Proposing a Criterion- and standard-based Assessment Framework for Cantonese

Operatic Singing with the Application of the New Bloom's Taxonomy

by

Luo Yue

A Thesis Submitted to

The Education University of Hong Kong

in Partial Fulfillment of the Requirement for

the Degree of Doctor of Education

December 2021

Statement of Originality

I, Luo Yue, hereby declare that I am the sole author of the thesis and the material presented in this thesis is my original work except those indicated in the acknowledgement. I further declare that I have followed the University's policies and regulations on Academic Honesty, Copyright, and Plagiarism in writing the thesis and no material in this thesis has been submitted for a degree in this or other universities.

_____

Luo Yue (signed)

Abstract

Cantonese opera (known as *Yueju* 粵劇) is one of the Chinese music traditions that has been recorded on the UNESCO Representative List of Intangible Cultural Heritage of Humanities (2009), which is popular around the world and especially in its heartlands of south China: Guangdong and Guangxi Province, Hong Kong, and Macau. For decades, the transmission of Cantonese opera has faced considerable challenges within a rapidly developing contemporary society. Although teaching traditional art forms in schools and in the community has proven an effective way of ensuring the transmission of traditional culture, due to an absence of research in this specific area, the lack of valid assessment constrains both education and popularity among the younger generation. Based on a holistic theoretical framework stemmed from the revised Bloom's taxonomy and adopting a user-centered design approach, the primary purpose of this study is to propose a criterion- and standard-based assessment framework for teacher reference that supports the education and transmission of this traditional genre. Four research questions guide this study: 1) What are the design and practices of the existing graded examinations of Chinese opera? 2) For the sake of transmission and education, how can an ideal assessment framework for Cantonese operatic singing be formulated? 3) What is common practice in assessments of Cantonese operatic singing? and 4) To what extent does the proposed assessment framework facilitate the teaching and learning of Cantonese operatic singing?

The study consists of two phases. Phase I is divided into three parts. In the first part, the design and practices of two relevant assessment systems of *Xiqu* (戲曲 literally "theatre and

music") were analyzed, to identify common components, assessment focus, and potential issues in the two existing public *Xiqu* graded examinations. In the second part, six eminent Cantonese opera experts from Guangdong Province and Hong Kong were interviewed in a semi-structured setting to gather their views on common practice within Cantonese operatic singing assessment concerning the graded examinations. Another purpose of the semi-structured interview was to collect information to use in the formulation of an ideal assessment framework for Cantonese operatic singing concerning transmission and education. Finally, based on the findings of the previous two parts and the theoretical framework of the study, a criterion- and standard-based assessment framework is noted, from which assessment rubrics for four domains of Cantonese operatic singing are formulated. Phase II is an exploratory study comprised of a series of assessment trials and subsequent semi-structured interviews as an evaluation. Three teachers were invited to use the assessment rubrics to evaluate their students. Adopting stratified sampling, students from six categories: top students, average students, and underachievers in senior and junior grades were selected according to their daily academic performances to serve as assessment subjects in the exploratory study. Subsequently, another round of semi-structured interviews was conducted with each of the four participant teachers and 18 of the participant students to investigate to what extent the proposed assessment framework facilitated the teaching and learning of Cantonese operatic singing and what, if any, functional or practical problems remained to be solved.

The findings of contents analysis reflected that the design of the existing graded examinations in *Xiqu* is still developing. A common practice in the existing graded examinations of Chinese opera features: 1) dependence on subjective perception, rather than criterion- and standard-based assessment, 2) a focus on the assessment of performance skills and a lack of multidimensionality, and 3) an inclination to momentary judgment and a lack of

sustainability. Therefore, the thematic analysis in Phase I suggests that: 1) the proposed assessment framework should concentrate on student learning, 2) a criterion- and stand-based system should be employed, and 3) five supporting measures during the formulation. The findings of the thematic analysis in Phase II indicated that the characteristics of traditional assessments made during daily teaching and learning in Cantonese operatic singing are marked by: 1) opaqueness, 2) factitious scoring, 3) oral assessment, 4) incompleteness and non-normativity 5) generalization and implicitness, 6) a lack of objectivity and fairness, and 7) a lack of instructiveness. It was concluded that the new framework contributed to better assessments, which were 1) instructive and efficient 2) easy to start and user-friendly, 3) detailed and well-rounded, 4) normative and systematic, 5) clear, 6) documentable, 7) objective and fair, 8) transparent, 8) focused on process as well as scores, and 9) convincing.

By establishing a criterion- and evidence-based assessment framework, this study strives to set a blueprint for concrete assessment settings in the education and transmission of Cantonese operatic singing that helps ensure consistency, conservation, and vitality within an increasingly accountability-driven educational environment and the assessment-oriented world, which may further contribute to the inheritance, popularization, and development of this traditional art form. The evidence-based assessment may also help develop the concepts of teachers who have previously employed abstract assessment and enhance the quality of traditional music education.

Keywords: Cantonese opera, assessment, transmission, Bloom's Taxonomy, Chinese *Xiqu*

Acknowledgements

It is a genuine pleasure to arrive at the proper moment to write these acknowledgements. This thesis and the EdD program would not have been possible without the support and help of many people. I would like to express my sincere gratitude to all of them.

Foremost, I owe my deepest gratitude to my principal supervisor, Prof. Leung Bo Wah, not only for his guidance through each stage of the process, but also for inspiring my interest in the transmission of Cantonese opera. I will never forget his insightful feedback, which pushed me to sharpen my thinking and brought my work to a higher level. I will never forget his fatherly advice, which broadened my horizons. And I will never forget his enthusiasm and devotion to education and Cantonese opera, which inspired my sense of duty as a teacher and as a young researcher.

Many thanks must also go to Dr Yang Yang, Dr Koji Matsunobu, and other mentors at The Education University of Hong Kong. Their expertise and guidance kept me on the right track in the EdD program. Their wisdom and virtue set an example for me. Their encouragement and assistance helped me overcome difficulties that were beyond me.

And I am forever thankful for my husband, Dr Hang Su, who is my love and soulmate. Each time I was assailed by self-doubt or cried in low spirits, almost defeated by fatigue, he encouraged me to battle through. I am fortunate to have had his unstinting unconditional love and support throughout the entire thesis process. He was always there, wherever and whenever that was.

Incredibly special thanks to my baby for growing into a healthy little one inside me, even though sometimes I did not get enough rest for him while I drafted this thesis. He unknowingly experienced the process of thesis writing and my evolution into an EdD after his birth. He and this thesis are my adorable creations. I hope that one day he has enough interest to read through this thesis, which was born at nearly the same time as he was, so he can understand the effort we made together.

Last but not least, I would also like to acknowledge my other family members who endured this lengthy process alongside me, always offering support, understanding, and love. They deserve endless gratitude.

Table of Content

LIST OF ABBREVIATION

COSGEHK: Cantonese Operatic Singing Graded Examination in Hong Kong

HKAPA: Hong Kong Academy for Performing Arts

POGE: Peking Opera Graded Examination

HKACOS: Hong Kong Association of Cantonese Opera Scholars

COAHK: Cantonese Opera Academy of Hong Kong

CAAHK: Chinese Artist Association of Hong Kong

UWL: the University of West London

LIST OF FIGURES

List of Tables

CHAPTER 1

INTRODUCTION

Cantonese opera is one of the Chinese music traditions recorded on the UNESCO Representative List of Intangible Cultural Heritage of Humanities (2009). Like other Chinese theatre genres, Cantonese opera is a composite art form that features formulaic and distancing from reality. By singing, reciting, acting movement, and acrobatic fighting, Cantonese opera interprets stories with the synthetic effects of make-up, costumes, props, music. Most of the plots stem from Chinese history and famous Chinese classics and myths, which reflect China's traditional culture and philosophies. Cantonese opera combines literature, drama, *quyi* (曲藝), dance, and martial arts. The rich artistic value lies in its *changqiang* (唱腔), *hongdang* (行当), make-up, gestures, and *Gongche pu*（工车谱）. Apart from being a theatrical performing art, Cantonese opera is also of great social and cultural value as a ritual performance for religious activities.

1.1 Historical development of Cantonese opera

Researchers variously trace the history of Cantonese opera back to the Southern Song Dynasty (1127~1279) (Chan, 1999) or Ming Dynasty (1638~1644) (Leung, 1982). Cantonese opera experienced a series of evolutions before its unique characteristics emerged in the 1920s and 1930s. It originated in Foshan, matured in Guangzhou, and developed in Hong Kong (Wang & Liang, 1990).

During the late nineteenth and early twentieth centuries, with the migration of Cantonese colonies, Cantonese opera diffused into parts of Southeast Asia that included Indonesia, the Philippines, Vietnam, Thailand, and the Malayan Peninsula and Singapore, as well as into North America (Yung, 1989) and Europe (Leung, 1982).

But, as Guangdong and Hong Kong are the two most important arenas of this traditional genre, the narration of Cantonese opera development and the case studies of this research will focus on these two areas. A map of Hong Kong and Guangdong Province, including Foshan and Guangzhou, is given below.

Figure 1.1. Map of Guangdong Province and Hong Kong (d-maps, n.d.).

When tracing its history through the modern period, it can be argued that Cantonese opera reached its peak in the 1920s. More repertoire and diversified singing styles were created, and a large number of famous masters and sects emerged. At the same time, it absorbed the skills of various local operas and the martial arts of the Northern School (北派), which made the performing art of Cantonese opera more colourful and mature. The innovation of stage art and scene-setting were implemented, and the vocals, light, electricity, and other technology were introduced into Cantonese opera, which increased the charm of Cantonese opera, making it a local opera that features Lingnan (岭南) characteristics. Since then, Cantonese opera has spread to thousands of households, from Guangzhou, the economic and cultural centre of the Lingnan region, and the Pearl River Delta, to Guangdong province, Guangxi, Hong Kong and Macao, Shanghai, Taiwan and other places where Cantonese live, such as the United States, Canada and Southeast Asia (Lai, 2001). Hong Kong, which was once an island, grew into a flourishing metropolis after being ceded by Qing China and became a colony of the British Empire in 1842. The enormous number of residents needed mass entertainment. Perhaps excepting most of the city's Western denizens, Hong Kong residents shared an interest in Cantonese opera with the rest of Guangdong Province (Lai, 2010). Indeed, Cantonese opera became more and more popular in Hong Kong after the Qing Dynasty. Increasing numbers of large theatrical troupes from Guangzhou performed in public in Hong Kong. A type of large-scale theatrical troupe, named *Shenggang ban* (省港班), arose (Lai, 2010, p. 298). The word "s*heng* (省)" refers to the capital of Guangdong Province, namely Guangzhou, while "gang (港)" refers to Hong Kong. Therefore, this denotation is a mirror of Cantonese opera's two main centers: Guangzhou and Hong Kong.

The boom in Cantonese opera meant many modern theatres were built by theatrical companies - both in Guangzhou and Hong Kong (Lai, 2010). Theatrical troupes performed almost daily for two decades (Leung, 1982). Apart from the construction of theatres, theatrical companies also administered the business of performances, both in Guangzhou and Hong Kong, which forced the theatrical troupes to travel between the two regions. As the monopolization of theatrical companies increased, the markets of these two regions were unified and consolidated, marking a prosperous atmosphere (Lai, 2010).

However, Cantonese opera began to decline after 1928 because of the rise of movies and remained depressed until 1932 (Wang & Liang, 1990). The decline aggravated after the 1940s due to complex political and social factors (Yung, 1989). After the Cultural Revolution in mainland China and economic depression and community instability in Hong Kong, Cantonese opera continued an overall decline into the late 1960s and early 1970s (Schenk, 2009). Audience numbers dropped markedly, and practitioners either quit the profession or emigrated to other countries, thereby causing a decline in the number of performers (Leung, 2015b).

The base of Cantonese opera lay in Guangzhou before Japan occupied it in 1938. But, after World War II, with the transfer of populations, artists, and troupes, the base of Cantonese opera shifted to Hong Kong (Lai, 2010).

As Guangzhou and Hong Kong's markets were originally homologous, the connection between these two regions resumed after WWII. However, this situation changed fundamentally before very long. The regime changes in mainland China and the establishment of the People's Republic of China in 1949 stimulated dramatic social and political transformations that led to the development of Cantonese opera bifurcated between Guangzhou

and Hong Kong (Ng, 2015): On one hand, the market-driven model of Cantonese opera administration in Guangzhou was transformed by the socialization of literature and art policy, which included changes to the capitalistic administration pattern of Cantonese opera troupes (Lau & Sin,1995). Cantonese opera in Guangzhou was remodeled according to the national "*Xiqu* reform" (戲劇改革), including innovations to scenarios, personnel management, and regimes (Lai, 2001). To promote artistic levels of Cantonese opera, the Cantonese Opera Company (廣東粵劇院) was established in 1958 and served as an official organization.

As Hong Kong was a colony ruled by the British, Cantonese opera in the city retained its original capitalistic administration patterns. It was market-orientated and catered to trends in social culture to ensure its transmission and development. For this reason, Cantonese opera in Hong Kong had to make significant efforts to seek innovation, including integration with cinema and popular culture (Lai, 2010).

1.2 Contextual background of the study

With the contributions of prominent artists from one generation to another, Cantonese opera experienced remarkable development and a rise in status (from being perceived as rough, unformed entertainment to being appreciated as fine art), performance places (from temporary bamboo theatres in rural areas to giant theatres in urban areas), content (from frequent improvisation to scenario-based performances), identity (from being seen as a non-native art form to being recognized as a local cultural expression), and transmission (from traditional apprenticeships to a contemporary conservatory tradition). With the globalization of education, the training of Cantonese opera artists and the transmission of the traditional opera genre was

significantly influenced by the western conservatory tradition (Leung, 2018b). Assessment of learning, traditionally conducted by the master, was highly subjective and informal (Leung, 2015b).

Although the conservatory tradition and the community training became the primary transmission and education modes of Cantonese operatic singing after the fading of apprenticeships in the mid-twentieth century (Leung, 2018a), assessment has been underused in this field and its reliability and validity have been arguable. By contrast, policy decisions surrounding academic programs have increasingly been made based on data stemming from objective standardized assessments (Porter, 2002; Linn, 2003). However, the quality of music education programs is difficult to evaluate due to the lack of precise quantitative measurements (Hanna, 2007). Asmus (1999) argued that "without standardized means of assessment, music programs will continue to struggle in their fight for academic legitimacy in a policy environment in which accountability matters" (p. 19). Following traditional existing and developed conventions, the education and transmission of Cantonese operatic singing often encounter challenges or constraints within the accountability-driven educational environment. This study was initiated within such a contextual background.

Traditionally, the teaching and learning of Cantonese opera are typically conducted in the professional diploma program and amateur training. The teaching and learning of Cantonese opera have unique features in Hong Kong and Guangdong Province due to various social environments, which will be discussed in the following subsections.

1.2.1 Teaching and learning of Cantonese opera in Hong Kong

Professional teaching and learning of Cantonese opera in Hong Kong arise from the Hong Kong Academy of Performing Art (HKAPA) and the Cantonese Opera Academy of Hong Kong (COAHK).

HKAPA offers three distinct levels of study programs:

a) Diploma in Cantonese Opera Foundations. This one-year full-time program offers solid foundation training in the Cantonese Opera art form to students from two specialized areas, including performance and music.

b) Advanced Diploma in Cantonese Opera. This two-year full-time program offers courses to those who have graduated from the Cantonese Opera Diploma program or those who hold an equivalent qualification in one of two specialized areas: performance and music.

c) Bachelor of Fine Arts (Honors) Degree in Chinese Opera. The Academy plans to launch Hong Kong's (perhaps the world's) first Cantonese Opera Degree as part of its efforts to preserve the traditional art form (HKAPA, n.d.). Besides various aspects that lead to professional excellence, this four-year full-time program will also offer a solid and balanced education to students that will prepare them to meet future challenges that include satisfying the ever-increasing sophistication of opera audiences, securing resources and patrons, meeting forces of innovation and development, and introducing the genre to the global arena. The academy believes that "apart from the continuous enhancement of their proficiency in Chinese and English and the use of information technology they will also benefit from liberal arts studies and both general and in-depth academic courses on Chinese music, Chinese opera, theory of western music and the performing arts, and

traditional and modern stage management" (HKAPA, n.d.). Accordingly, graduating students will be awarded a BFA in Chinese Opera (Cantonese Opera Performance) or BFA in Chinese Opera (Cantonese Opera Music) after completing a four-year full-time study program.

Unlike HKAPA, COAHK, a professional Cantonese opera institution attached to the Chinese Artist Association of Hong Kong (CAAHK), is responsible for professional education, social education, and school education. It provides not only professional training courses for each new generation of Cantonese opera actors and accompanists but also attaches great importance to promoting the inheritance of Cantonese opera art within school music education by cooperating with schools and universities (Luo, 2018). Aiming to cultivate new generations of talent for Cantonese opera and promote the sustainable development of Cantonese opera, COAHK restructured its programs in line with market conditions and needs in 2009. The main programs include: "Training Courses for Young Cantonese Opera Actors and Actresses", various in-service training programs for actors and actresses in collaboration with practitioners, and programs for Cantonese Opera arts in partnership with other secondary, primary, and tertiary institutions, including Po Leung Kuk Affiliated Primary Schools, Junior Secondary Classes of Po Leung Kuk Affiliated Secondary Schools, The Cantonese Opera Course at TWGHs Lui Yun Choy Memorial College, The Education University of Hong Kong, and Hong Kong University of Science and Technology (see Figure 1.2).
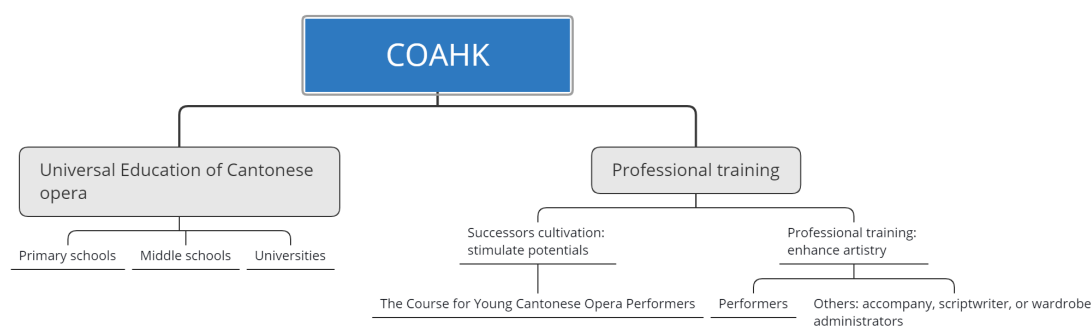
Figure 1.2 Curriculum of COAHK (CAAHK, n.d.)

Training Courses for Young Cantonese Opera Actors in COAHK feature formal study in class supplemented by an extracurricular internship. Aiming to cultivate a new generation for the Cantonese opera arts, this program enrolls teenagers aged 13 to 20. Senior actors teach formal courses in Cantonese opera circles, including singing skills, basic skills, gymnastic skills (毯子功 literally "skills on the mat"), pair practices in fighting with weapons (把子功), and performance skills. The timetables for the courses are fixed on the weekends. This four-year program includes six classes per week, which equates to 192 class hours throughout the year. Only students with attendance rates above 70% obtain graduating certificates. Diplomas are granted when the four-year program is complete. The students are required to take part in internships as an informal study outside of class time. Luo (2018) stated that, although the teaching and learning modes of Training Courses for Young Cantonese Opera Actors in COAHK have already transferred from apprenticeships to the conservatory tradition, oral teaching has been retained. Despite being a part-time program, this four-year program helps teenaged amateurs to perfect and professionalize their Cantonese opera performance skills through immersive study one day per week.

Cantonese opera arts courses collaborate between COAHK and different secondary, primary, and tertiary institutions and feature regular Cantonese opera classes supplemented by multifaceted Cantonese opera activities. For example, as the first Hong Kong middle school to bring Cantonese opera into its formal curriculum, TWGHs Lui Yun Choy Memorial College, is one of the collaborating schools using COAHK's courses. Professional artists conduct this program alongside teachers from the middle school. Apart from routine Cantonese opera courses, this program also encompasses informal learning. The goals of this program are more diverse than those of the Training Courses for Young Cantonese Opera Actors. Students can gain a solid foundation in singing, reciting, acting movement, and acrobatic fighting. Coordinated campus activities featuring Cantonese opera, such as character doll-making and opera mask design competitions, provides the wider student body with opportunities to be initiated into the traditional genre.

Community-based organizations are another channel for the transmission of Cantonese opera in Hong Kong. There are numerous associations, workshops, or clubs that offer Cantonese opera or Cantonese operatic singing practice. Modes of teaching, the curriculum, and pedagogy have undergone years of development in these organizations (Leung, 2018b). Leung (2018b) has noted the teaching methods in a case study. They can be categorized as 1) oral instruction, 2) expository teaching, 3) hands-on guidance, 4) demonstration and imitation, 5) repetitive practice, and 6) interactive teaching.

1.2.2 Teaching and learning Cantonese opera in Guangdong Province

The professional teaching and learning of Cantonese opera are mainly implemented by The Guangdong Cantonese Opera School, and aim to provide an all-round, systematic, and standardized professional training. Leung (2018b) noted the course features from six aspects:

a) The course content encompasses professional training in Cantonese opera and cultural classes.

b) Disciplined schedules are formulated.

c) Boarding systems are adopted. Only on weekends are the students permitted to leave school.

d) Different teaching progression is implemented in each semester throughout the six-year secondary education.

e) Standardized teaching is employed.

Similarly to Hong Kong, amateur teaching and learning of Cantonese opera in Guangdong Province is conducted in community-based organizations and collaborative extra-curricular activities between The Cantonese Opera Company and various levels of schools, namely "introduce Cantonese opera into the campus" (粵劇進校園).

However, according to Huang (2019), the collaborative extra-curricular activities in Hong Kong and Guangdong Province feature different characteristics, especially with respect to their course content and teaching methods. The teaching in Guangdong Province concentrates on performance practice that involves almost every performance skill in every *hongdong* (行當)

and includes singing, reciting, acting movement, and acrobatic fighting. Teaching in Hong Kong focuses on singing and reciting and makes light of acting movement and acrobatic fighting. The performance of *sheng* (生 male role) and *dan* (旦 female role) is at the core of the collaborative courses in Guangdong Province, and the popularization of other *hongdong* is undervalued. The course content in Guangdong Province features positive value orientations and is more age-appropriate to young students. Apart from typical and straightforward extracts, such as "Ode to Lychee", there are new plays that have been created especially for children, such as "The New Eight Immortals Crossing the Sea" (新八仙過海). By contrast, the content of collaborative courses in Hong Kong is almost always selected from classical extracts, such as *Princess Cheungping* (帝女花), which can be difficult for children to understand. *Gongche Pu* (工尺譜) and *Banghuang* (梆黃 *aria* type music) are subsumed into the content surrounding popularization. However, these are not in the study range of even professional students in Guangdong Province. The teaching in Guangdong Province advocates diligent study and strenuous training, and young students must endure rigorous and severe training in basic skills that are aimed at cultivating their persistence and hard-working spirit beyond the interest in Cantonese opera. Collaborative courses in Hong Kong are inclined to be comparatively relaxed, which not only respects and protects children's interest in Cantonese opera but also fully encourages initiative and creativity. Huang (2019) indicated that extra-curricular collaborative courses in Hong Kong belonged to a well-rounded education, whereas those in Guangdong Province are skills-oriented. She also argued that, compared with that in Hong Kong, the teaching of Cantonese opera in Guangdong Province takes on more responsibility for cultural propaganda.

1.3 Research gap in Cantonese operatic singing

Research into this traditional genre has focused on its history and the musical ontology, such as performing skills. Until comparatively recently, empirical study of its education has been rare. However, it is a consensus view that education contributes to transmitting this traditional genre to the next generation and cultivating new audiences, which in turn is vital to preserving, popularizing, and transmitting Cantonese opera within contemporary society (Lyu, 2018). Teaching traditional music within formal education is considered to be an effective way of protecting traditional culture (Leung, 2015a). Leung also emphasized that the institutionalization of traditional music had been a critical issue about transmission. Increasing attention has been drawn to the education of Cantonese opera and its effects on transmission, with teaching and learning modes becoming hot issues. However, little interest has been shown in assessment, which is an indispensable part of Cantonese opera education. In Cantonese opera, the assessment of learning has traditionally been conducted by a master (Leung, 2015a), which is apt to be subjective. Although there are various assessments utilized in the daily teaching and learning in conservatory tradition and community training, their design and practice have not been proven to be reliable and valid. This might undermine education and transmission or fail to accommodate an increasingly accountability-driven educational environment and assessment-oriented world. This study aims to fill this gap.

1.4 Purpose of this study

Although assessing musical performance is common across numerous types of music education practice, valid assessment of Cantonese operatic singing is scarce. Given the

significance of assessment in contemporary education, this study attempts to adopt the revised Bloom's Taxonomy as a theoretical framework and user-centered design approach to formulate a criterion- and standard-based assessment framework for Cantonese operatic singing using two phases. It aims to provide a foundation for assessment settings in the daily teaching and learning of this traditional genre.

1.5 Significance of this study

The proposed criterion- and standard-based assessment framework will guide an assessment formulation that contributes to 1) improving the quality of learning and teaching with formative and summative assessments; 2) examining student performance of learning outcomes; and 3) defining and maintaining education and academic standards.

This study contributes to three aspects:

a) Assessment in education: Establishing an assessment framework lays a foundation and setting for the assessment of Cantonese operatic singing, which will help standardize and normalize assessment in this traditional genre, so that the education of Cantonese opera can be carried out more efficiently and effectively.

b) Traditional transmission: An efficient and effective education in Cantonese opera may further facilitate this traditional art form's inheritance, popularization, and development.

c) Social implications: With the establishment and development of the Guangdong-Hong Kong-Macao Greater Bay Area, Cantonese opera has the chance to become one of its cultural highlights. Cantonese opera encompasses a great deal of Cantonese culture. Promoting and popularizing this traditional art may enhance regional identity.

1.6 Overview of the thesis

In Chapter 1, the context of the study has been introduced. A research gap has been identified, the research value has been stated, and the construction of the thesis has been outlined. In Chapter 2, the existing literature is reviewed to introduce the research object, delineate multifaceted assessment issues in music education, and present a theoretical framework derived from the revised Bloom's Taxonomy. In Chapter 3, the research topic, research questions, and research design are followed by detailed illumination of the adopted research methods. In Chapter 4 and Chapter 5, the research results of Phase I and Phase II are disclosed. A discussion of the significance and implications of this research is provided in Chapter 6.

CHAPTER 2

LITERATURE REVIEW

Unlike Western operas, Cantonese opera, as one of the Chinese *Xiqu*, has unique characteristics and systems across diverse aspects. When exploring a proper mode of assessment for this traditional genre, it is essential to preserve its features whilst digging for elements that can be measured and assessed. Thus, the first part of this chapter will focus on the transmission of Cantonese opera from an educational perspective. Another focus will be exploring the assessment of Cantonese operatic singing. As it is a subset of music education, a series of components and modes of assessment systems will be discussed in the second part of the chapter. In the final part, a revised Bloom's Taxonomy of Educational Objectives, forming the theoretical framework of this study, will be presented.

2.1 Introduction of the research object

Chinese opera, or *Xiqu* (戲曲), came into being during the *Yuan* Dynasty (1271~1368) (Mai, 1958). *Xiqu* is mainly derived from folk songs, dances, and acrobatics (Lai, 2010). *Xi* (戲) in China refers to drama and movement, acting, acrobatic fighting, and burlesque. *Qu* (曲), on a large scale, refers to singing with a close relationship to linguistic tones (Lau & Xin, 1995).

The appellations of Cantonese opera have come into being since the 1920s and 1930s. They are called *yueju*（粵劇）and *daxi*（大戲）in Chinese. The word *yue* refers to the abbreviation for Guangdong Province, the womb of Cantonese opera. *Daxi* literally means

"grand opera", which not only mirrors physical grandeur but also reflects its leading role when compared to other kinds of smaller-scale opera that flourished in the Pearl River Delta. The word "Cantonese" refers to Guangdong Province and the dialect used in this traditional genre.

Cantonese opera, which is a representative Chinese opera, has been popular in southern China (including Guangdong and Guangxi Province, Hong Kong, and Macau) for over 300 years and enjoyment of the art has even spread to North America, Latin America, and Southeast Asia (Lai & Huang, 1988). In 2009, the form was placed on the UNESCO Representative List of Intangible Cultural Heritage of Humanities (UNESCO, n.d.). However, the transmission of Cantonese opera has been facing considerable challenges within contemporary society. Although traditional teaching is often considered the most effective way to ensure the transmission of traditional culture (Leung, 2015a), the lack of consistent and effective assessment in the teaching of Cantonese opera has hampered its education and generalization amongst younger generations. This study adopts the Taxonomy of Educational Objectives as a theoretical framework to investigate educational objectives that support the assessment of Cantonese operatic singing performance within education, and thereby facilitate its inheritance, popularization, and development.

2.1.1. Cantonese opera and Cantonese operatic singing

Cantonese opera features four major skills (四功): singing (唱), reciting (念), acting movement (做), and acrobatic fighting (打). Unlike Italian opera, which uses singing, music, and acting as the major media on the stage, Cantonese opera delivers stories by reciting and acrobatic fighting, making it a comprehensive traditional genre.

Cantonese opera has a close relationship with Cantonese operatic singing. The perennial separation of Cantonese opera, Cantonese operatic singing, and Cantonese music is a folk tradition. Cantonese opera emphasizes Cantonese operatic singing, which attaches more importance to singing and reciting across the four major skills (Lai, 2001). The contrast between Cantonese opera and Cantonese operatic singing is illustrated in Table 2.1 (Chan, 1997, p. 202):

Table 2.1. The contrast between Cantonese opera and Cantonese operatic singing

	Cantonese opera		Cantonese operatic singing
	Ritual Performance	**Theatrical Performance**	
Performance venue	Rural outdoor mashed theatre	Urban indoor theatre	Urban indoor teahouse/theatre
Showtime	Seasonal and periodic (during traditional festivals and religious rites)	Throughout the year	
Audience sources	Geographical, religious, and patriarchal organization	Citizens from different classes	
Nature of works	Whole plays, including the development of plots. Excerpts from plays are rarely seen.		Single Cantonese operatic songs present scenes but are less focused on storytelling.
Artistic focus	Synthesis of visual and auditory senses, including singing, reciting, acting, and acrobatics		Focus on the auditory senses, featuring singing and instrumental accompaniment

Cantonese operatic singing is derived from Cantonese opera but became gradually independent from Cantonese opera as its artistic style took shape in the 1920s and 1930s. From the perspective of music educational research, this study will focus on Cantonese operatic singing because it has the closest relationship to music and music education.

2.1.2 Transmission and education of Cantonese opera

Traditionally, teaching and learning in Cantonese opera occurred within an apprenticeship system (Lau & Xin, 1995). The tradition of apprenticeship in Cantonese opera was popular in the early twentieth century in Hong Kong before it faded in popularity (Leung, 2015a). With the globalization of education that took place in the twentieth century, the transmission model of Cantonese opera moved from apprenticeship to conservatory training and community troupes.

2.1.2.1 Apprenticeship

Like some other folk arts in East Asia, Cantonese opera was transmitted using apprenticeships and an oral approach before the mid-twentieth century (Leung, 2015a). Traditionally, signing a seven-year apprentice contract was the first step to bonding the relationship between apprentice and master. A formalization of the ties featuring an apprenticing ceremony would typically be witnessed by other practitioners. From then on, the apprentice would live with the master like a family member and learn Cantonese opera by observing the master's performances, practising, and proactively questioning through the edification derived from sharing everyday life with the master.

The characteristics of the apprenticeship in Cantonese opera can be summarized as follows:

Quasi-parental relationship (San, 1993): As the old Chinese saying goes, "a day as a teacher, a lifetime as a father". This kind of relationship between master and apprentice is derived from Confucian ethics, which might be challenging to make sense of in the modern world (Leung, 2015a). Although the apprentice contract would normally expire after seven

years, the master- apprentice relationship would often be lifelong (Leung, 2018b). Leung also noted that familial intimacy existed in the ideal apprenticeship. Unlike modern schooling conventions, the apprentice had no obligation to pay tuition fees. On the contrary, the master would pay for the whole of the apprentice's daily expenses as if bringing up a child of his own.

Oral approach: According to Cai (2010), traditional repertoires are considered to be inherited by people rather than being transmitted through documentation.

Informal learning (Yuen, 1993): Since the master was normally a renowned performer, who had little time and energy to teach, apprentices were expected to learn proactively by observation, imitation, and questioning (Leung, 2015a). In addition, within this lengthy "internship" period, the apprentice learned both consciously and unconsciously.

Within the apprenticeship, informal learning was emphasized, and apprentices gained immediate feedback from both daily instructions from the master and audiences' reactions to their performances that might take place with or without their masters on stage. Using contemporary concepts, this can be characterized as a formative assessment method (Leung, 2015a).

2.1.2.2 Professional training in conservatory tradition

Apprenticeships featured slow learning progress, which failed to cater to the demand for quicker study. Therefore, it has gradually faded as a mode of instruction and a conservatory tradition that has been adapted from the West has been increasingly adopted since the mid-twentieth century (Leung, 2015a).

Nowadays, the academy is the main cradle of future professional artists of Cantonese opera - both in Hong Kong and mainland China. Founded in 1984, The Hong Kong Academy for Performing Arts launched what is probably the world's first Cantonese Opera Bachelor's Degree in 2013. And every year, there are about 10 students graduate from HKAPA (HKAPA, n.d.). Guangdong Dance and Drama College are one of the very few institutions nurturing professional Cantonese opera artists in mainland China. It was founded in 2012 and was based on its predecessor, the Cantonese Opera School, which was founded in 1958 (Guangdong Dance and Drama College, n.d.). Both academies adopt specialized instructional approaches.

The characteristics of professional training within the conservatory tradition in Cantonese opera can be summarized as follows:

a) Teacher-student relationship: Unlike the intimacy of apprenticeships, teaching and learning take place in the classroom and follow an instructional timetable and modern, ethical relations in schooling (Leung, 2015a). Moreover, as students learn from different teachers, they inherit collectively held skills, rather than the artistry of a single sect.

b) Institutionalized approaches: With the globalization of education that occurred in the twentieth century, the training of future artists has been considerably influenced by Western institutions. This has included systematic documentation, chronological grading, hierarchical structures, and integration with formal education systems such as undergraduate programs (Leung, 2015a).

c) Formal learning: Training time scopes are definite, teaching is structured, and the curriculum is divided into discreet subjects such as basic training, singing, gymnastic training, repertoire study, and martial arts. Specialists are employed to teach each

individual subject, which is introduced at certain stages across the learning process. Graduates are granted certification and even degrees after teaching requirements and achievements are met. Apart from the feedback received during instruction, assessment is summative and in the form of final examinations.

2.1.2.3 Community training and learning

One of the major crises that Cantonese opera presently faces is the discontinuity of inheritance (Yuen, 2008). If Cantonese opera is to be preserved and inherited, the cultivation of practitioners and audiences is of great importance. Audiences are the medium and impetus for the development of Cantonese opera. The broader the audience market is, the broader the development that can be achieved. The transmission and development of Cantonese opera depend on its audiences (Lai, 2010). Thus, practitioners in both Hong Kong and mainland China have made significant efforts to explore new ways to transmit this genre to the public, especially to younger generations.

In Hong Kong, community-based organizations were established to offer courses for the younger generations. Leung (2018b) identified the Hong Kong Children and Juvenile Chinese Opera Troupe, which was founded in 1993, as one such organization. These organizations have developed their own modes of teaching, curriculum, and pedagogy. However, a lack of assessment and certification is detrimental.

The transmission of Cantonese opera in Guangdong province is subject to similar challenges found in Hong Kong. Using a strategy of "inviting the audience into the Cantonese opera theatre and presenting a performance to the public", the transmission of Cantonese opera

has been stable and significant since 2010 (Zhang & Ma, 2018). Cultivating this traditional genre's potential audience and training future artists and amateurs have been critical to the substantial transmission of Cantonese opera.

Some of the characteristics of the above three transmission modes of Cantonese opera are summarized and illustrated in Table 2.2.

Table 2.2. A comparison among the practice of apprenticeship, community training, and conservatory tradition in Cantonese opera (Leung, 2018b, p. 14)

Apprenticeship	Community training	Conservatory
Enrolment		
Introduced by the master; limited ways of entry	Open enrolment	Open enrolment
Start at a very young age (no age limit)	Start at a very young age (no age limit)	Mainly admit early youths (approx. 11 years old)
Semi-involuntary learning (related to the economic environment)	Voluntary learning (related to parents)	Voluntary learning (related to the economic environment and personal factors)
Enrolment means entering the field	Learn it as a hobby	Enrolment means entering the field
Teaching mode		
Informal learning mainly	Non-formal learning mainly	Formal learning mainly
Learning by doing	Learning by doing	Learning by doing/theoretical courses related to culture
One master being the major role model of learning; supplemented/supported by other specialists	No specified teacher; specialised teaching	No specified teacher; specialised teaching
Harsh training; more traditional	High teaching standard	High teaching standard
Learning period not regulated	Flexible learning progress	Learning period regulated
Cultural norm		
Close master–apprentice relationship	Loose master–apprentice relationship	Loose master–apprentice relationship
Other than learning and performing, apprentices need to understand and observe the cultural environment and human relations	Focus on learning and performing; emphasise on respecting the teacher and his/her teaching	Focus on learning and performing
Performance and employment		
Participate directly in professional productions, learning while performing	Public performances available but not related to employment	Establish connection with the field; internship available but might not be responsible for students' future employment

Apprenticeship is the traditional mode of transmission and education of Cantonese opera, and this transmission form was popular in the early twentieth century in Hong Kong before it declined. With the globalization of education in the twentieth century, the transmission model of Cantonese opera transformed from apprenticeship to conservatory training and community troupes.

