Art. #2480, 18 pages, https://doi.org/10.15700/saje.v44ns1a2480

Investigation and research on comprehensive quality assessment in the reform of new college entrance examination

Mengna Zheng and Chengwu Ruan
School of Educational Science, Anhui Normal University, Wuhu, China
rcw243312198@163.com

Comprehensive quality assessment is an assessment system that identifies and explores students' strengths. By examining the developmental progress made in pilot provinces that have implemented comprehensive quality assessment, valuable insights and guidance can be derived for other provinces preparing to adopt this assessment approach. In this study we conducted an investigation and research involving students, teachers, and middle managers from 12 high schools in 2 pilot areas, the Zhejiang and Jiangsu provinces. Using a combination of empirical and qualitative research methods, we aimed to explore the levels of satisfaction and recognition among different research subjects regarding the current implementation situation and to analyse issues in the current implementation process. According to the corresponding theory, an in-depth analysis is conducted from 4 dimensions to uncover the underlying causes of the problem, while proposing recommendations to enhance and advance comprehensive quality assessment. These recommendations are firstly to enhance the standard level of matching and integration in comprehensive quality assessment; secondly to augment the utilisation level of comprehensive quality assessment; thirdly to elevate the significance of comprehensive quality assessment results; and in the fourth place, to improve the credit rating system associated with comprehensive quality assessment.

Keywords: comprehensive quality assessment; higher education; new college entrance examination; ordinary high school; quality progress

Introduction

Background of the Study and Problem Statement

Significance of comprehensive quality assessment in the reform of the new college entrance examination

In the context of China's recent reform of the college entrance examination, comprehensive quality assessment is not only part of the reform of the mode of selecting talented individuals but also an essential transformation of their mode of training (U-Sayee & Adomako, 2021). It is also a part and highlight of the leading measures to reform the new round of the high entrance system (Cheng, Z 2023). Assessment is the transformation of the social needs for exam-oriented selection of talented individuals and one of the vital measures of the span between relying on "exam points" and paying attention to "individual development." By observing and recording students' growth process, the comprehensive quality assessment further analyses their development process, explores and cultivates their strengths, and meets their diverse development needs. It also serves as a talent reserve to provide for the needs of economic and social development and the cultivation of varied high-quality talented individuals (Kamyab, Gholami, Behdad & Jeihooni, 2023). In the face of the development of comprehensive quality assessment over the years, the common focus worldwide is to establish what kind of experts should be cultivated to adapt to the era of such rapid growth. The United Nations Educational, Scientific and Cultural Organization (UNESCO) report by the International Commission on Education for the 21st Century, published in 1996 (Gao, B & Cheng, 2013), included the concept of lifelong learning and the three stages of the core literacy of learning outcomes, published in 2012 (Modiba & Sefotho, 2019). Both put forward new requirements for the quality of citizens (Zhenguo, 2018). The three stages include self-cognition and attitude to learning, subject knowledge and thinking ability, practical ability and innovative spirit. Among them, the demand for comprehensive quality assessment in the new college entrance examination reform is the demand for the third stage of the core ability development of students (Wen & Song, 2022). Considering the needs that arise due to the rapid development of society, the significance of including the content of the comprehensive quality assessment in the new college entrance examination as a reference basis for the college entrance examination has emerged. At present, the Zhejiang and Jiangsu provinces have already started to implement the results of the comprehensive quality assessment as a reference basis for the admissions process of the college entrance examination system, so that comprehensive quality assessment can be genuinely incorporated into the admissions process (Yaqian & Yuqiang, 2022). The significance of the assessment results as a basis for talent selection is, firstly that it can encourage students to change their way of thinking about the assessment and make them pay attention to their development in different fields, instead of just focusing on the results of a core subject or a specific type of competition that can add points to the entrance examination. Secondly, it can develop students' multiple abilities, so that they can discover their strengths, develop their horizontal and vertical thinking, improve their diversified understanding, and enrich their creativity in the process of development (Ojo & Adu, 2018). Thirdly, through the effective execution of the 3-year assessment, students can enhance their sense of self-responsibility so that when they face the changes and challenges of the future society the assessment process will have prepared them for a more humanistic and socially practical culture, and not just theoretical knowledge from books.

Practical problems facing comprehensive quality assessment in high schools[i] in the pilot provinces of the new college entrance examination

Since the inception of the pilot reform of the new college entrance examination of China in 2014, along with the implementation of comprehensive quality assessment in the Zhejiang and Jiangsu provinces, certain issues have been identified during assessment and development, primarily those discussed below.

Issues involving students are that they passively lack understanding of the assessment process, which encompasses various dimensions such as ideology and morality, academic level, physical and mental health, interests and specialties, social practice, personal accomplishments, awards, and sports health (Wu & Ji, 2022). Moreover, they possess limited familiarity with the electronic recording platform used.

Teachers still play a dominant role in the assessment process and the following issues are noted. Firstly, they rely heavily on ranking regular exam scores as a reference for assessment, disregarding fair and objective process-based assessment of students' development. Secondly, some teachers lack the necessary professional competence to perform effective assessment and may resort to random assessments merely to complete superficial tasks during the implementation of the assessment process (Dong, 2020). Thirdly, China's long-standing reliance on an "exam-based" assessment system that places immense pressure on academic performance has been ingrained deeply in teachers' mindsets and now hinders their ability to shift their psychological role within the assessment process. The resulting psychological pressure also contributes to why teachers struggle to exert a positive guiding influence in this context.

The school administrators also face difficulties despite the fact that some schools have formed a professional assessment research group to show how to formulate an assessment plan suitable for the school at the beginning of the assessment; For example, how to actively guide students and teachers to participate in the actual implementation of the assessment process and how to solve the problem that school administrators find it difficult to complete or understand the assessment content (Zhenguo, 2018); how to record the assessment process and results, and how to convert the recorded text into scores according to the situation of the school.

Literature Review

One view is that students' comprehensive quality should be composed of both academic and non-academic ability (Cheng, Z 2023). A contrasting view is that comprehensive quality only refers to non-academic ability, and extensively includes all kinds of knowledge and abilities that can be examined using traditional paper-and-pencil examinations. However, in a general sense it refers to the sum of other qualities that cannot be reflected by subject tests (Cheng, L 2017). As for the essence of comprehensive quality assessment, a typical view is that comprehensive quality assessment is regarded as personality development assessment, and the essence of quality.

In studies on the dimensions of the comprehensive quality assessment system in China it was found that it includes the requirements of students' moral and civic qualities, learning and innovation ability, cooperation and communication, sports and health, and aesthetics and performance. The first view is that education needs to respect the uniqueness of individual development, and since the essence of quality education is personality development education, comprehensive quality assessment is essentially personality development assessment. Li Yanbing believes that the essence of comprehensive quality assessment is personality development assessment, which involves authenticity, process assessment, or internal assessment (Yanbing, 2016). Wang Hongxi positioned the essence of comprehensive quality assessment as a kind of appreciative assessment or appreciation assessment (Wang, 2019).

The major problems arising in the implementation and practice of comprehensive quality assessment are student satisfaction (Gao, X 2013), teacher resistance to implementation (Shao, 2013), a study of rubrics (Du, 2017), fairness in the implementation (Huiying, 2011), comprehensive quality assessment for independent admissions or incorporated into college entrance examination admissions (Zhao & Meng, 2013), web-based and management systems (Petalla & Tatlonghari, 2023).

