BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

Validity and Reliability of an Evaluation Instrument for English Curriculum System in Chinese Vocational Higher Education Institutions

***Yuping Qu*[1], *Arasinah Kamis*[2*], *Liping Jiang*[3]**

[1]*Faculty of Social Sciences and Liberal Arts, UCSI University, Kuala Lumpur, KL, Malaysia*

[2]*Faculty of Technical & Vocational, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim Perak, Malaysia*

[3]*School of Foreign Languages & International Business, Guangdong Mechanical & Electrical Polytechnic, Guangzhou, China*

ABSTRACT

The English curriculum is the basic part of the curriculum system of vocational higher education institutions (VHEIs) in China, and it is the public compulsory curriculum that students of all majors must study. It not only cultivates students' ability to learn and apply English but also lays an English foundation for students' future further learning and lifelong development. The purpose of this study is to validate the evaluation instrument for English curriculum system. This is a quantitative research and the questionnaire has adapted from previous scholars. A stratified random sampling technique has been employed to select 350 English lecturers from a total of 864 English lecturers at 20 vocational HEIs in Shandong Province, China. The results showed: i) Each dimension Clonbach alpha coefficient and McDonald's omega (Composite reliability) are more than 0.70. ii) For the convergent validity, each item factor loading >0.5, each dimension AVE >0.5. iii) For the discriminant validity, the square root of the AVE of the five dimensions is greater than the corresponding correlation coefficients between the two dimensions. Iv) Measurement model meet the fit indices, that CMIN $\chi 2 = 411.228$, with degrees of freedom (df) 314, CMIN/df = 1.310 (1-3), CFI = 0.979 (>0.9), GFI = 0.916 (>0.9), NFI = 0.916 (>0.9), and RMSEA = 0.031(<0.08). The results prove that this English curriculum implementation instrument has good validity and reliability. The instrument can measure the level of English curriculum systems in VHEIs in Shandong Province, China.

*CORRESPONDING AUTHOR:

Arasinah Kamis, Faculty of Technical & Vocational, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim Perak, Malaysia; Email: arasinah@ftv.upsi.edu.my

ARTICLE INFO

Received: 3 April 2024 | Revised: 28 April 2024 | Accepted: 1 May 2024 | Published: 4 September 2024

DOI: https://doi.org/10.30564/fls.v6i4.6599

CITATION

Qu Y., Kamis A., Jiang, L., 2024. Validity and Reliability of an Evaluation Instrument for English Curriculum System in Chinese Vocational Higher Education Institutions. Forum for Linguistic Studies. 6(4): 9-17. DOI: https://doi.org/10.30564/fls.v6i4.6599

The research findings showed there were five dimensions in the evaluation of English curriculum system. The research can be used as the basis for evaluation curriculum system for other researchers and education administrators.
Keywords: English curriculum system; CIPPO Evaluation Model; Reliability; Confirmatory factor analysis; Vocational Higher Education Institutions

1. Introduction

Vocational education has always attracted the attention of various countries, like Norway, Australia, the Netherlands, and the United States have taken measures to invest in its development (Göbel, 2019). Of course, China is no exception. The promulgation of a series of documents, such as the National Implementation Plan for Vocational Education Reform and the Action Plan for Improving the Quality and Excellence of Vocational Education (2020-2023), is better proof of this. The English curriculum is a common basic course in vocational education, with the aim of tools and humanities. To support the national implementation plan for vocational education reform and improve the quality of talent training, the General Office of the Ministry of Education issued the English curriculum standard for Vocational Higher Education Institutions (VHEIs) in 2021 (General Office of the Ministry of Education, 2021). This new standard provides guidance for English curriculum system and evaluation at VHEIs. In the context of new English curriculum standards, the system of English curriculum at VHEIs faces several challenges, the evaluation of the effectiveness of the curriculum system is one of them (Bai, 2021; Chen, 2022).