Traditionally, Cantonese opera was transmitted through informal settings which featured apprenticeship and aural approaches, and had the primary purpose of cultivating new professional artists. More recently, the transmission of Cantonese opera has shown a tendency towards Westernization and formalization (Leung, 2018b), which is most salient in schools and conservatories. The overview of Cantonese opera transmission in the context of contemporary music education is illustrated below:
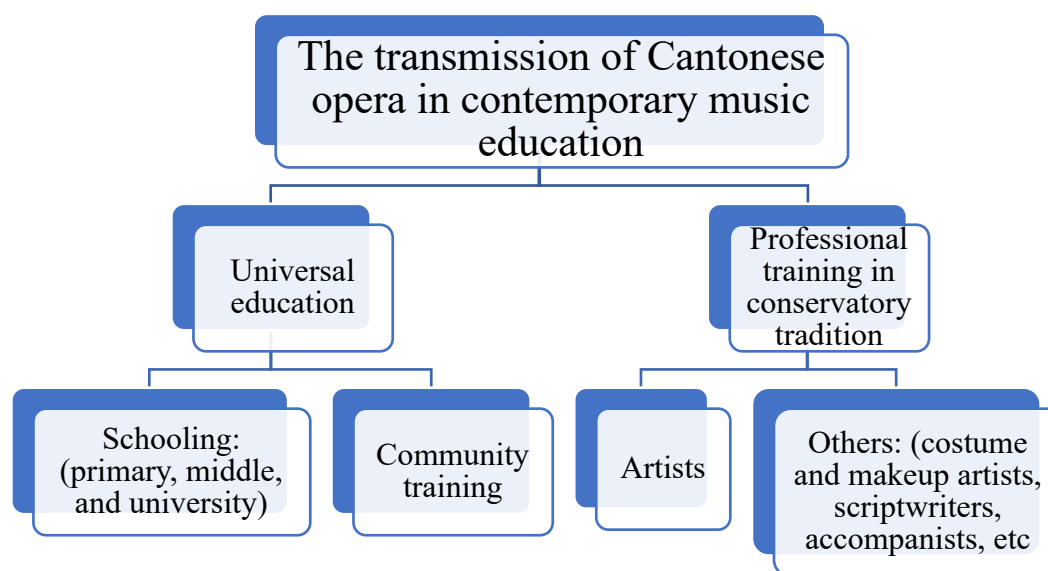
Figure 2.1 Overview of Cantonese opera transmission

Although the teaching and learning of traditional genres using Westernized approaches is commonly seen in many developing nations, critics have noted that Westernization may jeopardize traditional music (Nettl, 1985; Leung, 2018b). In mitigation, this study will respect the intrinsic features of Cantonese opera while adopting a contemporary theory and framework. Therefore, the features of Cantonese opera are not only elucidated upon in the first section of the literature review, but have also been used as a core value that guides this study.

Teaching and learning in formal settings, such as schools or conservatories, as well as contemporary methodologies, have been assimilated into the transmission of much traditional music (Leung, 2018b). Cantonese opera is no different.

Research into the transmission of traditional music overlaps with the disciplines of ethnomusicology and music education. Education, whether in formal, informal, or non-formal settings, is imperative to the preservation and transmission of traditional music (Leung, 2018b). Therefore, the next section of the literature review explores broader music education.

2.1.3 Assessment in Cantonese operatic singing

Assessment is tightly embedded within the learning process with curriculum and instruction (Brophy & Albert, 2008). About Cantonese opera, the limited research has focused on instruction and research into assessment is rare. After searching on the CNKI (China National Knowledge Infrastructure), a critical national research and information publishing institution in China, only one journal article and a master's thesis involved assessment in this field. Both were published in Chinese. The journal article only mentioned the assessment of an optional course delivered by South China Normal University, located in Guangzhou, "The

Appreciation of Art of Singing in Cantonese Opera". The master's thesis outlined three different models of assessment in Hong Kong, including the "Course for Young Cantonese Opera Performers" at the Cantonese Opera Academy of Hong Kong, which is attached to the Hong Kong Bahe Association, Cantonese opera art courses for junior high school students in Liu Yun Choy Memorial College, and an optional course named "Introduction to Cantonese Opera Art" at The Education University of Hong Kong.

In her journal article, Preliminary Study on the Construction of Cantonese Opera Appreciation Course for Professional Music Students in Normal University: A Case Study of Music College of South China Normal University, Wang (2019) argued that assessment in education should serve educational development and provide feedback beneficial to both student development and teacher instruction. However, the assessment was found to take the form of an essay, thought to be a conventional and one-sided assessment mode incapable of fully reflecting learning outcomes. Students are more inclined to demonstrate inquiry-based learning through performance on the stage. The author emphasized that a combination of practice and theory can be implemented within curricula and promote student development, and adapt to social needs. In addition, Wang (2019) mentioned that assessment in Cantonese opera was traditionally made by the master, who decided the extent to which an apprentice had learned. Traditional assessment was integrated with all four skills; singing, reciting, acting movement, and acrobatic fighting.

Luo (2018) outlined three different modes of assessment typically conducted in Hong Kong and provides more information about existing assessment practices. The Course for Young Cantonese opera Performers in the Cantonese Opera Academy of Hong Kong works to

cultivate new generations of Cantonese opera artists. Students who complete this four-year training course obtain an annual leaving certificate if their attendance rate does not slip below 70% within the year. No test is conducted. Mr Lui Hung-kwong, the program leader, explained that examinations were abolished, though they used to exist in the past. He attributed the abolition to three reasons: "Firstly, when the students enter the fourth year of the program, they are normally getting preparation for the matriculation examination. Thus, they might have to ask for a leave to take tutoring classes for other subjects or are occupied at school at the weekends, resulting in a considerable absence rate at the time of examinations. Secondly, since our teachers are all famous Cantonese opera artists with tight performance schedules, they might have to ask for leave even though the examination timetable is fixed. In sum, gathering teachers and students is tricky due to too many uncertainties. But the main problem is that certificates and examinations are less important in the Cantonese opera circles. What is valued most is whether you perform well or not. That is to say, whether be selected into a theatrical troupe or be popular with the audience is because of your actual performance level rather than your certificate or scores. But despite no examinations, both the teachers and the students know how one is doing from their performance during teaching and learning. The process is the most critical issue in this program. How much you gain depends on how much effort you have made. Especially there is a final show, which can be recognized as a test. As a result, the examination scores turn to be dispensable".

Luo (2018) indicated that for professional Cantonese opera performers in Hong Kong, real examinations are held in the market of Cantonese opera. One of the second-year students stated, "Normally we are also invited by some theatrical troupes to act in their play, by which we can

earn money—but acting as a protagonist or a walk-on all depend on your competence. Even in the performance within our school, important roles will not be given to one who is not good enough". This is why no criteria, standards, or examinations were set in the Course for Young Cantonese opera Performers - because to professional Cantonese opera performers, the "test" is in each performance rather than in the classroom and the criteria and the standards derive from their audience, instead of teachers.

Cantonese opera classes in junior high school aim to cultivate student interest in the art of Cantonese opera rather than training professional performers. Four aspects of educational functions that stem from Cantonese opera are fully exploited in the classes, including culture cultivation, art appreciation, kinetic training, and individual quality. The Cantonese opera classes in Liu Yun Choy Memorial College, the first school to introduce Cantonese opera into its formal curriculum, attach great importance to the student attendance to Cantonese opera related activities. Performances and assessments concentrate on the singing and reciting of Cantonese opera as examinations of music discipline.

In the final examination of singing and reciting, students work in pairs and sing a designated extract from Cantonese opera. The total possible score is 25, which is allocated equally into five dimensions equally: articulations, *changqiang*, pitch, rhythm, and attitude. The teachers make judgments and give comments based on these five dimensions. These comments usually feature euphemism and encouragement. For example, the comment for one of the students reads: "Inadequate progress with inaccurate pitch, which might be related with the music knowledge. More achievement can be expected with the effort to relevant music knowledge". Luo (2018) found that some students did not stop practising even during the class

break. One of the students told Luo (2018) that she spent much time practicing every night all week long for the examination, and she kept humming the melody even as she walked to class, which demonstrates the stimulation derived from educational assessment. Luo (2018) further indicated that, according to considerable experience and research, taking a test to record scores is proven to stimulate student motivation to learn. An examination is essential for junior school students with poor self-control or lack learning objectives. It contributes to familiarizing the students with the extracts taught in the class and enhances autonomic learning, thereby realizing the regulation of the control functions of educational assessment. Teachers can also learn more about their efficacy with reference to their student academic achievement and adjust the future educational design of Cantonese operatic singing. This facilitates the realization of the diagnosis of educational assessment.

The assessment approach of the Introduction to Cantonese Opera Art is as Table 2.3:

Table 2.3. The assessment approach of the Introduction to Cantonese Opera Art in The Education University of Hong Kong (Luo, 2018, p. 55)

Assessment Items	Weightings
1. The performance of movement and singing assessment in Cantonese opera	30%
2. Feedback report after watching a Cantonese opera (minimum wordcount: 1800 words): students are required to watch a Cantonese opera recommended by the mentor performed in a theater at one's own expense and write a feedback report no less than 1800 words. The students should discuss and analyze the performance effect and express personal opinions in-depth using the knowledge of Cantonese opera theory and skills learned in class.	40%
3. Written examinations in Cantonese opera theory, including short and long topics.	30%

Luo (2018) indicated that, although the students are not required to learn exactly how to perform Cantonese opera in the gesture performance and evaluation of *Xiqu*, the assessment

tasks provide opportunities for students to experience the four major skills of singing, reciting, acting movement, and acrobatic fighting. This enhances their perception of the performance art of Cantonese opera and combines theory with practice. The feedback report students write after watching Cantonese opera aims to examine the student competence towards the art of Cantonese opera and test whether the students can employ the Cantonese opera theory taught in the class. In addition to imparting knowledge, the focus of college courses is to improve critical thinking abilities by studying the theoretical knowledge of Cantonese opera art. The written examination of Cantonese opera theory evaluates the student mastery of Cantonese opera theory.

Luo (2018) stated that this course highlights the features of theoretical study from both the educational model and the assessment approach, which conforms to the college student physical and mental conditions. Thus, it enables students to absorb and internalize a significant amount of Cantonese opera theoretical knowledge within a brief period, maximally cultivating future audiences and stimulating student academic research potential at the same time.

After searching databases in English, only three journal articles about the assessment of Cantonese opera were found, all of which surrounded the Kinetic Chain Assessment and Learning computerized System for Cantonese Opera Movement. Only limited information about the outlines of the assessment can be extracted from these three articles. Mok et al. (2017) briefly mentioned that Cantonese opera is used to transfer through apprenticeship: "Masters of Cantonese opera teach their apprentices through demonstration and detailed one-to-one instruction, while the apprentices learn by observation, imitation, self-reflection, and practice. However, there is no standard curriculum or assessment for this kind of teaching and learning"

(p. 762). Leung (2018b) pointed out that apprenticeships have been replaced by the conservatory tradition since the mid-twentieth century. In this mode of transmission and education, assessments tend to be summative.

Up to now, the mainstream of research into Cantonese opera has focused on its history and musical ontology. Only a finite amount of literature can be found about its assessment, implying a research gap.

2.2 Assessment in music education

2.2.1. Overview of assessment in education

There is a prevalent misunderstanding that assessment simply means grading (Asmus, 1999). However, the assessment goes beyond that in today's educational world. This section will outline various aspects of assessment in education, teasing out helpful information for the current study.

2.2.1.1 Components of education

Stronge (2018) stressed that the alignment of curriculum, instruction, and assessment was of immense importance and integral to any instructional design of high quality. If these three critical educational components align with each other, the validity of assessment results is highly acceptable. If the three components are misaligned, this is not the case; the validity of assessment results is likely to be questioned. Even instruction of optimal quality will fail to achieve satisfactory assessment results. "When curriculum, instruction, and assessment are carefully aligned and working together, student learning is maximized" (p. 13). Alignment is "the deliberate linking of stated learning outcomes, teaching and learning activities, and

assessment tasks to promote consistency between what is learned and assessed" (Armstrong., 2008, p. 4). One key responsibility of educators is ensuring the alignment of learning outcomes with curriculum, instruction, and assessment (Biggs & Tang, 2007). This can be implemented by adopting the Taxonomy of Educational Objectives (see Section 2.3). Ramsden (2003) indicated that assessment "always defines the actual curriculum" (p. 182)". It is a "key driver of student activity" (Armstrong, 2008, p. 1). Thus, because of the correlation between assessment and other general education components, this study concentrates on assessing Cantonese operatic singing as a breakthrough point for further exploration of education in this field.

2.2.1.2 Definition of and correlations between assessment-related terms

Since the purpose of the current study is to propose a criterion- and standard-based assessment framework for Cantonese operatic singing, some common terms need to be defined. Their relationships need to be clarified before further discussion, for they are often used ambiguously and confusingly.

First of all, it is essential to make a clear distinction between assessment and evaluation, which is a common confusion. Assessment refers to the multifarious methods or tools used by educators to measure, evaluate, and record the academic achievements, skill acquisition, learning progress, or educational needs of students in education (Partnership, 2013). Assessment is an umbrella term that includes not only tests but also methods such as interviews, observation, and behavior monitoring. A test is considered to be a form of assessment. In other words, all tests are assessments, but not all assessments are tests (Kizlik, 2012).

Evaluation is "the process of making judgments about the level of student understanding or performance" (Marzano & Kendall, 2006, p. 126). Asmus (1999) argued that evaluation referred to "the collection and use of information to make informed educational decisions" (p. 24).

Making decisions in education is based on evaluation and there is an inextricable link between assessment and evaluation (Astin, 2012). Sound educational decision-making requires valid measurement, observation, and professional judgment (Hanna & Dettmer, 2004). A three-step process of this procedure is illustrated below (Hanna & Dettmer, 2004):

Assessment → Evaluation → Decision Making

Figure 2.2. Three-step Process of Educational Decision Making

Measurement means "assigning marks based on an explicit set of rules", whereas scoring and marking refer to "the numbers or letters assigned to assessments via the process of measurement" (Marzano & Kendall, 2006, p. 126). Asmus (1999) explained this term in the music context as "the use of systematic methodology to observe musical behaviors in order to represent the magnitude of performance capability, task completion, and concept attainment" (p. 21)

According to the definition, the logic of these terms can be teased out: assessment is the collection of data that can be used to make evaluations about students, whereas evaluations involve placement on a scale, measurement, using scores or marks (Marzano & Kendall, 2006).

Standards are "statements describing the level or quality of student performance in an assessment task" (Armstrong et al., 2008, p. 4). While criteria are "descriptive statements that provide learners and instructors with information about the qualities, characteristics, and aspects of a given learning task" (Royal Roads University, n.d.). Criteria and standards are both essential to assessment, but they are sometimes used interchangeably and often cause confusion (Sadler, 2005). Therefore, it is imperative to distinguish these two terms: Criteria often take the form of description, which can be implemented at various levels. Standards commonly take many forms and are written with brief but ample information about level and quality. Multiple assessors can use these comments to make reliable judgments of student achievement and effort within a specific assessment task. Although the description of standards usually encompasses criteria, criteria should not encompass standards (Armstrong, S., et al., 2008).

A rubric is commonly defined as a tool that articulates the expectations for an assignment by listing criteria, and describes levels of quality according to each criterion. (Andrade, 2000; Stiggins, 2001). Rubrics are useful tools that offer a basis for assessors to mark against and for students to aim at. Well-designed rubrics can also provide constructive feedback to students about their achievements or performances and indicate means of improvement (Armstrong, S., et al., 2008). It not only aids assessment of student work or academic achievement but also facilitates communication that provides students and teachers with an explicit outline of expectations by breaking down tasks into clear criteria and standards. Rubrics contain three

essential features: "evaluative criteria, quality definitions for those criteria at particular levels, and a scoring strategy" (Dawson, 2017. p. 3). Criteria are used to determine the level at which student work meets expectations. Comparability in scoring is achieved through the use of standardized rubrics, as well as training and moderation systems that enable scorers to use the same standards in consistent ways that result in reliable scores. Well-designed rubrics can mitigate student challenges and make the assessment task fairer and less subjective. Student prior awareness of how they will be marked and rated not only enhances learning but also encourages and guides self-direction (Hanna & Dettmer, 2004).

Assessment frameworks are structured conceptual maps that guide the design of the assessment and the content to be assessed (Jago 2009; National Assessment of Educational Progress, n.d.). This inevitably shapes assessment design (Wagemaker, 2020). Ainley & Schulz (2020) argued that "an assessment framework should be the main point of reference to understand how common elements of learning are defined and measured, and how comparable data are to be generated and analyzed" (p. 34).

2.1.1.3 Types of assessments in education

Educational assessment is both a qualitative and quantitative process of documenting student knowledge, skills, and attitudes (Iaquinta, 2013). There are several types of assessment in education. Knowing the types of assessment contributes to understanding how to make effective use of them. The most common are given below:

Formative and summative assessments are the most prevalent educational assessments. Formative assessment is used to "monitor student learning to provide ongoing feedback that

can be used by instructors to improve their teaching and by students to improve their learning" (Crook, 2001, p. 14). This is an assessment *for* learning. The purpose of summative assessments is to "evaluate student learning, academic achievement, and skill acquisition after a defined instructional period, such as at the end of a project, unit, course, semester, program, or school year" (Bhat & Bhat, 2019, p. 776). This is an assessment *of* learning. Students tend not to worry as much about formative assessments, because each assessment does not contribute much to their final grade, but summative assessments are often seen as particularly important (Burke, 2010).

Diagnostic assessments, also known as pre-assessments, are given at the beginning of an academic term to ascertain student strengths and weaknesses and identify any legacy skills and knowledge that they bring with them to instruction (Treagust, 1995).

Norm-referenced assessment and criterion-referenced assessment are oppositional concepts. Norm-referenced assessment compares student performance against fixed average norms that have been established with reference to a large number of representative individuals. Tests often "rely heavily on multiple-choice, with some short-written responses, and tend to be based on national standards as opposed to state and local standards" (American Intercontinental University, n.d.). The SAT is the most well-known norm-referenced assessment.

Cantonese operatic singing is not a good fit for this kind of assessment. By contrast, the criterion-referenced assessment compares student performance against pre-defined criteria (Asmus, 1999). This might be more suitable to the assessment formulation of Cantonese operatic singing.

Summative assessments and diagnostic assessments can be either norm-referenced or criterion-referenced. However, formative assessment is criterion-referenced by definition (Harlen & James, 1997).

Standard-based assessment relies on a set of pre-determined statements outlining different standards or levels of achievement in an assessment component, course, or program - normally defined in terms of the stated assessment criteria (Asmus, 1999). It enables students to know against which criteria they will be judged and the standards attached to every criterion (Armstrong et al., n.d.).

2.2.1.4 Roles of assessment in learning

As Gardner (2012) put it: "assessment in education must, first and foremost, serve the purpose of supporting learning" (p. 11). Assessment is "not only a measure of student learning but also a means to student learning" (Grant & Gareis, 2015. p. 8). Now that assessments play a significant role in music education, how can music teachers use assessments as effective tools to boost learning? First of all, the three roles for student assessment ought to be clarified (see Figure. 2.3):

Assessment of Learning	Assessment for Learning	Assessment as Learning
"Done to" the student	"Done for" the student	"Done by" the student
Traditional	Constructivist	Self-reflective
Centered on teacher	Centered on student	Centered on self
Students as passive learners	Students as active learners	Students as active learners
Summative: Occurs after instruction	Formative: Integrated with instruction	Formative: Integrated with instruction
Administrative control	Helps students learn	Helps students learn
Competitive	Collaborative	Personal
Limited implementation of feedback to inform future learning	Implementation of constructive feedback from others to inform future learning	Implementation of feedback from self to inform future learning
Limited communication between teacher and students and among peers	Increased communication between teacher and students and among peers	Increased communication between teacher and students and among peers
Norm-referenced; criterion-referenced	Criterion-referenced	Criterion-referenced; self-referenced

Figure 2.3 Roles for Assessment (Scott, 2012, p. 32)

"Assessment *of* learning serves a summative role by providing information upon which to judge the student performances at a terminal point in instruction and to provide a basis for grades" (Scott, 2012, p. 35). It is ingrained in the traditional assessments of Cantonese operatic singing, such as the audition assessments at the Children's Palace or the final assessments at the Cantonese Opera School, which tend to be competitive. Since assessment of learning is based on assessment data collected by the teachers mainly for the purpose of "administrative control", it is "done to" the student and is "centered on teachers". Since students function as "passive learners" in this assessment mode, there can only be "limited discussion about the assessment process" and "limited use of feedback from assessment to inform future learning". Assessment *of* learning can be acquired either using criterion-referenced assessment (comparing student achievement to preset criteria), or using norm-referenced assessment (comparing student achievement to their peers). In traditional Cantonese operatic singing assessments lack criteria:

"Assessment *for* learning represents a constructivist perspective in which students, as active learners, use assessment feedback to extend their current levels of understanding" (Scott, 2012, p. 32).

Assessment *for* learning plays a summative role and is criterion-referenced. From the constructivist perspective, assessment *for* learning is "done for" the student. The development of assessment tools is normally collaborative between teachers and students and "Centered on student". Students learn what they need to pay attention to and to what extent they are required to achieve. Thus, the development process helps students to learn. Assessment for learning contributes to "implementation of constructive feedback from others to inform future learning" and "increased communication between teacher and students and among peers". Through assessment for learning, students are provided with an opportunity to make further improvements by using the assessment information, which facilitates their academic achievement, increases positive attitudes toward learning, and boosts self-esteem.

"Assessment *as* learning provides avenues by which students may reflect on their learning and, through this self-reflection, gain a deeper understanding of music and music-making" (Scott, 2012, p. 35). Like assessment for learning, assessment as learning is formative. It facilitates "implementation of constructive feedback from others to inform future learning" and "[increases] communication between teacher and students and among peers". Assessment as learning is both criterion-referenced and self-referenced. "Done by" the students, assessment as learning helps students to learn. Guidance is needed for students to learn to assess themselves accurately and to use the assessment results to direct future progress. Through this process, students gain the initiative in their studies gradually as they develop skills and musicianship.

Students are expected to adapt and extend self-regulated learning beyond school and into life-long learning.

The above three roles of assessment play an important part in supporting and improving student learning, which embodies different purposes in education. Therefore, the assessment must be developed with explicit purposes in mind and optimized when applying the assessment framework.

2.2.1.5 Significance of assessment in teaching and learning

In the contemporary education context, teaching is a comprehensive concept consisting of three sections: curriculum, instruction, and assessment, which are inextricably intertwined (Duke, 1999). For students, assessment always "defines the actual curriculum" (Ramsden, 2003, p. 182), which makes it a key driver of student activity (Kizlik, 2012). Good assessments facilitate students to become more effective self-directed learners (Angelo & Cross, 1993). As assessment is inevitably related to instruction, it should be considered when planning and before instruction actually begins (Duke, 1999), otherwise good teaching cannot be expected (Mantie, 2019). To be more specific, instruction should be guided by assessment. According to "the framework for effective teaching" constructed by Stronge (2018), assessment is one of the critical aspects of effective teaching, which are illustrated below.

Figure 2.4 The framework for effective teaching (Stronge, 2018, p.12)

Assessment, whether it takes the form of a classroom quiz, a standardized test, or a nationwide exam, is intrinsically connected to instruction and curriculum. Suskie (2018) illustrated the importance of assessment and its correlation with teaching and learning in a continuous four-step cycle which promotes student learning as below:
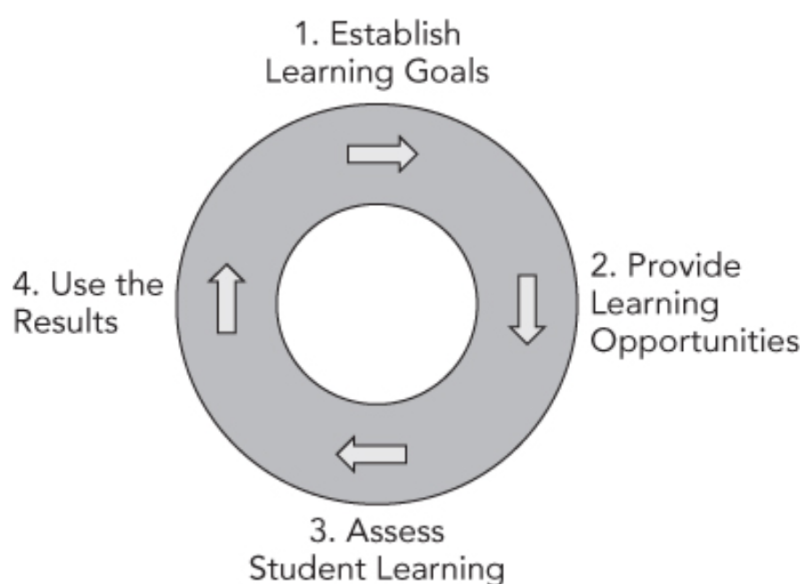
Figure 2.5. Teaching, Learning, and Assessment as a Continuous Four-Step Cycle

(Suskie, 2018, p. 2)

The above cycle demonstrates a four-step teaching-learning-assessment process:

a) Establish clear, observable expected goals for student learning.

b) Provide students with sufficient opportunities to achieve those goals.

c) Assess systematically how well student learning meets those goals.

d) Use the information obtained from the assessment to check and improve student learning.

There is also considerable evidence to show that assessment not only facilitates teacher instruction but also drives student learning (Wormald et. al, 2009). Assessment has positive effects on the quality of teaching and on the improvement of learning. If the ongoing assessment is implemented in the teaching practice, it can facilitate:

a) Greater achievement for students.

b) Improvement to teacher instruction and pedagogical decisions, making them more responsive to student learning.

c) Exhibition of greater concerns and academic emphasis in learning and classroom practices.

Better supervision of the adequacy of student learning, the identification of student needs regarding extra or different forms of instruction, and the modification of teaching practices will help to maximize student learning (Waugh & Gronlund, 2012; McMillan, Venable, & Varier, 2013).

As Brown (2005) asserted, "assessment is probably the most important thing we can do to help our students learn" (p.81). He later emphasized that, "we need to ensure that decisions

about assessment strategies are based on the best available evidence-based research on assessment, rather than on custom and practice or what is easy to do" (Brown, 2005, p. 88). Assessment is also an essential issue in music education (Mark & Madura, 2013). How to use assessments affects the efficacy of improvements in education and education programs (Brophy & Albert, 2008). Assessment in music education will be further discussed in the next section.

2.2.1.6 Principles of quality assessment

Armstrong (2008) concluded that a quality assessment should be a:

a) Valid assessment. For an assessment to be considered valid, it should be explicitly aligned with educational objectives and measure student achievement of those objectives.

b) Reliable assessment. Reliability in an assessment refers to the consistency of assessment over time, including across a cohort of students and between multiple assessors (Morgan et al., 2004). When assessment results remain consistent across multiple assessors, it is referred to as inter-rater reliability. When a single assessor makes consistent judgments across tasks, it is referred to as intra-rater reliability. Criterion- and standard-based assessments are articulated beforehand. Agreed-upon criteria and standards for an assessment task can be developed using discussions between assessors. These discussions are referred to as moderation, which is an overarching component of reliable assessment, especially when there is more than one assessor. Making assessments based on transparent criteria and standards also contributes to more consistent feedback for students from multiple assessors.

c) Fair assessment. A fair assessment can provide students with equitable opportunities to demonstrate their achievements in learning (Lam, 1995). The introduction of criteria and standards into the assessment task makes the assessment process more transparent for students. It might be time-consuming for teachers to articulate knowledge that is sometimes implicit and develop a criterion- and standard-based assessment. However, those efforts are rewarding, because time is saved when making assessments. It also contributes to the consistency of assessments. Well-developed criterion- and standard-based assessments can minimize controversy caused by assessment results as students understand the criteria and standards beforehand and the process is more transparent.

Concisely, the first principle can be achieved by employing Bloom's Taxonomy. The criterion- and standard-based assessment approach aids achievement of the last two principles.

2.2.2 Challenges facing assessment in music education

Traditionally, assessment has been correlated with grading, which music teachers have been involved with for generations, though it might not always have been implemented in a very valid way (Brophy & Albert, 2008). Despite there being a general consensus about music's inherent value in education (Hanna, 2007), and growing attention being given to assessment in education, music education often encounters challenges when it comes to assessment (Brown, 2005). For example, Lehman (2008) stated an obstacle to music assessment and pointed out the solution accordingly: music assessments are thought to be difficult, time-consuming, and costly. Some teachers suspect or even reject music assessment, asserting much music learning is highly subjective and no test can measure student perceptual

skills, creative ability, or love of music. The intangible quality of the highest level of musicianship may be difficult or even impossible to assess. Nevertheless, the skills and knowledge related to these phenomena can be assessed. These skills and knowledge are prerequisites for students to obtain high-level musicianship and gain achievement.

Although music has been incorporated into the curriculum of public schools for a long time (Mark and Gary, 1999), music teachers still find that they have to constantly rationalize the contribution of music to wider academic performance in order to defend their programs from budget cuts (Mishook & Kornhaber, 2006). Thus, a central challenge for music teachers is assessing student performance (DeLuca & Bolden, 2014). Decision-making surrounding resource allocation is normally based on observable and objective outcomes, rather than personal bias or perceptions. But, unlike other subjects such as mathematics or science, which are liable to be assessed objectively, music education is always judged subjectively. This lacks the precise quantitative measurements needed for quality evaluation, and thereby falls outside the scope of standardized testing (Hanna, 2007). Without reliable and valid assessment, music teachers will have to continue fighting for academic legitimacy in a society within which accountability matters (Asmus, 1999). Even though a great many countries have established national standards for music education, alignment between objectives and curricula is undeveloped (Colwell, 2003).

Student assessment is indispensable to music education. And music teachers are responsible for providing valid and reliable information about their students' academic achievements that pertain to multifarious learning outcomes (Scott, 2012). The teaching and learning of music inevitably involve cognitive, psychomotor, and affective domains. Therefore,

the revised Bloom's taxonomy is introduced and used to establish a holistic theoretical framework for this paper, with which music education outcomes can be translated into objective criteria and an assessment framework proposed to guide assessment settings within Cantonese operatic singing.

2.2.3 Value of using Criterion- and standard-based assessment

Although assessment is an effective tool to improve education, it might become adverse or even counterproductive if it is misused. Thus, to shape effective student learning and teaching practice and guarantee quality assessment, criteria- and standards-based assessment can be adopted in the formulation of assessments (Armstrong et al., 2008). When the criteria are identified in an assessment task, students gain a more explicit idea of what they are expected to do. And when the standards of the assessment criteria are publicized, students are offered the opportunity to know how they will be judged (Biggs, 2003). As Morgan et al., (2004) argued, "Clear standards of expected student performance have increasingly become recognized as the vital ingredient in effective assessment" (p. 4).

Criterion- and standard-based assessments: 1) are based on the academic achievement of the student, 2) establish clear guidelines for proficiency, which is also serve as the foundation of rubrics, and 3) compare each student performance to preset criterium- and standards, not the performance of other students (Armstrong, S., et al., 2008).

For students, criterion- and standard-based assessment provide goals to pursue and achieve as well as guidance for the learning and judgment of their work. For instructors, it is beneficial to use criterion- and standard-based assessments to ensure greater reliability when

multiple markers are engaged and better consistency over time for individual markers. For educational institutions, adopting criterion- and standard-based assessment helps to ensure the quality of education (Hendry, et al., 2012).

Due to the great value of criterion- and standard-based assessment, this study will employ it as the design rationale for the formulation of an assessment framework, which teachers can use as a tool to guide multifarious assessment settings in music education, thereby enhancing its efficiency and effect.

2.2.4 Value of assessment frameworks

Assessment frameworks not only define what metrics should be measured but also clarify how and why those metrics should be measured (Wagemaker, 2020). They are believed to be able to improve both validity and reliability in assessment and facilitate the formulation of robust assessment instruments (Pearce, 2018). With the application of assessment frameworks, assessment instruments can be established based on agreed-upon framework components, through which specific competencies are liable to be targeted (Pearce et. al, 2015). As Amin put it, "we need to articulate the purpose of the particular assessment with the greatest possible clarity in a manner that goes beyond its simple categorization as summative or formative. We must ask repeatedly what the real purpose of assessment is and be certain of its explicit, as well as implicit, agenda" (p. 6). This can be achieved by a well-articulated assessment framework (Pearce et. al, 2015).

One effective way to underpin consistent assessments within and across classes, grades, institutions, or organizations is to have a common assessment framework, within which

definitions and standards are explicitly articulated and agreed. The assessment framework can be used as a foundation and reference that a community can use to build feasible and desirable assessments. It also facilitates the reflective development of instruments, enhances accountability, and reduces bias during practice (Pearce et. al, 2015). Ainley & Schulz (2020) emphasized that "the process of reviewing literature and educational practice results in the formulation of research questions that become key elements of the assessment framework" (p. 28). This will influence the research design of the current study.

2.2.5 User-centered approach in designing an assessment framework

The user-centered design approach involves users in every step of the process of the design and evaluation of products (Johnson, 1998). Kahraman (2010) defined it as "a contemporary design approach which actively involves users at every stage of the design process for the development of more effective, efficient, and safe products" (p. 2071). Abras et al., (2004) went as far as to argue it was not only a general term of approach but a design philosophy.

The user-centered design process consists of three major phases: pre-design analysis, the creation of design solutions, and post-design evaluations (Evans, 2002). In the first phase, the needs, limitations, expectations, and suggestions of target users are collected and analyzed using various methods that yield data for the design of the assessment framework. This includes interviews, focus groups, questionnaires, and the like (Peerce et al., 2002). In the second phase, alternative design solutions are developed for the users' evaluation. Designers amplify their understanding by listening to user feedback or discussing alternative designs. This yields information that is not acquired from initial observations, interviews, and needs analyses. The

process is interactive and iterative and is shared between designers and users. This allows designs to be refined (Abras et al., 2004). The final phase of this design approach, post-occupancy evaluations, gives an evaluation from the perspective of users of the final version of the design (Abras et al., 2004). In this phase, designers collect information about users' satisfaction and any functional or practical problems by using questionnaires, focus groups, or interviews (Kahraman, 2010).

Essentially, the designer and users are co-creators (Abras et al., 2004). Instead of forcing users to change their habits to accommodate new products, user-centered design optimizes the usability of designed products by involving users as active participants in the design process (Johnson, 1998). Acting as a facilitator and mediator, the designer refers to users' needs and interests. This allows users to affect the design process and outcome (Abras et al., 2004). By keeping users' satisfaction in mind, this design approach heightens the usability, success, and performance of a product, so that it can be employed "with minimum effort and optimum efficiency" (Kahraman, 2010, p. 2071).

Abras et al. (2004) summarized the advantages and disadvantages of user-centered approach as below (see Table. 2.4):

Table 2.4. Advantages and Disadvantages of User-Centered Design (Abras et al., 2004, p. 11)

Advantages	Disadvantages
Products are more efficient, effective, and safe	Costly
Assists in managing users' expectations and levels of satisfaction with the product.	Time-consuming
Users develop a sense of ownership for the product	May require the involvement of additional design team members (i.e., ethnographers, usability experts) and a wide range of stakeholders
Products require less redesign and integrate into the environment more quickly	It May be difficult to translate some types of data into design
The collaborative process generated more creative design solutions to problems.	The product may be too specific for more general use, thus not readily transferable to other clients; thus more costly

2.3 Theoretical approach: Taxonomy of educational objectives

Cantonese operatic singing is a multifaceted art genre that involves knowledge, attitudes, and skills in performance. Thus, its assessments need to encompass each of the above aspects. Bloom's Taxonomy is a powerful tool to guide the formulation of an assessment method appropriate to Cantonese operatic singing. The subsequent sections will follow the logic of what-why-how: the first three subsections will outline what is meant by taxonomy, Bloom's Taxonomy, and its content. Subsection 2.3.4 (Function of the Bloom's Taxonomy of Educational Objectives) and 2.3.5 (Impact of the Bloom's Taxonomy of Educational Objectives in music education) will reveal why Bloom's Taxonomy has been introduced into the current study. The last three subsections will discuss how to establish an appropriate theoretical framework from Bloom's Taxonomy and how it is used in the current study.

2.3.1 Overview of the taxonomies

Taxonomy is "a classification system that establishes the hierarchy of the part to the wholes" (Hauenstein, 1998, p. 2). A hierarchy signifies that a lower-order layer is a prerequisite to a

higher-order layer, and the higher-order layer encompasses lower-order layers. Taxonomies are considered to be useful tools for stating objectives, building curricula, and constructing and testing evaluation procedures (Anderson & Sosniak, 1994). There are several similar taxonomies illustrated as below:
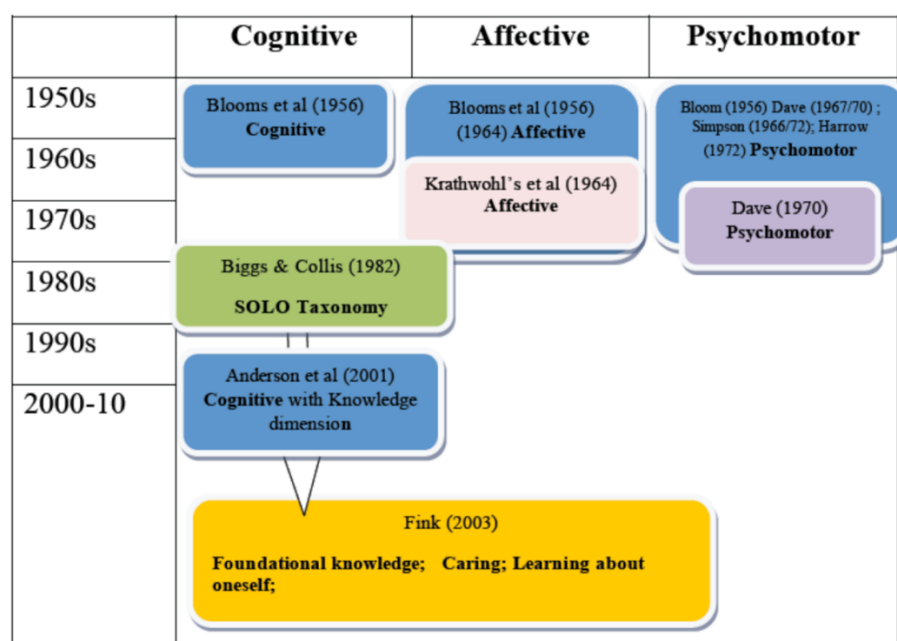
Figure 2.6. An Overview of Development of Taxonomies and their Domains
(O'Neill, et. al., 2010, p. 7)

As shown in Figure 2.6, SOLO Taxonomy, as proposed by Biggs and Collis in 1982, only covers the cognitive domain, whereas Fink's Taxonomy, which was presented in 2003, also includes the affective domain. However, the series of Bloom's Taxonomy has grown into the most comprehensive form, and includes cognitive, affective, and psychomotor domains. In 2001, Anderson et al revised the cognitive domain of Bloom's Taxonomy, mitigating original weaknesses and improving the practical limitations (Amer, 2006).

Bloom's taxonomy is considered to be "one of the most influential educational monographs in the past half-century" (Marzano & Kendall, 2006, p. 1) and is the most popular form around the world.

Thus, the current study will adopt Bloom's taxonomy as a basis and revise it according to new practical theory to form a theoretical framework that will better accommodate the research topic.