The assessment experiences of Western and East Asian countries are reflected in the following observations: The admission standard of American colleges and universities is to select outstanding students by considering various aspects of their comprehensive level and diversity. They do not only use a single test score as the admission criterion (Wang, 2019). Specifically, in the United States of America (USA), a comprehensive ability assessment and assessment system with pluralism and rich diversity is implemented, and the recruitment criteria referred to cannot be measured by achievement scores alone, which does not mean that academic test scores are not valued; academic performance is used as a process assessment benchmark (Freeks, 2015). Students' academic ability is examined through the content of the prescribed unified examination, and the academic results achieved by students at each stage of each semester for 3 years of high school are summarised as process assessment (Mohono-Mahlatsi & Van Tonder, 2006). In addition, the students' overall development is examined in a diversified and

multifaceted manner by combining interviews, letters of recommendation from teachers, application plans for admission, research proposals, and supporting documents of various extracurricular activities in which the student has participated (Wu & Ji, 2022).

To summarise, current research on comprehensive quality assessment focuses primarily on the understanding and theoretical analysis of relevant policy texts, with some scholars also exploring the psychological dimension through theories such as multiple intelligences. However, limited research exists on examining the actual implementation of the assessment. Therefore, in this study, two pilot areas with different situations were selected as the objects of investigation. Investigations were done from the perspective of the connotation of the comprehensive quality assessment system, the process of the comprehensive quality assessment activities, the structure of the comprehensive quality assessment, and the motivation and satisfaction of the participants in the assessment process. Through empirical research and in-depth interviews, combined with the theories of subjective education, assessment of multiple intelligences, Marx's theory of comprehensive human development and developmental assessment, we explored the real issues reflected in the implementation process of the policy in different regions, and offer corresponding suggestions for improvement.

Research Questions

Within the general scope of the study, the following objectives were set: to use empirical research to analyse the current satisfaction degree of the three role-players of comprehensive quality assessment (students, teachers and school administrators) with the implementation of the assessment; learn about the current status and situation of the pilot work of comprehensive quality assessment in the Jiangsu and Zhejiang provinces; analyse the real advantages and disadvantages; draw lessons from the excellent points, and put forward effective suggestions on the weak points to provide reference for other provinces to carry out assessment.

Methodology

Procedure

The survey was conducted from November 2016 to January 2019. The survey sample was drawn from 12 schools in two cities in the Jiangsu and Zhejiang provinces, since the Zhejiang province is the priority region where the assessment is being implemented, and the Jiangsu province is the region where the assessment is being pre-implemented. Therefore, the selection of study subjects in these two provinces was representative. According to the different positioning of schools, they were classified into three types: provincial model high schools, general high schools and integrated high schools. Three different groups (students, teachers and school-level administrators) were randomly selected as the research sample from the three different types of schools in the two provinces to complete questionnaires and participate in interviews. Provincial model high schools refer to those high schools with strong academic ability in academic level tests and a high overall level of final exam results, followed by particular category high schools that include particular subjects, for example, English, German, and French, which feature minor languages as separate tests to select students and train them for specialties. Generalist high schools refer to schools with slightly weaker academic abilities and a balanced training of students' academic performance and comprehensive practical skills. Integrated high schools refer to schools where students choose to enter vocational and technical high schools after junior high school, but still have the opportunity to take the college entrance exam.

Participants

A questionnaire survey was the main research method, involving three types of direct participants in comprehensive quality assessment (students, teachers, and school managers) to establish satisfaction with the assessment and present the results of the data analysis (see Table 1). With the feedback results of data analysis and the research of relevant educational theories, with this article we reveal the current status of the implementation of assessment in the Jiangsu and Zhejiang provinces, summarise the problems arising in the practice of assessment and analyse the causes. This was done to draw lessons from the good practices in the implementation of assessment, and put forward suggestions to attend to shortcomings.

Table 1 Composition of research objects

Type	Student group		Teacher group		Management groups		Total
Gender	Male	Female	Male	Female	Male	Female	
Key type	206	321	26	59	23	28	663
Ordinary type	179	210	31	21	35	21	497
Accommodating type	138	144	16	22	32	15	367
Total	523	675	73	102	90	64	1,527

A total of 1,880 questionnaires were received: 1,500 paper versions of the survey feedback questionnaires for students, 200 questionnaires for teachers, and 180 questionnaires for school-administrators. Among them were 1,198 valid questionnaires for students, 175 for teachers, and 154 for school-level administrators. In addition, using internet technology, 430 questionnaires were collected of which 358 were valid (see Table 2).

Table 2 Dimensional framework of the comprehensive quality assessment scale

Questionnaire subjects	Dimension of the questionnaire	Questionnaire topic
Students (independent variable) Gender, age, school location (1–5)	Value assessment	6, 9, 10, 11
	Content analysis	4, 5, 7, 8, 27
	Cognitive grade	15, 21, 22, 25
	Participation level	14, 28, 16, 17
	Credit rating	12, 13, 26, 18
	Action level	23, 24, 19, 20
Classroom teachers (Independent variable) Gender, age, teaching experience, education, title, position, location of teaching, type of school teaching (11–8)	Value assessment	14, 15, 9, 10
	Content analysis	32, 21, 22, 23
	Cognitive grade	17, 18, 19, 16
	Participation level	20, 25, 30, 13
	Credit rating	24, 11, 12, 31
	Action level	28, 29, 26, 27
School-level managers (Independent variable) Gender, age, teaching experience, education, title, position, location of teaching, type of school teaching (11–8)	Value assessment	14, 15, 9, 10
	Content analysis	32, 21, 22, 23
	Cognitive grade	17, 18, 19, 16
	Participation level	20, 25, 30, 13
	Credit rating	24, 11, 12, 31
	Action level	28, 29, 26, 27

Measures

This study was based on a concept in psychological research, namely psychological acceptance, first developed by Gordon W. Allport (1897–1967), an American personality psychologist, and a questionnaire based on the research of related scholars. The questionnaire for the three groups who participated in the survey is a targeted questionnaire based on the research of relevant scholars. The questionnaire covers six dimensions: value assessment, content analysis, cognitive level, participation level, credit level, and role level, with four questions for each dimension, making a total of 24 questions. The questionnaire was scored on a 5-point Likert scale, with 1 representing satisfaction, 2 representing basic satisfaction, 3 representing basic dissatisfaction, 4 representing complete dissatisfaction, and 5 representing great satisfaction, with higher scores representing higher student satisfaction with the overall implementation of the comprehensive quality assessment. In our study, Cronbach's α was 0.76. This study was conducted to explore the reliability of this questionnaire in depth and to indicate that the questionnaires for students, school-level administrators, and classroom teachers have good reliability and validity, and are suitable for the application of the survey on satisfaction with comprehensive quality assessment under the new college entrance examination reform.

Data Analysis

The survey data were compared with the alternative media theory data (3.0) of the comprehensive quality assessment scale for satisfaction degree, using the *t*-test, *f*-test and one-way analysis of variance to detect the difference of survey data among different types of schools, as well as the impact on students, teachers, and schools. By using these statistical analysis methods, we were able to examine the differences in survey data between different types of schools and the implications for students, teachers, and schools (see Tables 3–5).