The perspective of this study on curriculum system: Goodlad (1979) provides a comprehensive perspective on the concept of the curriculum. He divided the curriculum into five different levels: ideological curriculum, formal curriculum, perceived curriculum, operational curriculum, and experiential curriculum. In this research, the formal curriculum is the new English curriculum standard, which is determined by officials. The perceived curriculum is the understanding the receivers have of the new English curriculum standard, which is determined by VHEIs and lecturers (Jiang, 2022; Jiang & Liang, 2023). The operational curriculum is the actual teaching activity of the lecturer in the classroom. The experiential curriculum is how students perceive and experience through the curriculum. Based on the view of the Goodlad curriculum, this research curriculum system is considered the process of transforming the formal curriculum into an experiential curriculum. Therefore, the evaluation of curriculum system should evaluate the entire process, i.e. from preparation to operation to outcomes.

The evaluation function consists of formative and summative. The formative function provides information and curriculum improvement considerations, while the summative function gives consideration to the results of curriculum development (Scriven, 1996). This research curriculum evaluation results are used as a strategy for curriculum system improvement. The CIPPO evaluation model is refined by Sax from the CIPP evaluation model developed first by Stufflebeam, it contains context, input, process, product, and outcomes five dimensions (Imansari & Sutadji, 2017). The CIPPO evaluation model provides a comprehensive framework for evaluating curriculum system, and many researchers use this model to evaluate educational system. Kusmiyati et al. (2023) applied a qualitative method using the CIPPO model to evaluate the system of the English language education and training program at the Indonesian Navy Education Service for Indonesian Navy personnel. Widayanto et al. (2021) applied a quantitative method using the CIPPO model to evaluate the achievement of the implementation of training programs at UPT BLK Surabaya and to provide recommendations for the next training program. Purnawirawan et al. (2020) applied the CIPPO model in evaluating the performance of school for producing entrepreneurs programs in vocational high school. Therefore, the CIPPO evaluation model was considered a comprehensive evaluation model and applied to determine the effectiveness of the learning process (Purnawirawan et al., 2020; Hamid et al., 2022; Robiah et al., 2023). **Figure 1** shows the CIPPO model for evaluating curriculum system.

In light of China's ongoing vocational education reform and the recent implementation of new English curriculum standards, there emerges a pressing imperative to develop a comprehensive evaluation instrument that caters to the unique nuances of English curriculum systems within Vocational Higher Education Institutions (VHEIs). Presently, the

evaluation of English curriculum systems in VHEIs primarily relies on examination-based methodologies, which do not fully align with the holistic approach mandated by the new English standards (Bai, 2021). Internationally, scholars have recognized this disparity and endeavored to bridge the gap by focusing on the entire curriculum implementation process and devising evaluation instruments accordingly (Hamid et al., 2022; Widayanto, 2021; Imansari & Sutadji, 2017). Evaluation instruments based on the CIPPO evaluation model are widely accepted by scholars. Building upon this scholarly foundation, the current study endeavors to adapt and validate such evaluation instruments for the context of VHEIs in Shandong Province, China. Through rigorous testing of construct validity and reliability, this research aims to provide educators and policymakers with a robust tool for assessing the efficacy and quality of English curriculum systems. By furnishing insights and benchmarks derived from empirical evidence, this study aspires to not only inform but also guide future research initiatives and educational policy-making endeavors, both domestically within China and internationally.

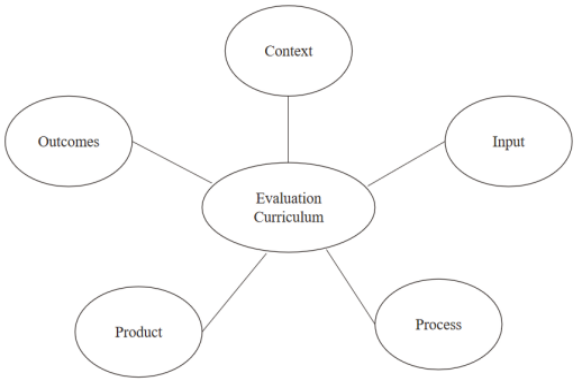

Figure 1. CIPPO model to evaluate curriculum system (Imansari & Sutadji, 2017).