2.3.2 Overview of Bloom's Taxonomy

Bloom's Taxonomy of educational objectives was published in 1956, initially by a team under the American academic and educational expert Benjamin S. Bloom. It gradually evolved into a set of three cumulative hierarchical models that can be used to categorize educational learning objectives into levels of specificity and complexity, including the cognitive domain (thinking), the affective domain (emotion/feeling), and the psychomotor domain (physical/kinesthetic) (see Figure 2.7).
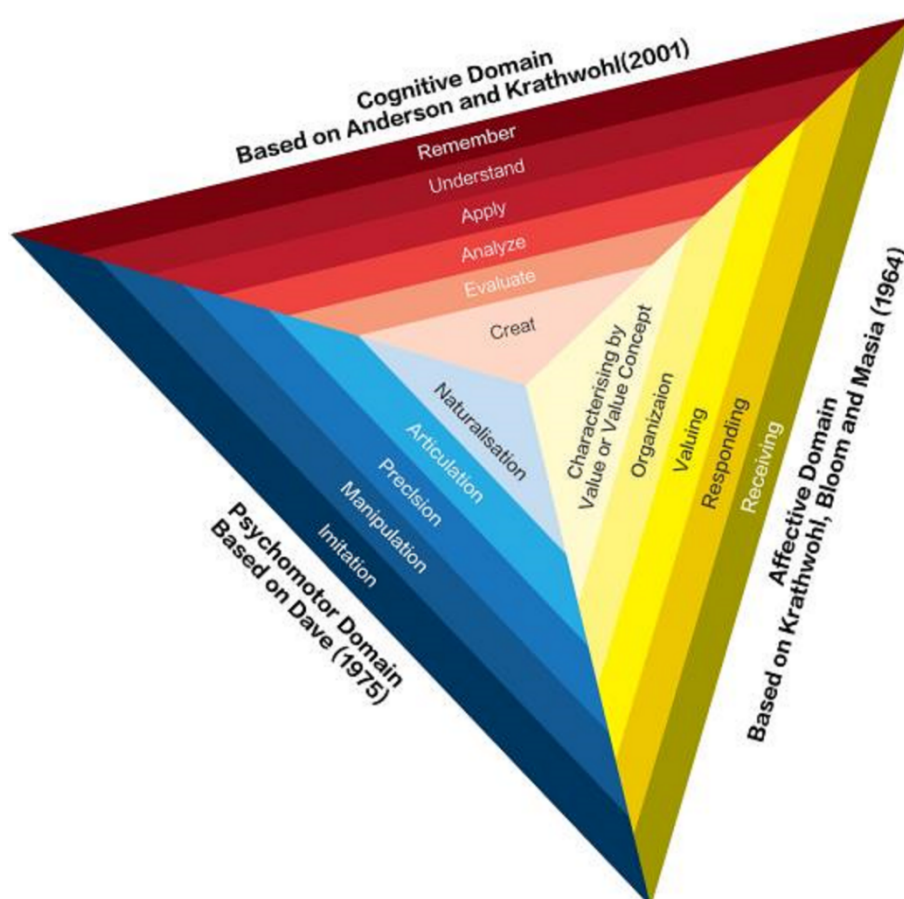
Figure 2.7. The pyramid model of the taxonomy of educational objectives

The taxonomy of educational objectives facilitates careful inspection and ultimately reinforces the connections between the aforementioned components of education: curriculum, instruction, and assessment (Airasian & Miranda, 2002). One purpose of the taxonomy is to disassemble an educational development process into smaller parts. This makes discussion of educational objectives easier, guides instruction, and allows useful assessment (Krathwohl, 2002).

2.3.3 The three domains

A domain is "a distinctly delimited sphere of knowledge or intellectual activity" (Hauenstein, 1998, p. 2). There are three domains in the taxonomy: cognitive, affective, and psychomotor.

2.3.3.1 Cognitive domain

The cognitive domain refers to "those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills" (Bloom et al., 1956). In 1956, Bloom and his associates initiated and published a widely accepted taxonomy for classifying learning objectives and assessment items within the cognitive domain in the book, *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain* (Bloom's et al., 1956; Anderson & Krathwohl, 2001). It was meant to be an analytical tool for the educational objectives of learning, instruction, and assessment. But the use of nouns to name each order of objectives had made Bloom's taxonomy unnecessarily abstract and harder to apply. In 2001, the revision framework was developed and published by Anderson, Krathwohl, and their colleagues in their book, *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives*, which emphasized "student-oriented, learning-based, explicit, and assessable statements of intended cognitive outcomes" (Anderson & Bloom, 2001, p. 23). Henceforth, Bloom's taxonomy is referred to as the original Bloom's Taxonomy, while Anderson and Krathwohl's new taxonomy is recognized as the revised Bloom's Taxonomy (Krathwohl, 2002). The current study will adopt the revised Bloom's Taxonomy.

2.3.3.2 Affective domain

The affective domain contains "objectives which describe changes in interest, attitudes, and values". In 1964, Bloom, David Krathwohl, and their associates published a second handbook, which focused on the affective domain. This included receiving/attending, responding, valuing, organizing, and characterizing (see Figure 2.8).
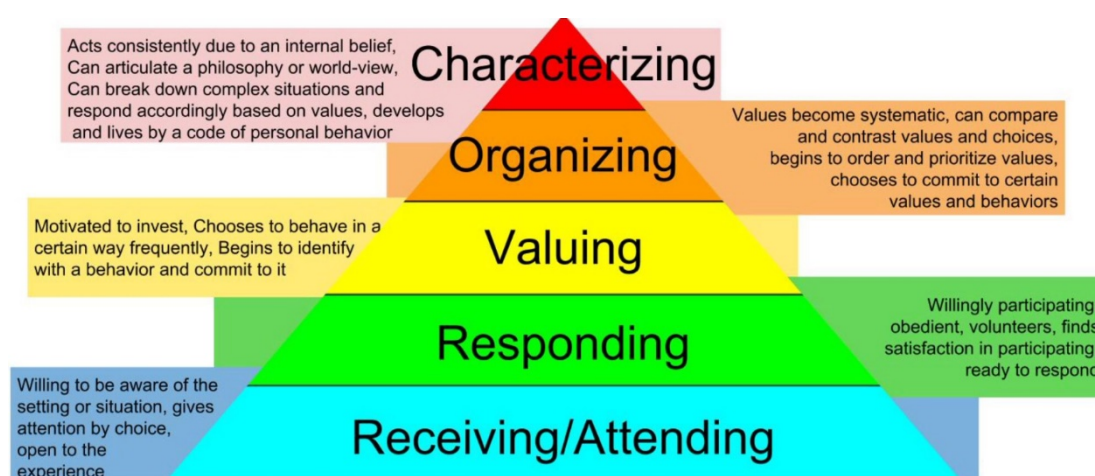
Figure 2.8. Affective Domain (Thiele, 2015)

Level 1 to Level 5 shows progressive levels of internalization in terms of interest, appreciation, attitude, value, and adjustment (see Figure 2.9).

Figure 2.9. Internalization continuum of affective domain (Adapted from Krathwohl, David R., Bloom, Benjamin S. and Masia, Bertrom B., *Taxonomy of Educational Objectives, Book 2: Affective Domain*)

Although the affective domain is considered to be more problematic, it is an indispensable complement to the cognitive and cannot be neglected (Lynch et al., 2010).

2.3.3.3   Psychomotor domain

The psychomotor domain is "the manipulative or motor-skill area" (Bloom et al., 1956, p. 7). A taxonomy for the psychomotor domain was never published by the original team, but Simpson (1966), Dave (1970), and Harrow (1972) continued the work (Krathwohl, 2002). This study adopts Dave's psychomotor domain (see Figure 2.10) as it is more suitable and precise when applied to music education.
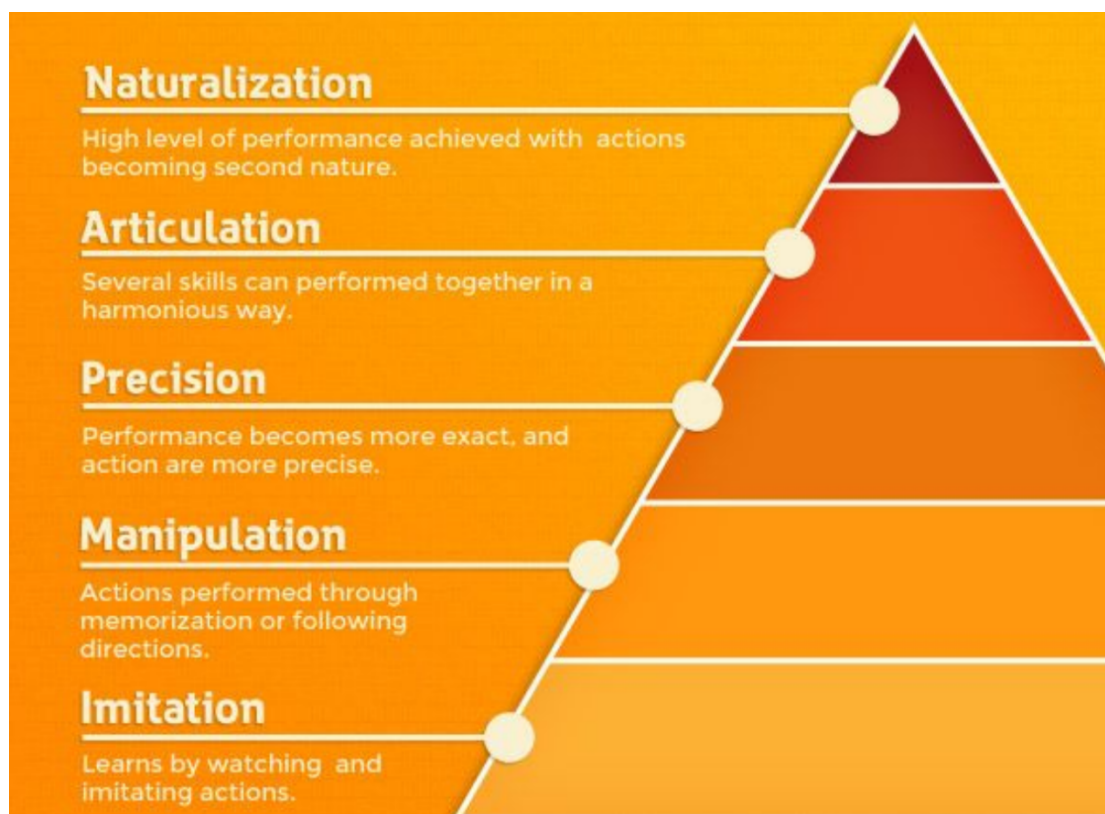
Figure 2.10. Dave's psychomotor domain (Dalto, 2014)

Bloom's taxonomy suggests a well-designed education should include learning objectives in each of the cognitive, affective, and psychomotor domains (Bloom et al., 1956).

2.3.4 Function of Bloom's Taxonomy of Educational Objectives

Bloom deemed the taxonomy could serve as (Krathwohl, 2002, p. 212):

a) A common language about learning goals to facilitate communication across persons, subject matter, and grade levels.

b) the basis for determining the specific meaning of the broad educational goals of a particular course or curriculum, such as those found in prevalent national, state, and local standards.

c) a means for determining the congruence of educational objectives, activities, and assessments in a unit, course, or curriculum, and

d) a panorama of the range of educational possibilities against which the limited breadth and depth of any particular educational course or curriculum might be contrasted.

Anderson (2001) offered six reasons for the contribution of categorizing objectives using the taxonomy as a framework to guide the classification (pp. 36-37):

a) it permits educators to examine objectives from the student's point of view.

b) it helps educators consider the panorama of possibilities in education.

c) it helps educators see the integral relationship between knowledge and cognitive processes inherent in objectives.

d) it helps educators deal with "instruction and assessment questions".

e) it makes more readily apparent the consistency, or lack of it, among the stated objectives for a unit, the way it was taught, and how the learning was assessed; and

f) it helps educators make better sense of the wide variety of terms that are used in education.

The taxonomy of educational objectives is a useful tool that helps to translate "intuitive understandings and tacit knowledge that are involved in music into an academic language that can be understood even by non-musicians" (Hanna, 2007, p.9). Making good use of this tool,

we can coordinate with the Cantonese opera experts to tease out suitable assessment criteria and standards, so that reliable assessments can be ensured. This will help this traditional genre to thrive in an assessment-oriented world.

2.3.5 Impact of Bloom's Taxonomy of Educational Objectives in music education

Prior to the taxonomy, music learning was considered to be subjective and esoteric. Essentially, it was assessed by applying standardized approaches common to other subjects, and assessments depended on implicit intuitions. As a result, assessment of music education has remained beyond the realm of standardized tests, which has made it difficult for the public to evaluate its academic value (Colwell, 2017).

Therefore, music education is in a predicament: on one hand, there is a consensus surrounding music's intrinsic value to education; on the other hand, decisions or policies surrounding resource allocations are usually based on observable outcomes and objectives, at the expense of personal feelings or perceptions (Linn, 2003; Porter, 2002).

Unlike other subjects, such as mathematics, science, literacy, and so on, which are prone to be taught and assessed objectively, assessment in music education normally lacks quantitative measurements that evaluate quality - especially in the realms of aesthetics and psychomotor ability. Hanna (2007) noted that the current form of assessments used within music education remains subjective and informal. Without standardized approaches to assessment, music education cannot win its struggle for academic legitimacy within a policymaking environment where accountability matters (Asmus, 1999).

Bloom's taxonomy greatly altered this predicament. Music educators can now "make their implicit notions of music learning explicit"(Hanna, 2007, p. 15). Using the revised Bloom's taxonomy, achievable learning goals can be created and a corresponding plan to the goals can be made. Educators can prepare reasonable lesson plans with well-defined goals (Sami, 2020).

Bloom's taxonomy is a useful tool for music assessment. It provides shared educational terminology for music educators to articulate and assess learning achievements. Music educators' tacit knowledge and intuitive notions of cognitive processes in music and its learning can be translated into a common academic language that can be comprehended even by non-musicians. Adopting widely accepted objective assessment criteria in music that have been based on Bloom's taxonomy elevates academic credibility from the perspective of policymakers and other stakeholders. This may enhance the wider attention given to music education (Hanna, 2007).

Hanna (2007) deemed Bloom's taxonomy to be a welcome addition to music education. But, compared with other subjects, the adoption of Bloom's in music education has been scarce. Specifically, its application in China is rarely seen. By searching in the most authoritative database of mainland China, CNKI (China National Knowledge Infrastructure), only one paper can be found, which pertains to piano playing. Needless to say, applying taxonomies (or any other Western educational framework or pedagogies) to the education of Chinese *Xiqu* is unprecedented.

2.3.6 A holistic application approach to Bloom's taxonomy

"Individuals learn as whole persons" (Hauenstein, 1998, p.1). In the learning process, people use their feelings (affective domain), brains (cognitive domain), and bodies (psychomotor domain). However, the taxonomies of these three domains are commonly treated as separate entities. This may "lack a unifying context, have different numbers of categories, are unparallel in categorical intent" (Hauenstein, 1998, p. 5), "often incompatible and often too difficult to apply in the classroom" (Hauenstein, 1998, p. ix).

To compensate for the above deficiencies, a holistic approach to the three domains of the taxonomy is necessary and imperative. This is especially true for music education, which embraces an integration of the cognitive, affective, and psychomotor domains of learning. Many of the complex cognitive facets of music learning inherently pertain to affective and psychomotor learning (Hanna, 2007). For example, the affective domain is essential for the development of musicality, thus it is vitally important to student musical development and achievement.

Hauenstein (1998) stated that no attempt had been seen to synthesize or consolidate these three domains to guide the learning process until he proposed a conceptual framework; a holistic approach to the traditional taxonomies which feature below.

A fourth domain, the behavioral domain, is also presented. Hauenstein (1998) suggested that in any given context, one does not merely know, but has affected what one knows. One does not merely do or perform but performs in accordance with knowledge and affect. One does not merely feel but feels in relation to their knowledge and experience. One behaves or acts in relation to what one knows and feels and can do. As the composite domain of the

traditional three domains, the behavioral domain is defined with five hierarchical categories: acquisition, assimilation, adaptation, performance, and aspiration (Hauenstein, 1998, p. 10):

a) Acquisition: the gaining of new information and content.

b) Assimilation: working the new knowledge into what is already known.

c) Adaptation: applying what is known of various situations or problems to one's skills and values.

d) Performance: producing as a matter of routine and accommodating new knowledge, skills, and values.

e) Aspiration: seeking to do better, to excel, in accord with one's beliefs and skills.

The cognitive dimension is further refined since Hauenstein (1998) argued that it was wrong to deem whatever transmitted is knowledge. He believed that what is being transmitted is others' knowledge, which is information/content for the student. Information/content (others' knowledge), which is external and may be input in the form of books, lectures, media, computers, and so on, will not become the student internal knowledge until it has been experienced, remembered, understood, and even applied.

The traditional taxonomies have been redefined to 1) reduce the number of categories and subcategories of the objectives, and 2) ensure parallel categorical intent and make a holistic approach possible. Because the traditional domains have up to 63 categories and subcategories of objectives (see Table 2.5), classroom teachers often find it rather difficult to use in teaching.

Table 2.5. Traditional Domains and Categories (Hauenstein, 1998, p. 6)

| Domains	Categories	Subcategories
Bloom, et al.,1056)	6	23
Affective (Krathwohl, et al. 1964)	5	13
Psychomotor (Simpson, 1966)	7	8
Totals	18	44
Grand Total	63	

Based on Hauenstein's (1998) work, the holistic taxonomy of educational objectives was refined and adapted, as shown in Table 2.6. The suggestion proposed by Anderson (2001) was adopted and verbs were used to present every educational objective. This was aimed at making the educational statement more explicit.

Table 2.6. Redefined Holistic Taxonomy of Educational Objectives

|	**Behavioral Domain**	**Affective Domain**	**Cognitive Domain**	**Psychomotor Domain**	
1.0	**Acquire** • Receiving • Remembering • Perception	**Receive** • Awareness • Willingness • Attentiveness	**Remember** • Identification • Definition • Generalization	**Perceive** • Sensation • Recognition • Observation • Predisposition	Short Term Objectives
2.0	**Assimilate** • Responding • Understanding • Simulation	**Respond** • Acquiescing • Complying • Assessing	**Understand** • Translation • Interpretation • Extrapolation	**Simulation** • Activation • Imitation • Coordination	
3.0	**Adapt** • Valuing • Application • Manipulation	**Value** • Accepting • Preferring • Confirming	**Apply** • Clarification • Solution	**Manipulate** • Integration • Standardization	
4.0	**Perform** • Believing • Evaluation • Articulation	**Believe** • Trusting • Committing	**Evaluate** • Analysis • Qualification	**Articulate** • Maintenance • Accommodation	Long Term Objectives
5.0	**Aspire** • Behaving • Synthesis • Mastery	**Behave** • Demonstrating • Modifying	**Synthesize** • Hypothesis • Resolution	**Mastery** • Perfection • Naturalization	

2.3.7 Theoretical framework

According to the above analysis and discourse, the revised traditional three domains can

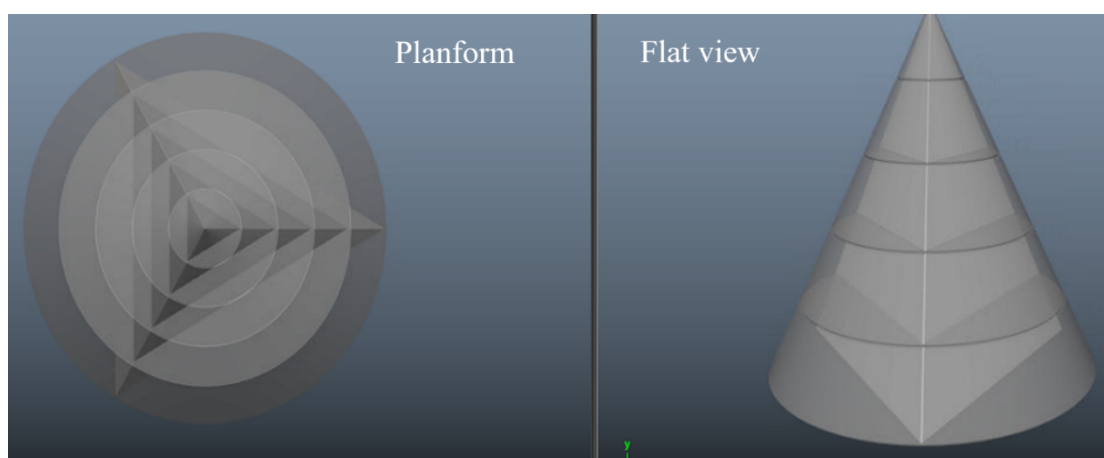be illustrated in a pyramid structure as below:
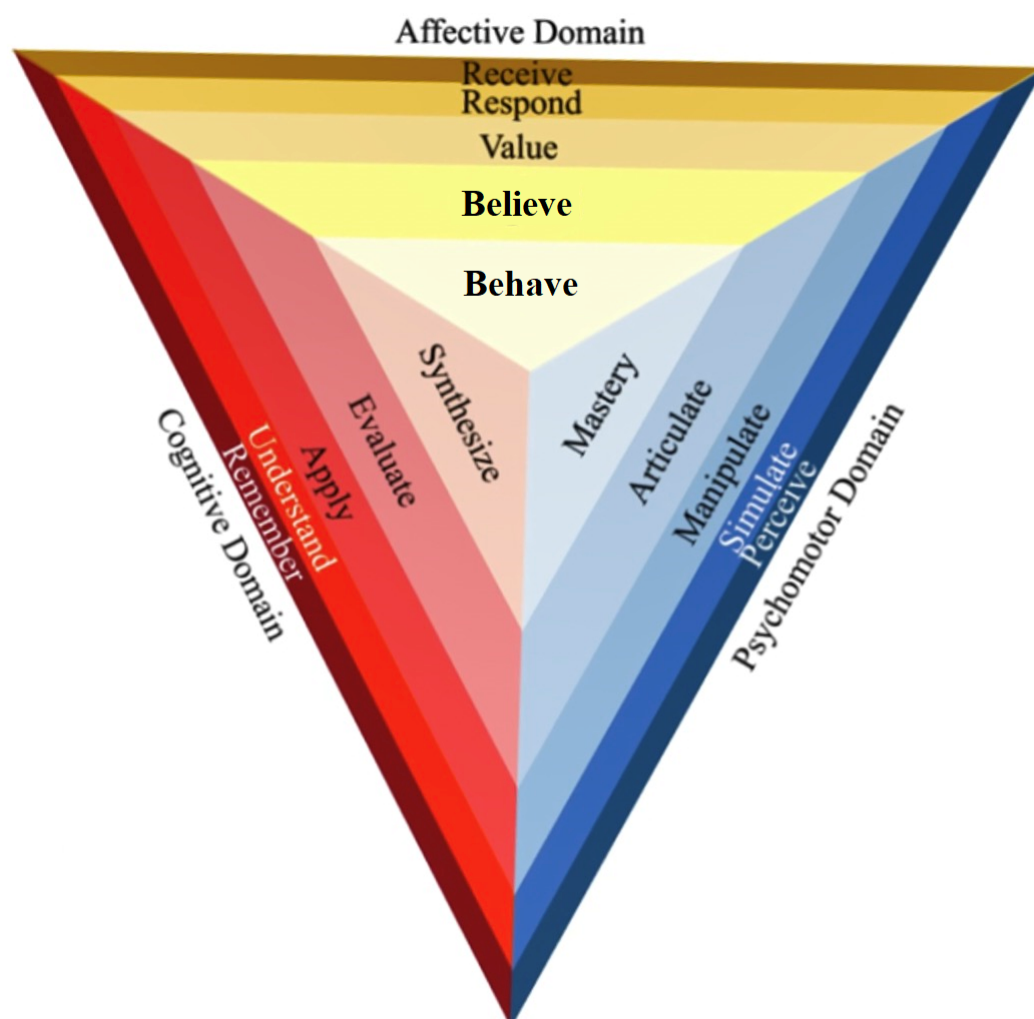
Figure 2.11. The Structure of Revised Three Domains

Four domains can be synthesized as below:
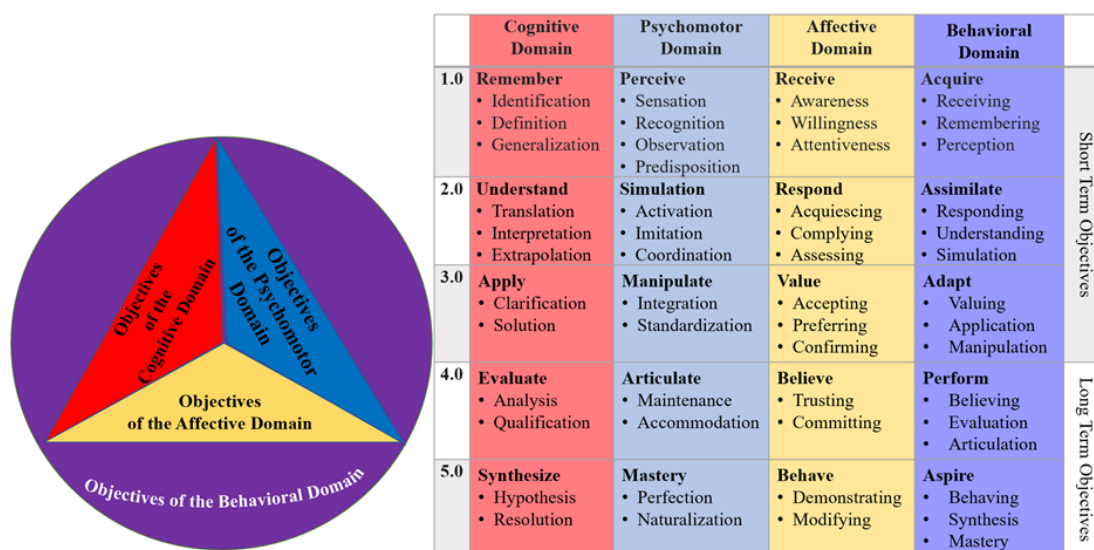
Figure 2.12. The Four Domains Composite Chart

There are five hierarchies of every domain shown in the chart. The three-dimensional structure of the composite chart can be diagrammatized as below:

Figure 2.13. The Three-dimensional Structure of the Four Domains

2.3.8 Application in the current study

Although this theoretical framework might be used in multiple fields or contexts, the current study will focus on discussing its utilization in the field of Cantonese operatic singing and within the educational context of Cantonese operatic singing. The cognitive and psychomotor domains can be adopted to assess corresponding aspects of Cantonese operatic singing. For example, the former is suited to assess intellectual subjects or content such as the history of Cantonese opera, the structure of Cantonese opera, and the like. The latter is fit for assessing those subjects that concern the utilizing and coordinating of motor skills, such as Cantonese opera singing skills. The assessment of the affective domain can be used to assess student psychology during learning. The composite domain and the behavioral domain can be used in the assessment of Cantonese operatic singing performance, which is a comprehensive display of the other three domains.

2.4 Summary

This chapter began by outlining the transmission and education of Cantonese opera. As the traditional mode of transmission and education, apprenticeship was replaced by professional training in the conservatory tradition and community training and learning from the mid-twentieth century onwards. This inevitably influenced the assessments used in this field. Subsequently, assessments of Cantonese operatic singing were narrated. Since research into assessments is scarce, only limited literature can be found. In the aforementioned literature, incomplete information was given, leaving a blank space for the future exploration of this issue and implying the necessity of the current study.

In the second section, multifaceted aspects of assessment within music education were outlined. The components of education were introduced, and the relationships between curriculum, instruction, and assessment were clarified. The importance of alignment among these three components was emphasized. The definitions of, and the relationships between, assessment-related terms were clarified. Types of assessments in education were introduced to aid a better understanding of the discussion of assessment. Three roles of assessment in learning were explained, including assessment *for*, *of*, and *as* learning. This set the foundations for the utilization of the proposed assessment framework. The significance of assessment in teaching and learning was elicited naturally after the first subsection. After the overview of the assessment in education, challenges to good assessment in music education were indicated.

To ensure the validity and reliability of assessment and ensure coverage of the multifarious domains of music education, a holistic theoretical framework stemming from the revised Bloom's Taxonomy will be established. This will allow music education outcomes to be translated into objective criteria so that an assessment framework can be proposed to guide assessment settings for Cantonese operatic singing. To shape effective student learning and teaching practice and guarantee quality assessments, the adoption of criterion- and standard-based assessment was recommended. Thus, in Section 2.2.5, after presenting the definition of and correlation between the criterion- and standard-based assessment and other relevant elements, the value of using criterion- and standard-based assessment was highlighted. The value of assessment frameworks was elaborated upon. To guarantee consistent, feasible, and desirable assessments and to facilitate the reflective development of instruments, enhance accountability, and reduce bias during practice, the formulation and utilization of an assessment

framework were proposed. Finally, to design a good assessment framework, a user-centered approach was selected, within which three major phases (pre-design analysis, the creation of design solutions, and post-design evaluations) were detailed and the advantages and disadvantages of this design approach were discussed.

In the third section, the theoretical framework of this study was reduced by a progressive introduction of Bloom's Taxonomy. Based on this theoretical framework and using the user-centered approach in design, criteria- and standards-based assessment framework can be formulated. This will allow an academic subject as multifaceted as Cantonese operatic singing to thrive in an increasingly accountability-driven educational environment and assessment-oriented world.

CHAPTER 3

METHODOLOGY

3.1 Research topic and questions

Previous studies of Cantonese opera mainly focus on its history, evolution, and ontology. Little attention has been given to the teaching and learning of Cantonese operatic singing. More specifically, how best to assess Cantonese operatic singing has remained unstudied. The research topic of this study is focused on the assessment of Cantonese operatic singing, with the research questions as follows:

1. What are the design and practices of the existing graded examinations of Chinese opera?

2. For the sake of transmission and education, how can an ideal assessment framework for Cantonese operatic singing be formulated?

3. What is common practice in assessments of Cantonese operatic singing?

4. To what extent does the proposed assessment framework facilitate the teaching and learning of Cantonese operatic singing?

3.2 Research design

Aiming to formulate a criterion- and standard-based assessment framework for Cantonese operatic singing, this study adopts a user-centered design approach, which includes three main stages: pre-design analysis, creating design solutions, and post-design evaluations. The first

two stages of this user-centered design approach were conducted in Phase I, and the final stage was conducted in Phase II. The whole research design can be illustrated in Figure 3.1:
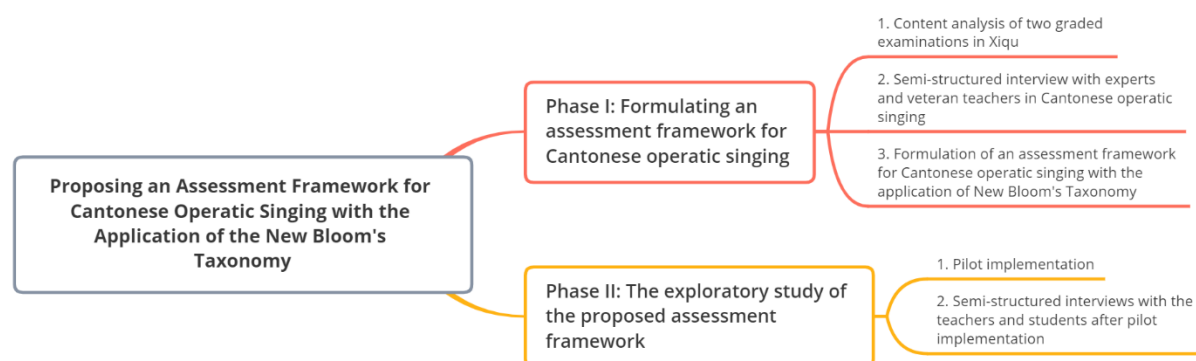
Figure 3.1. Research design of the current study

In Phase I, content analysis of two *Xiqu* graded examinations was designed to address the first research question: "What are the design and practices in the existing graded examinations of Chinese opera?" A series of semi-structured interviews with experts and teachers in Cantonese operatic singing was conducted to elicit answers to the second and third research questions. As an attempt to seek an answer to the last research question, "To what extent does the proposed assessment framework facilitate the teaching and learning in Cantonese operatic singing?" a range of pilot implementations was implemented.

3.3 Phase I: Formulating an assessment framework for Cantonese operatic singing

The purpose of Phase I is to formulate an assessment framework based on the theoretical framework of the current study by employing two steps: 1) analyzing existing representative

and authoritative assessments of *Xiqu* to learn from experience and practice and 2) conducting

a range of semi-structured interviews to collect opinions and suggestions from experts.

3.3.1 Content analysis of two graded examinations in *Xiqu*

Content analysis is an essential technique for "making systematic, credible, or valid and

replicable inferences from texts and other forms of communication" (Drisko & Maschi, 2016,

p. 7). Since Cantonese operatic singing is mainly transmitted using oral teaching and

emphasizes performance over formal education, few documents, especially educational

materials, are available to be acquired. The most authoritative references for assessment in

Cantonese operatic singing are the graded examinations in *Xiqu*. Therefore, in this study, two

representative graded examinations in *Xiqu* were selected to be the subjects of documentary

analysis.

Conducting content analysis of the existing relevant graded examinations at the beginning

of the study aims to 1) identify gaps that need to be filled and spot matters that might otherwise

be overlooked when formulating assessments for Cantonese operatic singing, and 2) elicit

questions that need to be resolved in subsequent interviews.

3.3.1.1 Data collection

Like some other folk arts in East Asia, Cantonese operatic singing was transmitted through

an apprenticeship with an oral approach up to the mid-twentieth century (Leung, 2015). Formal

curricula and assessments were absent (Leung, 2021). Leung (2015) indicated that, dominated

by the master, assessment of Cantonese operatic singing was apt to be subjective. Thus,

systematic and disclosed assessments are rare. As the only existing public assessment in

Cantonese operatic singing, the Cantonese operatic singing graded examination in Hong Kong is the most suitable reference for formulating a new assessment framework. To broaden data sources, another representative public assessment of *Xiqu*, the Peking opera graded examination, was also selected as an analysis subject.

3.3.1.2 Data analysis

The process of conducting a content analysis can be summarized as follows (Denscombe, 2014, pp. 283-284):

Step 1 Choosing an appropriate sample of data.

Step 2 Breaking down text into smaller component units of analysis.

Step 3 Developing appropriate categories for analyzing the data.

Step 4 Coding the units to fit the categories.

Step 5 Conducting frequency counts of the occurrence of the units.

Step 6 Analyzing the text from the basis of the unit frequencies and how they relate to other units in the text.

During analysis, software is helpful to unlock deeper insights and generate explicitly articulated, defensible findings that are underpinned by convincing evidence. As a result, NVivo was adopted during content analysis.

3.3.2 Semi-structured interviews with experts and veteran teachers of Cantonese operatic singing

Based on the findings from the content analysis in section 3.3.1, a range of semi-structured interviews were conducted. In the semi-structured interviews, several questions (see Appendix A) and topics were given as a prepared interview guide. These were usually open-ended, and the subsequent questions, as well as the wording, could be tailored to every interviewee and the responses given, with interviewer prompts and probes (Roulston, 2010). Data collected from the semi-structured interviews were analyzed using description, understanding, explanation, interpretation, and the drawing of conclusions (Gläser & Laudel, 2013).

3.3.2.1 Data collection

Hochschild (2009) indicated that the interview is a widely used instrument for data collection which allows interviewers to dig into issues and explore how and why people make connections between events, ideas, behaviors, and so forth, as well as how and why they frame their views in the way that they do. It is a powerful research instrument that can draw explanatory insight from observation, survey data, and content analysis (Cohen et al., 2018). According to Cohen et al. (2018), "…in the semi-structured interview, the topics and questions are given, but the questions are open-ended, and the wording and sequence may be tailored to each interviewee and the responses given, with prompts and probes" (p. 511). Striving to explore and formulate an assessment framework that accords with the essence and educational practice of Cantonese operatic singing, the authentic thoughts of practitioners, teachers, and experts were of significant importance to this study. Therefore, the semi-structured interview would be suitable for this study.

The findings and queries derived from the content analysis became the central questions of the semi-structured interview. For example, what is the basis of the hierarchy, categories, criteria of traditional opera assessment, and measurable elements of Cantonese operatic singing in mainland China and Hong Kong? How could assessment criteria be designed to conform with intrinsic characteristics and fulfill the purpose of transmission? Other problems were also discovered during the content analysis.

The procedure of planning and conducting semi-structured interviews can be summarized in three steps as follows (Whiting, 2008):

The first step is the identification of a "good informant," whose qualities include 1) knowledge of the topic, 2) the ability to reflect and provide detailed experiential information about the area under investigation, and 3) a willingness to talk. So, purposive sampling was adopted to ensure information-rich interviewees were selected. Purposive sampling is a vital feature of qualitative research (Cohen et al., 2018), which facilitates access to "knowledgeable people" by reference to their expertise, professional roles, experience, and so on (Ball, 1990). To collect authoritative opinions and representative information for this study, reputational case sampling (Teddlie & Yu, 2007) was adopted, which subsumes to purposive sampling, to select key informants from Guangdong and Hong Kong. The interviewees were chosen rigorously by their expertise, professional roles, and experience as listed in Table 3.1 below:

Table 3.1 The profiles of the selected interviewees

Subject	Location	Expertise	Age
Interviewee 1 (I-1)	GZ	● Famous artist ● The winner of the Plum Blossom Award 梅花奖（the supreme prize of Chinese Xiqu）	~ 50
Interviewee 2 (I-2)	GZ	● Dean of the Department of Drama of Guangdong Dance and Drama College ● One of the persons in charge of the Cantonese Opera Graded Examination	~ 50
Interviewee 3 (I-3)	HK	Cantonese opera performer	~ 60
Interviewee 4 (I-4)	HK	● Former performer in a Cantonese opera troupe in Guangzhou ● A senior teacher in a Cantonese opera institute in HK	~ 35
Interviewee 5 (I-5)	GZ	● Graduated from the Department of Drama of Guangdong Dance and Drama College ● A capable young teacher in GD ● A performer who is active in GD, HK, and Macau	~ 30
Interviewee 6 (I-6)	HK	● Experienced performer ● Competent teacher in HK ● One of the examiners of the Cantonese Operatic Singing Graded Examination	~ 40

The second step was interview preparation. Before the interview, a checklist was drafted so that all the key points would be approached in the interview. The purposes of the interview were also defined. The current study aims to: 1) investigate the previous or traditional practices and experience pertinent to assessment of Cantonese operatic singing, 2) explore the underlying reasons behind current practices in the assessment of Cantonese operatic singing, and 3) elicit professional opinions or suggestions of appropriate assessments of Cantonese operatic singing. Furthermore, interviews included: 1) following up on topics or ideas, 2) probing for further

information or responses, 3) asking respondents to specify and provide examples, 4) directly or indirectly asking for information, and 5) interpreting respondent replies. To balance interviews effectiveness and efficiency, the approximate length of each interview was kept within 1.5 hours. Finally, to ensure the interviewees' rights, consent forms and information sheets for participants of different age groups were drafted according to the guidance of the Education University of Hong Kong (see Appendix B and Appendix C). All of the above were manifested in the semi-structured interview protocol of Phase I (see Appendix A).