Table 3 Student personal basic information form

		f	%	Effective %	Cumulative %
Gender	Male	565	47.2	47.4	47.4
	Female	627	52.3	52.6	100.0
Missing	System	6	0.5		
	Senior	510	42.6	42.7	42.7
Year	Sophomore	454	37.9	38.0	80.7
Age	Senior	230	19.2	19.3	100
Missing	System	4	0.3		
	Key schools	595	49.7	50.6	50.6
School type	Ordinary schools	556	46.4	47.3	97.9
	Township schools	10	2.1	2.1	100
	Private schools	15	98.2		82.4
Missing	System	22	1.8		
	Total	1,192	100	100	——

Table 4 Teacher and head teacher form personal basic information form

		f	%	Effective %	Cumulative %
Education	Secondary school	6	3.8	3.8	3.8
	College	29	18.2	18.2	22.0
Background	Undergraduate	77	48.4	48.4	70.4
	Master's degree	39	24.5	24.5	95.0
	Doctoral student	8	5.0	5.0	100
Gender	Male	72	45.3	45.3	45.3
	Female	87	54.7	54.7	100
Age	20–29	49	30.8	30.8	30.8
	30–39	25	15.7	15.7	46.5
	40–49	39	24.5	24.5	71.1
	50 and above	42	26.4	26.4	100
Teaching period	1–5 years	29	18.2	18.7	18.7
	6–10 years	31	19.5	20.0	38.7
	11–15 years	21	13.2	13.5	52.3
	16–20 years	27	17.0	17.4	69.7
	21 years and above	47	29.6	30.3	100
Missing	System	4	2.5		
Position	Junior faculty	38	23.9	23.9	23.9
	Intermediate teachers	40	25.2	25.2	49.1
	Senior teacher	58	36.5	36.5	85.5
	Special grade teachers	20	12.6	12.6	100
School type	Key schools	46	28.9	28.9	28.9
	Ordinary schools	71	44.7	44.7	73.6
	Township schools	15	9.4	9.4	83.0
	Private schools	17	10.7	9.3	100
	Total	175	100	100	——

Table 5 Personal information sheet of school management

		f	%	Effective %	Cumulative %
Education	Secondary school	0	0	0	0
	College	0	0	0	0
Background	Undergraduate	86	55.8	55.8	51
	Master's degree	66	41.5	41.5	41.2
	Doctoral student	2	1.25	1.25	100.0
Gender	Male	90	58.4	58.4	56.3
	Female	64	41.6	41.6	100.0
Age	20–29	12	7.79	7.79	5.6
	30–39	71	46.1	46.1	46.1
	40–49	79	51.2	51.2	51.2
	50 and above	4	2.5	2.5	100.0
Teaching period	1–5 years	0	0	0	0
	6–10 years	13	8.4	8.4	8.4
	11–15 years	58	37.7	37.7	46.1
	16–20 years	0	0	0	0
	21 years and above	83	53.9	53.9	100.0
Position	Junior teacher				
	Intermediate teacher	71	46.1	46.1	46.1
	Senior teacher	73	47.4	47.4	47.4
	Special grade teacher	10	15.4	15.4	100.0
School type	Key schools	64	40.2	40.2	39.8
	Ordinary schools	72	45.2	45.2	41
	Township schools	6	3.77	3.77	2.6
	Private schools	12	7.5	7.5	100.0
Position	School leaders	86	55.8	55.8	55.8
	School middle-level cadres	68	44.2	44.2	100.0
	Total	154	100	100	——

Table 6 Independent sample *t*-test for comparison of different dimensions of students in the two provinces

	Zhejiang province		Jiangsu province			
Name of dimension	*M* ± *SD*	Mean difference in each dimension	*M* ± *SD*	Mean difference in each dimension	*t*	*p*
Value assessment	11.075 ± 3.473	2.768	9.611 ± 2.334	2.402	11.276	0.000
Content analysis	9.573 ± 2.406	2.393	9.163 ± 2.177	2.290	2.828	0.000
Cognitive grade	10.903 ± 3.156	2.725	11.201 ± 3.044	2.800	-3.368	0.001
Participation level	9.132 ± 2.249	2.283		2.167	-1.175	0.241
Credit rating	9.320 ± 1.796	2.33		2.359	-0.244	0.808
Action level	8.221 ± 1.315	2.055	8.484 ± 1.673	2.121	-2.920	0.004
Total satisfaction score	58.224 ± 7.877		38.459 ± 3.304		4.668	0.000
Average score of satisfaction	9.704 ± 1.312		6.409± 1.715			

Table 7 Comparison of different dimensions of teachers in the two provinces independent sample *t*-test

	Zhejiang province		Jiangsu province			
Name of dimension	*M* ± *SD*	Mean difference in each dimension	*M* ± *SD*	Mean difference in each dimension	*t*	*p*
Value assessment	10.664 ± 3.243	2.666	11.019 ± 3.744	2.754	-1.779	0.007
Content analysis	9.345 ± 3.159	1.938	7.754 ± 1.657	2.336	-16.360	0.000
Cognitive grade	9.645 ± 2.387	2.411	8.800 ± 2.254	2.2	4.080	0.000
Participation level	8.735 ± 2.067	2.183	9.516 ± 2.330	2.379	-3.495	0.001
Credit rating	10.825 ± 2.865	2.706	10.871 ± 3.399	2.717	-0.176	0.861
Action level	8.548 ± 1.743	2.137	7.793 ± 1.292	1.948	4.350	0.000
Total satisfaction score	57.762 ± 9.067		60.354 ± 7.276		-9.247	0.000
Average score of satisfaction	9.627 ± 1.511		10.059 ± 1.212			

The satisfaction analysis of senior high school leaders

Table 8 Independent sample *t*-test for the comparison of different dimensions of university-level managers in the two provinces

	Zhejiang province		Jiangsu province			
Name of dimension	*M* ± *SD*	Mean difference in each dimension	*M* ± *SD*	Mean difference in each dimension	*t*	*p*
Value assessment	11.238 ± 3.161	2.809	9.859 ± 3.395	2.464	-4.875	0.000
Content analysis	11.088 ± 2.197	2.772	8.711 ± 2.056	2.177	9.389	0.000
Cognitive grade	10.577 ± 3.239	2.644	8.861 ± 3.733	2.215	7.390	0.000
Participation level	11.183 ± 2.960	2.795	12.283 ± 2.220	3.070	-3.947	0.000
Credit rating	10.427 ± 3.002	2.606	10.333 ± 3.752	2.583	0.403	0.668
Action level	9.205 ± 2.073	2.301	8.777 ± 1.841	2.194	1.932	0.055
Total satisfaction score	63.72 ± 11.449		58.824 ± 8.532		4.456	0.000
Average score of satisfaction	10.620 ± 1.908		9.804 ± 1.422		0.742	0.000

Table 9 *f* test of student satisfaction in third-class schools in Zhejiang province

Province	Zhejiang province					
Name of dimension	Provincial demonstration type ① *M* ± *SD*	Common type ② *M* ± *SD*	Accommodating type ③ *M* ± *SD*	*f*	*p*	Pairwise comparison
Value assessment	14.140 ± 2.321	11.795 ± 1.375	7.290 ± 2.160	608.557***	0.000	①>②>③
Content analysis	11.275 ± 1.370	10.435 ± 1.676	7.010 ± 1.572	427.693***	0.000	①>②>③
Cognitive grade	14.640 ± 1.456	10.140 ± 1.459	7.930 ± 1.488	1084.657***	0.000	①>②>③
Participation level	10.530 ± 1.546	10.785 ± 1.698	13.430 ± 2.183	154.045***	0.000	②<③①<③
Credit rating	10.355 ± 1.283	8.420 ± 1.544	9.185 ± 1.941	73.034***	0.000	①<② ②>③
Action level	8.315 ± 1.167	8.290 ± 1.158	8.060 ± 1.568	2.296***	0.000	
Total satisfaction score	69.255 ± 4.018	59.865 ± 3.6884	52.905 ± 4.657	785.105***	0.102	①>②>③

Table 10 *f* test of student satisfaction in third-class schools in Jiangsu province