2. Method

This study adopts a quantitative research design coupled with a survey methodology to delve into its research objectives. The questionnaire adapted from evaluation instruments developed by previous scholars. Prior to data collection, the content validity of the instrument was tested by experts. After data collection, the collected data will be analysed using SPSS and AMOS software. Confirmatory factor analysis (CFA) is used to evaluate the construct validity (convergent and discriminant validity) of the instrument, Cronbach's alpha (α) and McDonald's omega (Composite reliability) are used to evaluate the reliability of the instrument.

2.1 Sample and data collection

The targeted population comprises full-time English lecturers employed in Vocational Higher Education Institutions (VHEIs) across Shandong Province, China. To ensure a representative sample, this research narrows its focus to 20 VHEIs, selected randomly from the total pool of 86 such institutions within Shandong Province. Employing the "Small Sample Techniques," as proposed by Krejcie and Morgan (1970), a sample size of 269 was determined using their table-based sample size formula. However, to guarantee the acquisition of adequate and valid data, a total of 350 questionnaires will be distributed and collected during the study. Utilizing a stratified random sampling technique, 350 English lecturers will be selected from the pool of 864 English lecturers across the 20 chosen VHEIs. This meticulous approach aims to ensure the representativeness and reliability of the data collected, thereby enhancing the robustness of the study's findings.

2.2 Instrument

This study used an adapted questionnaire instrument. The questionnaire was divided into two parts. The first part focused on gathering demographic information about the respondents, such as gender, age, teaching experience, and teacher position, and contained 6 items. The second part focused on collecting information about the curriculum systems and contained 27 items. These 27 items were adapted from the seminal work of Hamid et al. (2022) and the English curriculum Standards for Vocational Higher Education Institutions (VHEIs) outlined in 2021, and were carefully crafted to capture in a multifaceted dimensions to the evaluation of English curriculum systems. These dimensions, meticulously tailored to the specific context of VHEIs, encompass five crucial aspects: context, input, process, product, and outcomes. Within this framework, each dimension is represented by a set number of items, with context comprising 5 items, input 5 items, process 6 items, product 6 items, and outcomes 5 items, totaling 27 items overall. To

gauge respondents' perceptions effectively, the instrument employs a refined five-point Likert scale, offering a nuanced spectrum of responses ranging from Strongly Agree (SA) to Strongly Disagree (SDA). Each level on the Likert scale is assigned a corresponding weight to accurately capture the degree of agreement expressed by the respondents: SA-5, A-4, N-3, DA-2, and SDA-1. Through this meticulously structured instrument, respondents are empowered to convey their nuanced perspectives on each item, thereby facilitating comprehensive data collection and analysis.

2.3 Validity and reliability of the instrument

Content validity: The purpose of content validity is to test whether the content of the questionnaire meets the purpose and requirements of the study. Multiple judges rating method, statistical method and test specification method are the evaluation methods for content validity (Rubio et al., 2003). In this study, the multiple judged rating method was used, and the researcher consulted experts in the field of study and judged each item. Three Chinese experts have been invited to participate in this study. Expert J is in the field of questionnaires, Expert L is in the field of English language education in VHEIs, and Expert Z is in the field of curriculum. The experts gave advice on the clarity of the presentation of the questions and the content setting of the items of the questionnaire. Through the synthesis and summarization of expert suggestions, the instrument was formed. Good content validity was determined through a rigorous procedures.

Structural equation modeling (SEM) refers to a set of statistical techniques used for estimating the magnitudes and directions of presumed causal effects in quantitative studies. It is commonly used in social sciences, psychology, and other fields to analyze complex relationships among variables. The method of SEM consists of three distinct families of techniques or approaches to causal inference. They are covariance-based SEM (CB-SEM), partial least squares SEM (PLS-SEM), and nonparametric SEM (Kline, 2023). This research used the CB-SEM, which includes a measurement model and structural model, the measurement model estimates relationships between the observed variables (indicator) and latent variables (Stein, 2012). Confirmatory factor analysis (CFA) is most commonly used to validate measurement models (Brown, 2015). In this study, the researcher used CFA to verify the construct validity of the instrument.