The ultimate step was conducting the interview. Besides audio recordings, writing down thoughts and feelings before, during, and after the interviews, and a reflective was thought to be beneficial. More questions were teased out and raised after the content analysis of Phase I.

As a device to ensure that the interviewees' rights were protected, the transcripts were sent back to the interviewees for their verification and confirmation. They were allowed to make revisions, including deleting any part of the interviews. Data were analyzed only after the transcripts were confirmed.

3.3.2.2 Data analysis

Thematic analysis is a systematic method of identifying, organizing, and extracting insightful themes across data sets (Clarke & Braun, 2014). Not only obvious or semantic meanings can be reported, latent meanings, assumptions, and ideas can also be interrogated using this method (Gareth et al., 2017). Braun and Clarke (2006) outlined and illustrated a six-step approach to thematic analysis:

Step 1: Familiarization with the data

This step encourages the researcher to become intimately familiar with the content of their data sets, thereby be more likely to notice valuable information that might be relevant to the research questions (Braun & Clarke, 2006). Thus, textual data should be read and reread critically, actively, analytically. During this process, making notes is imperative, including writing comments, underlining valuable information, and annotating transcripts (Gareth et al., 2017). The memory and comprehension gained from this phase facilitate good coding and analysis (Krippendorff, 2013).

Step 2: Generating initial codes

Codes refer to "the most basic segment, or elements, of the raw data or information that can be assessed in a meaningful way regarding the phenomenon" (Boyatzis, 1998, p. 63). Codes are not restricted to the participant meanings. They can also provide an interpretation of the data sets. Therefore, coding can be conducted at both the semantic level and the latent level of the data content. With the familiarity of the data sets, a systematic analysis will be defined through coding. "Coding is a typical categorizing strategy in qualitative research" (Maxwell, 2012, p. 133). Cohen et al. (2018) indicated that "coding is the process of breaking down segments of text data into smaller units, and then examining, comparing, conceptualizing and categorizing the data" (p. 668). By coding the data, frequencies and patterns are revealed. From coding, which is the initial step of qualitative data analysis, data can be subsumed into categories, themes, trends, and patterns, and the relations, similarities, and differences between themes can be identified (Cohen et al., 2018). In addition to manual work, coding can also be

performed via software programs (Clarke & Braun, 2014) such as NVivo, MAXQDA, and the like.

Step 3: Searching for themes

A theme refers to "something important about the data about the research question and represents some level of patterned response or meaning within the data set" (Braun & Clarke, 2006, p. 82). This step involves identifying overlaps and similarities between codes and sorting the relevant coded data into identified themes (Gareth et al., 2017). Meanwhile, analysts should start to consider relationships between themes and how they work together, trying to tell a whole story about the data (Clarke & Braun, 2014). In this process, visual representations can be employed to aid the sorting of different codes into themes, which helps analysts to reflect on the relationships between codes, themes, and even various levels of themes. These visual representations include mind-maps, tables, computer-assisted analysis systems, or even old-fashion methods such as organizing pieces of paper containing code names (Helene & Yardley, 2004). At the end of this phase, a table or a thematic map outline can be generated (Braun & Clarke, 2006).

Step 4: Reviewing potential themes

This step involves an iterative process of reviewing and refining the themes within the coded data and the entire data set for two purposes: 1) to ascertain whether the themes align with the data set, and 2) to code any data missed in previous coding stages within themes. These

processes can be repeated *ad infinitum* until a satisfactory thematic map has been devised. At the end of this phase, a distinct division of themes and their inter-relationships, together with an overall story told by these themes about the data should emerge (Braun & Clarke, 2006).

Step 5: Defining and naming themes

This step involves defining and refining the themes, thereby analyzing the data within them, and deciding what to call each theme. It should be ascertained that every theme is applicable to the research questions and the other themes.

Step 6: Producing the report

The report aims to provide "a concise, coherent, logical, non-repetitive and interesting account of the story the data tell-within and across themes" (Braun & Clarke, 2006, p. 93). This story should go further than merely describing the data and making a good argument with the research questions in mind (Clarke & Braun, 2014).

3.4 Phase II: Exploratory study of the proposed assessment framework

Often termed the feasibility study, the exploratory study is a vital step to yielding evidence for deciding whether and how to advance a full-scale effectiveness study (Hallingberg, Britt, et al., 2018). Before a large-scale study to validate the proposed assessment framework could be conducted, an exploratory study was implemented in Phase II, which consisted of two components: a pilot implementation and a series of semi-structured interviews.

3.4.1 Pilot implementation

Pilot implementation is an activity undertaken within the context of systems development and implementation to evaluate a system in a realistic setting and thereby learn about the fit between the system and its organizational usage situation together with any changes necessary before full-scale deployment (DTU, n.d., p. 6). Bansler and Havn (2010) indicated that a pilot implementation could be used:

a) To evaluate the usefulness and usability of the system,

b) To improve the system design based on user feedback, practical user experience, and observed results (e.g., productivity or quality data),

c) To identify necessary or desirable changes in work organization and processes,

d) To provide input for formulating implementation strategies and plans, based on users' reactions to the pilot" (p. 639).

Therefore, before the full-scale study, a pilot scheme was implemented with the purpose of:

a) identifying potential deficiencies in the proposed assessment framework for further improvement and determining the feasibility of the study,

b) testing whether the introduction of the study, the instruction of the procedure, and the description in the assessment rubrics could be clearly understood by the participants,

c) evaluating the application process of assessments based on the proposed assessment framework, including the time taken to make the assessment and the participants' willingness to participate in the study,

d) testing the assessment instruments, including whether the questionnaires were asking intended questions, whether the rubrics could assess the intended subjects, and whether the assessment design based on the theoretical framework of the study covered the aspects that needed to be assessed during teaching and learning.

The procedure of the pilot implementation was as below:

Step 1. Teachers were required to use the proposed assessment rubrics derived from the assessment framework to assess the six kinds of representative students: top students, average students, and underachievers in both senior grades and junior grades.

Step 2. Students were required to complete the online questionnaire about student learning statuses and psychology.

Although the proposed assessment framework could be applied to every subject that confirmed a specific domain, in the current study's questions and answers (Q&A), Cantonese operatic singing skills and the drama performance were selected as the representative subjects of the cognitive domain, the psychomotor domain, and the performance domain.

3.4.1.1 Data collection

Similarly to Phase I, purposive sampling was used to locate eligible participants who were the most representative (Cohen et al., 2018). Unlike Phase I, only frontline teachers were

recruited to conduct the exploratory study, thereby providing first-hand information and opinions. Thus, four teachers from the Cantonese Opera School, who represented professional education, and one teacher from Children's Palace, representing community training, were rigorously selected with reference to their expertise, professional roles, and experience. To collect diverse data, the participants included senior and junior teachers in both the exploratory study and the semi-structured interviews. Stratified sampling (Arnab, 2017) was adopted to select six categories of students in the exploratory study: top student, average student, and underachiever in both senior grades and junior grades, who were selected by their teachers according to their daily academic performance. Besides the four teachers, 18 of the 24 students being assessed were randomly chosen to be student interviewees.

The teachers received "The introduction of the current study and the instructions of the exploratory study" (see Appendix D), the assessment toolkits in Excel form (see Appendix E), and an Online Survey of Student Perception on Learning (see Appendix F). Explanations and elaboration were also offered. Consensus-based comprehension was ensured before the assessment was implemented. Data were collected and analyzed after the pilot implementation of the assessment derived from the proposed assessment framework. Interviews with the four teachers and 18 students were conducted. Finally, the conclusion and discussion were written after analyzing the interviews.

3.4.1.2 Organizing and presenting data

To organize and present the data collected from the pilot implementation, tabulation was used to reduce the substantial amounts of data into more manageable sizes. Tabulation is a

process of arranging data logically and systematically into rows and columns that makes it easier for researchers to draw reasonable interpretations and conclusions (Cohen et al., 2018). These orderly arranged data are one of the foci of the subsequent semi-structured interviews.

3.4.2 Semi-structured interview with the teachers and students after the pilot implementation

A range of semi-structured interviews was conducted in conjunction with the pilot implementation to collect comprehensive information. The teachers who had taken part in the pilot implementation, and some of their students who had been assessed, were invited to be the interviewees to provide opinions and suggestions for future improvements.

3.4.2.1 Data collection

In the semi-structured interview, twelve questions (see Appendix H) and topics were given as a prepared interview guide, usually open-ended, and the subsequent questions, as well as the wording, was tailored to every interviewee and the responses given, with prompts and probes (Roulston, 2010).

The procedure of planning and conducting the semi-structured interviews was identical to the semi-structured interview of Phase I.

3.4.2.2 Data analysis

Data collected from the semi-structured interviews were analyzed through description, understanding, explanation, interpretation, and conclusions (Gläser & Laudel, 2013). The strategies of data analysis have been delineated in detail in section 3.3.2.2.

3.5 Summary

The purpose of this chapter was to outline the research method adopted in the current study to answer the research questions. To formulate a criterion- and standard-based assessment, two phases were implemented. The overview of the procedure, purposes, and methods of these two phases are demonstrated in Table 3.2:

Table 3.2 Overview of the two phases in this study

Procedure	Purpose	Methods
Phase 1: Formulating an assessment framework for Cantonese operatic singing		
Step 1: Analysis of two relevant *Xiqu* graded examinations	To identify gaps that need to be filled and spot matters that might otherwise be overlooked when formulating a new framework. To develop the fundamental interview questions of step 2.	Content analysis
Step 2: Semi-structured interview	To find out previous or traditional practices and experience pertinent to the assessment in Cantonese operatic singing. To dig deep into the underlying reasons behind current practices in the assessment of Cantonese operatic singing. To elicit professional opinions or suggestions of assessment.	Multiple semi-structured interviews and thematic analysis
Step 3: Formulation of assessment framework	To establish a criterion- and standard-based assessment framework for Cantonese operatic singing	Not applicable

Phase II:	**The exploratory study of the proposed assessment framework**	
Step 1: Pilot implementation	To evaluate the usefulness and usability of the system.	Not applicable
	To improve the system design based on user feedback, practical use experience, and observed results (e.g., productivity or quality data).	
	To identify necessary or desirable changes to work organization and processes.	
	To provide inputs for formulating implementation strategies and plans, based on user reactions to the pilot.	
Step 2: Semi-structured interview	To collect feedback and opinions of the proposed assessment framework from students and teachers	Multiple semi-structured interviews and thematic analysis

3.6 The ethical issue of this study

To ensure that no one was harmed or suffered adverse consequences as a result of the current study, the potential ethical issues involved were fully considered. The ethical review guidelines published by the Human Research Ethics Committee of The Education University of Hong Kong were followed to safeguard the following five principles of research ethics:

a) Minimize the risk of harm.

b) Obtain informed consent.

c) Protect anonymity and confidentiality.

d) Avoid deceptive practices.

e) Provide the right to withdraw.

All the participants in the current study had read the information sheet and signed the consent form. The templates are enclosed in Appendix B and Appendix C.

3.7 Limitations of this study and implications for future research

As with most studies, the current study's design was subject to potential limitations.

The first was the limited sample size of the interviews. Although interviewees were selected to be representative of the Cantonese opera circle of Guangdong and Hong Kong, the limited number of interviewees might impose restraints on the comprehensiveness of opinions. A larger-scale study should be conducted in the future.

The second limitation is concerned with the lack of some of the materials in content analysis. Although there was a benefit to students knowing the criteria and rubrics of an evaluation by common consent, similar to other graded examinations in mainland China, Peking opera graded examinations publish only limited information. Criteria and rubrics needed to be investigated through internal relationships with significant effort. Due to social conventions in China, more contacts and rapport were necessary to get access to concealed vital information. This could be rectified in a future study.

The third limitation lay in the determining of the weightings. There were two kinds of weightings: subjective and objective (Mehdi et al., 2021). The calculations and justifications demanded considerable work, and might even deserve a whole study. As a result, the current study is focused on the construct of the assessment framework, and the specific numerical values of the weightings were left for further study. Weightings were a necessary design in the proposed assessment framework, equal weightings were allocated to indicate their existence.

CHAPTER 4

PHASE I: FORMULATING AN ASSESSMENT FRAMEWORK FOR CANTONESE OPERATIC SINGING

4.1 Content Analysis of two graded examinations

In this section, two representative and authoritative assessments of Chinese *Xiqu* were analyzed. The analysis subjects and their contexts follow:

Peking Opera Graded Examination

Initiated in 2017, the Peking opera graded examination is the first and only traditional operatic graded examination officially approved by the Ministry of Culture. Hosted by the National Peking Opera Company (國家京劇院), Peking opera graded examination carries both authority and representativeness. The examination system for the Peking Opera was established in 2015 under the auspices of the National Peking Opera Company, which is an embodiment of Peking opera's assessment. As a practical examination of Peking opera, POGE formulates its examinations according to the distinction of roles which are the most prominent in *Xiqu*: *chouhang* (丑行), *daomadan/wudan* (刀馬旦/武旦), *hualian* (花臉), *laodan* (老旦), *laosheng* (老生), *qingyi/huadan* (青衣/花旦), *xiaosheng* (小生), and *wusheng* (武生). POGE implements a progressive approach to examination design. Singing and reciting are the primary test items in the elementary grades (i.e., Grade 1-2). The holistic performance of Peking opera dominates the advanced grades (i.e., Grade 7-8), and is characterized by an increasing requirement for progression in the four major skills (NPOC, n.d.).

Analyzing the National Peking Opera grading system was of significance to the proper exploration of Cantonese opera assessment dimensions and assessment content.

Cantonese operatic singing graded examination in Hong Kong

The Cantonese Opera Graded Exam, which is held by the University of West London and the Hong Kong Association of Cantonese Opera Scholars, made its debut in Hong Kong in July 2018. The Cantonese Opera Graded Exam in Hong Kong is still being improved. The exam is divided into eight grades. A syllabus of the Graded Examination in eight grades have been published (Hong Kong Association of Cantonese Opera Scholars, n.d.).

The content analysis of the two subjects, POGE and COSGEHK, was focused on the essential components of graded examinations: 1) grade division, 2) test items, and 3) rubrics. Rubrics consist of three basic components: criteria, standards, and a scoring strategy (Reddy & Andrade, 2010). Consequently, three components will be compared and discussed separately in this section.

4.1.1 Content analysis of POGE

4.1.1.1 Grade division

POGE is divided into nine grades. Grades 1~3 are subsumed into the elementary level, Grades 4~6 are subsumed into the medium level, and grades 7~9 are subsumed into the advanced level. Apart from the progressive difficulty, POGE demonstrates no concrete and consistent requirement and reflection of these three levels in other aspects such as test items

and standards. Further exploration of this issue was thought necessary for subsequent interviews.

4.1.1.2 Test items

A test item is a specific task that candidates are asked to perform (Research & Theory, n.d.). The test items in POGE are unfixed and vary arbitrarily according to distinct roles and grades.

In general, there are ten kinds of test items in POGE: 1) singing, 2) reciting, 3) performance of the selected piece, 4) knowledge of Peking opera, 5) acting movement and acrobatic fighting, 6) basic skills, 7) posture and fighting with weapons, 8) imitated performance, 9) illustration of the role's costume and makeup, and 10) the elaboration of the skills that the role used in the play. The focus on the examination can be revealed by conducting frequency counting of these test items (see Table 4.1). Among all the test items in POGE with a total number of 150 (see Appendix G: Summary of Test Items of POGE), "singing," "reciting," and "inclusive performance" were the top three test items, which appeared 46, 43, and 25 times and accounted for 30.6%, 28.7%, and 16.7% of all test items, respectively. Interestingly, "singing" and "reciting" appear intensively in elementary grade tests, whereas "inclusive performance" is examined in the advanced grades.

Table 4.1. The result of frequency count in test items of POGE

Test items	Frequency	Percentage
Singing	46	30.6%
Reciting	43	28.7%
Performance of the selected play	25	16.7%
Knowledge of Peking opera	15	10%
Acting movement and Acrobatic fighting	6	4%
Basic skills	6	4%
Posture and fighting with weapons	4	2.7%
Imitated performance	3	2%
Illustration of the role's costume and makeup	1	0.7%
The elaboration of the skills that the role used in the play	1	0.7%
Total	150	100%

The top three test items appear throughout the roles, reflecting a universally recognized emphasis on the singing, reciting, and holistic performance of selected repertoires. In moderate grades, varied test items were detected for separate roles, revealing dissention of the disposition in these grades. For example, "knowledge of Peking opera" is the most noticeable test item of this kind, and appears 15 times but only in three roles, *Wusheng* (武生 literally "military male role"), *Laosheng* (老生 literally "elder male role"), and *Daomadan/wudan* (刀馬旦/武旦 literally "military female role"). The way of evaluating the "knowledge" is also worth exploration.

Although "knowledge of Peking opera" is one of the test items in grades 3 to 9 in *Wusheng* (武生), no detailed information or content was provided. The specific content of "knowledge of Peking opera" in *Laosheng* and *Daomadan* is given in Table 4.2. Apart from sharing the

same questions with Grade 1, there are one or two questions in every grade, except for the

missing information in grades 8 and 9 in *Laosheng*, for unknown reasons.

Table 4.2. Translation of contents of "knowledge of Peking opera" in *Laosheng* and *Daomadan*

Grades	*Laosheng*	*Daomadan*
Grade 1	1. Knowledge about four major roles (四大行當)	
	2. Four skills and five methods (四功五法)	
Grade 2	The categories of *laosheng*[1].	What roles are included in *dan*?
Grade 3	The main *shengqiang* (聲腔) and *banshi* (板式) of *laosheng* (老生).	What are the four commonly used *luogu* (鑼鼓) in *daomadan*?
Grade 4	What are the 13 categories of rhyme (字韻的十三徹)?	Illustrate the costume of the play "*Muguiying* (穆桂英) and *Husanniang*" (扈三娘)
Grade 5	1. The *luogu* (鑼鼓 percussion) used before singing. 2. Four commonly used *qupai* (曲牌)	What are the function and impacts of "*qiba* (起霸) [2]"?
Grade 6	Three major presentation features and ten colors in *Xiqu*.	Introduce the play of "*Dachushou* (打出手)" that you know.
Grade 7	Introduce the skills of the role that you impersonated.	Introduce the role analysis and technical key points of the role that you impersonated.
Grade 8	Not stated.	What are the features of performance in *Xiqu*?
Grade 9	Not stated.	Why is *kung fu* recognized as the premise of creating a role?

The breadth of knowledge is questioned due to the limited number of question(s) examined

at every grade. But the questions in *Daomadan* present a progressive tendency from concrete

---

[1] This is translated directly from the original text: "京劇常識：老生的分類".

[2] One of the patterns of *Xiqu* performance.

to abstract and from detailed to macroscopic comprehension of *Xiqu*, which reflects the depth of knowledge.

Besides "knowledge of Peking opera," the occurrence frequency of the remaining test items examined in moderate grades is restricted to single digits. What should be assessed in moderate grades needed to be discussed?

In general, the disposition of test items in every role demonstrates an increasing tendency of complexity and difficulty: from basic skills, such as "singing" and "reciting," to moderate levels such as "imitated performance", and eventually to "performance of selected excerpt", which synthesizes of four major skills of *Xiqu*. It accords with the hierarchy of the learning process in the cognitive, psychomotor, and behavioral domains. Based on the theoretical framework of the current study, and to examine whether the criteria of POGE are valid, the theoretical framework presented here, which is based on the taxonomy of learning, will be introduced into this content analysis. Codes are defined in the 20 elements of four diverse domains in the taxonomy of learning shown in Table 4.3.

Table 4.3. Codes for content analysis in criteria

Behavioral Domain	Affective Domain	Cognitive Domain	Psychomotor Domain
B1: Acquire	A1: Receive	C1: Remember	P1: Perceive
B2: Assimilate	A2: Respond	C2: Understand	P2: Simulation
B3: Adapt	A3: Value	C3: Apply	P3: Manipulate
B4: Perform	A4: Believe	C4: Evaluate	P4: Articulate
B5: Aspire	A5: Behave	C5: Synthesize	P5: Mastery

Based on the connotation of each hierarchy in the behavioral domain, coding can be conducted as below. This has been reviewed by the author's supervisors:

"Singing" and "reciting" belong to the behavioral domain and rank in the third level among the five hierarchies, coded as B3.

"Imitated performance" is a test item that requires students to imitate a specified performance, in which they are supposed to present a series of abilities: acquiring information and content from the specified performance (coded as B1), assimilating their acts in accordance with what they acquire from the specified performance (coded as B2) by adapting their existing knowledge and skills (coded as B3) to re-perform (coded as B4).

"Performance of selected piece" is a test item used in advanced grades. Students are expected to show their advanced knowledge of practice, proficiency of skills, and internalization of sensitivity, creativity, and artistry, which was coded as B5.

The corresponding relationships between coding and test items in POGE are illustrated in Table 4.4.

Table 4.4. Coding of test items in POGE

Test items	Coding
Singing	B3
Reciting	B3
Performance of the selected piece	B5
Knowledge of Peking opera	C1
Acting movement and Acrobatic fighting	P3
Basic skills	P3
Posture and fighting with weapons	P3
Imitated performance	B1, B2, B3 and B4
Acting movement imitation.	P1 and P2
The elaboration of the role's costume and makeup	C4
The elaboration of the skills that the role used in the play	C4

The involved coding and the occurrence frequency derived from running coding in NVivo

12 are given in Table 4.5.

Table 4.5. The involved components and the occurrence frequency in POGE

Coding	Occurrence Frequency
B1	4
B2	4
B3	83
B4	6
B5	12
C1	15
C4	2
P1	1
P2	1
P3	25

For convenience of analysis, the ranking and the number of coding references are illustrated in Figure 4.1.
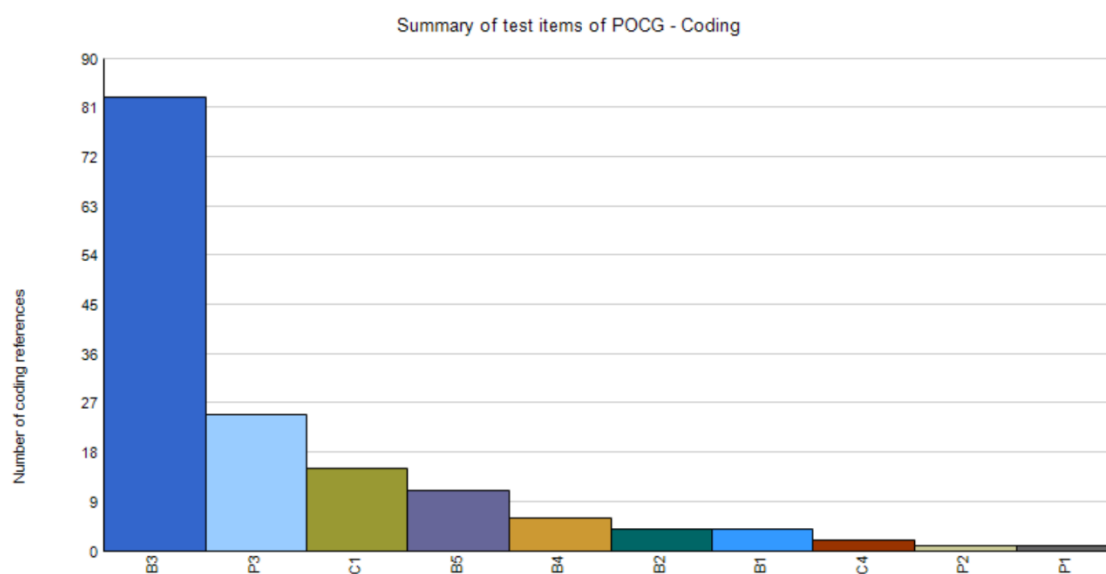
Figure 4.1. The ranking and the number of coding references

Using Table 4.3, the involved components and the occurrence frequency in POGE were revealed. Eight components from three domains out of four in the theoretical framework were covered, including behavioral, cognitive, and psychomotor domains. The behavioral domain gets full coverage from the lowest to highest hierarchical components. In the cognitive domain, the range shrinks to two components. One is in the lowest hierarchy, the other is in the fourth hierarchy of components. Three hierarchies in the psychomotor domain are incorporated. By contrast, not a single hierarchy in the affective domain was embodied in the test items.

As shown in Figure 4.1, the most remarkable component is B3 ("Adapt" in the behavioral domain) with 83 occurrences, exceeding half of the full coverage of all components in the three domains. Although P3 ("Manipulate" in the psychomotor domain) is the second most frequently tested component. Its number of occurrences shrinks sharply to 25. And C1

("Remember" in the cognitive domain) occurred 15 times. The remainder recorded single-digit occurrences.

Based on the above analysis, several characteristics are revealed:

Firstly, from the perspective of coverage in breadth and depth, POGE fails to embody any component of the affective domain. However, this domain is vital to student growth and development in music (Hanna, 2007). The most widely used assessment instruments for the affective domain are self-reported scales and observational checklists in formative assessment (Cate, & De Haes, 2000). As a kind of summative assessment, it might be difficult to implement assessment in the affective domain by using a graded examination. Hence, further exploration of the formation of this domain constitutes a gap that needs to be filled.

For the three domains involved in the current study (apart from the behavioral domain, which all the hierarchies covered), there are omissions in the cognitive domain and the psychomotor domain, partly due to the disposition of the test items. For example, "knowledge of Peking opera" concentrates on examining definitions that are subsumed into "understanding" (coded C1). Candidates do not have an opportunity to present their individual opinion or statements until Grade 9, which is subsumed into "evaluation" (coded C4) in the cognitive domain. Another example lies in the psychomotor domain. It is embodied in the test item "imitation", which reflects "perceive", "simulate", and "manipulate" within this domain. Adjusting the test items and how questions are asked may aid full coverage and balance in the dispositions of examination content, which might require a widely accepted assessment framework to regulate. This is the kernel of the current study.

Only one or two knowledge points are asked at each grade of POGE, which is too limited to reliably check relevant knowledge. Increasing the number of knowledge points might better examine candidates' "knowledge of Peking opera".

4.1.1.3 Rubrics

POGE provides no complete rubrics to the public, as some of the essential components are missing. Acknowledging the limited information disclosed by POGE to the public, the components of its rubrics can be classified as below:

Criteria

The criteria published by POGE are non-uniform and incomplete, which can be divided into three kinds:

1. Without any criteria. Such as the criteria of *xiaosheng* (小生), *qingyi/huadan* (青衣/花旦), *daomadan/wudan* (刀馬旦/武旦), and *laodan* (老旦).

2. With some criteria. Such as the criteria of *laosheng* (老生), *hualian* (花臉), and *chouhang* (丑行).

1) *Laosheng* (老生): Only the first two grades are provided with criteria, whereas the rest are missing:

Grade 1: Accurate in pitch and rhythm (*banyan*) (音準、節奏（板眼）準確).

Grade 2: a) sing or speak in a clear and rich tone (字正腔圓), b) Accurate in rhythm.

2) *Hualian* (花臉) lacks criteria except for the first two grades. Grades 1 and 2 share the same criteria: "Accurate in pitch and rhythm (*banyan*) (音準、節奏（板眼）準確)"

3) *Chouhang* (丑行): Various criteria are assigned to every grade listed in Table 4.6.

Table 4.6. Criteria of *chouhang* (丑行)

Grade	Criteria
Grade 1	Reciting: pronunciation and articulation, emotion and rhythm Singing: pronunciation, articulation, rhythm, and *xingqiang* (行腔melismatic singing)
Grade 2	Reciting: the standardized score of posture and movement are included in the content
Grade 3	The score of rhythm, facial expression, characteristics, the feature of the role in performance are included
Grade 4	Singing, reciting, acting movement, acrobatic fighting, hand gesture, expression in the eyes, gesture and movement, hair wiping, and the gait should achieve the standard
Grade 5	Reciting: pronunciation and articulation, expression and rhythm Singing: pronunciation, articulation, rhythm, and *xingqiang* (melismatic singing)
Grade 6	Singing, reciting, acting movement, and acrobatic fighting are required to be: 1) integrated both internally and externally, 2) technically achieve the standard, 3) in a humorous style, and 4) without offensiveness nor a bad taste
Grade 7	Singing, reciting, somersault, dance, acrobatic fighting, bearing, emotion and rhythm (念、做、翻、舞、打、神、情、節奏)
Grade 8	Singing, reciting, somersault, dance, acrobatic fighting, bearing, emotion and rhythm (念、做、翻、舞、打、神、情、節奏)
Grade 9	Nil.

With this limited information, problems are revealed in three roles:

1) In the criteria of *laosheng* (老生), terms are arbitrarily used. "字正腔圓" and "音準、節奏" descriptions in Eastern and Western culture, respectively.

2) In the criteria of *hualian* (花臉), the first two grades revealed the same criteria. Without standards, no discrimination can be detected between these two hierarchical grades.

Furthermore, confusions occurred between some terms in the criteria of both *laosheng* (老生) and *hualian* (花臉). Such as, "節奏（板眼）準確". *Banyan*（板眼）refers to the pulse in *Xiqu* instead of rhythm (節奏).

3) In the criteria of *chouhang* (丑行), more characteristics can be perceived:

a) Descriptions of criteria in grades 1 and 5 are duplicated.

b) Some of the descriptions can only be recognized as assessment orientation rather than normative criteria. These include the descriptions in grades 2, 3, 4, 7, and 8.

c) Standards are missing in most of the grades. The descriptions of grades 1~5 and grades 7~8 delineate criteria but omit standards. Both candidates and assessors know what to do, but do not know to what extent they are required to do so.

Identically to the first kind, this kind of criterion also fails to fulfill the requirement of quality assessment.

3. With complete criteria, such as the criteria of *wusheng* (武生) listed in Table 4.7.

Table 4.7. Criteria of *wusheng* (武生)

Grade	Criteria
Grade 1-3	Singing skills: Accurate in pitch, sing or speak in a clear and rich tone, fluent in singing, as well as express emotion through voice.
	唱功：音準、字正腔圓、行腔流暢、以聲寄情。
	Reciting skills: clear articulation, mature and smart performance, and rhythmic accuracy.
	念功：吐字清楚、醇厚挺拔、節奏分明。
Grade 4-6	Singing is expected to be 1) accurate in pitch, 2) in a clear and rich tone, 3) with appropriate energy, 4) with even breath, and 5) with appropriate expression.
	唱功：音準、字正腔圓、勁頭適度、用氣均勻、韻味醇厚。
	Reciting is expected to be 1) clearly articulated, 2) with dulcet timbre, 3) robust and ringing, and 4) sonorous and forceful.
	念功：吐字清晰、音色動聽、挺拔清脆、鏗鏘有力。
Grade 7-9	Singing and reciting are expected to be 1) flowing and smooth, 2) dulcet and beautiful, 3) with inflection and intonation, and 4) vary in tempo and dynamic.
	唱念：行雲流水、悅耳動聽 、頓挫抑揚、疾徐輕重。

The above criteria feature in:

1. Explicit distinctions of components in *Xiqu* performance.

2. Most of the descriptions of criteria and standards conform to the Chinese culture and aesthetics, which conveys the requirement of *Xiqu* more vividly and faithfully.

3. Explicit distinctions of the standard levels in the preliminary (grades 1~3), medium (grades 4~6), and advanced (grades 7~9) levels, which manifest progressive requirements.

4. Measurable or descriptive subcomponents are proposed for each component. For example, in "singing skills", candidates are required to be accurate in rhythm, sing or speak in a clear and rich tone, be fluent in singing, and express emotion through voice. These cover every aspect of this domain, including technique, musicality, and emotion.

Despite the above outstanding features, two flaws cannot be neglected:

1. There are no distinctions of the standard levels in every grade of the preliminary, medium, and advanced levels.

2. Some of the descriptions, such as "醇厚挺拔" (mature and smartly performed), are vague and are apt to be ambiguous. Different assessors might have a different understanding of such words. Adopting words and terms of consensus, or clarifying the meaning of words and terms, is imperative. This is discussed in the interviews.

Standards

POGE employs three-level standards: distinction, pass, and fail, which may not sufficiently discriminate between student performance. Because "distinction" and "pass" do

not constitute a continuum, there should be more detailed division. For example, "merit" is usually used between "pass" and "distinction".

Scoring strategy

POGE reveals no scoring strategy to the public at all.

4.1.2 Content analysis of COSGEHK

4.1.2.1 Grade division

COSGEHK divides its graded examinations into eight grades for two reasons. The eight-grade system is widely accepted in the western world and is internationally recognized (HKACOS, n.d., p.3). However, only implicit progressive distinctions are shown between grades in COSGEHK.

4.1.2.2 Test items

COSGEHK presents a unified and normative design of test items, including singing and reciting performances, which consists of *Shuobai* (說白), *Shuochang* and *Qupai* (說唱與曲牌), and *Banqiang* (板腔), a Q&A of basic knowledge in Cantonese opera, sight-singing, and an aural test.

1. Content analysis under the theoretical framework of the current study.

According to "codes for content analysis in criteria" in Table 4.3, coding can be conducted using Table 4.8.

Table 4.8. Coding of test items in COSGEHK

Test items	Coding
Singing and Reciting	B3
Q&A of the basic knowledge in Cantonese opera	C1, C2, and C3
Sight-singing	C1, C2, C3, B1, B2 and B3
Aural test	B1, B2 and B3

The specific practice of each test item plays a vital role in the coding. "Singing" and "reciting" the optional piece from the repertoire require the embodiment of the skills adaption that belongs to "adapt" in the behavioral domain. In the "Q&A of the basic knowledge in Cantonese opera", candidates are required to answer questions raised by the examiners from the specified terminologies of every grade, including brief explanations of the terms and theories of Cantonese operatic singing relevant to the piece performed by the candidate. From Grade 6, reading the work in *zhongzhou qiang* is also embraced as a requirement. Therefore, this test item involves "remember" (C1), "understand" (C2), and "apply" (C3) in the cognitive domain. In "sight-singing", candidates are expected to sing a melody after reading a *gongchepu* that they have not seen before. This embraces previous knowledge of *gongchepu,* which is subsumed into "remember" (C1) and "understand" (C2) in the cognitive domain, and sing on sight according to the *gongchepu,* which involves "acquire" (B1), "assimilate" (B2) and "adapt" (B3) in the behavioral domain, and "apply" (C3) in the cognitive domain. Finally, in the aural test, candidates are required to hum the melody played by an instrumentalist, with *fakou* (發□) chosen by the candidate, and identify the name of *luogudian* (鑼鼓點) played by an instrumentalist (exclusively for grades 3 to 8). Hence, the aural test is believed to be relevant to "acquire" (B1), "assimilate" (B2), and "adapt" (B3) in the behavioral domain.

By conducting a content analysis, it became evident that COSGEHK concentrates on the first three hierarchies of the behavioral and cognitive domains, and did not embrace the psychomotor or affective domains.

2. Coverage of breadth and depth of knowledge.

Among all the test items in COSGEHK, "Q&A of the basic knowledge" is noticeable, and is also examined in some of the roles of POGE. COSGEHK emphasizes "knowledge" and examines it from Grade 1 to Grade 8. From grades 1 to 5, candidates are required to give the definitions of the terms used in Cantonese operatic singing theory that are relevant to the piece they performed. Based on these requirements, candidates of grades 6 to 9 are required to read out the specific terms in *zhongzhou qiang*.

The concrete content of "knowledge of Cantonese opera" has been translated and is listed in Table 4.9.

Table 4.9. The translation of contents of "knowledge of Cantonese opera"

Grades	COSGEHK	
	Pinghou (平喉)	*Zihou* (子喉)
Grade 1	1. *Ding* (叮) and *ban* (板)	
	2. *Luogu* (锣鼓)	
	3. *Shuobai* (说白)	
	4. *Bailan* (白榄)	
Grade 2	1. *Changqiang* (唱腔)	
	2. *Shibai* (詩白)	
	3. *Zhengban* (正板) and *Diban* (底板)	
	4. *Zhengding* (正叮) and *Diding* (底叮)	
Grade 3	1. *Yibanyiding* (一板一叮)	

	2.	*Yibansanding* (一板三叮)
	3.	*Liushuiban* (流水板)
	4.	*Sanban* (散板)
	5.	*Jibenzi/Zhengzi/Shizi* (基本字/正字/實字)
	6.	*Miezaizi* (孭仔字)
	7.	*Koubai*(口白)
	8.	*Tuobai* (托白)
	9.	*Gongchepu*(工尺譜)
	10.	*Banmian* (板面)
	11.	*Guoxu*(過序)
	12.	*Tutou/ Zhizhuan/ Tuxu*(禿頭、直轉、禿序)
Grade 4	1.	*Xiankou* (線口)
	2.	*Zhengxian* (正線/合尺線）
	3.	*Fanxian* (反線/上六線）
	4.	*Yifanxian* (乙反線/乙反調式）
	5.	*Yinbai* (引白)
	6.	*Langlibai* (浪裡白)
	7.	*Yunbai* (韻白)
	8.	*Jie* (介)
	9.	*Lingchang* (另場)
Grade 5	1.	*Dahou/Baqiang* (大喉/霸腔)
	2.	*Guanghua* (官話/中州音)
	3.	*Guqiang* (古腔/唱官話）
	4.	*Bangqiang* (幫腔/合唱/齊唱）
	5.	*Bangqiangti* (板腔體/曲牌體)
	6.	*Dadiao/Xiaodiao* (大調/小調)
	7.	*Baodu* (爆肚)
	8.	*Yizixingqiang* (依字行腔)
	9.	*Wenziluoqiang* (問字攞腔)
	10.	*Fakou* (發口/法口)
	11.	*Zhuanfakou* (轉發口/轉法口)
	12.	*Shengmu/Yunmu* (聲母/韻母)
	13.	*Pingsheng/Zhesheng* (平聲、仄聲)
	14.	*Shangju/Xiaju* (上句、下句)
	15.	*Shou/shoujue* (收、收掘)

	16. *Ju or dun in Bangqiang* (板腔之句、頓)
Grade 6	1. Yingtou (影頭)
	2. *Shouying* (手影)
	3. *Zhuangban* (撞板)
	4. *Shixian* (食線)
	5. *Xiandi/Xianmian* (線底/線面)
	6. *Yayun* (押韻)
	7. *Shengdiao: nine tones in Cantonese* (聲調：粵語九聲)
	8. *Shuochang* (說唱)
	Other requirements (read in *Zhongzhouyin/Guanghua* (以中州音/官話唸出) ：
	Gongzhu (公主, literally "princess"), *Dangzhen* (當真, literally "are you sure?"), *Guoran* (果然), *Xiaqu* (下去), *Zhidao* (知道), *Youli* (有禮), *Baliao*(罷了), *Zhege* (這個), *Haodan* (好膽), *Jiangjun* (將軍), *An* (俺), *Chuan* (傳)
Grade 7	1. *Jiaotou* (叫頭, literally "calling")
	2. *Diao* (調, literally "pitch")
	3. *Diaoshi* (調式, literally "mode")
	4. *Qushi* (曲式, literally "musical form")
	5. *Zhuanquzhuanqiang* (專曲專腔, literally "specific tunes from the specific piece")
	6. *Pengmian* (棚面, literally "surface of the awning" which refers to the music ensemble)
	7. *Getan* (歌壇, literally "singing clubs")
	8. *Geling* (歌伶, literally "singers")
	Other requirements (read in *Zhongzhouyin/Guanghua* (以中州音/官話唸出) ：
	Huilailiao (回來了), *Rucishuo* (如此說), *Wansui* (萬歲), *Ganfan* (敢煩), *Lingming* (領命), *Shenme* (甚麼), *Dawang* (大王), *Qingzuo* (請坐), *Aijia* (哀家), *Liangliang* (娘娘), *Gaotui* (告退), *Laofu* (老夫)
Grade 8	1. Yingxiongbai (英雄白)
	2. *Qiangkou* (腔口)
	3. *Xueqiang* (薛腔)
	4. *Maqiang* (馬腔)
	5. *Xingqiang* (星腔)
	6. *Fanqiang* (凡腔)
	7. *Xinmaqiang* (新馬腔)
	8. *Fangqiang* (芳腔)
	9. *Nvqiang* (女腔)
	10. *Kaishan* (開山)

11. *Shouben* (首本)

12. *Shigongxian* (士工線/調弦)

13. *Chewuxian* (尺五線/調弦）

Other requirements (read in *Zhongzhouyin/Guanghua* (以中州音/官話唸出) ：

Kanjiangqilai (看將起來), *Yaochabuyao* (要茶不要), *Budeweiling* (不得違令), *Niyaoting* (你要聽), *Niyaojiang* (你要講), *Kenaoye* (可惱也), *Fumaye* (駙馬爺), *Najiulai* (拿酒來), *Buhaoliao* (不好了)*, Shabailiao* (殺敗了), *Daihouzhe* (侍候著), *Nalizou* (哪裡走), *Yayimen* (衙役們), *Suiwolai* (隨我來), *Xiezhushang* (謝主上), *Shizuiliao* (食醉了), *Wenfangcihou* (文房侍候)

Unlike the extremely limited knowledge points at every grade of POGE, the amount of knowledge evaluated in COSGEHK is greater, reflecting a broader coverage. Candidates in COSGEHK are randomly spot-checked against the examiner's prescribed knowledge listed in the above table.