Province	Jiangsu province					
Name of dimension	Provincial demonstration type ① *M* ± *SD*	Common type ② *M* ± *SD*	Accommodating type ③ *M* ± *SD*	*f*	*p*	Pairwise comparison
Value assessment	11.055 ± 2.492	9.730 ± 1.812	8.050 ± 1.532	114.877	0.000	①>②>③
Content analysis	11.310 ± 1.467	8.160 ± 1.491	8.020 ± 1.710	284.2967	0.000	①>②
Cognitive grade	14.495 ± 1.378	11.100 ± 1.710	8.010 ± 1.385	936.1387	0.000	①>② ②<③
Participation level	14.335 ± 1.977	8.325 ± 1.922	12.630 ± 2.441	424.0757	0.000	①>② ②<③
Credit rating	8.920 ± 1.419	10.595 ± 1.540	8.525 ± 1.428	112.7687	0.000	①<② ②>③
Action level	8.410 ± 1.511	8.600 ± 1.594	8.430 ± 1.892	0.7787	0.460	
Total satisfaction score	68.525 ± 4.402	56.510 ± 3.837	53.665 ± 4.586	676.861		①>②>③

Comparative analysis of satisfaction degree of different types of high school teachers

Table 11 *f* test of teacher satisfaction in third-class schools in Zhejiang province

Province	Zhejiang province					
Name of dimension	Provincial demonstration type ① *M* ± *SD*	Common type ② *M* ± *SD*	Accommodating type ③ *M* ± *SD*	*f*	*p*	Pairwise comparison
Value assessment	14.660 ± 2.054	8.717 ± 1.511	8.807 ± 1.442	202.615	0.000	①>②
Content analysis	7.820 ± 1.814	7.603 ± 1.418	7.846 ± 1.742	0.335	0.716	①>③
Cognitive grade	11.100 ± 1.515	10.792 ± 1.894	7.076 ± 1.081	109.810	0.000	①>③
Participation level	10.640 ± 1.561	8.320 ± 1.626	7.326 ± 1.451	61.225	0.000	①>②>③
Credit rating	13.980 ± 1.743	8.264 ± 1.388	10.403 ± 1.774	158.471	0.000	①>② ②<③
Action level	10.000 ± 1.385	7.832 ± 1.565	7.884 ± 1.338	37.810	0.000	①>②>③
Total satisfaction score	68.200 ± 3.917	51.528 ± 3.495	49.342 ± 3.054	440.557	0.000	①>②
Average score of satisfaction	11.366 ± 0.652	8.588 ± 0.582	8.224 ± 0.509	73.426	0.000	

Table 12 *f* test of student satisfaction in three types of schools in Jiangsu province

Province	Jiangsu province					
Name of dimension	Provincial demonstration type ① *M* ± *SD*	Common type ② *M* ± *SD*	Accommodating type ③ *M* ± *SD*	*f*	*p*	Pairwise comparison
Value assessment	15.860 ± 1.829	9.056 ± 1.433	8.365 ± 1.692	317.563	0.000	①>②>③
Content analysis	9.000 ± 1.714	13.528 ± 2.606	14.384 ± 1.931	93.958	0.000	①<②<③
Cognitive grade	10.740 ± 1.468	8.584 ± 2.239	7.153 ± 1.289	55.838	0.000	①>②>③
Participation level	11.200 ± 1.628	7.094 ± 1.078	10.365 ± 1.749	107.475	0.000	①>②, ②<③
Credit rating	14.640 ± 2.388	10.509 ± 1.395	7.615 ± 1.598	188.612	0.000	①>②>③
Action level	7.0800 ± 1.428	7.886 ± 1.250	7.884 ± 1.338	0.295	0.745	
Total satisfaction score	68.520 ± 4.162	56.657 ± 3.872	55.766 ± 3.615	193.924		①>②, ①>③
Average score of satisfaction	11.42 ± 0.693	9.443 ± 0.645	9.294 ± 0.602	32.320		

In the process of statistical data analysis to test the degree of satisfaction of different types of high school teachers with the actual implementation of the comprehensive quality assessment, we conducted the relevant single-factor analysis of variance and LSD for the satisfaction values of teachers in two provinces and three types of schools. From the analysis results, it can be concluded that the total scores of the satisfaction values of teachers in two provinces and three types of schools were insignificant ($p > 0.05$). Their average satisfaction with comprehensive quality assessment was low, indicating that teachers in all types of schools were not satisfied with the state and level of assessment in the implementation process.

In the dimension of value assessment, the satisfaction values of teachers in the three types of schools in the Zhejiang province were statistically significant ($f = 114.877$, $p < 0.05$), and the performance was ①>②. The satisfaction values of teachers in the provincial model and general type schools were also significant, and the satisfaction values of teachers in general type and accommodation type, provincial model and accommodation type were not significant after comparison. In the dimension of cognitive grade, there was no significant difference in the satisfaction of teachers in the three types of schools ($f = 109.810$, $p > 0.05$), indicating that teachers' understanding of the content of comprehensive quality assessment was insufficient. In the dimension of content analysis, the satisfaction of teachers in provincial demonstration and accommodation schools was significant ($f = 0.335$, $p > 0.05$), which was ①>③. In the dimension of valuing the level of participation, the satisfaction of teachers in three types of schools in the Zhejiang province had significant differences ($f = 61.225$, $p > 0.05$), and the performance was ①>②>③. In the credit rating, the satisfaction value of teachers also had a significant difference ($f = 158.471$, $p > 0.05$), which was ①>②>③. In the level of action ($f = 37.810$, $p > 0.05$), there was a difference in the teachers' satisfaction between provincial demonstration schools and common type schools, but there was no significant difference between teachers' satisfaction in the two types of schools, which was ①>② (see Tables 6–12).

Table 13 *f* test of satisfaction degree of school leaders in third-class schools in Zhejiang province

Province	Zhejiang province					
Name of dimension	Provincial demonstration type ① *M* ± *SD*	Common type ② *M* ± *SD*	Accommodating type ③ *M* ± *SD*	*f*	*p*	Pairwise comparison
Value assessment	12.666 ± 2.259	13.667 ± 2.256	7.883 ± 1.678	78.992	0.000	①>③ ②>③
Content analysis	12.800 ± 2.461	10.083 ± 1.429	10.383 ± 1.427	26.309	0.000	①>② ②<③
Cognitive grade	13.950 ± 1.545	10.916 ± 1.393	6.866 ± 1.346	245.367	0.000	①>②>③
Participation level	10.450 ± 2.235	13.866 ± 2.396	9.233 ± 2.028	46.602	0.000	①<②>③
Credit rating	13.116 ± 1.541	11.250 ± 1.536	6.916 ± 1.429	178.805	0.000	①>②>③
Action level	10.833 ± 1.392	8.933 ± 1.821	7.850 ± 1.773	32.644	0.000	①>②>③
Total satisfaction score	73.815 ± 4.827	68.715 ± 3.983	49.131 ± 4.236	527.728	0.000	①>②>③
Average score of satisfaction	12.302 ± 0.804	11.452 ± 0.663	8.188 ± 0.706	87.954	0.000	

Table 14 *f* test of student satisfaction in third-class schools in Jiangsu province

Province	Jiangsu province					
Name of dimension	Provincial demonstration type ① *M* ± *SD*	Common type ② *M* ± *SD*	Accommodating type ③ *M* ± *SD*	*f*	*p*	Pairwise comparison
Value assessment	14.233 ± 1.711	14.200 ± 1.911	8.316 ± 2.190	121.846	0.000	①>③ ②>③
Content analysis	8.083 ± 1.305	10.666 ± 1.580	7.383 ± 1.595	53.222	0.000	①<②②>③
Cognitive grade	13.433 ± 1.788	6.216 ± 1.728	6.933 ± 1.990	187.747	0.000	①<②②>③
Participation level	10.716 ± 1.303	12.650 ± 1.715	13.483 ± 2.494	22.342	0.000	①>②②<③
Credit rating	14.916 ± 1.576	8.516 ± 2.213	7.566 ± 1.681	188.604	0.000	①>②>③
Action level	8.750 ± 2.038	9.116 ± 1.832	8.466 ± 1.599	3.664	0.014	
Total satisfaction score	70.131 ± 3.698	61.364 ± 4.573	52.147 ± 4.671	258.015	0.000	①>②>③
Average score of satisfaction	11.688 ± 0.616	10.277 ± 0.762	8.691 ± 0.788	43.002	0.000	

Through analysing and testing the data related to the satisfaction of school leaders of different types of high schools in two provinces to test the satisfaction of school leaders on actual implementation of comprehensive quality assessment, we respectively conducted a one-way variance analysis and a further post-factor test of the relevant satisfaction values. From the data, it can be concluded that in the two provinces the overall satisfaction of the administrators on the assessment of the implementation of provincial demonstration high schools was high, but the overall satisfaction of the integrated schools was relatively low, showing a significant state.