The construct validity includes convergent validity and discriminant validity. For the convergent validity, it is determined by the factor loading(acceptable > 0.5 and good ≥ 0.7) and Average Variance Extracted (acceptable ≥ 0.5) (Hair et al., 2010). For good discriminant validity, the square root of each construct's AVE should have a greater value than the correlations with other latent constructs (Fornell & Larcker, 1981).

To measure internal consistency, Cronbach's alpha (α) is the most commonly used index for assessing internal consistency or reliability, the value of Cronbach's Alpha was classified (≥0.9 highest, $0.7 \leq \alpha < 0.9$ good, $0.6 \leq \alpha < 0.7$ acceptable) (Sürücü & Maşlakçı, 2020; Jiang & Al-Shaibani, 2022). Creswell (2018) also thinks the optimal range is 0.7 to 0.9 . At the same time, the McDonald's omega (Composite reliability, acceptable ≥ 0.7) value is a complement to the reliability (Hayes & Coutts, 2020).

CFA was also used to test for goodness-of-fit between research data and hypothesized models. The assessment of model fit was based on multiple criteria. The researcher used 5 indices for the evaluation of model fit, these being: $\chi 2/df$, CFI, IFI, TLI, and RMSEA (Hair et al., 2006; Hair et al., 2010).

A small sample survey will conduct before the survey. A pilot study has been conducted to test the instrument's reliability. 30 English lecturers from VHEIs have been invited to participate. **Table 1** shows the result of the instrument's reliability. Through the pilot study, the reliability value of each dimension and the total is more than 0.7, which indicates that the instrument has good reliability and can continue to be distributed.

2.4 Data Analysis

In the data analysis phase, Confirmatory Factor Analysis (CFA) was employed as a robust method to assess the validity of the instrument utilized in this study. Specifically, the researcher conducted rigorous examinations to ascertain both convergent and discriminant validity. Convergent validity, which scrutinizes the relationship between various measures of the same construct, was evaluated against established benchmarks. Factors such as factor loading (FL), Average Variance Extracted (AVE) were meticulously assessed to ensure satisfactory convergent validity (Hair et al., 2010).

Table 1. Result of the instrument's reliability (Pilot Test).

	Total of items	Alpha Cronbach Value
Context	5	0.890
Input	5	0.901
Process	6	0.792
Product	6	0.949
Outcomes	5	0.859
Total	27	0.949

On the other hand, discriminant validity, aimed at determining whether measures from distinct constructs are distinct from one another, was meticulously scrutinized. This was achieved by comparing the square root of the AVE against the correlation between the constructs, as outlined by Fornell and Larcker (1981). Furthermore, to gauge the overall fit of the measurement model, five key indices were meticulously evaluated: $\chi 2/df$ (ideally between 1 and 3), Comparative Fit Index (CFI) (>0.9), Incremental Fit Index (IFI) (>0.9), Tucker-Lewis Index (TLI) (>0.9), and Root Mean Square Error of Approximation (RMSEA) (<0.08). These established benchmarks, as outlined by Hair et al. (2006; 2010), were leveraged to ensure the robustness and adequacy of the measurement model. Through this rigorous analytical framework, the validity and reliability of the instrument were meticulously scrutinized, providing a solid foundation for the subsequent data interpretation and analysis.

3. Results and discussion

In this study, 328 valid questionnaires were collected through the Questionnaire Star online questionnaire platform. A total of 27 items were measured for the validity and reliability of each dimension.

The reliability of the questionnaire was verified by measuring the Cronbach alpha by SPSS (version 26) software. Confirmatory factor analysis (CFA) was used to measure the instrument's construct validity by AMOS (version 24) software. The results of the reliability and the CFA analysis are discussed in the subsection.