In addition to the breadth, coverage of the depth of knowledge is worthy of note. Krathwohl (2001) proposed six increasingly complex levels of the cognitive process of knowledge: remember, understand, apply, analyze, evaluate, and create. As learning progresses, students are expected to go beyond factual knowledge and be assessed on more than simply their ability to recall or recognize (Lambert & McCombs, 1998). Therefore, as learning progresses, assessment in high grades should emphasize a higher hierarchy of the cognitive process of knowledge. Although COSGEHK covers more relevant knowledge in every grade than POGE, its modes of testing reveal no hierarchical differences between ascending grades.

4.1.2.3 Rubrics

COSGEHK publishes relatively complete rubrics, including criteria, standards, and scoring strategies. These are represented separately below (see Tables 4.10-4.16). These tables were translated from an original text in Chinese. They were retained in the original format so that the content and presentation mode could be analyzed. To ensure accuracy, an expert translator checked the translation.

Table 4.10. Rubrics of Grade 1~4 (Hong Kong Association of Cantonese Opera Scholars, p. 21)

Distinction (85-100%): Candidates are required to show sophisticated singing skills, high accuracy, fluency, excellence, and stability in timbre, pitch, and articulation, an outstanding sense of *Xiqu,* and mode of interpretation, as well as confidence and certainty in singing. A complete understanding of *Gongchepu* is also required. With the ascending of grades, personal style, perceptions in interpretation, and the ability in expression and charms to the audiences are expected.	Merit (75-84%): Candidates are required to show some extent of singing skills, accuracy, fluency, excellence, and stability in timbre, pitch, and articulation, appropriate sense of *Xiqu,* and mode of interpretation as well as confidence in singing. A thorough understanding of *Gongchepu* is also required. With the ascending of grades, personal style, perceptions in interpretation, and the ability in expression and charms to the audiences to some extent are expected.	Pass (65-74%): Candidates are required to show some singing skills and reasonable fluency but with some hesitation. The timbre, pitch, and articulation must meet the acceptable standards in general. An appropriate sense of *Xiqu* and mode of interpretation should be detected occasionally. The singing reflects slight confidence. A basic understanding of *Gongchepu* is also required. With the ascending of grades, slight personal style, and perceptions in interpretation, as well as the ability in expression and charms to the audiences are expected.
Fail (approximate to pass) (55-54%): Singing is partially correct but without influence and accuracy in singing skills. Candidates make a great effort to meet the standards in timbre, pitch, and articulation. The sense of *Xiqu* and the mode of interpretation need improvement. Little understanding of *gongchepu*. Both	Fail (0-54%): Fail to present fluent singing, within which more incorrectness than correctness is detected. The performance in timbre, pitch, and articulation are far below the requirement of a certain grade. No fully understanding of *gongchepu*. And lack of personal style and perceptions of interpretation.	

personal style and perceptions of interpretation fail to standards. No charms to the audience can be detected.	

Table 4.11. Rubrics of Grade 5~8 (Hong Kong Association of Cantonese Opera Scholars, p. 22)

Distinction (85-100%): Candidates are required to show sophisticated singing skills, extremely high accuracy, fluence, excellence, and stability in timbre, pitch, and articulation, mature sense of *Xiqu* and mode of interpretation, as well as confidence and certainty in singing, showing fully control of the overall interpretation on the performance. An all-sided understanding of *Gongchepu* is also required. Within grades 5~8, candidates are expected to convey a personal style pertaining to the performance and present abilities to reach for audiences and strike a chord.	Merit (75-84%): Candidates are required to show good singing skills, accuracy, fluency, excellence, and stability in timbre, pitch, and articulation, appropriate sense of *Xiqu,* and mode of interpretation, as well as confidence in singing. A good understanding of *Gongchepu* is also required. Within grades 5~8, candidates are expected to convey a personal style that pertains to the performance, and present abilities in reaching for audiences, and strike a chord to some extent.	Pass (65-74%): Candidates are required to show certain singing skills, fair accuracy, and reasonable fluency but with some hesitation sometimes. The timbre, pitch, and articulation should meet the acceptable standards in general. An appropriate sense of *Xiqu* and mode of interpretation can be detected now and then. The singing reflects fair confidence. A certain understanding of *Gongchepu* is also required. The candidates are expected to convey the elements of *Xiqu* as well as the abilities in reaching for audiences and strike a chord
Fail (approximate to pass) (55-54%): Singing is partially correct but without influence and accuracy in singing skills. Candidates need to strive for achieving the standards in timbre, pitch, and articulation. The sense of *Xiqu* and the mode of interpretation need improvement. Only a little understanding of *gongchepu* is noted. Limited abilities in conveying *Xiqu* elements. And no competence in reaching for the audience can be detected.		Fail (0-54%): Fail to present fluent singing, within which more incorrectness than correctness is detected. The performance in timbre, pitch, and articulation are far below the requirement of a certain grade. No full understanding of *gongchepu*. And lack of personal style and perceptions of interpretation, as well as the competence in reaching for the audience.

Furthermore, COSGEHK published a series of criteria and the corresponding weighting, as shown in Table 4.12.

Table 4.12. The global weighting of COSGEHK (Hong Kong Association of Cantonese Opera Scholars, p. 12)

Singing	Q&A	Sight-singing	Aural test
80%	10%	4%	6%

Table 4.13. Criteria and weightings for Singing in COSGEHK (Hong Kong Association of Cantonese Opera Scholars, p. 15)

Criteria	Weightings			
	Grades 1~2	Grades 3~4	Grades 5~6	Grades 7~8
Singing skills: <u>Candidates are required to be skillful in</u> <u>*changqiang* (唱腔 literally "sung melody"),</u> conforming to the scores and showing 1) accuracy in pitch, pulse, rhythm, projection, and articulation, 2) the control of breath, dynamics, tempo, timbre, and 3) high/low pitches, as well as 4) the fluency and stability of singing.	87.5%	75%	65%	56.25%
Interpretation (sense of *Xiqu*): <u>Candidates are required to present</u> personal interpretation technique when singing, exhibiting <u>exquisite</u> emotion and charm that evokes echoes and present various characteristics of *shuobai* (說白 literally "speech"), *qupai* (曲牌 literally "fixed tunes") and *qushi* (曲式 literally "musical form")	6.25%	12.5%	12.5%	16.25%
Knowledge in Cantonese operatic singing: Candidates are required to manifest <u>some understanding in the styles of scores, the</u> <u>mode of expression, and musical symbols in the scores, such as</u> <u>*luogudian* (鑼鼓點 literally "percussion patterns"), *jiekou* (介口</u> <u>literally "movements"), and so on.</u>	3.25%	7.5%	12.5%	15%
Expression ability: <u>Based on the essentials of the above three</u> <u>criteria, candidates are required to</u> manifest an explicit conception of how to express emotions and appeal to the audience, as well as present an appropriate presence.	2.5%	5%	10%	12.5%
Total	100%	100%	100%	100%

Table 4.14. Criteria and weightings for Q &A in COSGEHK (Hong Kong Association of Cantonese Opera Scholars, p. 16)

Criteria	Weightings			
	Grade 1~2	Grade 3~4	Grade 5~6	Grade 7~8
Knowledge in Cantonese operatic singing: Candidates are required to provide correct answers <u>when asked by the examiners</u>, presenting comprehension of musical symbols in the scores, <u>such as *luogudian* (鑼鼓點)*, jiekou* (介口)*, and so on*</u>, as well as the influences that come from this knowledge. Candidates of grades 5~8 are required to understand the play's background and its scenario.	100%	100%	60%	60%
Interpretation (sense of *Xiqu*): exclusive for the candidates of grades 5~8 only. Candidates are required to show appropriate understanding and sensitivity to the play that is performed, <u>including the symbols in the scores, style, and holistic performance.</u>	×	×	20%	20%
Expression ability: exclusive for the candidates of grades 5~8 only. Candidates are required to demonstrate accurate and appropriate expression ability, showing the musical symbols, forms, and style that they grasp.	×	×	20%	20%
Total	100%	100%	100%	100%

Table 4.15. Criteria and weightings for sight-singing in COSGEHK (Hong Kong Association of Cantonese Opera Scholars, p. 17)

Criteria	Weightings			
	Grade 1~2	Grade 3~4	Grade 5~6	Grade 7~8
Singing skills: Candidates are required to sing out the *gongchepu* (工尺譜) that has not been previewed clearly, fluently, and steadily after a short preparation, and count the *dingban* (叮板) simultaneously. The number of *ban* (板) cannot exceed two in grades 1~2 and four in grades 3~8. And the candidates in grades 5~8 are required to sing out the melody and lyrics and count the *dingban* (叮板) simultaneously.	50%	50%	25%	25%
Interpretation (sense of *Xiqu*): Candidates are required to present personal interpretation technique when singing, exhibiting exquisite emotion and charm, which evokes, echoes and presents various characteristics of *shuobai* (說白)*, qupai* (曲牌) *and qushi* (曲式)	×	×	25%	25%
Knowledge in Cantonese operatic singing: Candidates are required to manifest some understanding in the styles in scores, mode of expression, and musical symbols in the scores, such as *luogudian* (鑼鼓點)*, jiekou* (介口), and so on.	25%	25%	25%	25%
Expression ability: Based on the essentials of the above three criteria, candidates are required to manifest an explicit conception of how to express emotions and appeal to the audience, as well as present an appropriate presence.	25%	25%	25%	25%
Total	100%	100%	100%	100%

Table 4.16. Criteria and weightings for the aural test in COSGEHK (Hong Kong Association of Cantonese Opera Scholars, p. 18)

Criteria	Weightings			
	Grade 1~2	Grade 3~4	Grade 5~6	Grade 7~8
Sense of *Xiqu*: exclusive for the candidates of grades 5~8 only. Candidates are required to present an interpretation that conforms to the sense of *Xiqu*.	×	×	33%	33%
<u>**Knowledge in Cantonese operatic singing:**</u> 1. Candidates are required to hum (adopting the projection of one's choice, such as "ya", "la", "wu" and so on) the melody played by the instrumentalist, retain the accuracy of the pitch and rhythm, as well as the stability of the pulse by counting the *dingban* (叮板). Every excerpt will not exceed two *ban* (板). 2. Candidates are required to distinguish *luogudian* (鑼鼓點) according to the excerpt played by the instrumentalist. Candidates of grades 3~5 are required to distinguish the *luogudian* (鑼鼓點) according to the excerpt played by the instrumentalist, whereas those of grades 6~8 are expected to indicate the name of the *luogudian* (鑼鼓點) and recite the pithy formula of the *luogudian* (鑼鼓點).	100%	100%	67%	67%
Total	100%	100%	100%	100%

Although the above criteria, standards, and rubrics are relatively complete, flaws can be detected with scrutinization:

1. Flaws in criteria

Clear assessment criteria not only enable learners to know what they are expected to do to demonstrate achievement of the learning outcomes but also grounds to make judgments about student performance (Armstrong et al., 2008). Hughes (2007) proposed the following, which facilitate the formulation of clear criteria:

1) Try to keep descriptions of the criteria as concise and neutral as possible.

2) Use concrete verbs that refer to observable behaviors that will signal to students more clearly what they should do.

3) If possible, separate criteria so that each deal with only one behavior.

4) Remove unnecessary detail (detail will be added to the standards).

5) Limit the number of criteria to a manageable number. Be realistic about how many behaviors students can competently demonstrate in an assessment task and how many criteria assessors can juggle when grading.

According to the above notes, some flaws in the criteria that deviate from Hughes' guidance can be detected:

1) Unconcise and promiscuous description

Some of the descriptions in the criteria are not concise and are even mixed with standards or instructions relating to the test items. These are underlined in Tables 4.13-4.16. As indicated in the literature review, assessment criteria are descriptive statements that provide learners and instructors with information about different elements in the assessment task.

2) Overlapping descriptions

a) In the test of singing, "manifest some understanding in the styles in scores, mode of expression and musical symbols in the scores, such as *luogudian* (鑼鼓點)*, jiekou* (介口), and so on" in the component of "Knowledge in Cantonese operatic singing" is an example of "conform to the scores" in the "singing skills" components. Besides, there is some overlap between the component of "Interpretation (sense of *Xiqu*)" and that of "expression ability".

Specifically, "exhibiting exquisite emotion" in the former and "manifest an explicit conception of how to express emotions" in the latter, are the same.

b) In the Q&A test, the same subcomponent, "show the understanding of the symbols in the scores as well as style", is present in both the components of "Interpretation (sense of *Xiqu*)" and "expression ability".

c) In the sight-reading test, the same subcomponent, "the ability that sings out the *gongchepu* (工尺譜) that has not been previewed", exists in both the components of "singing skills" and "expression ability".

3) Failure to align with the assessment objectives.

When assessment and the stated objectives are not well aligned, the validity of the assessment is questionable (Airasian & Miranda, 2002). At the same time, the requirement of "presenting the understanding of melody, rhythm, pitch, and tempo" in assessment objectives fails to embody the criteria.

2. Flaws in weightings

Weighting is an algorithm used to emphasize certain parts of assessments when calculating an overall score. A percentage or parameter is assigned to parts of the assessment based on their importance.

Although COSGEHK formulated four test items, singing, Q&A, sight-singing, and aural, aiming to examine all sides of a candidates' Cantonese operatic singing, improper weighting may significantly impair the significance, or even threaten the importance of the existence of other test items. For example, "singing", which has a weighting of 80%, exceeds the pass mark,

which is 64%. Thus, anyone who gets a score of 80 or above in "singing" will easily get a pass for the whole examination. This jeopardizes the significance of the other three test items.

3. Flaws in presentation

Alongside flaws in descriptions, the criteria, standards, and rubrics of COSGEHK are dispersed across different tables. Information is entwined in these tables, which increases the difficulties associated with reading and applying them.

4.1.2.4 Grading and evaluation

COSGEHK adopted five-level evaluations for global assessment, with explicit scoring for the gradings: "Distinction: 85~100%, Merit: 75~84%, Pass: 65~74%, Fail (but approximate to pass): 55~64%, and Fail: 0~54%".

4.1.3 Comparison of POGE and COSGEHK

4.1.3.1 Grade division

COSGEHK divides its graded examinations into eight grades. POGE formulated a hierarchy of nine grades and specified that grades 1~3 are subsumed into the elementary level, grades 4~6 are subsumed into the medium level, and grades 7~9 are subsumed into the advanced level. Neither examination board gives an explanation of the grounds for the grade divisions.

4.1.3.2 Test items

With regard to construction, COSGEHK publishes a unified and normative design of test items. In contrast, no regularity can be seen in the test items of dissimilar roles in POGE.

With regard to content, the test items in COSGEHK and POGE show distinctive characteristics of diversity and the breadth and depth of knowledge.

1. Diversity

COSGEHK and POGE present differences in the diversity of test items. COSGEHK have formulated a diversified design that consists of four items: singing, reciting, sight-singing, and aural testing. Contrarily, the test items in POGE are limited and unfixed, and vary arbitrarily according to separate roles and grades.

2. Coverage in the breadth of examination scope

Among all the test items, the formulation of "knowledge" is noticeable. COSGEHK emphasizes "knowledge" and examines it from Grade 1 to Grade 8. While in POGE, three out of eight roles include "knowledge of Peking opera" documented as a test item. They are *laosheng*, *daomadan*, and *wusheng*. Among these three roles, "knowledge of Peking opera" is restricted to some of the grades in *laosheng* and *wusheng*. Only *daomadan* has "knowledge of Peking opera" assessed from grades 1 to 9. The incomplete disposition of "knowledge of Peking opera" reflects the arbitrariness of the examination formulation.

3. Coverage in the breadth of operatic knowledge

Operatic knowledge is assessed more in COSGEHK than in POGE, reflecting broader coverage. Candidates in COSGEHK are randomly spot-checked for the examiner's prescribed

knowledge as listed in the above table. In the POGE, only one or two knowledge points are covered at every grade, which is too limited to check the relevant knowledge with validity. Therefore, candidates can easily recite any answer to one or two specified knowledge points. Under the system, knowledge of Cantonese operatic singing cannot be examined authentically and has become formalistic.

4. Coverage of the depth of operatic knowledge

In addition to breadth, coverage of the depth of operatic knowledge is also worthy of note. Krathwohl (2001) proposed six increasingly complex levels of the cognitive process of knowledge: remember, understand, apply, analyze, evaluate, and create. With learning progress, students are expected to go beyond factual knowledge and be assessed on more than simply recall or recognition (Lambert & McCombs, 1998). Therefore, the assessments for high grades should emphasize the hierarchy of the cognitive process of knowledge.

Although COSGEHK covers more relevant knowledge in every grade than POGE, its testing modes reveal no hierarchical differences in examinations. On the contrary, a remarkable pattern emerges in *daomadan* within the finite amount of content. Unlike in COSGEHK, the examination of relevant knowledge in *daomadan* exhibits a progressive vein: from detail to the entirety and from perception of the superficial layer to that of the profound layer. COSGEHK does, though, cover more relevant knowledge at every grade than POGE, but the testing mode reveals no hierarchical differences.

5. Coverage in the breadth of learning domains

By comparing the coding of the test items, it can be seen that POGE involves all the hierarchies of the behavioral domain (coded as B1~B5), the first and fourth hierarchy of the

cognitive domain (coded as C1 and C4), and the first three hierarchies in the psychomotor domain (coded as P1~P3). While COSGEHK involves the first three hierarchies in the behavioral domain (coded as B1~B3) and the first three hierarchies in the cognitive domain (coded as C1 and C3), but not a single hierarchy in the psychomotor domain. The affective domain is not involved in either POGE or COSGEHK. However, four components are covered in these two graded examinations, i.e., B1, B2, B3, and C1 (see below).
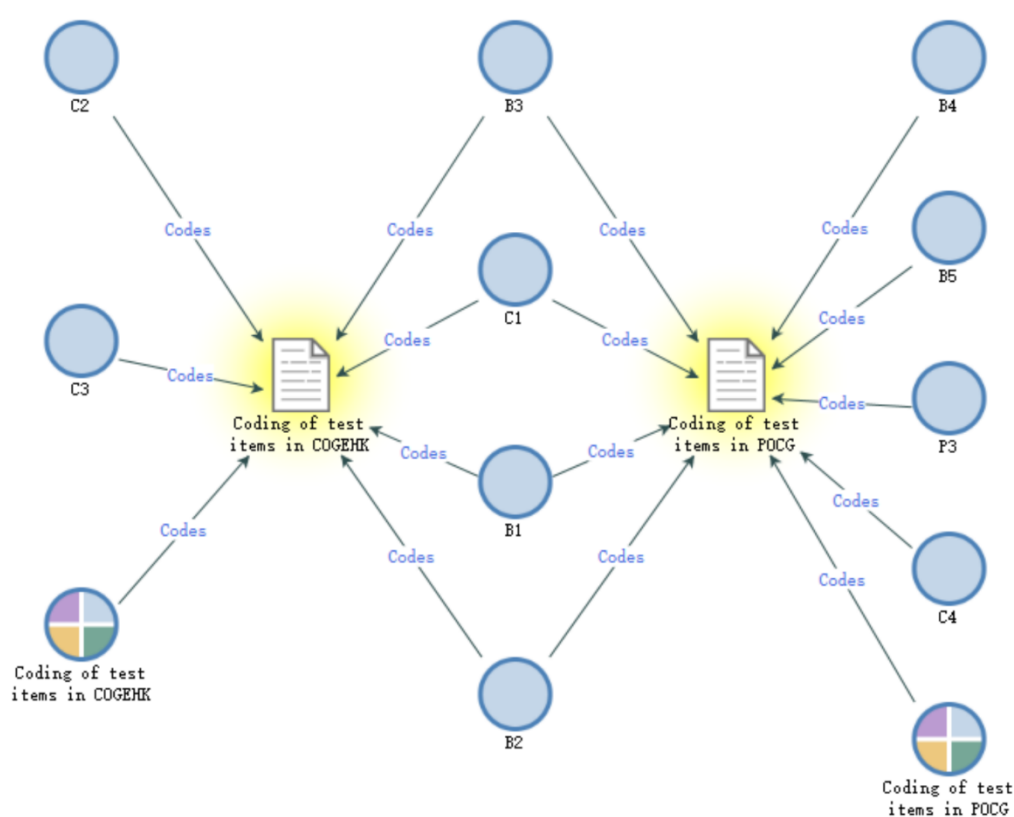
Figure 4.2. Diagram of components of the learning domains in POGE and COSGEHK

Besides the coverage, the depth of every component involved can be compared to reveal the foci of these two graded examinations. To ensure equipotent comparison, the original counting of occurrence frequency will be transferred into the ratio listed in Table 4.17.

Table 4. 17. The comparison figures of components of learning domains involved in POGE and COSGEHK

	B1	B2	B3	B4	B5	C1	C2	C3	C4	P1	P2	P3
COSGE HK	15.38 %	15.38 %	23.08 %	0.00 %	0.00 %	15.38 %	15.38 %	15.38 %	0.00 %	0.00 %	0.00 %	0.00 %
POGE	2.63 %	2.63 %	54.61 %	3.95 %	7.24 %	9.87 %	0.00 %	0.00 %	1.32 %	0.66 %	0.66 %	16.45 %

To better analyze the components, a visualized diagram can be exported from NVivo as shown in Figure 4.3.
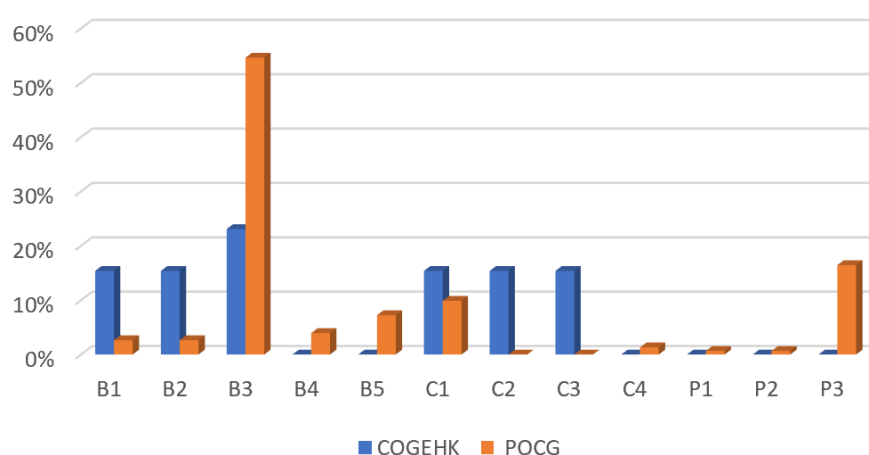
Figure 4.3. The comparison diagram of components of learning domains involved in POGE and COSGEHK

It is evident from the diagram that the examination formulated by COSGEHK involves only the behavioral and cognitive domains and focuses only on lower-order thinking (i.e., the first three hierarchies), with a balanced ratio ranging from 15.38% to 23.08%.

By contrast, POGE covers more domains than COSGEHK, but with misdistribution. The most noticeable component that is being examined in POGE is B3, which accounts for more

than a half (54.61%). The second most conspicuous component is P3, which accounts for 16.45%. Other components are documented as below 10%.

This will be an essential reference during the discussions in the interview, from which interview questions will be derived.

4.1.3.3 Rubrics

Rubrics are helpful tools that offer the basis for assessors to mark and for students to aim at. Well-designed rubrics can also provide constructive feedback to students about their achievements or performances and indicate what to do for improvement (Armstrong, S. et al., 2008).

COSGEHK proposed a series of rubrics, but their reasonableness needs further demonstration. On the other hand, POGE fails to publish any complete rubrics. Experts were consulted about the reasons for this and their replies are summarized below:

1. Unlike the COSGEHK, which focuses on Cantonese operatic singing and reciting, POGE is an examination of the holistic performance of Peking opera, which features the integration of the four major skills. The assessors would be distracted if they were required to mark while watching a candidate's performance.

2. *Xiqu* is a holistic arts genre, which should not be divided into separate parts.

3. The aesthetics of *Xiqu* are abstract, and there are diverse styles, i.e., *meipai* (梅派 Mei's sect), *chengpai* (程派 Cheng's sect), which cannot be clarified.

4. Rubrics are not part of the tradition of *Xiqu.*

Assessors prefer to give a holistic, abstract, and implicit evaluation of overall performance, based on their subjective perspectives, which may severely jeopardize the quality of assessment. Identifying the reasons for the missing assessment content might facilitate a confrontation with existing problems and lead to the seeking of solutions, rather than holding to the stereotype and refusing to improve. Thus, it was decided the topic would be discussed with experts in the interview.

Furthermore, the descriptions of rubrics need to be decided. Generally, there are two kinds: holistic and analytic rubrics.

Holistic rubrics embed a one-dimension criterion in a global description that judges overall student performance, rather than its separate parts (Mertler, 2000) (see Table 4.18).

Table 4.18. Template for holistic rubrics (Mertler, 2000, pp. 1-2)

Score	Criteria
5	Demonstrates a complete understanding of the problem. All requirements of tasks are included.
4	Demonstrates a considerable understanding of the problem. All requirements of tasks are included.
3	Demonstrates a partial understanding of the problem. Most requirements of tasks are included.
2	Demonstrates little understanding of the problem. Many requirements of tasks are missing.
1	Demonstrates no understanding of the problem.
0	No response/task not attempted.

In contrast, analytic rubrics are in two dimensions, with separate levels of performance described for each criterion (Mertler, 2000). An example from ABRSM is provided in Table 4.19).

Table 4.19. Marking criteria of ABRSM (ABRSM, n.d., p. 42)

Grades 1–8	Pieces Pitch	Time	Tone	Shape	Performance
Distinction 27–30	• Highly accurate notes and intonation	• Fluent, with flexibility where appropriate • Rhythmic character well conveyed	• Well projected • Sensitive use of tonal qualities	• Expressive, idiomatic musical shaping and detail	• Assured • Fully committed • Vivid communication of character and style
Merit 24–26	• Largely accurate notes and intonation	• Sustained, effective tempo • Good sense of rhythm	• Mainly controlled and consistent • Good tonal awareness	• Clear musical shaping, well-realised detail	• Positive • Carrying musical conviction • Character and style communicated
Pass 20–23	• Generally correct notes • Sufficiently reliable intonation to maintain tonality	• Suitable tempo • Generally stable pulse • Overall rhythmic accuracy	• Generally reliable • Adequate tonal awareness	• Some realisation of musical shape and/or detail	• Generally secure, prompt recovery from slips • Some musical involvement
Below Pass 17–19	• Frequent note errors • Insufficiently reliable intonation to maintain tonality	• Unsuitable and/or uncontrolled tempo • Irregular pulse • Inaccurate rhythm	• Uneven and/or unreliable • Inadequate tonal awareness	• Musical shape and detail insufficiently conveyed	• Insecure, inadequate recovery from slips • Insufficient musical involvement
13–16	• Largely inaccurate notes and/or intonation	• Erratic tempo and/or pulse	• Serious lack of tonal control	• Musical shape and detail largely unrealised	• Lacking continuity • No musical involvement
10–12	• Highly inaccurate notes and/or intonation	• Incoherent tempo and/or pulse	• No tonal control	• No shape or detail	• Unable to continue for more than a short section
0	• No work offered	• No work offered	• No work offered	• No work offered	• No work offered

Holistic rubrics are more suitable to assess: 1) complex tasks involving higher-order thinking, 2) tasks that are unlikely to separate one criterion from another, and 3) the globality of student performance in a task. Therefore, holistic assessment can only provide general or limited guidance for assessors and students (Armstrong, S. et al., 2008).

On the contrary, analytic rubrics push students to obtain more detailed information for preparation before an assessment and provide more feedback after the assessment (Armstrong, S. et al., 2008). They are considered to be easy for assessors to use and contribute to generating more precise and objective evaluations of student performance (Biggs, 2003).

Due to the respective characteristics, it can be seen that holistic rubrics are preferable for summative assessment, whereas analytic rubrics are particularly useful in formative assessment (Mertler, 2000). Although COSGEHK and POGE adopt the manner of a holistic assessment, whether a performance should be scored holistically or analytically should be decided before

formulating rubrics (Airasian, 2000 & 2001). Therefore, this topic was selected for inclusion in the interviews. The components of the rubrics in POGE and COSGEHK were compared as below:

Criteria

Apart from *wusheng* (武生), most roles in POGE lack criteria. Even in *wusheng* (武生), whether the proposed criteria cover all of the critical components of the assessment needs to be confirmed in interviews with multiple experts and with classroom observation of daily teaching.

On the other hand, COSGEHK publishes a series of criteria of *wusheng* (武生). As descriptions in the criteria of COSGEHK are not concise and are mixed with standards or instructions for test items, the details that did not belong to the criteria were removed. The components being examined in COSGEHK were extracted and are shown in Table 4.20:

Table 4.20. Components extracted from the criteria of the singing in COSGEHK

Singing skills	1) the application of *changqiang* (唱腔) 2) conformity with scores 3) accuracy in pitch, pulse, rhythm, projection, and articulation 4) the control of breath, dynamics, tempo, timbre, and high/low pitches 5) the fluency and stability of singing
Interpretation (sense of *Xiqu*)	1) personal interpretation technique 2) exhibit emotion and charm which evokes echoes 3) present various characteristics of *shuobai* (說白)*, qupai* (曲牌) *and qushi* (曲式)
Knowledge of Cantonese operatic singing	manifests understanding of the styles, scores, modes of expression, and musical symbols
Expression ability	1) manifests an explicit conception of how to express emotions and appeal to the audience. 2) present appropriate presence

Table 4.21. Components extracted from the criteria of the Q&A in COSGEHK

Knowledge in Cantonese operatic singing	1) provide correct answers.
	2) present comprehension of musical symbols in the scores
	3) show some understanding of the background of the play and its scenario (exclusive for candidates of grades 5~8)
Interpretation (sense of *Xiqu*)	show appropriate understanding and sensitivity.
Expression ability	1) demonstrates accurate and appropriate expression ability.
	2) present the musical symbols, forms, and style that grasped

Table 4.22. Components extracted from the criteria of the sight-reading in COSGEHK

Singing skills	fluently, and steadily
Interpretation (sense of *Xiqu*)	show a sense of *Xiqu* (exclusive for the candidates of Grade 5~8 only)
Knowledge in Cantonese operatic singing	1) manifest appropriate understanding in the *gongchepu* (工尺譜) and the symbols in the scores
	2) demonstrate an understanding of the melody, phrases, dynamics, forms, structures, and other knowledge of Cantonese operatic singing.
Expression ability	1) sing out the *gongchepu* (工尺譜)
	2) communicate with audience effectively

Table 4.23. Components extracted from the criteria of the aural test in COSGEHK

Sense of *Xiqu*	present an interpretation that conforms to the sense of *Xiqu* (exclusive for candidates of grades 5~8 only)
Knowledge in Cantonese operatic singing	1) retains the accuracy of the pitch and rhythm.
	2) retain the stability of pulse

As the most authoritative graded public examination in mainland China, it is perplexing that POGE fails to publish complete and normative rubrics or even criteria. Although some

components of assessment criteria in COSGEHK are listed out in tables 4.20-4.23, the criteria might be incomplete or not presented well, which necessitates improvement. Thus, identifying the components for the criteria should be a focus of any assessment. This received considerable attention when the assessment of Cantonese operatic singing was formulated for the pilot implementation in Phase II.

Standards

Standards in the rubrics of both graded examinations are incomplete and implicit. Even those with explicit standards all use a single standard. Compared to multiple standards, a single standard is easier to build but may make assessors spend more time providing feedback to students or even tempt them to skip this crucial step within the assessment. Multi-level performance standards provide more details about requirements, which offers students more guidance about to what extent they are expected to perform. Three-level standards (meets expectations, exceeds expectations, and below expectations) are a simple scheme that is easier to construct, administer, and employ (Mueller, 2006). But this type of standard is considered to be insufficient to discriminate between student performances. Therefore, five-level-standards are proposed: fail, pass, credit, distinction, and high distinction. To develop such a standard, Armstrong, S., et al., (2008) suggested starting with the identification of a threshold level, normally the "pass" level and then describing progressively higher and lower levels of performance. Figure 4.4 illustrates an example of a five-level-standard:

> *Criterion: Shows evidence of reading and makes reference to literature*
>
> *Standards*
>
> Pass: Relies primarily on recommended or standard texts with minimal or no reference to other sources.
>
> Credit: As well as meeting standard for pass, body of essay refers to some relevant literature and uses recommended or standard texts.
>
> Distinction: As well as meeting standard for credit, integrates ideas from wide reading of relevant literature and recent research.
>
> High distinction: As well as meeting standard for distinction, offers new understanding on ideas from literature. Distinguishes between the quality of sources and uses novel but relevant sources.
>
> Fail: Shows little or no evidence of reading recommended or standard texts; relies primarily on non-reviewed internet source; no reference to literature.

Figure 4.4. An example of a five-level-standards (Armstrong, S., et al., 2008, p. 24)

Therefore, how many levels of standards should be used in assessment was marked as one of the questions that should feature in the interview.

Weightings

Weighting is an important process that involves adjusting data by algorithm or values to reflect the significance of each part of an assessment. Therefore, weighting strategies in multiple criteria play a key role in the whole evaluation and decision-making process (Marichal & Roubens, 2000).

However, POGE provides no weightings. Whether weightings should be introduced into the assessment that is being formulated was marked as a question to be discussed in the interviews. The validity of the weighting strategies in COSGEHK is questionable. So, if weightings are considered to be necessary to the assessment, opinions and data need to be collected. There are numerous methods that help a researcher to identify the weightings within criteria, but the challenge is to choose an appropriate method.

Grading and evaluation

How many levels an evaluation should be divided into "depends on the ability of the assessment task to make fine distinctions reliably and the degree to which fine discrimination is required" (Hughes, 2007). COSGEHK adopts five-level evaluations for global evaluation: distinction, merit, pass, fail (but approximate to pass), and fail. POGE employs three-level evaluations: distinction, pass and fail. The former exhibits reasonable distinctions between levels, especially with its use of "fail (but approximate to pass)", which is an infrequent but encouraging evaluation of candidates who almost pass the examination. The latter have only rough divisions of the candidates' performance. It is beyond all doubt that "distinction" and "pass" are not part of a continuum, between which there should be some levels. A rough standard level of evaluation is a disadvantage to providing valid feedback to candidates.

Another characteristic that is worthy of note is grading. COSGEHK elucidates explicit scoring for the grading: "Distinction: 85~100%, Merit: 75~84%, Pass: 65~74%, Fail (but approximate to pass): 55~64%, and Fail: 0~54%". POGE mentions nothing about it, even cut-off scores, which is not reasonable. "The cut-off scores are often pulled from the air" or set to

an arbitrary number or percentage that is being used in other school districts. Traditionally, 50% to 70% is the most acceptable cut-off score in most institutions (Barman, 2008, p.1). For example, COSGEHK defines 64% as its cut-off score, whereas AMRSM states 66% is the cut-off score for its music theory examination.

4.1.4   Summary

The content analysis of POGE and COSGEHK focused on graded divisions, test items, and rubrics. Although strengths and weaknesses are manifest in both POGE and COSGEHK, the findings of the content analysis indicate that the design of the existing graded examinations of *Xiqu* is immature and disputable:

1. No explanations of grade divisions were provided by either of these two graded examinations.

2. With regard to the different dimensions of test items, COSGEHK stands out for its diversity, breadth of examination scope, and breadth of operatic knowledge in the assessment design. POGE presents a more convincing coverage of depth of operatic knowledge and breadth of learning domains.

3. POGE fails to publish complete and normative rubrics, which is perplexing. By contrast, COSGEHK features a normalized assessment formulation with complete rubrics, as well as a comparatively comprehensive set of test items and content. More specific content analysis has been calculated on the components of rubrics and comparisons have subsequently been made. These include criteria, standards, weightings, grading, and evaluation in POGE and COSGEHK.

The content analysis raised numerous questions that deserved further exploration in the subsequent semi-structured interviews:

1. What do you think of the assessment of Cantonese operatic singing? Can it be identified, measured, assessed, or described? Can you indicate what can be identified, measured, assessed, or described respectively?