The analysis shows that in the dimension of value assessment, the satisfaction values of school leaders of the provincial model, common type and accommodating schools in the Zhejiang province were significant ($f = 78.992$, $p < 0.05$) and show that ①>③ ②>③. The satisfaction values of school leaders of provincial model and ordinary schools were not significant in the two comparisons. In the dimension of content analysis, all three were significant ($f = 26.309$, $p < 0.05$), and show that ①>② ②<③, and the satisfaction values of school leaders of provincial model and ordinary schools were not significant in the two comparisons. In the dimension of content analysis, all three were significant ($f = 26.309$, $p < 0.05$). In the comparison of ①>② ②<③, the satisfaction values of school leaders in the provincial model and ordinary schools were not significant in the two comparisons, indicating that school leaders in the provincial model and ordinary schools did not rate the value of the assessment itself in the new college entrance examination reform. In the dimension of content analysis ($f = 26.309$, $p < 0.05$), it shows that ①>② ②<③. The satisfaction values of school leaders in ordinary schools in the Zhejiang province were not significant in the two comparisons. In the dimension of cognitive grade ($f = 245.367$, $p < 0.05$), it shows that, all three were significant (①>②>③); in the dimension of participation level ($f = 46.602$, $p < 0.05$), ①<②>③, the satisfaction value of school leaders in common type schools was the most significant. In the dimension of participation level, the satisfaction value of school leaders in common type schools in the Zhejiang province was the most significant. In the dimension of participation level, the satisfaction values of school leaders of common type and provincial model schools in the Zhejiang province showed higher values than the accommodating type, indicating a higher degree of importance to assessment. However, in a two-by-two comparison, the satisfaction values of school leaders of provincial model schools were higher than the other two types of schools, and all three categories showed significant status. In the dimension of credit level ($f = 178.805$, $p < 0.05$), all three showed significance as ①>②>③. In the dimension of action level ($f = 32.644$, $p < 0.05$), it shows ①>②>③, which is also significant (see Tables 13–14).

Table 15 Paired sample *t*-test for students

		Comparison of dimensions of different types of schools in Zhejiang province and Jiangsu province			
		Zhejiang province	Jiangsu province		
	Name of dimension	*M* ± *SD*	*M* ± *SD*	*t*	*p*
Provincial demonstration type	Value assessment	14.140 ± 2.321	11.055 ± 2.492	14.084	0.000
	Content analysis	11.275 ± 1.370	11.310 ± 1.467	-0.231	0.818
	Cognitive grade	14.640 ± 1.456	14.495 ± 1.378	1.025	0.307
	Participation level	10.530 ± 1.546	14.335 ± 1.977	-21.189	0.000
	Credit rating	10.355 ± 1.283	8.920 ± 1.419	9.978	0.000
	Action level	8.315 ± 1.167	8.410 ± 1.511	-0.678	0.499
	The total score of the degree of recognition	69.255 ± 4.018	68.525 ± 4.402	2.086	0.000
Common type	Value assessment	11.795 ± 1.375	9.730 ± 1.812	13.479	0.000
	Content analysis	10.435 ± 1.676	8.160 ± 1.491	13.236	0.000
	Cognitive grade	10.140 ± 1.459	11.100 ± 1.710	-5.951	0.000
	Participation level	10.785 ± 1.698	8.325 ± 1.922	14.041	0.000
	Credit rating	8.420 ± 1.544	10.595 ± 1.540	-14.377	0.000
	Action level	8.290 ± 1.158	8.600 ± 1.594	-2.288	0.023
	The total score of the degree of recognition	59.865 ± 3.688	56.510 ± 3.837	8.707	0.000
Accommodating type	Value assessment	7.290 ± 2.160	8.050 ± 1.532	-3.764	0.000
	Content analysis	7.010 ± 1.572	8.020 ± 1.710	-6.739	0.000
	Cognitive grade	7.930 ± 1.488	8.010 ± 1.385	-0.547	0.585
	Participation level	13.430 ± 2.183	12.630 ± 2.441	3.687	0.000
	Credit rating	9.185 ± 1.941	8.525 ± 1.428	3.926	0.000
	Action level	8.060 ± 1.568	8.430 ± 1.892	-2.053	0.000
	The total score of the degree of recognition	52.905 ± 4.657	53.665 ± 4.586	-1.724	0.09

Table 16 Paired sample *t*-test for teachers

Name of dimension		Comparison of dimensions of different types of schools in Zhejiang province and Jiangsu province			
		Zhejiang province	Jiangsu province		
		M ± *SD*	*M* ± *SD*	*t*	*p*
Provincial demonstration type	Value assessment	14.660 ± 2.054	15.860 ± 1.829	-2.788	0.008
	Content analysis	7.820 ± 1.814	9.000 ± 1.714	-3.759	0.000
	Cognitive grade	11.100 ± 1.515	10.740 ± 1.468	1.214	0.231
	Participation level	10.640 ± 1.561	11.200 ± 1.628	-1.680	0.099
	Credit rating	13.980 ± 1.743	14.640 ± 2.388	-1.584	0.120
	Action level	10.000 ± 1.385	7.0800 ± 1.428	7.587	0.000
	The total score of the degree of recognition	68.200 ± 3.917	68.52 ± 4.162	1.241	0.221
Common type	Value assessment	8.717 ± 1.511	9.056 ± 1.433	-1.267	0.000
	Content analysis	7.603 ± 1.418	13.528 ± 2.606	-14.634	0.000
	Cognitive grade	10.792 ± 1.894	8.584 ± 2.239	5.128	0.000
	Participation level	8.320 ± 1.626	7.094 ± 1.078	5.008	0.000
	Credit rating	8.264 ± 1.388	10.509 ± 1.395	-8.467	0.000
	Action level	7.832 ± 1.565	7.886 ± 1.250	-0.202	0.841
	The total score of the degree of recognition	51.528 ± 3.495	56.657 ± 3.872	-10.393	0.000
Accommodating type	Value assessment	8.807 ± 1.442	8.365 ± 1.692	1.506	1.506
	Content analysis	7.846 ± 1.742	14.384 ± 1.931	-18.796	0.138
	Cognitive grade	7.076 ± 1.081	7.153 ± 1.289	-0.330	0.000
	Participation level	7.326 ± 1.451	10.365 ± 1.749	-9.489	0.742
	Credit rating	10.403 ± 1.774	7.615 ± 1.598	8.794	0.000
	Action level	7.884 ± 1.338	7.884 ± 1.338	0.837	0.000
	The total score of the degree of recognition	49.342 ± 3.054	55.766 ± 3.615	-7.090	0.407