3.1 Convergent validity and reliability

According to Hair er al. (2010) a good convergent validity is determined by checking the factor loading of each item (>0.5), the Average variance extracted (AVE) of each dimension (>0.50). **Table 2** shows the result of the instrument's convergent validity and reliability. Referring to **Table 2**, the factor loading of each item ranges from 0.660 to 0.818, each factor loading has exceeded the cut-off value of 0.5. The AVE values are within 0.511 and 0.604, each AVE value has exceeded the cut-off value of 0.5. Thus the evaluation instrument for English curriculum system has good convergent validity. The CR for all dimensions are within 0.855 and 0.892, each CR has exceeded the cut-off value of 0.7 and Cronbach alpha values ranged from 0.853 to 0.892, each α value higher than 0.7. Thus the evaluation instrument for English curriculum system has adequate reliability and the internal consistency reliability is well.

3.2 Discriminant validity

According to Fornell and Larcker's (1981) criteria, discriminant validity is assessed by comparing the square root of the Average Variance Extracted (AVE) for each construct with its correlations to other latent constructs. A solid discriminant validity exists when the square root of a construct's AVE surpasses its correlations with other latent constructs.

Table 3 presents the outcomes of the evaluation instrument's discriminant validity. As evident from the table, the square roots of the AVEs for all five dimensions—context, input, process, product, and outcomes—are higher than the correlation coefficients between any two of these dimensions. This finding underscores the strong discriminant validity of the evaluation instrument. In essence, it demonstrates that the five dimensions are distinct from each other, highlighting significant differences among them. Therefore, the evaluation instrument possesses good discriminant validity, ensuring that each dimension—context, input, process, product, and outcomes—is measured independently without undue overlap or confusion with the others.

Table 2. Result of the instrument's convergent validity and reliability.

Construct	Item	Internal reliability Cronbach alpha (α)	Factor loading	Composite reliability1 (ω)	Average variance extracted2 (AVE)
context	context1	0.884	0.758	0.884	0.604
	context2		0.747		
	context3		0.779		
	context4		0.805		
	context5		0.795		
input	input1	0.853	0.701	0.855	0.541
	input2		0.739		
	input3		0.762		
	input4		0.737		
	input5		0.735		
process	process1	0.862	0.661	0.862	0.511
	process2		0.675		
	process3		0.780		
	process4		0.692		
	process5		0.775		
	process6		0.699		
product	product1	0.892	0.738	0.892	0.580
	product2		0.781		
	product3		0.773		
	product4		0.746		
	product5		0.749		
	product6		0.783		
outcomes	outcomes1	0.861	0.721	0.863	0.560
	outcomes2		0.723		
	outcomes3		0.807		
	outcomes4		0.818		
	outcomes5		0.660		

Note: [1]Composite reliability = (square of the summation of the factor loadings)/(square of the summation of the factor loadings) + (summation of the error variances)
[2]Average variance extracted = (summation of the square of the factor loadings)/(summation of the square of the factor loadings) + (summation of the error variances)

3.3 Criteria for Fit Index

To access the suitability of the measurement model, χ2/df, CFI, IFI, TLI, and RMSEA fitness index are used for analysis (Hair et al., 2010; Hair et al., 2006). According to Hair (2010), Bentler (1990), Marsh and Hocevar (1985), CMIN/DF value less than 5 indicate an adequate model. According to Hair (2010), Byrne and Van De Vijver (2010) RMSEA considered adequate if less than 0.08. According to Bagozzi and Yi (1988), Bentler (1990), Hair et al. (2006), and Schumacker, R. and R.G. Lomax(2010) the values of CFI, IFI,TLI considered adequate if more than 0.9. Referring to the **Table 4**, it shows the result of measurement model indices. All fit indices exceeded the suggested value, CMIN/DF= 1.310 (<5), RMSEA=0.031 (<0.08), TLI=0.976, IFI=0.979, CFI=0.979(>0.90). These all indices show a structural model has achieved fit indices and fit the data collected. **Figure 2** shows the result of measurement model of English curriculum system.