2. What do you think of these two graded examinations? Any advantages and disadvantages from your perspective?

3. How can we ensure diversified skills and knowledge are tested in the examination? (Show "Figure 4.3. The comparison diagram of components of learning domains involved in POGE and COSGEHK" to the interviewees and explain the implications, thereby leading them to tease out an answer to this question.)

4. How many grades should the examination be divided into? What is the examination goal of every grade? How should different grades be differentiated?

5. Given the construction and diversity, what test items should be included in the examination?

6. In professional and traditional terms or plain language, what should be adopted when writing the description of criteria?

7. How many levels should be adopted for the assessment standards?

8. Is it necessary to introduce weightings into the assessment? If so, how do we set up weightings to meet the requirements of specific assessments?

9. Which type of rubrics should be employed: holistic or analytic rubrics?

10. Is an assessment framework necessary for Cantonese operatic singing?

4.2 Thematic analysis of semi-structured interview

4.2.1 Code system

Using MAXQDA, nodes were defined using keywords for text coding of the semi-structured interviews, which were designed to collect expert opinions of the assessment and suggestions for the new assessment's formulation. The opinions of the experts unavoidably revealed their attitudes, which can be based on the discussion topics that emerged from the content analysis of the two *Xiqu* graded examinations, as well as on other derived questions that were raise in the semi-structured interview. 37 nodes can be extracted from the interviews, which represent the interviewees' attitudes. These can be grouped into four broader categories: concern, criticize, advocate, challenge, and believe (see Table 4.24).

Table 4.24. Code system of five categories and their family nodes for the interview transcription coding (Interviewees' attitude)

1.	Concern
	Gongche pu is dying out in mainland
	Unbalance between skills and theory
	The disappearance of creativity and personalization
	No sects are established in the past few decades
	Fairness
	Teacher competence
	Might be too hard to meet the requirement
	Implementation
2.	Criticize
	Lack of educational ideology or knowledge

		Lack of criteria
		Lack of pedagogies
		Lack of assessment and feedback
		Rarely revealed criteria
		Assess by feelings
		Lack of sustainability in assessment
	Arbitrary in assessment	
	Neglect of the traditions or relative knowledge	
	Focus on skills only	
3.	Advocate	
	Performance quality is more important than the difficulty coefficient	
	Comprehensiveness in question types	
	Different notations should co-exist	
	Relevant knowledge facilitates the skills	
	Guidance from the perspective of education	
	Traditions should be valued	
4.	Challenge	
	Lack of normative education in Cantonese operatic singing	
	Norms are hard to be fixed	
	Differences between HK and mainland China	
5.	Believe	
	Rubrics are the kernel of assessment, feedback, or instructions	
	Assessment is significant and necessary	
	Assess according to criteria is more reliable than feelings	
	Assessment can enhance the education in Cantonese opera	
	Comprehensive education contributes to the transmission	
	Concrete or fundamental elements are assessable	
	Traditional notation can express the flavor of *Xiqu*	
	Xiqu is hard to be assessed	

On the other hand, 33 nodes can be developed from the interview transcriptions to present

interviewees' suggestions for the assessment formulation, including "finding out the elements

that can be or cannot be assessed", "abstract elements can be assessed with flexibility", "teachers/assessors' training", "assessment design", "terminology or common language", "questions' types should be diversified", "assessing student learning is less controversial", "standards", "weights", and "rubrics" (see Table 4.25).

Table 4.25. Code system of nine categories and their family nodes for interview transcription coding (Suggestions for the assessment formulation)

1. Find out the elements that can be or cannot be assessed
Assessable
Breath skills and voice projection
Fluency
Pronunciation and Articulation
Pulse and Rhythm
Pitch
Elements that are hard to be assessed
Timbre
The artistic interpretation
Individual personality on the stage
Intonation
The flavor of Cantonese operatic singing
Sense of tableau
Emotion
Character images
2. Abstract elements can be assessed with flexibility
3. Terminology or common language
Should regulate and generate the terms
Should co-exist
4. Question types should be diversified
5. Standards
6. Weightings
7. Rubrics

Elaborate rubrics are needed
Should be simple and user-friendly
The layout of the rubrics
The feasibility and necessity to use analytical layout
Feedback should be combined with analytical layout
8. Assessment design
Compatible with the local needs
Based on reliable and suitable framework or guidance
Unified with feedback
Pre-announced
Pre-established
Integrated with pedagogy
Modes
9. Assessing student learning is less controversial
10. Teachers/assessors' training

4.2.2 Thematic networks and their connections

The thematic networks and the relations between the nodes that were elicited from MAXQDA are illustrated in Figure 5.1:
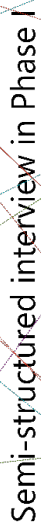
Figure. 4.5. Thematic networks of the semi-structured interview in Phase I

In the above networks, five categories that were extracted from the experts' attitudes disclosed both their opinions and some of the problems they associated with Cantonese operatic

singing assessment. Potential solutions were mostly provided in "suggestions for the assessment formulation". The arrows in the networks show the connections between "problems" and potential "solutions". Different colors are used to represent the connections derived from the different categories of expert attitudes: 1) blue arrows denote connections of "concern", 2) purple arrows denote connections of "criticize", 3) red arrows denote connections of "believe", 4) green arrows denote connections of "advocate", and 5) yellow arrows denote connections of "challenge".

During the interviews, experts indicated various key points about the assessment of Cantonese operatic singing. The connections can be teased out as below:

1) The interviewees believe that "assessment is significant and necessary", because "assessment can enhance the education in Cantonese opera".

2) The interviewees agreed that conducting assessments "according to criteria is more reliable than feelings" and "rubrics are the kernel of assessment".

3) The interviewees deemed that "*Xiqu* is hard to be assessed", especially in the artistic dimension and at senior levels, such as artists or masters, but "concrete or fundamental elements are assessable". When proposing suggestions for the formulation, interviewees indicated that the researchers should "find out the elements that can be and cannot be assessed". "Elements that are hard to be assessed" included "timbre", and "the artistic interpretation", such as "individual personality on the stage", "intonation", "flavor of Cantonese opera singing", "sense of tableau", "emotion", and "character images". For those aspects, the interviewees suggested that "abstract elements can be assessed with flexibility". The "assessable" elements

included "pitch", "pulse and rhythm", "pronunciation and articulation", "fluency", and "breath skills and voice projection".

4) Compared with assessing student performance from the perspective of artistry, "assessing student learning from the perspective of art education is less controversial".

5) The interviewees believe that "comprehensive education contributes to transmission", which might be a solution to "no sects have been established in the past few decades" that was noted in concerns.

6) One of the interviewees admitted that "traditional notation can express the flavor of *Xiqu*", which may become a solution to one of the elements in "the artistic interpretation" - the "flavor of Cantonese opera".

Among the multiple key points in the category of "concern", important connections can be elicited:

1) "The dying out of *Gongche pu* in mainland China" is a consequence of "unbalances between skills and theoretical knowledge", which leads to "disappearance of creativity and personalization", thereby resulting in "no sects are established in the past a few decades". This can be improved with the guidance of assessment, because assessment is considered to be one of the components of teaching and learning most closely related to curriculum and instruction. Therefore, the interviewees proposed that "question types should be diversified", which may be one way to guide assessment formulation. That is, question types should not only emphasize skills. If theoretical knowledge can be incorporated into assessments, the attention and

emphasis given to this knowledge will be invigorated. Creativity and personalization might be regenerated, and *Gongche pu*, as well as other characteristics of Cantonese opera, might be better handed down.

2) "Implementation" of the assessment will affect its "fairness", which makes "teacher/assessor training" extremely significant.

3) The interviewees in mainland China were concerned that it "might be too hard to meet the requirements" of the COSGEHK due to "teacher competence", which highlights the importance of "teacher/assessor training".

Among the multiple key points in the category of "criticize", important connections were elicited. These are given below:

1) The interviewees criticized multiple problems that exist in the teaching and learning of Cantonese operatic singing, including "assess by feeling", "lack of assessment and feedback", "lack of criteria", "rarely revealed criteria", "lack of sustainability in assessment", and "lack of pedagogies". This was commonly seen as a "lack of educational ideology or knowledge" in the field of education and transmission, thereby resulting in "arbitrary assessment". Therefore, an expert suggested that, when formulating the assessment, it should be "based on certain framework or guidance".

2) The interviewees also indicated that learners nowadays show a tendency to "focus on skills only" and "neglect the traditions or relative knowledge". Consequently, they advocated "teacher training", not only in the cultivation of multifaceted professional knowledge but also

in the understanding of education. Teachers should be able to emphasize the importance of traditions or relative knowledge as well as have the competence to integrate and penetrate knowledge into their students during teaching, instead of only coaching skills. On the other hand, "question types should be diversified" when formulating the assessment. As the assessment is considered to be a beacon of instruction and curriculum, it should arouse the attention of learners and teachers to different aspects of Cantonese operatic singing, instead of on skills only.

Among the multiple key points in the category of "advocate", important connections can be elicited as below:

1) The interviewees repeatedly emphasized that, to improve the education quality of Cantonese operatic singing, "guidance from the perspective of education" is imperative. Therefore, they suggested, the formulation of the assessment should be "based on a reliable and suitable framework or guidance".

2) The interviewees argued that "performance quality is more important than the difficulty coefficient of the repertoire".

3) Two interviewees in HK advocated that "different notations should co-exist". To generalize this notion to the teachers and students, "teacher training" is necessary. Besides, since assessment is a beacon of instruction and curriculum, "question types should be diversified" when formulating the assessment - to guide the orientation of instruction and curriculum.

4) With regard to the transmission of Cantonese operatic singing, the interviewees noted that "traditions should be valued". The habit of use in notations needs to be guided. Therefore, "teacher training" is imperative. They also noted that "question types should be diversified" to incorporate the content into the assessment and motivate both teaching and learning.

5) Compared to the convention of overemphasizing skills in Cantonese operatic singing, most of the interviewees agreed that "relevant knowledge", such as Cantonese opera structure, *luogu dian*, and the like, "facilitates the skills". Therefore, to motivate teaching and learning, "question types should be diversified" when formulating the assessment.

Besides various opinions and personal attitudes, the interviewees also provided some suggestions for the formulation of the assessment of Cantonese operatic singing. Some pertain to the concrete content of the assessment formulation, while the others concern the principles and good practice of the assessment formulation:

1. With regard to the concrete content of the assessment formulation, the interviewees suggested researchers should "identify the elements that can and cannot be assessed". The interviewees reached a consensus that "pitch", "pulse and rhythm", "pronunciation and articulation", "fluency", and "breath skills and voice projection" could all be assessed. The interviewees proposed various aspects that included "timbre" and "artistic interpretation" as being difficult to assess. Artistic interpretation includes aspects such as "individual personality on the stage", "intonation", "flavor of Cantonese opera singing", "sense of tableau", "emotion", and "character images".

2. With regard to the principles and good practice for the assessment formulation, the interviewees proposed:

1) "Abstract elements can be assessed with flexibility".

2) "Teacher/assessor training" should be valued.

3) Compared with assessing student performance from the perspective of artistry, "assessing student learning from the perspective of art education is less controversial"

4) For "the design of assessment", interviewees noted that good practice should be considered, including:

    a) Compatibility with local needs

    b) A base of frameworks or guidance

    c) Unification with feedback

    d) Assessment rubrics being pre-established and pre-announced

    e) Assessment is aligned with educational objectives in teaching

    f) Various modes for assessment

5) Four of the six interviewees (I1, I2, I4, and I5) agreed that terminology, jargon, and common language "should co-exist". Terminology and jargon have evolved to express authentic meaning in Cantonese opera, and common language facilitates understanding - especially for novices. Since controversy and misunderstanding surrounding terminology and jargon is quite common, the interviewees attached importance to "regulating the terms/jargon", especially when they were used in descriptions of the assessment.

6) To guide daily instruction and curriculum, "question types should be diversified" in the assessment.

7) The interviewees reached a consensus that fail, pass, merit, and distinction are suitable standard divisions to use in the assessment of Cantonese operatic singing.

8) Only two interviewees understood the importance and application of weightings. The rest of the interviewees showed little understanding of it and placed little value on it. This might have something to do with their experience and background as performers, rather than teachers. As a result, when dealing with assessment, relevant educational training is necessary.

Concerning the rubrics, the interviewees suggested that:

1) "Elaborate rubrics are needed".

2) "Rubrics should be user-friendly", which would help to alleviate their concerns about the "implementation" of the assessment.

3) It is feasible and necessary to use an analytical layout when formulating the educational assessment. Feedback should also be combined with the analytical layout.

4.2.3 Summary

After the thematic analysis of the semi-structured interviews, the answers or the solutions to the second and the third research questions can be summarized as below.

Common practice in assessments of Cantonese operatic singing include:

1. Dependence on the subjective perception of assessors, rather than on criteria-based assessment.

2. A focus on the assessment of performance skills, together with a lack of multidimensionality in assessment.

3. Momentary judgment and lack of sustainability.

With regard to transmission and education, the following key points should be considered important and taken into consideration when formulating an ideal assessment framework for Cantonese operatic singing:

1. Educational philosophy and knowledge ought to be valued in the assessment formulation for Cantonese operatic singing. Teacher/assessor training is recommended to enhance educational ideology and knowledge towards a better understanding and practice.

2. Attention should be attached to the theoretical knowledge alongside the current focus on skills. This will make a profound impact on the transmission and education of Cantonese operatic singing. A better balance between skills and theoretical knowledge should be achieved so that learners inherit a solid foundation. This should enable learners to be more creative and able to display their personalities when performing.

3. Since most of the interviewees believed that *Xiqu* was difficult to assess, assessing student learning from a perspective of art education should be less controversial than assessing performances from a perspective of artistry.

4. To avoid arbitrariness, a reliable and suitable guidance or framework, such as Bloom's Taxonomy, should be introduced into the formulation.

5. User-friendly design is the kernel of an assessment framework. It should be welcomed by teachers and students and generalized more easily.

4.3 Formulation of an assessment framework for Cantonese operatic singing

A new assessment framework was formulated based on the theoretical framework developed in this study. According to Bloom's Taxonomy, first proposed by American educational psychologist Benjamin Bloom in 1956, human learning can be divided into three domains, which are illustrated in Figure 4.6:
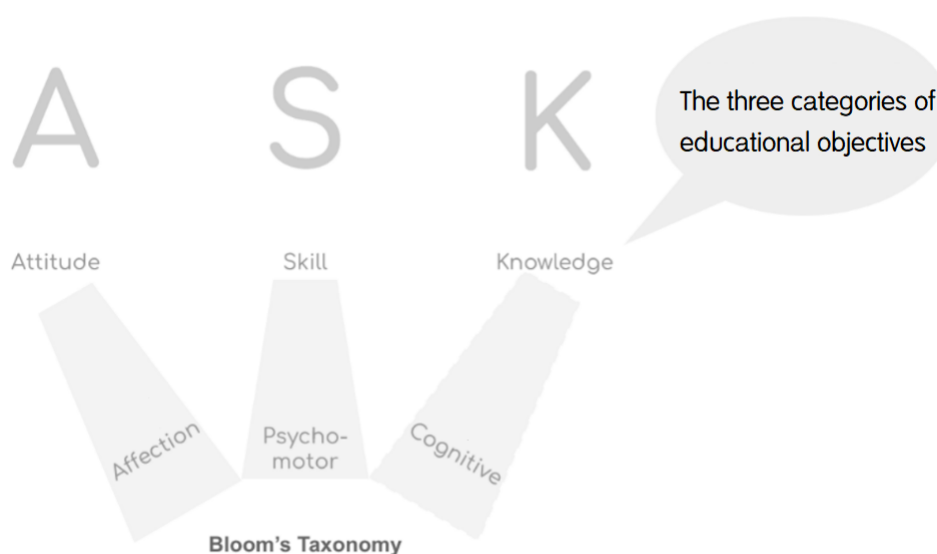
Figure 4.6. The original Bloom's Taxonomy

"Competence" is a comprehensive development of all three domains (see Figure 4.7).

Figure 4.7. The mechanism of all the domains in Bloom's Taxonomy

Therefore, the comprehensive output is based on overall behavior which is comprised of the cognitive, psychomotor, and affective domains. Based on the above rationale, an attempt was made to formulate an assessment framework according to human study progression in every domain of Cantonese operatic singing. The interpretation of every domain and its application in the specific subject of this traditional genre is illustrated in the subsequent subsections.

4.3.1   Intellectual abilities and skills (Cognitive Domain)

This domain involves intellectual abilities and skills, which may manifest themselves in open-ended questions and answers about relevant knowledge. Teachers can judge student cognitive levels and the ability of relevant knowledge points from student answers, which range from a low to a high hierarchy, including:
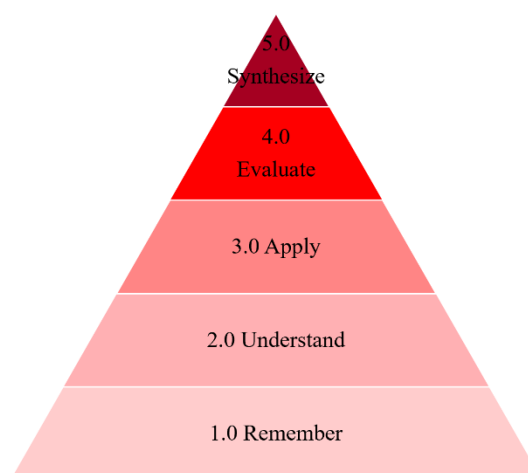
Figure 4.8. The hierarchies of the cognitive domain

Teachers are requested to use the rubrics when assessing students, based on an understanding of the above hierarchies. In this study, three questions were set to be assessed in the pilot implementation of the proposed assessment system. For further subdivisions, each level from 1.0 to 5.0 was divided into four levels that were represented by grids. The more to the right, the higher the level that was represented. For each line from 1.0 to 5.0, teachers were required to select only one of the grids according to student performance. Weightings were set as equal divisions to concentrate focus on the overall structure of the assessment system. Explorations of weightings will be one of the main foci of future studies.

To avoid human factors and ensure objectivity, algorithms were adopted in the rubrics. Since every line was assigned a hundred points and there were 20 grids in each line, a grid in the same line was assigned five scores (see Table 4.26).

Table 4.26. The Assessment Rubric of Cognitive Domain

Standards \ Items	Low — 1.0 Remember				2.0 Understand				3.0 Apply				4.0 Analyze				5.0 Synthesize — High				Weightings
Q1	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	1/3
Q2	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	1/3
Q3	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	1/3

Score: 0.00

As algorithms were applied to the rubrics, automatic calculations were triggered by teacher selections and a score would be generated automatically at the end. An example student assessment result is given below:

Table 4.27. An example of the assessment results of the cognitive domain

Standards \ Items	Low — 1.0 Remember				2.0 Understand				3.0 Apply				4.0 Analyze				5.0 Synthesize — High				Weightings
Q1	○	○	○	○	○	○	○	○	○	○	○	○	○	◉	○	○	○	○	○	○	1/3
Q2	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	◉	○	○	○	○	1/3
Q3	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	◉	○	○	○	1/3

Score: 78.33

4.3.2 Physical abilities and skills (Psychomotor Domain)

This domain involves the manipulative motor-skill area, which may manifest itself in physical abilities and skills. The progressive hierarchy of this domain is:
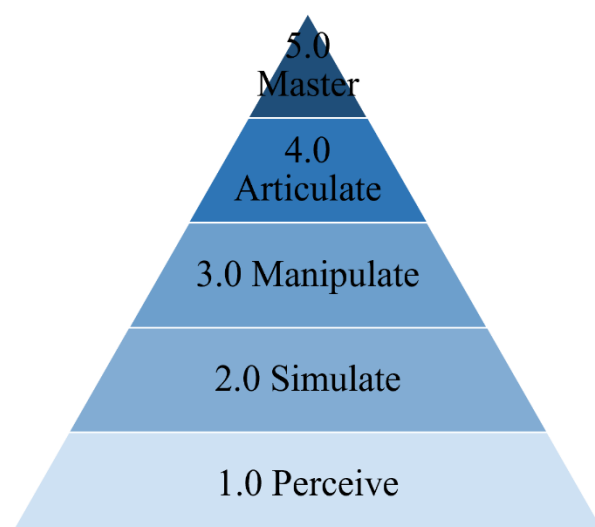
Figure. 4.9. The hierarchies of the cognitive domain

Although the following assessment rubrics might be employed in any subject that involves physical abilities and skills, Cantonese operatic singing skill was chosen as the subject of the pilot implementation, to verify the assessment rubrics of the psychomotor domain. According to the consensus drawn from the previous semi-structured interviews with veteran teachers and experts, an assessment rubric for Cantonese operatic singing skills was established as given below:

Table 4.28. The Assessment Rubric of Cantonese Operatic Singing Skill

Standards / Items	Low ─────────────────────────────► High																			Weightings
	1.0 Perceive				2.0 Simulate				3.0 Manipulate				4.0 Articulate				5.0 Master			
Intonation	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	20%
Rhythm	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	20%
Breath	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	20%
Articulation	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	20%
Timber	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	20%

Score: 0

4.3.3 The mentality (Affective Domain)

Affective characteristics are so imperative that they will influence future behavior, thereby affecting the individual and society as well (Reid, 2006). Bloom (1976) estimated that "25% of student performance is determined by affective characteristics" (p. 1). It is believed that successful students must possess not only rational cognitive study strategies but also the sound affective will to use these strategies (Weinstein, et. al, 1998). Ongoing assessment in the affective domain, such as student attitudes, values, interests, and so on, is essential for improving the ability of the educational experience and academic achievement. However, the assessment of the affective domain is often neglected in developmental education and the practice of routine assessment of student affective constructs is still rare (Saxon, et.al, 2008). Since no framework can currently be used to measure the extent of affective domain development (Stephens & Ormandy, 2019), it was decided to incorporate the affective domain in the assessment framework formulation of the current study.

There are three main dimensions of the affective domain: 1) self-reported data, 2) recorded data, and 3) observational data (Geisert, 1972). Since using multiple methods to collect

information on multiple characteristics from multiple sources is deemed to be particularly important in affective assessment (Oakland, 1997), each of the above three main dimensions was incorporated.

1. Self-report data

Since affective characteristics are personal attributes that can often only be determined through self-reporting (Davies, 2020), this is the preferred data collection approach in the collection of affective assessment data.

As shown in Figure 4.10, multiple instruments can be adopted when conducting an affective assessment, such as the Likert scale, semantic differential scale, Q-SORT, and the like (Hall, 2011; Geisert, 1972). The most widespread practice in the affective instrument is to "ask respondents to select one response from an ordinal series of response options" (McCoach, et. al, 2013, p. 40), so that a consistent "frame of reference for all the respondents" can be provided (Weisberg et al.1996, p. 84). Therefore, the current study employed a 5-point Likert scale as the instrument of student self-reporting. As respondents prefer close-ended questions to open-ended questions (Dillman et al. 2009), close-ended questions were used to collect self-reporting data and subsequent observational data.
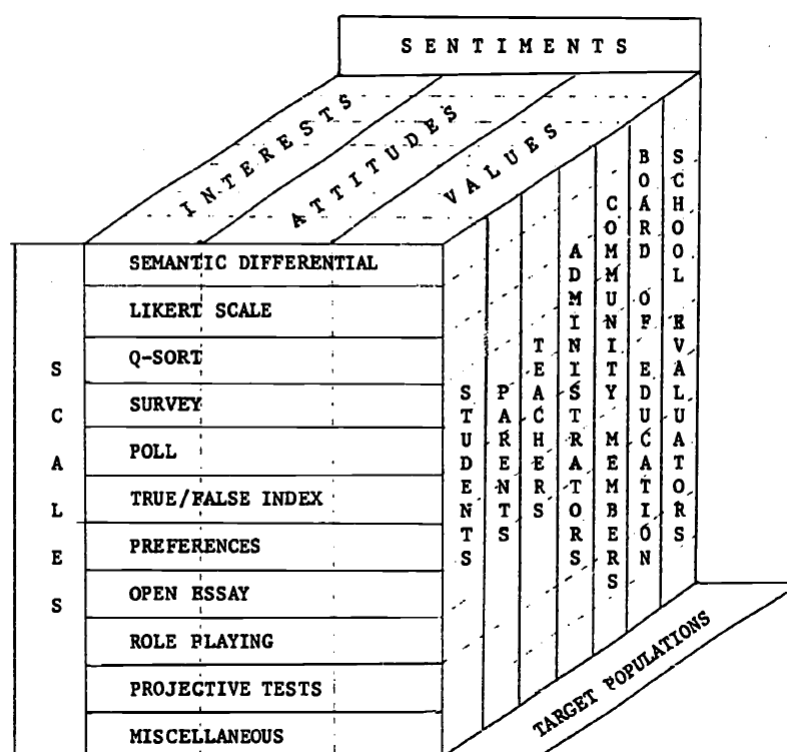
Figure 4.10. The dimensions of a self-reported data pool (Geisert, 1972, p. 16)

Psychologists have identified numerous constructs that reflect affective characteristics. These include attitudes, self-efficacy, values, motivation, and interest (McCoach, et. al, 2013). McCoach, et. al (2013) stated that "in general terms, attitudes were described as feelings toward some object; self-efficacy was referred to as a self-appraisal of capability; values reflected enduring beliefs; motivation can denote both external and internal states that drive us in a particular direction, and interests reflected preferences for particular activities" (p. 25). Thus, the questions surrounding self-reporting for assessment of the affective domain in the current study were developed around these five affective constructs.

The questions referred to some of the examples in The Post-experimental Intrinsic Motivation Inventory (Center for Self-Determination Theory, n.d.). Given the various positions

of professional students and amateurs, questions vary slightly accordingly. The question statements in tables 4.29 and 4.30 received approval from teachers who took part in the pilot implementation. To ensure reliability, cross-checking with enough questions was considered (Reid, 2006). Consequently, two related questions that were either positive or negative aimed at one affective construct. The congruent relationships between affective constructs and the sequence numbers of questions were 1) Value: 1 & 8; 2) Attitude: 2, 10; 3) Interest: 3, 6; 4) Motivation: 4,7; 5) Self-efficacy: 5, 9 (see Table 4.29).

Table 4.29. Online Survey of Student Perception on Learning

Standards / Criteria	Strongly agree ————————► Strongly disagree				
	1	2	3	4	5
1. I think it is important for my career to learn this course well.					
2. I have a feeling of powerlessness when studying this course					
3. I am very interested in learning this course.					
4. I am willing to spend more time and energy on this course.					
5. I am confident that I can master the knowledge and skills of this course.					
6. This course cannot hold my attention at all.					
7. To learn this course well, I have taken the initiative to consult teachers or predecessors.					
8. I think this course is useful for the transmission of traditional Chinese culture, which gives me a sense of mission and pride.					
9. I am not satisfied with my performance in this course.					
10. I enjoy the class.					

Despite the multiple advantages of self-reporting, an overarching disadvantage is its credibility (Paulhus & Vazire, 2007). Participants might distort their responses (Paulhus 1991). And some might even be concerned about how the instrument designers coded or measured deviations generated by respondents unknowingly (Blanton & Jaccard 2006; Greenwald et al., 2003; Wilson et al., 2000). Therefore, multiple sources for the assessment of the affective domain were necessary.

2. Recorded data

Recorded data are widely collected in public schools. This includes absenteeism, tardiness, homework performance, discipline, and the like. Since this dimension has been routinely applied in the Cantonese Opera School, it was decided to follow existing practice in the second dimension of assessment, rather than establishing new practices.

3. Observational data

The third dimension is observational data collected from a target population and evaluated by teachers based on the hierarchies of the affective domain. This included 1.0 Receive, 2.0 Respond, and the like. To collect the same information from every question from every student and teacher, the questions corresponded with those in the student self-report questionnaire. Thus, the congruent relationships between affective constructs and the sequence numbers of the questions were 1) Value: 1 & 8; 2) Attitude: 2, 10; 3) Interest: 3, 6; 4) Motivation: 4, 7; 5) Self-efficacy: 5, 9 (see Table 4.30).

Table 4.30. Observational Assessment of Student Attitude and Psychology in Learning (For professional students)

Grade:_____   Name:_____

Standards / Criteria	Lower-level →→→ Higher-level				
	1.0 Receive	2.0 Respond	3.0 Value	4.0 Believe	5.0 Behave
1. Regarding the importance of this subject, the student's behavior is:					
2. When confronted with learning difficulties thereby receiving encouragement and instruction of solutions from the teacher, the student's behavior is:					
3. When the teacher stimulates the student's interest in learning in certain approaches, the student's behavior is:					
4. When encouraged to devote more time and energy to the course, the student's behavior is:					
5. Regarding the confidence to master the knowledge and skills of this course, the student's behavior is:					
6. In the class, the student manifests himself as:					
7. Regarding the extra effort that may be required to learn the course well, the student's behavior is:					
8. Regarding the relevance of Cantonese opera to the inheritance of traditional Chinese culture, the student's behavior is:					
9. Regarding the performance in the class, the student's behavior is:					
10. Regarding the emotions in the class, the student's behavior is:					

Since student age in Children's Palace ranges from seven to 14, images were introduced into the questions as Likert Scale substitutions (see Table 4.31).

Table 4.31. Observational Assessment toward Student Attitude and Psychology in Learning (For amateur students)

Name: Grade:

1	I think this is an important course.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
2	I have a feeling of powerlessness when studying this course.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
3	I would describe this course as very interesting.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
4	I put a lot of effort into this.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
5	I am confident that I can master the knowledge and skills of this course.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
6	I thought this was a boring course.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
7	It was my family who wanted me to learn Cantonese opera rather than my own decision.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
8	I believe learning this course could be beneficial to me.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
9	I am satisfied with my performance in this course.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐
10	I enjoyed the class.	😃 ☐	🙂 ☐	😐 ☐	🙁 ☐	☹️ ☐

The answers to the questions were translated into the corresponding numbers of the

Likert Scale when exporting the data from the online survey as below:

Figure 4.11. Corresponding relationships between the images and numbers in Semantic Different Scale

With regard to the data obtained from the three dimensions, Reid (2006) doubted that it is

possible to conduct absolute measures of affective characteristics. Only comparison makes

sense. Similar contentions about the measurement of affective characteristics are ongoing

(Saxon, et.al, 2008). Bennett et al. (2001, p. 834) indicated seven areas of controversy in

measurement: (a) lack of precision over key definitions of terms; (b) poor design of instruments

and individual response items within instruments; (c) failure to address matters of reliability

and validity appropriately; (d) inappropriate analysis and interpretation of data; (e) lack of

standardization of instruments; (f) failure to draw on ideas from psychological theory; and (g)

failure to formulate the research concerning the theory of data collection tools. As a result, it

was decided not to measure the psychometric properties of the data obtained from the three

dimensions. Therefore, unlike the other domains, the affective domain was not counted into

the results either positively or negatively. The purpose of the affective assessment in the current

study is to collect information on student learning psychology so that teachers are able to infer

any relationship between observed behaviors and student statements. In that manner, timely concerns or interventions for academically marginal students are able to be triggered.

4.3.4 Performance assessment (Behavioral Domain)

This domain embraces an integration of cognitive, affective, and psychomotor domains, which were used to assess a Drama class. The progressive hierarchy of this domain is given in Table 4.12:
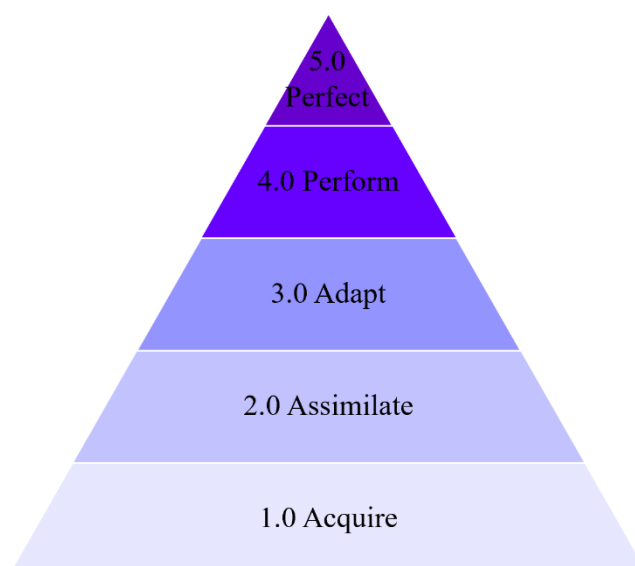
Figure 4.12.   The hierarchies of Performance Domain

According to the consensus surrounding the items of the previous semi-structured interviews with veteran teachers and experts, an assessment rubric for the drama class was established as below:

Table 4.32. Assessment Rubric of Drama Class

Items / Standards	Low →→→ High					Weightings
	1.0 Acquire	2.0 Assimilate	3.0 Adapt	4.0 Perform	5.0 Perfect	
Technical requirements						
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Pulse & Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Cantonese pronunciation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Changqiang	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Artistic requirements						
Artistic expression in line with the repertoire	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
The appearance in costume and makeup	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6

Score: 0

4.3.5 Summary

Based on the theoretical framework, as well as on the insight derived from the content analysis of the two graded examinations in *Xiqu* and a thematic analysis of the semi-structured interviews, an assessment framework that included four domains was formulated. To examine its feasibility and identify potential problems and deficiencies, a small-scale exploratory study was implemented in Phase II.

Chapter 5

PHASE II: EXPLORATORY STUDY OF THE PROPOSED ASSESSMENT FRAMEWORK

5.1 Introduction

"Exploratory studies, often termed pilot and feasibility studies, are a key step in assessing the feasibility and value of progressing to an effectiveness study" (Hallingberg，2018, p. 1). This helps provide pivotal information and lay the foundation for large-scale research, offering multiple advantages that include testing of proposed research designs and processes, identifying potential problems, and educating or familiarizing researchers with different techniques or instruments related to the study (Swedberg, 2020). Thus, in Phase II, an exploratory study to seek participant feedback about the proposed assessment framework proposed in Phase I was conducted. The procedures were as below:
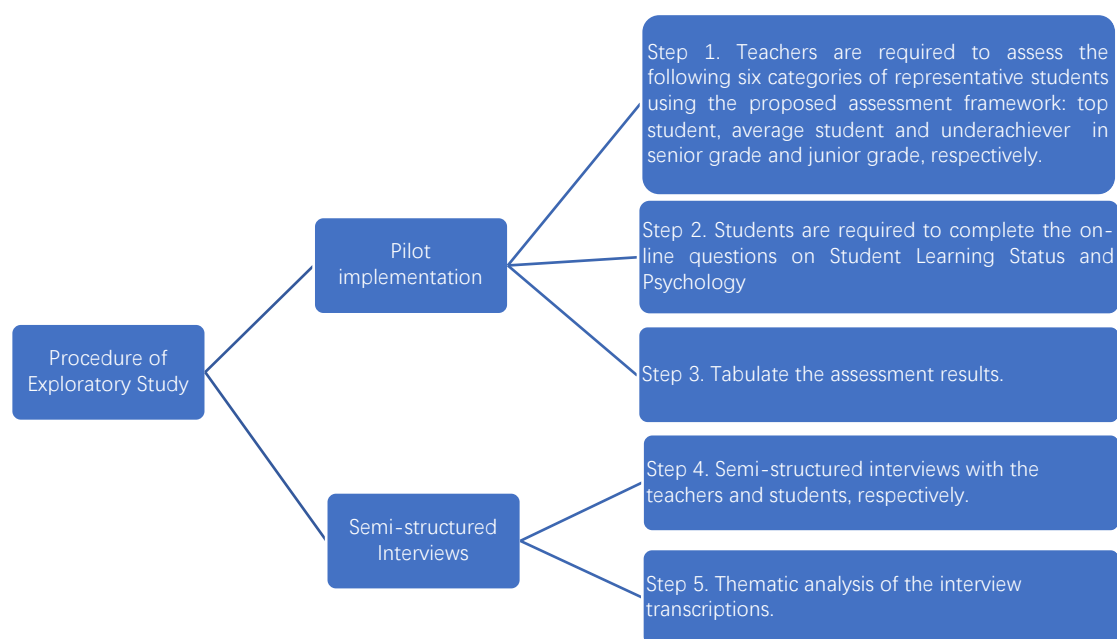
Figure 5.1. The procedure of the exploratory study in Phase II

The entire exploratory study incorporated two parts: pilot implementation and semi-structured interviews. The first part consists of three steps. In Step 1, teachers were required to assess the following six categories of students using the proposed assessment framework: top students, average students, and underachievers in both senior grades and junior grades. In Step 2, students were required to answer online questions about Student Learning Status and Psychology, as their self-reporting of the affective domain. In Step 3, the assessment results, which would constitute the foci of the follow-up semi-structured interviews, were tabulated for the participants' review. With the results collected from the pilot implementation, a series of semi-structured interviews with teachers and students were conducted (see Appendices H). The teachers and students were asked to reflect on their opinions of the assessment results and the process of assessment. After the pilot implementation, the semi-structured interviews were transcribed and the thematic analysis was implemented to record first-hand users' views.

5.2 Pilot implementation

5.2.1 Participants

To examine the feasibility of the proposed assessment framework, several teachers were invited to be participants. Elderly teachers proved reluctant to join the pilot implementation. Only one middle-aged teacher from the Children's Palace (code name: T1) and three younger teachers from the Cantonese Opera School (code name: T2, T3, and T4) showed a strong interest and were willing to join the study. These teachers were required to assess six categories of students: top student, average student, and underachiever in senior grade and junior grade. Teachers selected these six categories of students based on their daily academic performance. Thus, 24 students were assessed with the application of the proposed assessment framework and joined the online Survey of Student Perception on Learning to engage in self-reporting of the affective domain.

The teachers were encoded as T1, T2, T3, and T4. Accordingly, their students were encoded S1 to S6 and attached to their teacher code names (see Table 5.1). For example, the first student of T1 was encoded T1-S1.

Table 5.1. Codes for teacher and student participants in Phase II

Location	Teacher	Student
Children's Palace	T-1	T1-S1
		T1-S2
		T1-S3
		T1-S4
		T1-S5
		T1-S6
Cantonese Opera School	T-2	T2-S1
		T2-S2
		T2-S3
		T2-S4
		T2-S5
		T2-S6
	T-3	T3-S1
		T3-S2
		T3-S3
		T3-S4
		T3-S5
		T3-S6
	T-4	T4-S1
		T4-S2
		T4-S3
		T4-S4
		T4-S5
		T4-S6

5.2.2 Procedure

Twenty-four students were assessed comprehensively based on their academic performance in the Q&A of Cantonese opera general knowledge, Cantonese operatic singing

skills, and drama performance. This corresponded to the cognitive domain, psychomotor domain, and behavioral domain. With regard to the affective domain, both student online self-reports and teacher observational assessments were conducted. All the assessment results were tabulated and distributed to the students and teachers to be reviewed before the subsequent semi-structured interviews.