Table 17 *T*-test for paired samples of school leaders

Name of dimension		Comparison of dimensions of different types of schools in Zhejiang province and Jiangsu province			
		Zhejiang province	Jiangsu province		
		M ± *SD*	*M* ± *SD*	*t*	*p*
Provincial demonstration type	Value assessment	12.666 ± 2.259	14.233 ± 1.711	-4.692	0.000
	Content analysis	12.800 ± 2.461	8.083 ± 1.305	11.942	0.000
	Cognitive grade	13.950 ± 1.545	13.433 ± 1.788	1.727	0.089
	Participation level	10.450 ± 2.235	10.716 ± 1.303	-0.853	0.397
	Credit rating	13.116 ± 1.541	14.916 ± 1.576	-6.359	0.000
	Action level	10.833 ± 1.392	8.750 ± 2.038	5.866	0.000
	The total score of the degree of recognition	73.815 ± 4.827	70.131 ± 3.698	4.508	0.000
Common type	Value assessment	13.667 ± 2.256	14.200 ± 1.911	-2.580	0.012
	Content analysis	10.083 ± 1.429	10.666 ± 1.580	-1.955	0.051
	Cognitive grade	10.916 ± 1.393	6.216 ± 1.728	18.659	0.000
	Participation level	13.866 ± 2.396	12.650 ± 1.715	3.317	0.002
	Credit rating	11.250 ± 1.536	8.516 ± 2.213	6.825	0.000
	Action level	8.933 ± 1.821	9.116 ± 1.832	-0.472	0.639
	The total score of the degree of recognition	68.715 ± 3.983	61.364 ± 4.573	8.015	0.000
Accommodating type	Value assessment	7.883 ± 1.678	8.316 ± 2.190	-1.314	0.194
	Content analysis	10.383 ± 1.427	7.383 ± 1.595	9.767	0.000
	Cognitive grade	6.866 ± 1.346	6.933 ± 1.990	-0.202	0.841
	Participation level	9.233 ± 2.028	13.483 ± 2.494	-9.146	0.000
	Credit rating	7.566 ± 1.681	6.916 ± 1.429	2.419	0.019
	Action level	8.466 ± 1.599	7.850 ± 1.773	1.972	0.053
	The total score of the degree of recognition	50.397 ± 4.236	50.881 ± 4.671	-3.835	0.000

Paired-sample *t*-tests were conducted on the value assessment, content analysis, cognitive grade level, participation level, credit level, and action level dimensions of each type of school for students in the Jiangsu and Zhejiang provinces. The results (see Table 15 for details) indicate that there was a significant difference between the Zhejiang and Jiangsu provinces in the total anterior and posterior side scores of the value assessment dimension of provincial model schools, and the scores of the Zhejiang province in the value assessment were significantly different than that of the Jiangsu province ([(14.140 ± 2.321) vs. (11.055 ± 2.492), t = 14.084, $p < 0.001$]). In the dimension of content analysis, the scores of the Jiangsu and Zhejiang provinces for the provincial model type schools were not significant, while the scores of the Jiangsu province in this dimension were higher than those of the former, as shown by ([11.275 ± 1.370] vs. [11.310 ± 1.467], t = -0.231, $p > 0.05$). In the cognitive level dimension ([14.640 ± 1.456] vs. [14.495 ± 1.378], t = 1.025, $p > 0.05$), the student satisfaction values of provincial model type schools in both provinces were close to a tie, and neither was significant. In the participation level, the provincial teacher type high schools in the Jiangsu province were significantly higher than those in the Zhejiang province ([10.530 ± 1.546] vs. [14.335 ± 1.977], t = -21.189, $p < 0.05$). In the credit level, students' satisfaction in provincial model high schools in the Zhejiang province were significantly higher than the latter ([10.355 ± 1.283] vs. [8.920 ± 1.419], t = 9.978, $p < 0.05$); while the opposite values were found in the action level dimension, where both were not significant ([8.315 ± 1.167] vs. [8.410 ± 1.511], t = -0.678, $p > 0.05$) and the satisfaction of students in provincial model high schools in both provinces was lower than the theoretical mean value.

In the comparison of the satisfaction values of students in the two provinces in the common type schools, students in the common type schools in the Zhejiang province had significantly higher scores in the dimension of value assessment than in the Jiangsu province ([11.795 ± 1.375] vs. [9.730 ± 1.812], t = 13.479, $p < 0.05$). In the dimension of content analysis, both provinces were equally significant and close to the theoretical mean, with the Zhejiang province significantly higher than the Jiangsu province ([10.435 ± 1.676] vs. [8.160 ± 1.491], t = 13.236, $p < 0.05$). In the dimension of cognitive level, the scores of students in common type schools in the Jiangsu province were significantly higher than those in the Zhejiang province ([10.140 ± 1.459] vs. [11.100 ± 1.710], t = -5.951, $p > 0.05$]). In the dimension of participation level, the scores of students in common type schools in the Jiangsu province were significantly lower than those in the Zhejiang province ([10.785 ± 1.698] vs [8.325 ± 1.922], t = 14.041, $p < 0.05$). In the dimension of credit rating, the satisfaction value of common type students in the Jiangsu province was significantly higher than that in the Zhejiang province ([8.420 ± 1.544] vs. [10.595 ± 1.540], t = -14.377, $p < 0.05$). In the dimension of action level, the scores were closer and insignificant, although the satisfaction values of students in general type schools in the Jiangsu province were slightly higher compared to those in the Zhejiang province; the scores of both provinces still show low values compared to the theoretical mean ([8.290 ± 1.158] vs [8.600 ± 1.594], t = -2.288, $p > 0.05$) (see Tables 16–17).

Results and Discussion

The development of comprehensive quality assessment in ordinary senior high schools has undergone several changes. The purpose was to evaluate the situation in education and training, which take scores as the only assessment criterion, and to guide students to a more thorough and diversified development path. However, after more than 10 years, the results of each reform have been unsatisfactory, making it difficult to function in the selection of talented individuals in the college recruitment process. In the context of the reform of the new college entrance examination, as an important part of the recruitment and admission process, we investigated and analysed the implementation of the new comprehensive quality assessment. The findings of the study reveal three types of problems that the experimental zones in the two provinces faced when implementing comprehensive quality assessment. Firstly, the overall satisfaction with the assessment was low among the three groups studied, but the implementation of the assessment in the provincial model high schools was relatively good. Secondly, the specific contents and standards of the assessment were still at a relatively uncertain stage, which lead role-players to have limited understanding and implementation of the comprehensive quality assessment. Thirdly, the assessment results of each school were not consistent with the assessment standards for university admission, which makes the influence of the assessment results weak and, therefore, makes it difficult to promote the development of students' diversified abilities.

Relatively Low Overall Satisfaction with Assessment Implementation Among Three Types of Survey Respondents

After analysing the satisfaction surveys and interviews held with students, teachers, and school leaders, it can be inferred that the overall satisfaction with the comprehensive quality evaluation in the Jiangsu and Zhejiang provinces was relatively low. This situation was specifically reflected in the following. In terms of the dimension of value assessment, both students and school leaders in the two provinces exhibited a favourable disposition towards the intrinsic worth of evaluation within the

context of the reform of college entrance examinations. The pre-implementation stage of the comprehensive quality assessment in the two provinces was, however, characterised by instability, resulting in low satisfaction among teachers and students. This instability indicates an early stage of development (Horn, 2006). Regarding content analysis it was found that although school administrators in the Zhejiang province consistently expressed satisfaction with the content of the evaluation, other survey participants showed lower levels of satisfaction. This suggests that both provinces still had a relatively unclear understanding of the cognitive concepts associated with the content, standards, and scope of the evaluation. The satisfaction with the cognitive rating dimension was evaluated based on both the expected value of satisfaction with the theoretical value and students' actual scores in high school entrance examinations, which impacted enrolment decisions and contributed to a satisfactory state of basic admissions. Regarding the participation level, the three categories did not exhibit a positive reflection and the overall satisfaction remained at a low level.