The reliability and validity of the instrument are tested. The findings of the study showed that the CR and Cronbach alpha values were satisfactory. This confirms that the instrument is reliable, stable and consistent. The findings of the study showed that convergent validity, discriminant validity and model fit were satisfactory. This confirms that the instrument items are reasonably designed to effectively reflect the research objectives. Meanwhile, it also justifies the evaluation of curriculum system according to thee CIPPO

Table 3. Result of the instrument's discriminant validity.

	Context	Input	Process	Product	Outcomes
context	0.777				
input	0.560***	0.735			
process	0.622***	0.631***	0.715		
product	0.632***	0.507***	0.669***	0.762	
outcomes	0.580***	0.473***	0.577***	0.628***	0.748

Note: Diagonals represent the square root of the average variance extracted, while the other entries represent the Correlations Coefficients between the respective construct.

Table 4. Criteria for Index Fit.

Index	Authors	Suggested value	Measurement
CMIN			411.228
DF			314
CMIN/DF	Hair (2010)	<5	1.310
RMSEA (the root mean square error of approximation)	Hair (2010)	<0.08	0.031
TLI	Schumacker and Lomax (2012)	>0.9	0.976
IFI	Hair et al. (2006)	>0.9	0.979
CFI	Bentler (1990)	>0.9	0.979

evaluation model. This findings is consistent with Kusmiyati et al. (2023), Widayanto et al. (2021), Hamid et al. (2022). Overall, after preliminary testing, this instrument is feasible to measure the system of English curriculum.

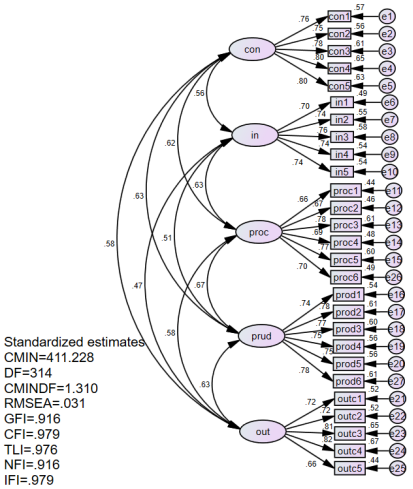

Figure 2. Measurement model of English curriculum system.

4. Conclusion and managerial implications

In conclusion, the validation of the instrument for measuring the English curriculum system represents a significant achievement in this study. Comprising five dimensions encompassing context, input, process, product, and outcomes, this instrument offers a comprehensive framework for assessing the effectiveness and efficiency of English curriculum systems in Virtual Higher Education Institutions (VHEIs). By providing a systematic approach to evaluation, it is hoped that this instrument will serve as a valuable tool for educators and policymakers alike in gaining insights into the strengths and areas for improvement within English curriculum systems. It is important to recognize that the curriculum system transcends beyond mere classroom teaching, encompassing various facets such as contextual factors, input mechanisms, teaching processes, and ultimately, student outcomes (Jiang, 2022). This underscores the holistic nature of curriculum development and implementation.

The validated English curriculum evaluation instrument introduced in this study serves as a valuable tool for vocational VHEI administrators. It provides a standardized, reliable method to assess the effectiveness of the English curriculum, guiding informed decisions for curriculum enhancement. The five evaluation dimensions pinpoint specific areas needing attention, enabling targeted improvements. Furthermore, this instrument fosters benchmarking and best practice sharing among VHEIs. Its validity and reliability bolster institutional accountability and transparency. In summary,

this evaluation instrument is a concise yet powerful tool for administrators to optimize their English curriculum systems.

However, it is essential to acknowledge a few limitations within this research endeavor. Firstly, the instrument's applicability and validity may vary across different educational contexts and settings, as the study primarily focused on VHEIs. Secondly, while the dimensions included in the instrument provide a comprehensive framework, there may still be additional factors or nuances within English curriculum systems that were not fully captured. Future research could delve deeper into these areas to enhance the instrument's comprehensiveness and applicability across diverse educational environments.

