Z (0.18)	I (0.05)	I_1	100	$S_9 = I_n$
		I_2	0	
	J (0.06)	J_1	100	$S_{10} = J_n$
		J_2	80	
		J_3	60	
		J_4	0	
	K (0.06)	K_1	100	$S_{11} = K_n$
		K_2	80	
		K_3	60	
		K_4	0	

2.5. Reliability and validity of the assessment scale

To test the reliability and validity of the assessment scale, 112 participants were included in the trial test. Among them, 13.3% were male and 86.7% were female. All the participants were informed that their personal information will be anonymous and kept confidential for research purposes only. Besides, they had right to refuse to participate in this trial test. The test was carried out using questionnaires, in which each questionnaire contains three experimental activities, related to all the indexes. The data were analyzed by using the Statistical Package for Social Sciences (SPSS) software. Cronbach's alpha was 0.797, thus demonstrating the reliability of the assessment scale [19]. Kaiser-Meyer-Olkin (KMO) and Bartlett's tests were performed to assess the validity of the assessment scale. The KMO value was found to be mediocre at 0.689.

3. Conclusion and limitations

In summary, an assessment scale for evaluating the experimental design ability of elementary science teachers was designed based on primary trait analysis. The first-level indexes were identified as ability of making assumptions, ability of planning experiments, and ability of explanations, reflections, and improvements. Eleven second-level indexes and their corresponding descriptions were also identified. The weights of indexes were determined, and the score for each second-level index description was assigned. This assessment scale was verified to be reliable and valid based on the results of the trial test. Overall, the assessment scale can be applied to evaluate the experimental design ability of elementary science teachers.

There are several limitations in this study; however, these limitations should be regarded as opportunities to design and develop future research. The participants enrolled in the trial test were only from the small scale of elementary schools based on convenience sampling, and most of them are female teachers. Future studies can use sampling to limit the sampling errors. The researchers' inadequacy in the comprehension of the analysis is also a limitation. Despite this, it is believed that more significant conclusions will be put forward with the continuing efforts of researchers in the future.

Funding

Introduce Talent Foundation of Wenzhou University (135010120719) and College Students Innovative Entrepreneurial Training of Wenzhou University (JWSC2019023).

Disclosure statement

The authors declare no conflict of interest.

References

[1] Rudolph JL, 2020, The Lost Moral Purpose of Science Education. Science Education, 104: 895-906.

[2] Yilmaz RM, 2020, Effects of Using Cueing in Instructional Animations on Learning and Cognitive Load Level of Elementary Students in Science Education. Interactive Learning Environments, 2020: 1-15.

[3] Kaya E, 2018, Argumentation in Elementary Science Education: Addressing Methodological Issues and Conceptual Understanding. Cultural Studies of Science Education, 13: 1087-1090.

[4] Ha S, Kim M, 2020, Challenges of Designing and Carrying Out Laboratory Experiments About Newton's Second Law. Science & Education, 29: 1389-1416.

[5] Burron G, Pegg J, 2021, Elementary Pre-Service Teachers' Search, Evaluation, and Selection of Online Science Education Resources. Journal of Science Education and Technology, 30: 471-483.

[6] Cruz-Guzman M, Garcia-Carmona A, Criado AM, 2020, Proposing Questions for Scientific Inquiry and the Selection of Science Content in Initial Elementary Education Teacher Training. Research in Science Education, 50: 1689-1711.

[7] Cruz-Guzman M, Garcia-Carmona A, Criado AM, 2017, An Analysis of the Questions Proposed by Elementary Pre-Service Teachers when Designing Experimental Activities as Inquiry. International Journal of Science Education, 39: 1755-1774.

[8] Zimmerman C, Glaser R, 2003, A Follow-up Investigation on the Role of Cover Story on the Assessment of Experimental Design Skills, Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California, Los Angeles, 19.

[9] Walvoord BE, Anderson VJ, 1998, Effective Grading: A Tool for Learning and Assessment, Jossey-Bass Publishers, San Francisco, CA.

[10] Walvoord BE, Bardy B, Denton J, 2007, Closing the Feedback Loop in Classroom-Based Assessment, Assessing Student Achievement in General Education, 64-70.

[11] Farmer DW, 1993, Course-Embedded Assessment: A Teaching Strategy to Improve Student Learning. Assessment Update, 5(1): 8-11.

[12] Baughin JA, Brod EF, Page DL, 2002, Primary Trait Analysis: A Tool for Classroom-Based Assessment. College Teaching, 50(2): 75-80.

[13] Liu D, Wang L, 2012, Elements of Scientific Inquiry Ability Based on Analysis of Foreign Science Curriculum Documents. Chemistry Education, 33(9): 44-49.

[14] Cao Y, 2014, Status Quo Test and Analysis of High School Students' Physical Experimentation Ability, Henan Normal University.

[15] Guo J, Wu X, Tang Y, 2005, Evaluation Implementation Procedure of Chemical Experimental Ability in Middle School. Modern Primary and Secondary Education, 2005(11): 65-68.

[16] Tian C, Qin Q, 2018, Using PTA Scale to Assess Students' Physics Experimental Design Ability. Journal of Physics Teaching, 2018(10): 64-66.

[17] Ministry of Education, 2017, Primary Science Curriculum Standards for Compulsory Education, Ministry of Education of the People's Republic of China.

[18] Ministry of Education, 2011, Biology Curriculum Standards for Compulsory Education, Ministry of Education of the People's Republic of China.

[19] Taber KS, 2018, The Use of Cronbach's Alpha when Developing and Reporting Research Instruments in Science Education. Research in Science Education, 48: 1273-1296.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.