The Implementation of Provincial Model High Schools in Two Provinces was Relatively Effective

Comparative and differential analyses were conducted on three types of Chinese general high schools. The findings indicate that the majority of dimensions in the survey assessing evaluation implementation in provincial teacher training type high schools in both the Jiangsu and Zhejiang provinces showed a significant performance, as evidenced by *f*-tests with LSD post-hoc tests and two-pair tests. The evaluation and implementation of the provincial model high school and the other two types of schools have shown significant synthesis concerning effectiveness, indicating a favourable developmental trend. Specifically, in terms of the value assessment and content analysis dimensions, although all three types of high schools exhibited relative prominence in evaluations and demonstrated a consistent trend of high satisfaction among administrators, teachers, and students in each school, the development of general and inclusive schools still lacked equilibrium. Regarding the perceived rating, provincial model schools placed greater emphasis on evaluation and considered its implications carefully for students' future advancement compared to middle leaders and teachers in other types of schools. Additionally, students expressed higher satisfaction with the credibility of the evaluations themselves and demonstrated an increased trust in their outcomes (Mohono-Mahlatsi & Van Tonder, 2006). The common and accommodating high schools had relatively higher levels of student participation, despite the discrepancy in developing comprehensive competencies related to academic professionalism and overall quality compared to provincial teacher training high schools.

Content and Criteria of the Overall Evaluation of Comprehensive Quality were not Explicit Enough

The satisfaction of students, teachers, and school-level administrators with the content of the evaluation was relatively low when the different types of schools were compared. This indicates that role-players in the two provinces harboured doubts about the fairness of comparing evaluation results. The current focus in both provinces lies in the development of an online evaluation system to assess students' overall quality comprehensively. In the Zhejiang province, this initiative was implemented using the Entrance to the Information Management Platform for Comprehensive Quality Evaluation of the Zhejiang province's general high school students, while the Jiangsu province adopted the Application Platform For Comprehensive Quality Evaluation in the Jiangsu province (Smith, Julie & Gierdien, 2020). Although the two provinces exhibited distinct characteristics concerning dimensions and assessment methods on the platform from a value perspective, the Jiangsu province quantified the evaluation of exercise and health through scores while the Zhejiang province recorded students' development in this area through textual expressions based on content analysis (Smith et al., 2020). However, it was evident that due to the diverse approaches to presenting assessment outcomes, the process of evaluating students' materials across schools remained enigmatic, thereby impeding meaningful comparisons. In the long term, in the forthcoming admission process of the senior secondary school entrance examination, challenges will also arise in aligning the outcomes of comprehensive quality evaluation conducted by schools across different regions with the actual admission criteria (Ojo & Adu, 2018).

Recommendations and Conclusion

Recommendation 1: Perfect the Matching and Docking Standard of Comprehensive Quality Evaluation

The evaluation system should be designed to allow for regional disparities, as schools in different regions exhibit variations influenced by factors such as cultural diversity, geographical conditions, and levels of economic development (Pu, 2021). Therefore, the system of quality evaluation cannot fully develop a standardised evaluation criterion to assess students' comprehensive abilities, necessitating some consideration of regional disparities. Secondly, the design of the evaluation system should be tailored to the individuality of each student. The emphasis in the content of the assessment system should be on facilitating students' use of their strengths, rather than focusing solely on avoiding their weaknesses. Because it implies that the assessment of human development,

as previously mentioned, should prioritise comprehensive and unrestricted growth rather than pursuing a notion of balanced development in education (He, 2021). Thirdly, the formulation of the content of the evaluation system should also reinforce the school's distinctive features. This is because only a personalised school can cultivate individualised students who possess creativity, can transition smoothly from specialisation to specialisation, and ultimately achieve all-round development. In the fourth instance, the notion of evaluation necessitates continual updating to align with contemporary perspectives. It is imperative to perceive test-taking proficiency as an integral component of holistic development rather than a burdensome entity by default. The internal structure of the evaluation system needs to be adjustable, taking into account regional disparities and individual differences in education, while also reflecting its comprehensiveness and flexibility during implementation. The last aspect involves fostering a sense of creativity and innovation in the evaluation process during implementation. Different regions must achieve interactive information feedback in practice, rather than relying solely on mechanical upward and downward effects. Moreover, cultivating the evaluators' creative consciousness should be emphasised from various perspectives.

Recommendation 2: Improve the Usable Level of Comprehensive Quality Evaluation Functioning

For a considerable time, China has relied heavily on teachers to assume a pivotal role in executing the evaluation process, thereby granting them a certain level of authority in assessment. The teachers' subjectivity, such as personal preferences, easily influences their definition of students' judgements. This can manifest in the interpersonal halo effect, particularly when labelling certain students as good or bad based on specific incidents (Zhang, 2017). Consequently, it impacts the evaluation results directly. Teachers, being the primary evaluators, tend to generalise student evaluations and often employ vague language such as "studies hard and works actively", which hinders the reflection of individual differences in evaluation outcomes. It is evident that when the primary focus of evaluation shifts away from the student body, which is the main subject under assessment, the original purpose of evaluation becomes compromised. This factor has long impeded the widespread implementation of quality education. Therefore, to ensure that a comprehensive quality assessment plays an effective role, the principle of "who uses, who evaluates" should be firmly adhered to at all times. The main body of evaluation should include students' self-assessment, mutual evaluation among students, and even parental evaluation. By increasing the participation of all parties involved in the evaluation process, we can not only diversify the types of assessment but also promote fair and impartial evaluations. Simultaneously, students can assess their performance not only through declarative statements, such as written essays, but also by exploring alternative forms of expression such as poetry or prose, thereby enabling a more diversified evaluation process.

For high schools, there is a need to focus on the authenticity of the student assessment process and the reliability of the assessment materials. The evaluation of the school is mainly based on the students' self-statement reports and the discussion of relevance with other teachers. In this dominant evaluation system where responsibility is shared, students can learn to take responsibility for their own development and self-summary.

Moreover, each institution of higher education must prioritise the meticulous use of resources during the evaluation and admission process. As the direct beneficiaries of comprehensive quality assessment these institutions should shoulder corresponding responsibilities for enrolment sources and long-term institutional development.

Finally, high schools and institutions of higher education may also take on the responsibility of selecting talented individuals by adopting a system of evaluation by type, which can be tailored to the expertise they wish to recruit.

Recommendation 3: Improve the Function Level of the Comprehensive Quality Assessment Results

Staverbeam, an American educator, said that the most important intention of evaluation is not to prove, but to improve (Cheng, Z & Liu, 2016). The aim is to promote the development of students, emphasising them as subjects and their diversified, comprehensive, and integrated development. The purpose of evaluation is clearly to promote the all-round development of quality education (Horn, 2006) and we should focus equally on its developmental and selection functions, looking to find a balance between them.

Therefore, in schools in the provinces where the evaluation has been implemented, points can be assigned to the content of the comprehensive quality evaluation assessment and done without too high a value, aiming at steady progress (Yang, 2022). For regions yet to implement the new college entrance examination reform, it is advisable to establish a "soft link" between evaluation and college entrance examinations to foster a more ethical environment within society. The purpose of doing so is to help schools that have not yet implemented the evaluation to understand how the evaluation should be practical as soon as possible, and to help parents and students to gradually pay attention to the cultivation of children's comprehensive quality and ability in the process of understanding (Guan & Li, 2022).