In future studies, it would be beneficial to conduct longitudinal research to track the implementation of recommended changes based on the evaluation of English curriculum systems. This would provide valuable insights into the effectiveness of interventions over time and contribute to the ongoing improvement of English language education in VHEIs.

Author Contributions

All authors contributed to the conception, design, and data analysis of this study. Yuping Qu contributed to the acquisition and interpretation of data and drafted the manuscript. Arasinah Kamis contributed to the design of the study and critically revised the manuscript. Liping Jiang contributed to the interpretation and provided critical feedback on the manuscript.

Data availability

Some or all data, models, or code generated or used during the study are available from the corresponding author upon request.

Confict of Interest

There are no conflicts of interest involved in this research.

Funding

This work was supported by 2023 Guangdong Education Science Planning Project (Special program for Higher Education) (2023GXJK711).

References

Bagozzi, R.P., Yi, Y., 1988. On the evaluation of structural equation models. Journal of the Academy of Marketing Science. 16(1), 74–94. DOI: https://doi.org/10.1007/bf02723327

Bai, H., 2021. Dilemma and path selection of English curriculum standards for Higher Vocational Education. Vocational Education Research. (12), 60–64.

Bentler, P.M., 1990. Comparative fit indexes in structural models. Psychological Bulletin. 107(2), 238–246. DOI: https://doi.org/10.1037/00332909.107.2.238

Brown, T.A., 2015. Confirmatory Factor Analysis for Applied Research, Second Edition. Guilford Publications.

Byrne, B.M., 1998. Structural equation modeling with LISREL, PRELIS and SIMPLIS: Basic concepts, applications and programming. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Byrne, B.M., Van De Vijver, F.J.R., 2010. Testing for Measurement and Structural Equivalence in Large-Scale Cross-Cultural Studies: Addressing the issue of Nonequivalence. Interna- tional Journal of Testing, 10(2), 107–132. DOI: https://doi.org/10.1080/15305051003637306

Chen, Y., 2022. A study on the current problems of evaluation in public English teaching in higher vocational colleges. Journal of Hubei Open Vocational College. (13), 186–188.

Creswell, J.W., Creswell, J.D., 2017. Research design: Qualitative, Quantitative, and Mixed Methods Approaches. SAGE Publications.

Fornell, C., Larcker, D.F., 1981. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research. 18(1), 39–50. DOI: https://doi.org/10.1177/002224378101800104

General Office of the Ministry of Education. (2021, March 23). English curriculum standards for vo- cational higher education institutions (2021). Available from: https://www.gov.cn/zhengce/zhengceku/2021-04/10/content_5598801.htm

Göbel, J. (2019, November 11). "Discover your talents." deutschland.de. Available from: https://www.deutschland.de/zh-hans/topic/jingji/lian-heguojiaokewenzuzhicujinzhiyejiaoyu

Goodlad, J.I., 1979. Curriculum Inquiry: The study of Curriculum practice. Available from: https://ci.nii.ac.jp/ncid/BA05076058

Hair, J.F., Black, W.C., Babin, B.J., et al., 2006. Multivariate data analysis (6th ed.). Pearson Prentice Hall Upper Saddle River, NJ.

Hair, J., Black, B., Babin, B., et al., 2010. Multivariate Data Analysis: A Global Perspective (7th ed.). New Jersey, USA: Pearson Education Inc.