5.2.3 Assessment results

Although each domain might correspond to multiple subjects, the Q&A was chosen as the representative of the cognitive domain, Cantonese operatic singing skills represented the psychomotor domain, and drama performance represented the performance domain in the pilot implementation.

According to the teacher judgments, total points were calculated automatically, based on the previously set algorithms in Microsoft Excel. The assessment results of one of the students is given in tables 5.2 to 5.5 as an example. The rest of the assessment results are given in Appendix I.

Figure 5.2. An example of The Assessment Result of Cantonese Operatic Singing Skill

Code Name: T2-S6   Grade: Senior

Standards / Items	Low → 1.0 Remember	2.0 Understand	3.0 Apply	4.0 Analyze	5.0 Synthesize	Weightings
Q1					⊙	1/3
Q2				⊙		1/3
Q3				⊙		1/3

Score: 86.67

Figure 5.3. An example of The Assessment Result of Q&A

Code Name: T2-S6   Grade: Senior

Standards / Items	Low → 1.0 Acquire	2.0 Assimilate	3.0 Adapt	4.0 Perform	5.0 Perfect	Weightings
Technical requirements						
Intonation					⊙	1/6
Pulse & Rhythm				⊙		1/6
Cantonese pronunciation				⊙		1/6
Changqiang					⊙	1/6
Artistic requirements						
Artistic expression in line with the repertoire					⊙	1/6
The appearance in costume and makeup				⊙		1/6

Score: 84.17

Figure 5.4. An example of The Assessment Result of Drama Class

Code Name: T2-S6   Grade: Senior

Criteria \ Standards	Lower-level → Higher-level				
	1.0 Receive	2.0 Respond	3.0 Value	4.0 Believe	5.0 Behave
1. Regarding the importance of this subject, the student's behavior is:	○	○	◉	○	○
2. When confronted with learning difficulties thereby receiving encouragement and instruction of solutions from the teacher, the student's behavior is:	○	○	○	◉	○
3. When the teacher stimulates the student's interest in learning in certain approaches, the student's behavior is:	○	○	◉	○	○
4. When encouraged to devote more time and energy to the course, the student's behavior is:	○	○	○	◉	○
5. Regarding the confidence to master the knowledge and skills of this course, the student's behavior is:	○	○	○	◉	○
6. In the class, the student manifests himself as:	○	○	○	◉	○
7. Regarding the extra effort that may be required to learn the course well, the student's behavior is:	○	○	○	○	◉
8. Regarding the relevance of Cantonese opera to the inheritance of traditional Chinese culture, the student's behavior is:	○	○	○	○	◉
9. Regarding the performance in the class, the student's behavior is:	○	○	○	◉	○
10. Regarding the emotions in the class, the student's behavior is:	○	○	○	○	◉

Figure 5.5. An example of The Observational Assessment toward Student Attitude and Psychology in Learning

The teachers distributed the assessment results of four domains to every student so that they could be reviewed. The answers to the Online Survey of Student Perception on Learning were exported from the online backstage to the teachers, so that they could be reviewed (see Figures 5.6 & 5.7).

Teacher	Student	1. I think this is an important course.	2. I have a feeling of powerlessness when studying this course.	3. I would describe this course as very interesting.	4. I put a lot of effort into this.	5. I am confident that I can master the knowledge and skills of this course.	6. I thought this was a boring course.	7. It was my family who wanted me to learn Cantonese opera rather than my own decision.	8. I believe learning this course could be beneficial to me.	9. I am not satisfied with my performance at this course.	10. I enjoyed the class.
T1	T1-S1	4	1		4	5		2	5	2	5
	T1-S2	5	5		5	5		5	4	2	5
	T1-S3	4	2		5	5		1	4	2	5
	T1-S4	5	2		4	5		5	5	2	5
	T1-S5	5	2		4	5		2	5	1	5
	T1-S6	5	1		5	5		2	4	1	5

	Likert scale:	1	2	3	4	5
		Strongly disagree →				Strongly agree

Figure 5.6. Answers to Online Survey of Student Perception on Learning (For amateur students)

Teacher	Student	1. I think it is important for my career to learn this course well.	2. I have a feeling of powerlessness when studying this course	3. I am very interested in learning this course.	4. I am willing to spend more time and energy on this course.	5. I am confident that I can master the knowledge and skills of this course.	6. This course cannot hold my attention at all.	7. To learn this course well, I have taken the initiative to consult teachers or predecessors.	8. I think this course is useful for the transmission of traditional Chinese culture, which gives me a sense of mission and pride.	9. I am not satisfied with my performance in this course.	10. I enjoy the class.
T2	T2-S1	5	3	5	5	4	5	4	5	3	2
	T2-S2	4	1	5	5	4	1	5	4	2	5
	T2-S3	5	3	4	4	4	1	4	5	1	5
	T2-S4	4	4	5	4	4	3	4	5	4	3
	T2-S5	4	3	5	5	4	3	4	5	1	4
	T2-S6	5	4	5	5	5	1	4	5	2	2
T3	T3-S1	5	1	5	5	5	1	5	4	1	5
	T3-S2	5	1	5	5	5	2	4	5	1	5
	T3-S3	5	1	5	5	5	1	5	5	1	5
	T3-S4	5	1	5	5	5	1	5	5	2	4
	T3-S5	5	4	5	5	4	5	3	4	5	2
	T3-S6	5	1	5	5	5	1	5	5	1	4
T4	T4-S1	5	1	5	5	5	1	4	4	1	5
	T4-S2	5	1	5	5	4	1	4	5	2	5
	T4-S3	5	1	5	4	5	1	4	5	2	5
	T4-S4	5	1	5	5	5	1	5	4	1	5
	T4-S5	4	3	5	4	4	2	5	5	1	4
	T4-S6	5	5	5	5	5	3	3	3	4	3

	Likert scale:	1	2	3	4	5
		Strongly disagree →				Strongly agree

Figure 5.7. Answers to Online Survey of Student Perception on Learning (For professional students)

Unusual answers were highlighted for further interpretation. For example, one of the students (code name: T2-S1) chose 5 in the Likert Scale, denoting "strongly agree" as their answer to Q6: "This course cannot hold my attention at all", and chose 2, denoting "disagree" as their answer to Q10: "I enjoy the class". This seemingly reflected a disturbing learning status. The issue was later raised in the semi-structured interviews with the student and their teacher.

The total points allocated in the psychomotor domain, cognitive domain, and performance domain of each student were tabulated for teacher and researcher review. This is shown in Table 5.2:

Table 5.2. Summary sheet of pilot implementation

Location	Teacher	Student	Singing skills (Psychomotor Domain)	Q&A (Cognitive domain)	Drama Performance (Performance Domain)	Level of Usual Academic Performance	Grades
Children's Palace	T-1	T1-S1	50	55	55	Underachiever	Junior Grades
		T1-S2	51	63.33	58.33	Average	
		T1-S3	88	85	85.33	Top	
		T1-S4	75.83	85	78.33	Underachiever	Senior Grades
		T1-S5	80	80.33	80	Average	
		T1-S6	85.83	81.67	85	Top	
Cantonese Opera School	T-2	T2-S1	57	55	55	Underachiever	Junior Grades
		T2-S2	75	60	67.5	Average	
		T2-S3	85	61.67	76.67	Top	
		T2-S4	70	71.67	69.17	Underachiever	Senior Grades
		T2-S5	75.83	75	79.17	Average	
		T2-S6	85.83	86.67	84.17	Top	
	T-3	T3-S1	66.5	81	68.75	Underachiever	Junior Grades
		T3-S2	86.5	70	77.75	Average	
		T3-S3	82.5	73	84.5	Top	
		T3-S4	69.5	54	69.25	Underachiever	Senior Grades
		T3-S5	75.75	58	74.75	Average	
		T3-S6	89.75	67.25	84.5	Top	
	T-4	T4-S1	55	51.67	55	Underachiever	Junior Grades
		T4-S2	85.5	65.5	80.5	Average	
		T4-S3	89	71.67	88.33	Top	
		T4-S4	66.5	53	68	Underachiever	Senior Grades
		T4-S5	81.5	58.5	80.5	Average	
		T4-S6	88	66.67	85	Top	

From Table 5.2, it is clear that the assessment results are almost in line with the regular academic performance of students, except for 1) student T3-S1, whose assessment result of Q&A was much higher than average and top students at the same grade; and 2) student T3-S2

whose assessment result of Cantonese operatic singing skills was a little higher than the top student at the same grade. This unevenness became one of the topics discussed in the subsequent semi-structured interviews with Teacher No.3.

Whether the above assessment results were acceptable to the students and teachers, and whether there was a distinct gap between the assessment results and the normal academic performance of the students became a focus of the subsequent semi-structured interviews.

5.3 Thematic analysis of the follow-up semi-structured interview

5.3.1 Coding system

Rich insight can be derived from interviews (Reid, 2006). Using MAXQDA, nodes were extracted from the text coding of the semi-structured interviews that were designed to collect teacher and student opinions after the pilot implementation of the proposed assessment framework. The opinions included views on traditional assessment, the proposed assessment framework, the affective domain, their attitudes, and their suggestions.

During the interviews, teachers and students indicated that traditional assessments and the proposed assessment framework had both advantages and disadvantages. As a result, 34 nodes were extracted from teacher (T), amateur student (AS), and professional student (PS) opinions on traditional assessments, which formed two categories: advantages and disadvantages (see Table 5.3).

Table 5.3. The coding system of two categories and their family nodes for interview transcription coding (Traditional assessment)

Categories	Nodes	Number of nodes
1.	Advantages	
	Habitually practice	2
	The pass rate can be controlled and ensured	1
	Three-level promotion mechanism as a kind of assessment	1
2.	Disadvantages	
	Factitious scoring	
	T	
	Avoid failure rate	13
	Traditional ideology	4
	Protect student self-esteem	3
	PS	
	Avoid failure rate	6
	Protect student self-esteem	4
	Affected by human factors	7
	General and implicit	
	T	4
	AS	6
	PS	8
	Oral assessment rather than in the written form	
	T	7
	PS	19
	Apt to be forgotten	8
	AS	7
	Apt to be forgotten	6
	Incomplete and non-normative	
	T	2
	PS	7
	AS	4
	Total points only and without monomial points	
	PS	9
	T	2
	Lack of objectivity and fairness	
	T	8
	PS	10
	Lack of instructiveness	1
	Untransparent	3

The interviewees mentioned only three advantages of the traditional assessments:

a) It has been "habitually used", and does not demand any extra endeavor to learn to use, as would a new approach. One teacher stated that: "I think the traditional assessment is good for teachers, for we have already gotten used to it. Some teachers are not willing to take the trouble to learn a novel approach, especially the older teachers who are weak in operating computers and prefer to stick to the old way." (T-4)

b) "Pass rate can be controlled and ensured". In the professional education traditionally associated with Cantonese operatic singing, there is a series of mechanisms to guarantee the quality of teaching and learning. If a student fails an examination, they are later required to re-take a similar examination. If the student fails the re-take, they are offered a final opportunity to try again before graduation. If the student fails again, then they do not get a graduation certificate. Thus, to "protect" students and their future careers, evading this mechanism has become common practice. One of the teachers stated that:

> In our school, if a student gets an F, they are required to take part in make-up examinations until they eventually pass. Otherwise, they cannot graduate. So, unlike the scoring in competitions, even though the students do not do well in the examination, as long as they are not far too poor, we would still give them at least a passing score to ensure the pass rate. For example, the following table is the transcript of one of the classes that I taught in my school. (T-2)

Table 5.4. A transcript of professional student academic performance in a semester

No.	Name	Mid-term Grade	Final Grade	Total Grade
1	HJL	85	88	87
2	LXC	92	96	94
3	FLQ	92	95	94
4	YTR	86	88	87
5	ZZY	90	96	94
6	OYSQ	78	84	82
7	YYX	83	86	85
8	HZZ	90	92	91
9	LJM	60	60	60
10	LLD	78	82	80
11	HJJ	90	94	92
12	JZT	88	94	92

c) "The three-level promotion mechanism is a kind of practical assessment system" at Children's Palace. Starting with the regular class, students can be promoted to art ensemble, if they pass the examination, and they would be selected to take part in competitions if outstanding performance is achieved. One of the teachers mentioned that:

> We seldom give scores or comments in written form to students. Instead, we employ a three-level promotion mechanism, which we think is more practical to amateur students and more suitable to the teaching and learning of the Children's Palace (T1)

More disadvantages were pointed out by the interviewees, including:

1) Opaqueness. This disadvantage is particularly prominent from the perspective of the amateur student parents:

> We used to get the hang of my daughter's study situation by asking her or asking the teachers occasionally. But my daughter cannot make it clear. And the teacher is apt to say things are going well to reassure us. So, how my daughter is doing in her Cantonese operatic singing is untransparent to us. (T1-S4)

2) Factitious scoring. Since the Children's Palace offers stars or certificates of merit to amateur students as alternative scores, neither teachers nor amateur students mentioned

scoring. But in the professional education of Cantonese operatic singing, factitious scoring aims to ensure the pass rate and avoid re-taken examinations, which seems to be an open secret. Both the professional students and teachers attributed this practice to "protecting student self-esteem" and "avoiding failure rate":

> In traditional assessment, we would usually make some adjustments so that all of the students can pass the examinations. Otherwise, it would be vexatious for both students and teachers. What is more, the student hearts are very fragile nowadays. I am afraid it might hurt the student learning initiative if I give an F to them. (T2)

> I have not heard that there is anyone who ever got an F in the examinations before. Maybe it was because the teachers try not to hurt us. (T3-S5)

> The teachers also implied that the factitious scoring might come from traditional ideology: "It is even more overt in the assessment of competitions. The scoring is usually affected by who is whose student and who is whose master. It is a common phenomenon in the Chinese theatrical circle." (T4)

> Some of the professional students believed that human factors might be the reason for factitious scoring: "I think the scores that we got were relatively higher than our ability. It might be because the teachers were careful of our self-esteem and might thus lead to the involvement of personal emotions." (T3-S4)

3) Oral assessment rather than written form. Traditionally, Cantonese operatic singing was handed down in the oral form, so with regard to its assessment during teaching and learning: "We are not good at providing assessments in written form. Normally, the comments of student performance would be offered in oral form during the class." (T1)

> Both professional students and amateur students complained multiple times that the details of content were apt to be forgotten when they mentioned oral assessments: "The teachers usually give us comments about our singing orally in class, which I think is good. But since the comments are in oral form, it is hard for me to remember everything, especially after a while." (T4-S1)

4) Incomplete and non-normative. Both teachers and students mentioned that comments or assessments are given randomly whenever problems are detected in class. This is apt to be incomplete information. Since the assessment or scoring follows no previously announced criteria, normalization is hard to ensure:

To be honest, I am afraid the traditional assessment is not normative. Normally we would refer to the student daily performance when marking. For example, if a student does not perform as well as usual, we would still give her/him a better score, like Brownie points. (T2)

The teacher would only correct us whenever we encounter problems. But they would not point out our advantages and disadvantages completely. (T1-S4)

5) General and implicit. Traditionally, assessments in the teaching and learning of Cantonese operatic singing consist of daily oral comments in class and an impression-based total points score at mid-term or at the end of a semester. No monomial points are offered. Teachers, professional students, and amateur students mentioned the general and implicit assessments that were offered in the interviews:

Our teachers normally give us a total point at the end of the semester, but without any monomial points. So, we do not know even how we get this total point and what our problems are exactly. (T3-S3)

Besides daily oral comments during the class, we offer total points only, without any monomial points. I think monomial points come from subentries of assessment criteria. But no one ever bothers to dig into this part. And I think it is hard to do so. (T4)

In the traditional assessment, we usually got stars as an assessment or certificate of merit as an award for outstanding students, which was not as concrete and explicit as the new assessment that you proposed. With the results derived from your assessment framework, we can know our advantages and disadvantages, which is good for our further improvement. (T1-S3)

6) Lack of objectivity and fairness. This disadvantage might be a result of the aforementioned five disadvantages:

We usually got impression-based marking from the traditional assessment. So, I think those scores mixed with personal emotions. The scores we get might depend on the relationship with teachers (T3-S3)

Since there are no unified standards, some teachers are rigorous, whereas the others are loose when scoring. Consequently, some scores are relatively low, whereas some are relatively high, which makes the comparison amongst students unworkable and confuses the students: how exactly did I perform? (T4)

7) Lack of instructiveness. This disadvantage might be a result of the aforementioned six disadvantages: "An effective assessment should enable us to know the advantages and

disadvantages of learning so that we can make further improvements. But the traditional assessment in Cantonese operatic singing fails to hit this mark." (T1-S2)

Some 40 nodes about the proposed assessment framework can be developed from the interview transcriptions, within which two categories can be further delineated: advantages and disadvantages (see Table 5.5).

Table 5.5. Code system of two categories and their family nodes for interview transcription coding (The proposed assessment framework)

Categories	Nodes	Number of nodes
1.	Disadvantages	18
	Not found yet	14
	AS	5
	PS	9
	Increase teacher workload	4
2.	Advantages	133
	Others	5
	Focus on process as well rather than scores only	3
	AS	1
	PS	2
	Convincing	2
	Instructive and Efficient	17
	AS	1
	T	2
	PS	14
	Easy-start and user-friendly	10
	PS	8
	AS	2
	Detailed and All-sided	23
	PS	20
	AS	2
	T	1
	Normative and systematic	6
	T	2
	AS	2

PS	2
Clear	30
PS	18
AS	9
T	3
Documentable	9
PS	4
AS	4
T	1
Objective and Fair	30
PS	18
AS	6
T	6
Transparent	3
T	2
AS	1

All four teachers noted that the only disadvantage of the proposed assessment framework would be that "Teacher workload would be increased":

> It is fine if the assessment is only for the Cantonese operatic singing skill lesson. It would be daunting if I need to assess every subject that I teach in this way, for there are almost 20 students in a class, and I teach several classes. (T4)

> While when the students were asked about the disadvantages of the proposed assessment framework, "not found yet" became the common answer: "I cannot find any disadvantage for now. I think this new assessment framework is well-designed." (T1-S2)

The advantages were tabulated and are shown below:
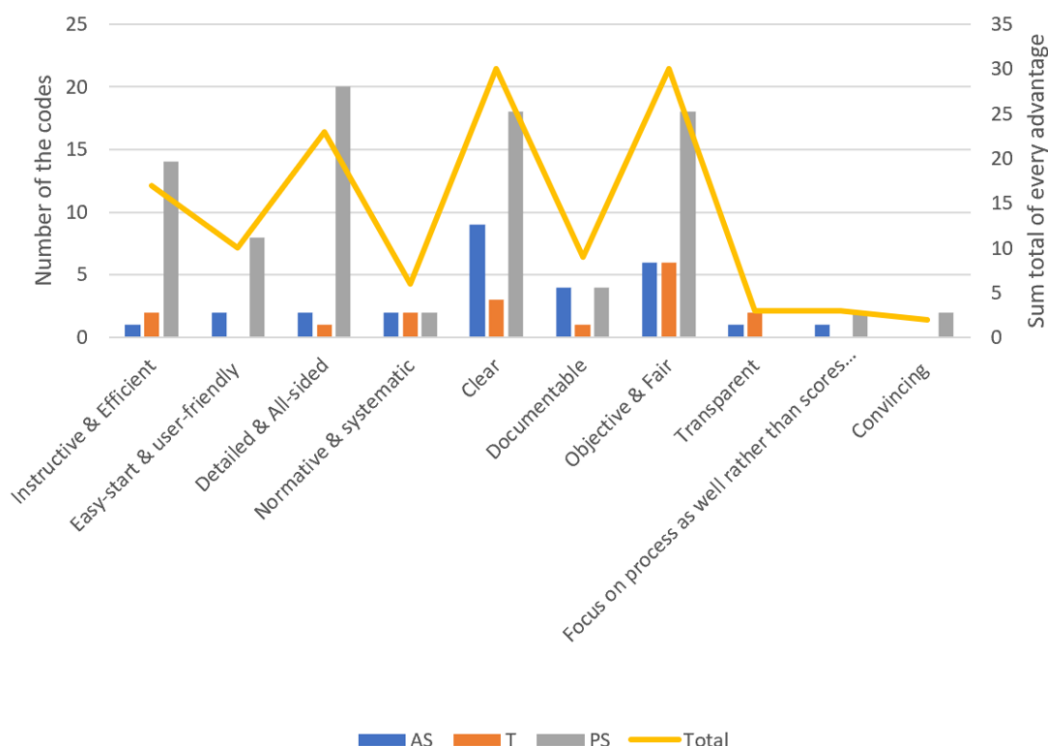
Figure 5.8. Comparison diagram of the proposed assessment framework's advantages

Figure 5.8 shows the most remarkable advantages are: "clear", "objective and fair", and "detailed and all-sided". The sum totals of these advantages were 30, 30, and 23, respectively.

> (Clear): The most outstanding advantage of the new assessment framework is that it enables students to realize their problems more clearly. (T4)
>
> (Objective and Fair): I think the results of the new assessment framework are objective. Because the "rhythm sensation" has always been my daughter's strength, for she has been learning jazziness since she was three and a half years old. And since we are native Cantonese, we speak Cantonese without an accent. That's why she is so good at "pronunciation", which is one of the criteria of assessment. But compared with those who have been immersed in Cantonese operatic singing since an early age, "the flavor" of Cantonese operatic singing is her weakness. Consequently, I think the assessment is objective, which authentic mirrors her advantages and disadvantages. (T1-S4)
>
> (Detailed and All-sided): The traditional assessment would only offer a total point, which is not as detailed as the new one. I think the new assessment framework is more all-sided, which covers subjects of different nature. Even the student learning status and psychology are included. (T2-S1)

"Instructive and efficient", "documentable", and "normative and systematic" were also frequently mentioned. The sum totals of these advantages were 17, 10, and 6, respectively.

(Instructive and Efficient): Merely a total point cannot disclose concrete problems in learning. The new assessment framework can not only generate monomial and total points but also provide progress bars from the perspective of learning for every criterion. Integrating with teacher oral comments in the class, the proposed assessment framework facilitates offering student explicit feedback or instructions on their academic performance. Even the laymen, such as the student parents, would also know which aspects the students need to improve. To students and their parents, the results derived from your assessment framework are more instructive, thereby facilitating their study efficiency. (T1)

(Documentable): Compared with oral assessment, the new assessment framework offers assessment results in written form, which helps us recall the details of every aspect. And we can make profiles with these assessment results for future inquiries or check-ups. (T2-S6)

(Normative and Systematic): It is the first time that we received such a systematic assessment since I started to learn Cantonese operatic singing. I believe that the assessment would be normalized if teachers use this system. (T1-S4)

By contrast, "Easy-start and user-friendly", "transparent", "focus on process as well rather than scores only", and "convincing" were sporadically indicated:

(Easy-start and User-friendly): At first, I worried that it would be time-consuming to learn to use the new assessment framework. But I found it quite simple and easy to use after I gave it a shot. (T4-S1)

(Transparent): The rubrics of the new assessment framework are in analytical form, which makes the assessment key points transparent to the students thereby contributing to their preparation for the examinations. (T2)

(Focus on process as well rather than scores only): The new assessment framework is based on the progress of learning, which enables us to concentrate more on the process of learning rather than the scores only. (T2-S3)

(Convincing): Since teachers make an assessment based on every criterion in the new assessment framework, and the total points were automatically calculated according to teacher judgment of every criterion, I think it would be more convincing with fewer human factors. (T2-S1)

Some 35 nodes were developed from the interview transcriptions to represent interviewees'

attitudes, including thinking little of the factitious scores, endorsements, concerns, a preference

for the assessment framework, and opinions on the assessment results (see Table 5.6).

Table 5.6. Code system of five categories and their family nodes for interview transcription coding (Attitudes)

Categories	Nodes	Number of nodes
1.	Thinking little of the factitious scores	
	PS	7
2.	Endorsements	
	T	
	Advocate for impartial assessment	1
	Better than the traditional assessment in use	4
	Easy to learn to use	3
	AS	
	Detailed assessment in written form	6
	A good supplement to the three-level promotion mechanism	2
	Instructive to study	7
	Facilitate the communication efficacy	2
	PS	
	The value of teacher comments in class	6
	Assessment in written form	1
	Formalism rather than pragmatism in traditional assessments	6
	Advocate for impartial assessment	6
	Instructive to study	14
3.	Concerns	
	AS	
	Not enough explanatory information for parents to understand	5
	Hard to persevere with the study due to opacity	1
	Lack of systematism	3
	High communication cost	2
	PS	
	Traditional assessment mainly focuses on scores	2
	Time is needed to adapt to the new assessment framework	2
	Doubt if the artistry can be assessed	1

T		
	Resisted by elder teachers	1
	Fail in scoring might hurt the students	2
	How to combine with the daily teaching	1
	Imperfection	1
	Practicability	1
	Whether fit for Cantonese opera	1
4.	Preference for the assessment framework	
	Contradiction	
	PS	1
	T	1
	Approval	
	T	2
	PS	12
	AS	8
	Combined use	
	AS	1
	T	1
	PS	1
5.	Opinions on the assessment results	
	Coincide	
	AS	5
	PS	7
	T	2
	Lower	
	T	2
	PS	9

With regard to the first category, several professional students "think little of the factitious

scores":

> I think the comments that we got from the teachers in the class were very useful. But the scores at the mid-term and the end of the semester were too formalistic that we do not take it seriously anymore. (T2-S5)

In the second category, "Endorsement", several professional students stated that they made

light of the scores:

> It is a tacit understanding amongst students that teachers would normally make some adjustments to our scores to ensure the pass rate and avoid make-up examinations. So,

it is easy to get a passing score. The scores are so factitious that we gradually do not mind it anymore and take it as a kind of formalism. (T3-S2)

Both teachers and professional students "advocate for impartial assessment":

For the good of students, I think we should demand them to reach specific standards and provide them with impartial assessments or scores to contribute to their study and development. (T2)

Factitious scores are apt to give us an illusion. It makes us hard to be aware of our real capacity and I do not know how to improve myself, which is not good for our study and future career development. That is why I prefer impartial scores. (T2-S4)

Both amateur students and professional students believed that the proposed assessment

framework is "instructive to study".

The progress bars in the rubrics of the new assessment framework indicated the levels of every criterion that I achieved. It is so explicit that I do not need to guess what my problems are anymore as I used to when being assessed using the traditional assessment. If I do not know what exactly my problems are, I cannot make any improvement no matter how hard I try. But now I know exactly what I need to pay more attention to with the indication derived from the new assessment framework, which is very instructive to my study. (T3-S2)

I did not know to what point and to what extent exactly that I would be assessed in the past, whereas now it becomes very clear that when I received the rubrics of the proposed assessment framework before the assessment is conducted. The rubrics are not only good guidance, with which I would know how to prepare for the examinations but also excellent instructions that indicate the problems and merit that I show in the examination, with which I would know how to improve in the future. (T1-S5)

The teachers endorsed the proposed assessment framework as "easy to learn to use" and

"better than the traditional assessment in use".

("Easy to learn to use"): At first, I worried that the proposed assessment framework was hard to learn to use. But after understanding the rationale and usage, I found it pretty easy to use. (T4)

I found it is pretty efficient to use the proposed assessment framework with an iPad or laptop when assessing students during the examinations. The workload can be dramatically reduced in this way. We used to write down the scores on paper and then input them into the computer system afterward, which is not as easy as this new assessment framework in use. (T2)

("Better than the traditional assessment in use"): I think the new assessment framework is critical to an assessment of high quality, whereas the traditional assessment only offers a general score to students, which is not as detailed as the new

one. I think the proposed assessment is better and is worth popularizing not only in Cantonese operatic singing but also in other art genres. (T2)

The professional students not only advocated for impartial assessment but also endorsed:

1) "Assessment in written form": "I think assessment in written form is better than that in oral form only, for the former would be more convenient for us to review our problems whereas the latter would be easily forgotten." (T3-S2)

2) "The value of teacher comments in class" in traditional assessment: "Although the scores in the traditional assessment have little reference value to us, teacher comments in class are always to the point, which is much more useful than the scores." (T2-S5)

3) It is "formalism rather than pragmatism in traditional assessment":

> We get two general scores in every semester: one is given in the mid-term; the other is at the end of term. But this kind of general score makes no sense, for it cannot indicate any specific problems for further improvement. So, it is rather formalistic to us. (T2-S5)

The amateur students and their parents endorsed:

1) "Detailed assessment in written form":

> The teachers would correct our problems whenever s/he found them rather than indicate all our problems and merit roundly. Besides, the traditional assessment is in oral form rather than written form, which leaves us no materials to thumb through or make a comparison with the past. Thus, I prefer a detailed assessment in written form so that my merit and demerit are clear at a glance. (T1-S5)

2) The proposed assessment framework is a "good supplement to the three-level promotion mechanism" at Children's Palace:

> The teachers used to offer comments to the students in the class once a week or to the parents in the parents' meeting, which is too limited for me to know my kid's study status. I think the new assessment framework enables the students to be aware of their merit and demerit, which is instructive to their studies. It is not contradictory with the

three-level promotion mechanism in the Children's Palace at all. On the contrary, it is a good supplement that facilitates students to stand out from the audition. (T1-S2)

3) The proposed assessment framework "facilitates the communication efficacy":

The new assessment framework is conducive for the parents to know more about their children's learning progress and status. Even if we want to further communicate with the teachers, we would be more well-founded and niche targeting, which will greatly increase the communication efficacy between parents and teachers. (T1-S2)

On the other hand, ten nodes were abstracted from the interviews that represented interviewees' opinions on the affective domain. These can be grouped into two categories: absent and necessary (see Table 5.7).

Table 5.7. Code system of two categories and their family nodes for interview transcription coding (Opinions on the affective domain)

Categories	Nodes	Number of nodes
1.	Absent	
	AS	5
	T	4
	PS	13
2.	Necessary	
	AS	3
	T	3
	PS	
	Approbatory	4
	Indifferent	2

Most of the students and teachers admitted that the affective domain is absent in the existing traditional assessment:

The traditional assessment mainly focuses on our academic performance, whereas concerns in our psychology or status of learning were rare. Sometimes the teachers would make comments about student behavior at the end of the class. But no specialized report of this aspect would be offered to the individuals periodically. (T1-S4)

Some of the interviewees believed that concerns about the affective domain were "necessary":

> It is definitely better to involve the affective domain, for there always are ups and downs in student learning status and psychology. I think we should keep a watchful eye on this aspect. Interventions should be initiated once dramatic fluctuation or an acute situation arises, which is very helpful for student growth and development. (T1-S6)

However, one teacher and some of the students doubted the necessity of involving the affective domain, which reflected a neglect and incomprehension of this aspect:

> I do not think the affective domain is necessary for my students. It should be used in universities instead of in polytechnic schools, such as our school. Because even though some of my students have problems in the psychology and status of learning, these problems would be solved or disappear by giving attention to in the class or consolations for a few days. I think you can involve this aspect in the senior class, whereas it is not necessary for junior classes. (T4)

Finally, another eleven nodes were developed from the interviewees' suggestions during the interviews, which were subsumed into three categories: AS and P (amateur students and their parents), T (teachers), and PS (professional students). See Table 5.8:

Table 5.8. Code system of two categories and their family nodes for interview transcription coding (Suggestions)

Categories	Nodes	Numbers of Nodes
1.	AS and P	
	Renaissance of Cantonese traditional culture	1
	Cultivation of both performers and audiences	2
	Systematisms and normativity in education	4
	Comparison sheet	1
2.	T	
	More explorations in details	1
	Generalize to more subjects	1
	Descriptions of standards can be more accurate	1
	Grids in rubrics can be more visualized	1
	Questions in AD should be simpler and more specific for kids	2
3.	PS	
	Assessment framework introduction should be easier and clearer	1
	Generalize to every subject	2

The amateur students and their parents proposed multiple macroscopic suggestions:

1) Renaissance of Cantonese traditional culture. The decline of regional culture is a non-negligible phenomenon that jeopardizes the development and transmission of traditional culture:

> I think in addition to education, the cultural atmosphere is also notable. Nowadays, a great number of children in Guangdong Province cannot speak Cantonese. Instead, they speak in Mandarin. Since the Cantonese language and culture are the roots of Cantonese operatic singing, which greatly affect the spread and transmission of this traditional genre. For example, my daughter's classmates do not even understand what she sings at all not to mention the appreciation of Cantonese operatic singing. (T1-S6)

2) Cultivation of both performers and audiences. The prosperity of an art genre demands excellent performers on the stage and savvy audiences off the stage, which makes a virtuous cycle of the spread and transmission of an art genre:

> I think the audiences of Peking opera have a better understanding of *Xiqu* than those in Cantonese operatic singing. The former applauds and acclaims just at the right point,

whereas the latter usually become exciting whenever the performers do some attractive action, such as somersaults. It is just like the old saying: "The insider knows the ropes, while the outsider just comes along for the ride." Most of the time, performers would have to cater to the audiences' aesthetic and preference, which gradually impacts the development of this art genre. Thus, I think the cultivation of audiences is as important as that of performers. Even if the amateur students will not become professional students or performers in the future, they will most likely grow into fans of *Xiqu*. So, I hope the authority attaches more attention to the cultivation of us, amateur students. (T1-S6)

3) Systematisms and normativity in education. Professional training in the conservatory tradition and community training has taken the place of the apprenticeship and become the main educative and transmissive approach of contemporary society. However, the amateur students and their parents noted that community training is not systematic or normative:

> As far as I know, the teaching of Cantonese operatic singing lacks a system. Unlike other subjects in school, there is no textbook or even syllabus in Cantonese operatic singing. My daughter has been learning for so many years. Yet, we have not ever seen a textbook or syllabus. The teachers give lessons merely based on their own experience. (T1-S3)

4) Comparison sheet. One of the amateur student parents advocated making comparisons, so that learning progress could be evident and overt:

> A new assessment framework is a good approach. If only comparison sheets can be tabulated in the future, which will be a useful indicator of the student progress and further facilitate the teaching and learning. After all, as parents, we need something to prove if we should let our children continue learning. (T1-S4)

The teachers proposed some suggestions based on their professional experience. First, they advocated for "more explorations in details". The teachers affirmed the feasibility and usefulness of the proposed assessment framework after the pilot implementation, but they hoped that more details could be developed to adapt the assessment to Cantonese operatic singing:

In general, I think this new assessment framework is workable. But more explorations in the details are necessary, such as the weightings, criteria, and so on. (T3)

Agreed. I have drafted general criteria for the pilot implementation at present. After all, the criteria and weightings must be the same to make comparisons amongst partitioners during the pilot implementation. But teachers are encouraged to revise the description of criteria or adjust the weightings according to their needs or understandings in ordinary use. (Interviewer)

That would be fantastic. (T3)

Secondly, there were suggestions to "generalize the proposed assessment framework to more subjects". To evaluate the structural function of this new assessment framework, the rubrics of every domain were applied to one representative subject within the current study as a pilot implementation. For example, the psychomotor domain was used to assess Cantonese operatic singing, though it could be applied to essential technique classes or gesture training classes as well: "I think this new assessment framework is fine for now. But whether is good enough needs more explorations. I suggest generalizing it to more subjects, such as essential technique, and the like." (T3)

Thirdly, there was an expectation that "descriptions of standards can be more accurate". Due to the particularity of terms in *Xiqu*, and subtlety of translation between English and Chinese, the descriptions of standards need to be further polished to accommodate the practice:

I think some of the descriptions of standards in the rubrics are not very appropriate. For example, I think "1.0 Acquire", "2.0 Assimilate" and "3.0 Adapt" are farfetched to be used in describing the learning hierarchies of "the appearance in costume and makeup". It would be better if this kind of description can be more suitable for *Xiqu*. (T2)

Fourthly, the researchers were advised that the "grids in rubrics can be more visualized". Since the proposed assessment framework employed analytical rubrics, several subentries of assessment criteria were included. To make the assessment more precise, every hierarchy was divided into four grids (see Figure 5.9 as an example). Teachers were required to select one of

the grids in every subentry to reflect the student level achieved during their performance. Thus, from the perspective of user-friendliness, it would be more efficient and convenient if the grids were more visual, so that teachers were able to make judgments and select grids with a glance.

Standards / Items	Low → High					Weightings
	1.0 Acquire	2.0 Assimilate	3.0 Adapt	4.0 Perform	5.0 Perfect	
Technical requirements						
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Pulse & Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Cantonese pronunciation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Changqiang	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Artistic requirements						
Artistic expression in line with the repertoire	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
The appearance in costume and makeup	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
					Score:	0

Figure 5.9. An example of rubrics of the proposed assessment framework

> I think the rationale of this new assessment framework is advanced. I also understand that it would be more precise to subdivide every hierarchy into four incremental sublevels. I think it would be better to make the grids more visualized or convenient to be selected during the assessment. Otherwise, teachers will not bother to use it in practical application or daily teaching. (T4)

Finally, there was an observation that "questions in affective domain should be simpler and more specific for kids". A portion of amateur students are children who cannot comprehend survey questions that are too abstract. Consequently, the questions in the affective domain should be drafted based on the children's understanding.

> I think the survey questions of the affective domain are too complicated or abstract to children. After all, children's comprehensive ability cannot compare with that of adults. The questions can be more concrete. For example, when probing the kids' interest in

Cantonese operatic singing, we can ask "how many times have you watched Cantonese operatic singing in the last three months?" (T1)

Similarly to the teachers, some professional students also suggested: "generalize the proposed assessment framework to more subjects": "After the pilot implementation, I found that this new assessment framework provides us with explicit, concrete, and fair assessments. I hope that it can be employed in every subject in our school." (T2-S6)

One of the professional students noted that "the introduction of the proposed assessment framework should be easier and clearer": "I think the introduction is a bit too academic, which is hard for us to understand. If only it can be written more straightforwardly." (T3-S6)

5.3.2 Thematic networks and their connections

The thematic networks and the relations between the nodes that elicited from MAXQDA are illustrated in Figure 5.10:

Figure. 5.10. Thematic networks of the semi-structured interview in Phase II
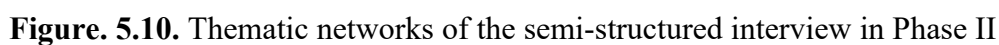
Three categories were encompassed within this schematic, including traditional assessments, the proposed assessment framework, and attitudes. Three types of relations were proposed amongst 445 nodes, including "associated with" (denoted by the yellow arrows), "influenced by" (denoted by the red arrows), and "solved by" (denoted by the blue arrows):

Eight pairs of interactions were detected in the first kind of relationships ("associated with"):

1) "Total points only and without individual points" might result in "factitious scoring" in the traditional assessment.

2) The teachers implied that "traditional ideology" was one of the reasons for "factitious scoring". Professional students also believed that "factitious scoring" was "affected by human factors", of which "traditional ideology" was one.

3) Both teachers and professional students believed that two reasons for "factitious scoring" were "protecting the student self-esteem" and "avoiding failure rate".