Conclusion

Against the background of the new college entrance examination reform, we adopted a combination of empirical and qualitative research methods, aimed to reveal the actual implementation of comprehensive quality assessment in the Jiangsu and Zhejiang provinces from the perspective of implementing assessment. The research shows that the current problems with the implementation of the assessment in the two provinces were mainly in understanding the cognitive concepts related to the content, criteria, and scope of the assessment, which were still relatively unclear to the relevant role-players in the two provinces. Differences exist between the ability to cultivate different types of high school teachers and the development of comprehensive competencies in terms of academic professionalism and comprehensive qualities. The consistency of the results of the comprehensive qualities assessment in the schools of the different districts deviates from the actual problem with admission criteria and related issues encountered in the course of piloting the e-platform.

In the face of these problems and based on subjective education theory, multiple intelligence assessment theory, Marx's comprehensive development of human beings theory, developmental assessment theory, and other related educational theories, we put forward some suggestions for perfecting the matching and docking standards of comprehensive quality assessment to improve the availability of the assessment function level, and the function level of assessment results. It is hoped that these specific recommendations will provide a point of reference and assist other provinces that are about to implement the assessment.

Data Availability Statement

The original contributions presented in this study are included in the article/supplementary material; possible inquiries may be directed to the corresponding author.

Acknowledgements

Many thanks to Professor Ruan Chengwu for his long-term support and strong guidance. This research was supported by the National Social Science Fund project (CN) (Grant Number No.: 21BKS018).

Authors' Contributions

MNZ conceived and designed the analysis, collected the data, performed the analysis, and wrote the manuscript. CWR supported the data collection, reviewed the included studies, and assisted in data analysis. Both authors contributed to the article and approved the submitted version.

Conflict of Interest

We have no conflict of interest or competing interest to disclose. We confirm that it is our own work and it has not been published before and is not under simultaneous review elsewhere.

Notes

i. The term "high school" is used is used in short for ordinary high school.

iii. DATES: Received: 23 April 2023; Revised: 26 June 2024; Accepted: 26 September 2024; Published: 31 December 2024.

References

Cheng L 2017. The alienation of comprehensive quality assessment in high school and its solution strategies. *Journal of Guizhou Normal University*, 31:70–72.

Cheng Z 2023. Problems and optimization of regional promotion of comprehensive quality assessment for junior high school students. *Teaching and Management*, 1:67–69.

Cheng Z & Liu Z 2016. Evaluation of high school students' Subjectivity: Based on the perspective of subjectivity development theory. *Journal of the Chinese Society of Education*, 5:33–37+55.

Dong X 2020. Consensus and hidden concern in the implementation process of comprehensive quality evaluation. *Educational Development Research*, 40:28–41.

Du F 2017. *A study on the application of student comments in the comprehensive quality assessment of high school students*. Henan, China: Henan University.

Freeks FE 2015. The influence of role-players on the character-development and character-building of South African college students. *South African Journal of Education*, 35(3):Art. # 1086, 13 pages. https://doi.org/10.15700/saje.v35n3a1086

Gao B & Cheng R 2013. Research on the education in China. *Educational Research and Review (Secondary Education Teaching)*, 34:58–66.

Gao X 2013. *A study of student satisfaction with comprehensive quality assessment in high school*. Nanjing, China: Nanjing Normal University.

Guan X & Li H 2022. Regional promotion of comprehensive quality evaluation of ordinary senior high school students. *Journal of the Chinese Society of Education*, 6:105.

He L 2021. *Research on implementation problems and countermeasures of high school comprehensive quality evaluation based on PDCA cycle*. Changsha, China: University of Hunan.

Horn G 2006. Education solutions to improve the employability of senior high school learners. *South African Journal of Education*, 26(1):113–128. Available at https://www.sajournalofeducation.co.za/index.php/saje/article/view/64/69. Accessed 25 October 2024.

Huiying C 2011. The role of personalized comments and the method of writing them: An example of teachers' comments in the comprehensive quality

assessment of general high school students in Beijing. *Educational Measurement and Evaluation (Theory Edition)*, 1:32–35.

Kamyab A, Gholami T, Behdad K & Jeihooni AK 2023. An application of a series of theory-based educational intervention based on the health belief model on skin cancer prevention behaviors in female high school students. *Heliyon*, 9(6):e17209. https://doi.org/10.1016/j.heliyon.2023.e17209

Modiba MR & Sefotho MM 2019. Life orientation teacher training needs in career guidance at rural high schools. *South African Journal of Education*, 39(Suppl. 2):Art. #1538, 8 pages. https://doi.org/10.15700/saje.v39ns2a1538

Mohono-Mahlatsi L & Van Tonder F 2006. The effectiveness of mentoring in the Distance Teacher Education Programme at the Lesotho College of Education: Student teachers' and tutors' perceptions. *South African Journal of Education*, 26(3):383–396. Available at https://www.sajournalofeducation.co.za/index.php/saje/article/view/88/43. Accessed 25 October 2024.

Ojo OA & Adu EO 2018. The effectiveness of Information and Communication Technologies (ICTs) in teaching and learning in high schools in Eastern Cape Province. *South African Journal of Education*, 38(Suppl. 2):Art. #1483, 11 pages. https://doi.org/10.15700/saje.v38ns2a1483

Petalla MB & Tatlonghari JKC 2023. Preventive practices against fraudulent imitation of information on social media platforms among senior high school students. *Asian Journal of Education and Social Studies*, 38(2):8–14. https://doi.org/10.9734/ajess/2023/v38i2820

Pu R 2021. *Research on school support system of high school students comprehensive quality evaluation*. Shanghai, China: East China Normal University.

Shao M 2013. *Teacher resistance to comprehensive quality assessment of high school students and its dissolution*. Georgetown, TX: Southwestern University.

Smith CR, Julie C & Gierdien F 2020. The integration of semiotic resources and modalities in the teaching of geometry in a Grade 9 class in a South African high school: The four cases of congruency. *South African Journal of Education*, 40(2):Art. #1682, 12 pages. https://doi.org/10.15700/saje.v40n2a1682

U-Sayee CR & Adomako EB 2021. Supervisory practices and challenges faced by senior high school principals in Greater Monrovia, Liberia: Implications for quality education. *Heliyon*, 7:e06895. https://doi.org/10.1016/j.heliyon.2021.e06895

Wang H 2019. Comprehensive quality assessment for high school students: Misinterpretation and clarification. *Chinese Journal of Education*, 8:65–67.

Wen J & Song S 2022. Regional organization and management practice of general high school comprehensive quality evaluation. *Basic Education Curriculum*, 50(1):55–61.

Wu Z & Ji J 2022. Use comprehensive quality evaluation to leverage the reform of ordinary high school education mode. *People's Education*, 9:44–47.

Yanbing L 2016. The essence of comprehensive quality assessment. *Educational Research and Review (Secondary Education Teaching)*, 6:91.

Yang Y 2022. How to integrate the comprehensive quality evaluation of senior high schools with the college entrance examination in China—localization research of comprehensive quality evaluation in senior high schools from the perspective of Sino-American comparison. *Journal of Educational Research and Policies*, 4:45–48.

Yaqian Z & Yuqiang Z 2022. The dilemma and breakthrough of comprehensive quality evaluation in high school enrollment. *Educational Measurement and Evaluation*, 50(1):64–69.

Zhang Z 2017. *Research on comprehensive quality evaluation of ordinary high schools under the background of big data*. Shanghai, China: East China Normal University.

Zhao H & Meng W 2013. Comprehensive quality evaluation into the college entrance examination admission system implementation measures. *Hebei Education (General Edition)*, 17:42–43.

Zhenguo Y 2018. Exploration in reform context: Improving Gaokao system with Chinese characteristics. *Journal of East China Normal University (Educational Sciences)*, 36(3):1–12+166. https://doi.org/10.16382/j.cnki.1000-5560.2018.03.001