Hamid, M.R., Sutaman, S., Natsir, M., et al., 2022. The development of an evaluation instrument for the implementation of the Arabic language curriculum in Islamic High School. Jurnal Al-Bayan, 14(1), 242–257. DOI: https://doi.org/10.24042/al-bayan.v14i1.10303

Hayes, A.F., Coutts, J.J., 2020. Use Omega Rather than Cronbach's Alpha for Es- timating Reliability. But... Communication Methods and Measures. 14(1), 1–24. DOI: https://doi.org/10.1080/19312458.2020.1718629

Imansari, N., Sutadji, E., 2017. A concep- tual framework curriculum evaluation Electrical engineering education. International Journal of Evaluation and Research in Education. 6(4), 265. DOI: https://doi.org/10.11591/ijere.v6i3.7624

Jiang, L., 2022. Effective Utilization of Computer-Mediated Communication Technology in Network-based Foreign Language Teaching. Wireless communications and mobile computing. 1–7. DOI: https://doi.org/10.1155/2022/1048311

Jiang, L., 2024. Factors influencing EFL teachers' implementation of SPOC-based blended learning in higher vocational colleges in China: A study based on grounded theory. Interactive Learning Environments. 32(3), 859–878. DOI: https://doi.org/10.1080/10494820.2022.2100428

Jiang, L., Al-Shaibani, G.K.S., 2024. Influencing factors of students' small private online course-based learning adaptability in a higher vocational college in China. Interactive Learning Environments. 32(3), 972–993. DOI: https://doi.org/10.1080/10494820.2022.2105901

Jiang, L., Liang, X., 2023. Influencing factors of Chinese EFL students' continuance learning intention in SPOC-based blended learning environ- ment. Education and Information Technologies. 28(10), 13379–13404. DOI: https://doi.org/10.1007/s10639-023-11734-4

Kline, R.B., 2023. Principles and practice of structural equation modeling. Guilford publications.

Krejcie, R.V., Morgan, D.W., 1970. Determining sample size for research activities. Educational and Psychological Measurement. 30(3), 607–610. DOI: https://doi.org/10.1177/001316447003000308

Kusmiyati, N., Hamidah, K., 2023. CIPPO model Evaluation on the English Language training program at the Indonesian Navy Education Services. IJHCM (International Journal of Human Capital Management) (Online). 7(1), 104–114. DOI: https://doi.org/10.21009/ijhcm.07.01.8

Marsh, H.W., Hocevar, D., 1985. Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. Psychological Bulletin. 97(3), 562–582. DOI: https://doi.org/10.1037/0033-2909.97.3.562

Purnawirawan, O., Chintya, P.P., Sholihah, M., 2020. The application of CIPPO Evaluation Model in evaluating the performance of school for Producing Entrepreneurs programs in Vocational High School. Proceedings of the International Conference on Science and Education and Technology (ISET 2019). DOI: https://doi.org/10.2991/assehr.k.200620.075

Robiah, R., Hendarman, H., Hidayat, R., 2023. Evaluasi Program Literasi Anak dengan Pendekatan Model CIPPO. Murhum. 4(1), 528–539. DOI: https://doi.org/10.37985/murhum.v4i1.262

Rubio, D.M., Berg-Weger, M., Tebb, S., et al., 2003. Objectifying content validity: Conducting a content validity study in social work research. Social Work Research. 27(2), 94–104. DOI: https://doi.org/10.1093/swr/27.2.94

Schumacker, R.E., Lomax, R.G., 2012. A Beginner's Guide to Structural equation modeling: Third Edition. Routledge.

Scriven, M., 1996. Types of evaluation and types of evaluator. Evaluation Practice. 17(2), 151–161. DOI: https://doi.org/10.1016/s0886-1633(96)90020-3

Stein, C.M., Morris, N.J., Nock, N.L., 2012. Structural equation modeling. Statistical human genetics: Methods and protocols. 495–512.

Sürücü, L., Maşlakçı, A., 2020. VALIDITY AND RELIABILITY IN QUANTITATIVE RESEARCH. Business & Management Studies: An International Journal (Online). 8(3), 2694–2726. DOI: https://doi.org/10.15295/bmij.v8i3.1540

Widayanto, L.D., Soeharto, S., Sudira, P., et al., 2021. Implementation of the Education and Training Program seen from the CIPPO Perspective. Journal of Education Research and Evaluation (Online). 5(4), 614. DOI: https://doi.org/10.23887/jere.v5i4.36826