4) Because teachers "avoid failure rate" intentionally, "the pass rate can be controlled and ensured".

5) The teachers and the students stated that the proposed assessment framework was "objective and fair", which might be associated with one of its advantages, "convincing".

6) Although amateur students and professional students are in different modes of education, they all appreciated that the proposed assessment framework was conducive to study.

7) Some professional students endorsed "the value of teacher comments in class" about their academic performance, which might have an association with their advocacy of a "combined use" of the traditional and the proposed assessment framework.

8) Two of the main disadvantages of the traditional assessments, "lack of objectivity and fairness" and "lack of instructiveness" have close relationships with multiple disadvantages, such as "total points only and without individual points", "factitious scoring", "incomplete and non-normative", "oral assessment rather than in written form", and "general and implicit". Some practices were mentioned by the interviewees that might result in a "lack of objectivity and fairness in the traditional assessment" and a "lack of instructiveness" (see Figure. 5.11):
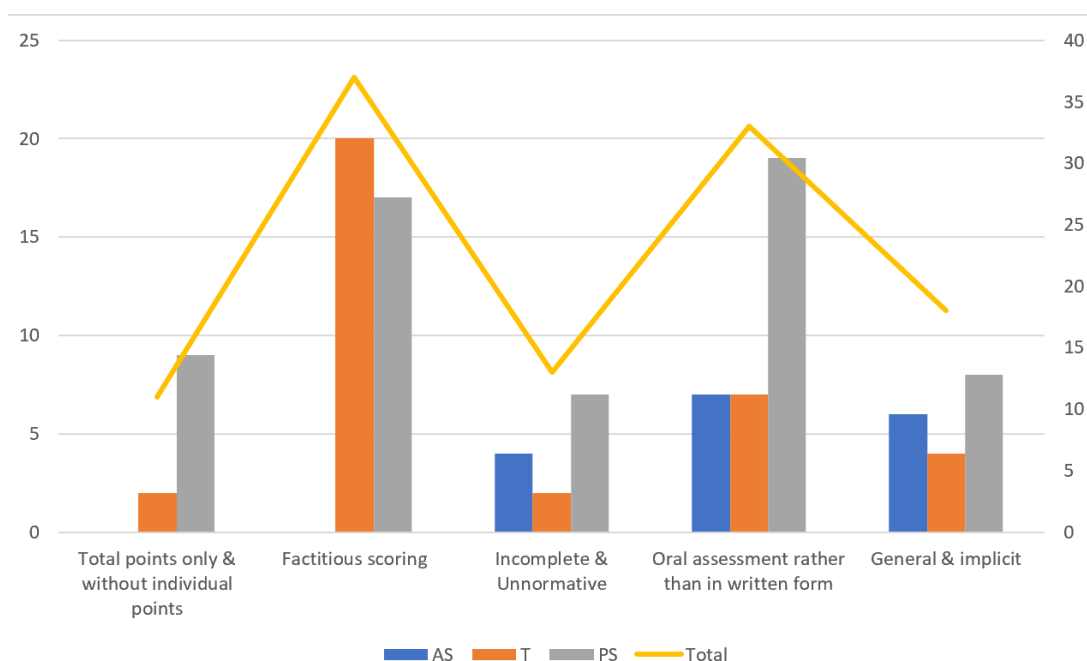
Figure 5.11. The practice that might result in the lack of objectivity and fairness

As we can see in the above schematic, "factitious scoring" is the most salient problem in the professional education of Cantonese opera. It was mentioned 37 times in the interviews. From both the perspectives of teachers and professional students points, it helps to "avoid

failure rate" (mentioned 13 times by teachers and six times by professional students) and will not "hurt the student self-esteem" with scores below the pass rate (mentioned 13 times by teachers and six times by professional students). The teachers also let slip that it might derive from their "traditional ideology", whereas the professional students believed that factitious scoring was "affected by human factors", such as bias. Besides, some of the teachers (mentioned twice) and professional students (mentioned nine times) pointed out that only "total points are given" within which "no individual points are included". By contrast, no amateur students or their parents mentioned it. They did not receive points but stars or merit certificates as an alternative.

Furthermore, the teachers (mentioned seven times), professional students (mentioned 19 times), and amateur students (mentioned seven times) admitted that "oral assessment rather than in the written form" is prevalent. Both the professional students (mentioned eight times) and the amateur students (mentioned six times) asserted that oral assessments are apt to be forgotten. There is also a widely held belief amongst the interviewees that the process and results are "incomplete and non-normative" as well as "general and implicit".

In the second kind of relationships ("influenced by"), nine pairs of relationships were perceived:

1) Getting to know (and learning to use) a newly developed assessment framework demands extra time and endeavor. The proposed assessment framework requires teachers to assess specific points rather than give a total point score, and is therefore more detailed. Thus, teachers believed that it would "increase their workload", which

might result in "resistance of using the new assessment approach, including this new assessment framework".

2) The multiple advantages of the proposed assessment framework might be the main reason for some of the interviewees' endorsements. These included "better than the traditional assessment in use", "easy to learn to use", and "a good supplement to the three-level promotion mechanism" at the Children's Palace.

3) One of the interviewees' endorsements, "easy to learn to use", might be influenced by "easy-start and user-friendly", which was one of the noted advantages of the proposed assessment framework.

4) One of the interviewees' endorsements, "instructive to study" might be influenced by "instructive and efficient", which was one of the noted advantages of the proposed assessment framework.

5) One of the interviewees' endorsements, "detailed assessment in written form" might be influenced by "detailed and all-sided", which was one of the noted advantages of the proposed assessment framework.

6) "Contradiction", one of the interviewees' options regarding "preference for the assessment framework", might be influenced by one of their concerns: "time is needed to adapt to the new assessment framework".

7) "Lack of objectivity and fairness", one of the disadvantages of the traditional assessment, might affect interviewees' multiple endorsements and concerns, including "thinking little of the factitious scoring" and "advocating for impartial assessment".

The scores that some of the average student and underachievers' received under the proposed assessment framework were "lower" than their daily ones, and did not include "enough explanatory information for parents to understand" Cantonese operatic singing and their children's academic performance in this area.

8) One of the teacher endorsements, "advocate for impartial assessment", might derive from "factitious scoring", which is one of the noted disadvantages of the traditional assessments.

9) "Untransparent", which was noted by amateur students and their parents to be one disadvantage of traditional assessments, may reflect some of their concerns, including "not enough explanatory information for parents to understand" and "high communication cost".

In the third kind of relationships ("solved by"), four pairs of relationships were noticed:

1) The professional students expressed concerns that "the traditional assessments mainly focus on scores", which might be solved by the proposed assessment framework. One of its noted advantages was that it "focused on process … rather than scores only".

2) Another professional student concern was that "time is needed to adapt to the new assessment framework". This might not be a problem since "easy-start and user-friendly" was considered to be one of the proposed assessment framework's advantages.

3) "Transparency", one of the proposed assessment framework's advantages, contributes to one of the traditional assessment's disadvantages, "opacity", and mitigates some of

the amateur students and their parents' concerns, such as "facilitate the communication efficacy", "high communication cost", and "hard to persevere with the study due to opacity".

4) "Lack of instructiveness" was deemed to be one of the traditional assessment's disadvantages, which might be solved if the proposed assessment framework is used, as "instructive and efficient" was believed to be one of its advantages.

5.4 Summary

This chapter reports that a proposed assessment framework, formulated according to the outcomes of Phase I, was put into a pilot implementation. Subsequently, the opinions of teachers, professional students, and amateur students (and their parents) were collected via a series of semi-structured interviews. Instructive thematic networks and their connections were extracted with thematic analysis. The findings revealed that the disadvantages of traditional assessments of Cantonese operatic singing outweighed the advantages. These disadvantages included "lack of instructiveness", and "lack of objectivity and fairness". After the pilot implementation of the newly proposed assessment framework, most participants agreed that, when compared with traditional assessments, the new assessments offered better assessments that featured multiple advantages. Participants thought the new assessments were more explicit and transparent. Interviewees expressed their attitude towards the assessments, proposed insightful suggestions, and gave their opinions of the assessment of the affective domain. The

findings of this chapter responded to the last research question. Further insights into Phase I and Phase II will be discussed in Chapter 6. A conclusion will also be given.

Chapter 6

DISCUSSION AND CONCLUSION

6.1 Introduction

The previous chapters presented the results and findings of this study. In this chapter, the discussion of the results and conclusion will be given. This chapter is divided into seven sections. Section 6.2 provides a discussion of the findings of the two studies. Section 6.3 describes the significance of the study. Section 6.4 presents the implications of the study. Section 6.5 discusses the limitations of the study. Section 6.6 offers some recommendations for further research. Finally, section 6.7 draws the study to a close-by delineating some conclusions.

6.2 Discussion of Findings

Cantonese operatic singing experienced a series of evolutions and took shape across the 1920s and 1930s (Wang & Liang, 1990). Mainstream research has concentrated on its history and musical ontology instead of its education. Attention to the assessment of Cantonese operatic singing is even scarcer. The primary goal of this study is to propose and pilot an assessment framework for Cantonese operatic singing for the sake of its education and transmission. Four research questions were posed during the process of approaching this goal:

1. What are the design and practices of the existing graded examinations of Chinese opera?

2. For the sake of transmission and education, how can an ideal assessment framework for

Cantonese operatic singing be formulated?

3. What is current common practice in assessments of Cantonese operatic singing?

4. To what extent does the proposed assessment framework facilitate the teaching and learning of Cantonese operatic singing?

Surrounding these research questions, a range of content analysis, semi-structured interviews, and thematic analysis was conducted in two phases, within which multiple insights were revealed.

6.2.1 Design and practices in the existing graded examinations of Chinese opera

To respond to the first research question, a content analysis of two representative graded examinations in *Xiqu* was implemented. The content analysis subjects of the current study, POGE and COSGEHK, are still in development. The contents analysis of these two publicly graded examinations in Phase I suggest that COSGEHK features a normalized assessment formulation with comparatively comprehensive test items and content. No regularity could be perceived in POGE's construction or its test items and no complete rubrics are published for the general public. There might be two main reasons for this state of affairs.

There are different views and beliefs in Western and Chinese culture. Wang et al. (2021) indicated that "comparing with the West, which is excelled in using logical and analytical thinking modes and utilizing rational cognition, China is far better at using dialectical and holistic thinking modes and applying intuitive comprehension" (p. 1). That is, Chinese people generally prefer holistic thinking, where a whole background is considered and tend to follow

the principle of integrity (Nisbett et al., 2001). This traditional ideology affected thinking and acting when formulating the assessment framework. Thus, the salient difference between COSGEHK and POGE lies in the design of the rubrics: in cooperation with UWL, COSGEHK adopted an analytical format with separated assessment components as criteria, following Western analytical thinking that features a Westernized system of logic. Designed by genre artists, POGE employs a holistic format in its incompletely published criteria. Therefore, unlike COSGEHK, which assesses discreet components of Cantonese operatic singing, POGE attempts to assess an entirety. Significantly, the details of professional practitioners or a mature artist's performance are typically blended into one harmonious whole and are hard to express without reference to holistic Chinese thinking. This is embodied in the descriptions of the criteria in POGE. "Criteria are statements that identify the key characteristics or attributes of student performance in an assessment task" (Armstrong, S. et al., 2008, p. 5). If assessors and candidates cannot accurately comprehend an assessment, it necessarily becomes invalid. Some of these descriptions, such as "醇厚挺拔" (mature and smartly performed) in POGE, are representatives of the "intuitive comprehension" used to describe the performance of *Xiqu* from a traditional Chinese perspective. These descriptions are so hard to understand explicitly that different assessors or candidates might have different interpretations, thereby leading to ambiguity or misunderstanding. Consequently, a long-term lack of explicit ground in assessments has resulted in subjective perception rather than criteria- and standards-based assessment, which is the most prominent practice across existing grade examinations of *Xiqu*.

A lack of educational philosophy and knowledge is another reason for the non-normative and unsystematic graded examination design. Before the conservatory mode came into being,

apprenticeship was the primary transmission mode of Cantonese opera. This dated from the early twentieth century (Leung, 2015), and featured informal learning (Yuen 1993) and master-centered teaching (Leung, 2015). Rather than formal learning in class, apprentices learned by observing and imitating a master's performances (Leung, 2014). In this traditional mode of transmission, nurturing professional performers was the main goal of professional education, and practitioners tended to concentrate on skill acquisition (Leung, 2015), rather than other aspects, like education-related knowledge and competencies, not to mention assessment training. For example, most of the teachers admitted that they were not clear about the different roles of assessment and could not use the assessment *of, for,* and *as* learning distinctively and accurately. As performers and educators typically have different perspectives, their understanding of assessment and other educational issues consequently differs. The designers' lack of educational philosophy and knowledge might have led to the lack of modern education ideology in the assessment design, resulting in a deficiency.

By contrast, COSGEHK, formulated by the Hong Kong Association of Cantonese Opera Scholars and the University of West London (UWL), is based on UWL's century-old experience of examination design and administration. It shows innovation in the amalgamation of the traditional art genre with modern education. A long-tested assessment framework is a structured blueprint for the study program's learning outcomes along with details of how achievement of the outcomes can be measured. This contributes to improving the validity and reliability of the assessment (Pearce et al., 2015). UWL's assessment framework enabled the test developers of COSGEHK to create robust assessment instruments more easily and efficiently. The comparatively normative COSGEHK featured cooperation with education

experts when formulating the assessment of the traditional genre, rather than only focusing on artistry. In contemporary studies, education is a comprehensive concept consisting of three sections: curriculum, instruction, and assessment, which are inextricably intertwined (Duke, 2009). As an essential component of education, assessment demands educational expertise. For example, Stronge (2018) stressed that the alignment of curriculum, instruction, and assessment was of immense importance, and is integral to any instructional design of high quality. Alignment is "the deliberate linking of stated learning outcomes, teaching and learning activities, and assessment tasks to promote consistency between what is learned and assessed" (Armstrong., 2008, p. 4). One of the critical responsibilities of educators is to ensure the alignment of learning outcomes with curriculum, instruction, and assessment (Biggs and Tang, 2007). Cooperation between experts drawn from the Cantonese operatic singing circle and experts drawn from the education circle facilitates the bringing together of the best of both sides. It might be an effective way to formulate a robust, standardized assessment in a short time.

Although the design of COSGEHK is comparatively normative, there is a critical flaw that challenges its rationality. Initially, the test item design of COSGEHK seems to be comprehensive, as it includes singing, Q&A, sight-singing, and oral tests, which aids an assessment of the candidates' competence. However, the allocation of scores might break down the seemingly comprehensive examination design: singing accounts for as much as 80% of the total points available in COSGEHK, with the rest of the points won from the Q&A (10%), sight-singing (4%), and oral tests (6%). The pass mark is set at 65%. That is to say, candidates could pass the exam by focusing on the singing set pieces. This makes the other three

examination items dispensable. As a result, this graded examination is an achievement-based assessment, rather than one based on competencies. The repertoire that candidates can use for their singing test is pre-set and can be prepared for, whereas the Q&A, sight-singing, and oral tests are examined on the spot and cannot be prepared in advance. This could be argued to jeopardize the rationality of COSGEHK's design. Details and specific practices should be given attention when formulating an assessment framework, to ensure function and efficacy. By contrast, ABRSM used to set cut-off scores for each part of its graded examination before it was reformed, including recital, sight-reading, scales and arpeggios, and oral test. The recital could be prepared beforehand, whereas the other examination aspects could not. These were good tests of candidate performance, achievement, and musical competence. Candidates were required to exceed the cut-off score of every part to pass the full examination. This measure contributed to ensuring the test of every detail in this examination had to be fully implemented, rather than evaded.

6.2.2 To formulate an ideal assessment framework for Cantonese operatic singing

The content analysis of the two graded examinations of *Xiqu* was salutary for formulating a new assessment framework. It was acknowledged that the firsthand information and opinions from the experts and teachers of this traditional genre should be valued. As a result, a series of semi-structured interviews were conducted. With regard to transmission and education, findings from the thematic analysis of the interview transcripts prompted a focus on the assessment of student learning in art education. They revealed that a criterion- and stand-based

system should be employed to formulate an ideal assessment framework for Cantonese operatic

singing, and five supporting measures for the formulation were extracted.

6.2.2.1 Orientation of the assessment subjects

In the semi-structured interviews, teachers, students, and their parents admitted that

"assessment is significant and necessary in education". Effective assessments focus on

essential learning goals, which should be clear to both students and instructors (Suskie, 2018).

But *Xiqu* is thought to be very personalized, which is also part of its charm and accords with

the pursuit rooted in Chinese philosophy: "harmony without uniformity". Experts noted that

since there are various sects exist in this traditional genre, it is hard to agree standards which

surpass those extant, asserting that the "standards lie in the audiences' preference thereby

impossible to be unified". Thus, it might lead to a general misunderstanding prevalent in the

Cantonese operatic singing circle: "as an art, *Xiqu* is hard to be assessed". Some participants

noted that the fundamental components of Cantonese operatic singing could be assessed with

flexibility and that student learning from the perspective of art education was more workable.

However, since 'flexibility' is hard to define, it is easy to provoke controversy. The current

study aimed to formulate an assessment framework for the education and transmission of

Cantonese operatic singing. Therefore, it was decided to propose an assessment framework for

student learning from the perspective of art education rather than from the perspective of

artistry.

Nevertheless, there is an almost universal misconception that student learning in music

cannot be assessed like other subjects, such as science and mathematics (Hanna, 2007). Music

learning involves a comprehensive interweaving of cognitive, affective, and psychomotor domains. The current study's theoretical taxonomy framework is valuable for categorizing teaching and learning into different educational domains and dividing standards of learning progression into progressive hierarchies. Thus, with the application of the current study's theoretical framework, assessment criteria could be teased out within corresponding categorized educational domains, and assessment standards from the perspective of student learning could be identified. Anderson et al. (2002) noted that Bloom's Taxonomy is a special kind of framework, which can be used to classify statements of what is expected of students as a result of instruction. It is often depicted in the form of a pyramid, illustrating "cumulative hierarchy: hierarchy because the classes of objectives were arranged in order of increasing complexity, and cumulative because each class of behaviors was presumed to include all the behaviors of the less complex classes" (Krietzer et al., 1994, p. 66). The understanding of each increasingly specific category was assumed to be a prerequisite to proficiency in each increasingly complex one (Krathwohl, 2002, p. 213). There was a requirement that prior skills or abilities must be mastered before the next learning goal was approached. One purpose of a taxonomy is to break this overall development process into smaller parts. It makes it easier to discuss educational goals, construct achievement metrics, and evaluate individual achievement. The theoretical framework of the current study that stems from Bloom's Taxonomy was helpful guidance when formulating an assessment framework for student learning in Cantonese operatic singing.

The orientation of the assessment subjects is essential to formulating an ideal assessment framework for Cantonese operatic singing. Since the research for this study focuses on learning

rather than performance, the proposed assessment framework concentrates on student learning

for the sake of the education and transmission of this traditional genre. Consequently, the

theoretical framework derived from Bloom's Taxonomy was established to guide the

formulation. These are the essential foundation for the assessment framework formulation,

which partially responds to the second research question.

6.2.2.2 Adoption of the criterion- and standard-based assessment

A salient finding that emerged from the content analysis indicated that assessors depend

on subjective perception rather than a criterion- and standard-based assessment. The expert

opinions extracted from the semi-structured interview verified this finding. Many teachers

evaluate art based on vague personal thinking or preference without regard to specified criteria

that help determine academic growth and learning (Day, 1985). This tends to be highly

subjective and based on personal aesthetic biases. Several researchers have indicated that this

subjectivity can be reduced if explicit assessment criteria that are compatible with teaching

objectives are developed at the outset of the lesson planning process (Gruber, 1994; Hardiman

& Zurnich, 1981; Lowenfeld & Brittain, 1975; Rush, 1987, 1989). In addition to being assessed

by feeling, the testing of Cantonese operatic singing lacks assessment feedback and easily

available published criteria. Criterion- and standard-based assessment is the key to obtaining a

reliable assessment.

Criterion- and standard-based assessment: 1) is based on the academic achievement of the

student, 2) establishes clear guidelines for proficiency, which is also the foundation of rubrics,

and 3) compares each student performance to preset criteria and standards, not to the

performance of other students (Armstrong, S. et al., 2008). For students, criterion- and standard-based assessment provides goals to pursue and achieve, which acts as guidance for their learning and judgment of their work. For instructors, it is beneficial to use criterion- and standard-based assessments to ensure better reliability of marking when multiple markers are engaged and better consistency over time for individual markers. For schools or educational institutions, adopting criterion- and standard-based assessment ensures quality of education (Hendry et al., 2012). To establish robust criterion- and standard-based assessments that align with curriculum, instruction, and assessment, Bloom's Taxonomy of Educational Objectives was introduced into the current study as a theoretical framework. The proposed assessment framework was formulated in this manner.

6.2.2.3 Supporting measures for the assessment framework formulation

The conservatory mode had taken the place of the apprenticeships to become the primary transmission mode of Cantonese opera by the early twentieth century (Leung, 2015). Alterations to ideology and understanding in contemporary education are not as prominent as they are in modalities such as the learning mode. Traditional assessment was passed down from one generation to the next, thereby continuously affecting assessment practices in teaching and learning. Research or alteration was rarely seen. Thus, despite the disadvantages of the traditional assessments disclosed by the interviewees in the semi-structured interviews, it has remained the first choice for most teachers or practitioners in this realm. The elderly teachers attached little importance to assessment and also showed conservative views towards trying a new assessment approach. Even the younger teachers who took part in the pilot implementation

exuded the custom of master-centered teaching that is one of the characteristics of apprenticeships (Leung, 2015). One of the teachers spoke candidly that the proposed assessment framework might not be effective for students; the traditional approach might be better for teachers because it had been habitually used. Some experts admitted they were "trained as performers rather than teachers" and "teaching and performing are not the same things". Therefore, when they became instructors, a deficiency in pedagogical training manifested itself. An amateur student's parent bluntly stated that she thought the teachers gave lessons merely based on their experience. No course outline or textbooks were seen, which also marked a close relationship with the previous experience of teachers as performers, rather than instructors. They also knew little about assessments because of a lack of pedagogical training. For example, what are the categories of assessments? How are they applied in different contexts or for different purposes? How are assessment results best interpreted? Since assessment is an indispensable component of education and one of the critical qualities of effective teaching (Stronge, 2018), making effective use of it can considerably boost teaching and learning. Besides using assessment, teachers are often inevitably involved in the formulation of assessments. Therefore, as the participants in (and implementors of) assessment, assessor and teacher training is of significant importance to formulating an ideal assessment of Cantonese operatic singing. This should attract attention. Colwell (2003) claimed that "if we cannot provide a well-defined and rigorous education in assessment for these teachers, we should curtail all discussion about the assessment in art" (p. 17).

The Chinese philosophy behind traditional apprenticeships emphasizes industrious learning over a lengthy period (Leung, 2014) but neglects systemization and efficiency in

teaching. Practitioners of Cantonese operatic singing are inclined to practice their artistry rather than strive for excellence in educational practice in the past and the present. This might be because of their identification as performers, rather than educators, which stems from their cultivated experience. As a result, teaching in Cantonese operatic singing is usually based on personal experience without pedagogy and does not make full use of assessment as one of the components in education. In the master-apprentice relationship, the masters (usually great performers) were apt to assess according to their subjective experience, without employing explicit criteria. This might even include corporal punishment (Leung, 2015). Apprentices were inclined to obey masters, yield to their authority, and accept assessments whether they were right or wrong or explicit or implicit. However, nowadays, students receive contemporary education and are used to institutionalized practices and modern educational ideologies. Burke (2010) indicated that today's teachers are expected to be assessment literate and comprehend a complete picture of the preparation and practice of assessments. Outdated ideologies and assessment training make it difficult to meet student needs or gain their endorsement, which might harm the effect of any assessment. For example, compared to the "master-centered" apprenticeship, "student-centered" instruction is one of the most salient perspectives of contemporary education (Krahenbuhl, 2016). Thus, nowadays, the consideration and decision of strategies in education should be based on a student-centered perspective, rather than a master-centered perspective. Since assessment is an essential component of education, and demands educational expertise, teachers of Cantonese operatic singing who were initially trained as performers and are lacking pedagogical training might find it difficult to assess accurately and making full use of the process. Introducing educational philosophies and

knowledge that conform to contemporary education to this traditional genre (thereby cultivating educational ideology) should be valued, so that instructors can fully comprehend the significance of assessment in education. This will contribute to the willingness to take part in assessment formulation and to carry out new assessment frameworks faithfully.

Besides valuing the introduction of educational ideology into this traditional genre, maintaining its traditional artistic characteristics is also essential. As a branch of *Xiqu*, Cantonese operatic singing has unique features. Designing an assessment framework that conforms with aesthetics and philosophy that are rooted in the East and also conform to the educational rationales of contemporary society is important. The comparatively successful experience of the design of COSGEHK suggests that cooperation between artists in Cantonese operatic singing and experts in education is a smart move. Moreover, with respect to the roots of Cantonese operatic singing, the theoretical framework proposed in the current study can be applied to the process of teasing out assessment components and ensuring these components are articulated in common educational terminology that can be understood by musicians and non-musicians alike and both the Eastern and Western world. That is another reason Bloom's Taxonomy was introduced into the current study.

A user-centered design approach was adopted in the current study to formulate an ideal assessment framework for the Cantonese operatic singing, which included three stages: pre-design analysis, the creation of design solutions, and post-design evaluations (Evans, 2002). In the research design of the current study, pre-design analysis was defined by the content analysis of two representative graded examinations of *Xiqu*, together with semi-structured interviews with experts and teachers of Cantonese operatic singing. Based on these findings and the

study's theoretical framework, an assessment framework for Cantonese operatic singing was formulated. Finally, a pilot implementation of the proposed assessment framework and subsequent semi-structured interviews with teachers and students were conducted to collect qualitative information regarding satisfaction and any problems that remained to be solved.

User-friendly design is the driving force behind innovation (Rubin, 1994), and this was mentioned by several experts during the semi-structured interviews. Computer-aided assessment rubric toolkits were designed to formulate user-friendly assessments for Cantonese operatic singing. During the formulation process, algorithms were applied to the rubrics. Scores were generated after automatic calculations triggered by teacher selections in the grids to avoid human biases, ensure objectivity, and increase the efficiency of the assessment.

To formulate an ideal assessment framework for Cantonese operatic singing from the perspective of transmission and education, several aspects should be given attention to: 1) valuing the introduction of educational ideology into this traditional genre, 2) cooperation between artists in Cantonese operatic singing and experts in education, 3) application of the theoretical framework derived from Bloom's Taxonomy, 4) user-centered design approach, and 5) computer-aided assessment rubrics toolkits. Thus, in the current study, the rubrics of the assessment framework were based on the theoretical framework. The criteria were formulated with Cantonese operatic singing artists. The formulation process followed a user-centered approach, and a range of computer-aided assessment rubrics toolkits was developed to serve as the kernel of the assessment framework. These measures fully answered the second research question.

6.2.3 Characteristics and derivational problems of traditional assessments in Cantonese operatic singing

The findings from the thematic analysis of the series of semi-structured interviews in Phase II revealed the traditional assessments of Cantonese operatic singing featured 1) opaqueness, 2) oral rather than written assessment 3) incomplete and nonnormative assessment, and 4) general and implicit assessment. These features brought with them considerable disadvantages, including a lack of objectivity and fairness and a lack of instructiveness.

The above characteristics discredited the authority and authenticity of assessment in this field. Some professional students stated that they thought little of the scores derived from the traditional assessments. One of the teacher interviewees asserted that the unconvincing assessment of Cantonese operatic singing might be one of the reasons for its ailing transmission and development. During the interviews, a salient phenomenon emerged: both professional students and teachers mentioned that scores would usually be factitiously adjusted to ensure a certain pass rate. Otherwise, the students who failed would be required to re-take exams. If they failed again, students would be given a last opportunity right before graduation. If they failed once again, students would not qualify for graduation. Initially, this was a quality assurance mechanism of education programs. As assessment is a key approach to quality assurance, the factitious scoring jeopardized this vital mechanism. "Educational accountability underscores the need for reliable assessment and evaluation to support innovations in curriculum design, instructional methods, program funding, and the appraisal of student achievement" (Gruber, 2008, p. 42). Thus, reliable assessment is indispensable to educational accountability, and helps Cantonese opera to exist and thrive in an increasing accountability-

driven educational environment and assessment-oriented world. This should further contribute to the inheritance, popularization, and development of this traditional art form.

6.2.4 Feedback after the pilot implementation

Feedback after the pilot implementation was collected during the semi-structured interviews of Phase II to respond to Research Question 4. The responses were positive, affirming that the proposed assessment framework facilitate the teaching and learning of Cantonese operatic singing across multiple aspects.

6.2.4.1 Positive post-design evaluation

Both professional and amateur students, as well as their parents, stated that the new assessment framework is more instructive than traditional methods, which achieves the overarching purpose of assessment. Gardner (2012) indicated that: "assessment in education must, first and foremost, serve the purpose of supporting learning" (p. 9). Most of the interviewees endorsed the efficacy of the proposed assessment framework, including its 1) objectivity and fairness, 2) instructiveness and efficiency, 3) detail and well-roundedness, 4) normativity and systematism, 5) user-friendliness, 6) clarity, 7) convincingness, and 8) focus on process, rather than on scores. These advantages are imperative to a newly proposed assessment framework, especially from an educational perspective. The theoretical framework of the revised Bloom's Taxonomy features cumulative hierarchical divisions that represent the levels of specificity and complexity of educational learning objectives in different domains. It contributes to the disassembling of the educational development process into smaller parts. Thus, the hierarchies of the educational development process were adopted as assessment

standards. In the progress bar for each criterion, a gradient color replaced scores to represent student academic achievement (see Figure 6.1). This weakens the excessive attention paid to scores and guides students to concentrate more on study progress. Besides, the preset algorithms in the rubrics enabled the system to automatically calculate final scores according to teacher judgments by using the grids within the rubrics. This contributed to restrained factitious scoring and provided fairer results. Some teachers believed the new assessment framework was worthy of being applied to every subject in Cantonese opera and even other artistic genres.

Figure 6.1. An example of the assessment results of cognitive domain

6.2.4.2 Computer-aided rubrics facilitate user-friendly and efficiency

The teacher interviewees also declared that the new assessment framework was easy to learn and better than the traditional assessments across multiple aspects, which was a confirmation of the user-friendly design rationale. This turned out to be different from the teacher concerns noted before the pilot implementation. The computer-aided assessment rubrics were considered effective and efficient once a participant learned how to use them. The teachers who were willing to participate in the pilot implementation were younger teachers

aged 25 to 40. By contrast, older teachers tended to reject or be reluctant to try (or even consider) an innovative approach. Weakness in computer operation was another reason older teachers responded negatively to computer-aided assessment. Thus, being more willing to try new methods and accept new auxiliary technology, younger teachers are more likely to be the main supporters of the innovative assessment framework proposed in this study.

6.2.4.3 Improved assessment feedback

Feedback is an essential component of assessment (Burke, 2010). Constructive feedback to students about their achievements or performances, together with an indication of how to pursue improvement, can be obtained from well-designed rubrics (Armstrong et al., 2008). Traditional assessments in Cantonese operatic singing are often given orally, following a traditional approach that was apt to be forgotten and harm its instructiveness. As time passed, teachers and students began not to take it seriously. Thus, the traditional assessments have gradually become formalized, and their effectiveness has been doubted by teachers, students, and parents. On the contrary, both professional and amateur students noted that the assessment feedback that resulted from the proposed framework, which is in a written form, clarified their academic achievement from a perspective of learning progress - every detailed criterion was documentable and easily reviewable. An appreciation of the written format of the assessments in the rubrics was shown by participants as instructive to their study, especially when combined with the conventional oral feedback of daily teaching. Although professional students, amateur students, and their parents prefer assessments to be in written form, teachers think little of it. The Children's Palace even abolished student assessment reports some years ago, asserting that

no one took it seriously because of unconvincing assessment feedback. The Children's Palace's original assessment had collapsed and was replaced by a three-level promotion mechanism. This mechanism has its a significance and was endorsed by some amateur student parents. But a salient problem cannot be ignored: all the parents in the study declared that they did not know the assessment criteria of auditions, and that the audition process was opaque as it is conducted behind closed doors. Current literature related to assessment advocates that it should be transparent, and that criteria ought to be explicit to teachers, students, parents, and other stakeholders from the outset (Brown and Glasner, 1999; Thorpe, 2000). Students benefit when they understand the criteria and receive regular, descriptive feedback on their progress toward their goals. But, in current common practice, instead of convincing assessment results and overt assessment processes, students and parents can only rely on the authority of teachers. After the pilot implementation, parents stated that the pre-established and pre-announced assessment rubrics enabled them to better engage in student learning and provide appropriate support. Even in the context of community training, where trainees are mainly amateur performers or fans, explicit and visible outcomes of training are expected. The amateur students and parents believed that the new assessment framework could not only supplement the existing three-level promotion mechanism at the Children's Palace but might also be applied to auditions. They felt that this might provide more explicit and transparent assessment feedback as well as fairer assessment results that contributed to decision making (see Figure. 6.2). Assessment only becomes valid if assessors match evidence of achievement against criteria (Gibbs and Rowntree, 1999).

Figure 6.2. Three-step Process of Educational Decision Making (Hanna and Dettmer, 2004)

6.2.4.4 Different applications in different educational modes

After the fading away of apprenticeships, the conservatory tradition and community training became the mainstream for the transmission and education of Cantonese operatic singing. Although the primary purpose of modern conservatories is to nurture professional musicians and performers (Leung, 2014), it is not the primary pursuit of community training. Since assessment should reveal how well students have achieved educational objectives (Anderson, 2001), should assessment be identical when assessing the artistic education of these two modes? Since they have different nurturing purposes and different educational objectives, assessments should not be identical. Thus, some amateur student parents noted different assessment demands than professional students. One of the parents stated that "unlike mathematics or English, there is seldom normative assessment in Cantonese operatic singing. The only judgment is her achievement in audition or outcomes in competition. But as an extra-curricular activity, what I value most is that my daughter can have class, communicate, and take part in the competitions with peers that share the same passion for Cantonese operatic singing. This kind of experience is more important than the scores or performance in Cantonese operatic singing per se." This is a dilemma that often emerges in the field of community training, such as at Children's Palace. Should amateur student achievements be documented in the interests of art itself or for the sake of how arts education impacts success in other areas? During

the interviews, one amateur student parents raised doubts. The initial purpose of her sending her daughter to Cantonese operatic singing classes at the Children's Palace was not to have her train to be a professional practitioner of the future, but to expand her interest in the arts, Cantonese culture, history, and create an opportunity for her to communicate with like-minded children. The three-level promotion mechanism in the Children's Palace might be able to demonstrate a student's progress in Cantonese operatic singing performance, but what about other expectations or original purposes?

Unlike those for professional students, amateur student music assessments must incorporate achievements far beyond music, such as "outcomes related to socialization, cultural values, character traits, and moral education that include respect, hard work, punctuality, and honesty" (Colwell, 2003, p. 17). A revelation comes from The National Coalition for Core Arts Standards (NCAS) that was released by the National Core Arts Standards (NCAS) in 2014. Based on philosophical foundations that lead to lifelong goals, the standards consist of principle common values and expectations for learning in arts education. These include the arts as communicative and creative personal realization of culture, history, and as a means to wellbeing and community engagement across the five arts disciplines (dance, media arts, music, drama, visual arts) (see Figure 6.3.).

Philosophical Foundation	Lifelong Goals
The Arts as Communication	
In today's multimedia society, the arts are the media, and therefore provide powerful and essential means of communication. The arts provide unique symbol systems and metaphors that convey and inform life experience (i.e., the arts are ways of knowing).	Artistically literate citizens use a variety of artistic media, symbols and metaphors to independently create and perform work that expresses and communicates their own ideas, and are able to respond by analyzing and interpreting the artistic communications of others.
The Arts as Creative Personal Realization	
Participation in each of the arts as creators, performers, and audience members enables individuals to discover and develop their own creative capacity, thereby providing a source of lifelong satisfaction.	Artistically literate citizens find at least one arts discipline in which they develop sufficient competence to continue active involvement in creating, performing, and responding to art as an adult.
The Arts as Culture, History, and Connectors	
Throughout history the arts have provided essential means for individuals and communities to express their ideas, experiences, feelings, and deepest beliefs. Each discipline shares common goals, but approaches them through distinct media and techniques. Understanding artwork provides insights into individuals' own and others' cultures and societies, while also providing opportunities to access, express, and integrate meaning across a variety of content areas.	Artistically literate citizens know and understand artwork from varied historical periods and cultures, and actively seek and appreciate diverse forms and genres of artwork of enduring quality/significance. They also seek to understand relationships among the arts, and cultivate habits of searching for and identifying patterns, relationships between the arts and other knowledge.
Arts as Means to Wellbeing	
Participation in the arts as creators, performers, and audience members (responders) enhances mental, physical, and emotional wellbeing.	Artistically literate citizens find joy, inspiration, peace, intellectual stimulation, meaning, and other life-enhancing qualities through participation in all of the arts.
The Arts as Community Engagement	
The arts provide means for individuals to collaborate and connect with others in an enjoyable inclusive environment as they create, prepare, and share artwork that bring communities together.	Artistically literate citizens seek artistic experience and support the arts in their local, state, national, and global communities.

Figure 6.3. Philosophical foundations and lifelong goals of NCAS (NCCAS, 2014)

It is a good example of how to forge a divide between arts and assessment-based learning.

Wilkerson (2019) noted that "NCAS is a resource for arts educators to improve pedagogy and

assessment in the arts; it is also a tool to advocate for the importance of arts in education" (p.

226). Referencing the philosophy of NCAS or other applicable philosophies, the concrete

criteria of the proposed assessment framework can be adjusted to meet the needs of amateur students that have multi-dimensional purposes for learning Cantonese operatic singing, aside from ambitions to become future professional artists. After all, from the perspective of the education and transmission of this traditional genre, Cantonese operatic singing requires not only excellent performers but also audiences who are knowledgeable and give appropriate responses. An ideal assessment framework is not only able to provide dependable and convincing assessment, but also ought to fit the diverse needs of diverse types of learners with different demands.

6.2.4.5 Detailed considerations in practice

Although no major problems with the proposed assessment were indicated by the interviewees, a minor flaw was spotted during the pilot implementation. Before the pilot implementation, a teacher from the Children's Palace expressed her concerns about student understanding of the questions included in the Online Survey of Student Perception on Learning. The amateur students at the Children's Palace are aged six to 13, whereas the professional students in the Cantonese Opera School are aged 13 to 20. With respect to the vocabulary and language levels of the younger students, the questions given to the amateur students were simplified and options were denoted by different facial expressions with corresponding scores (see Figure. 6.4).
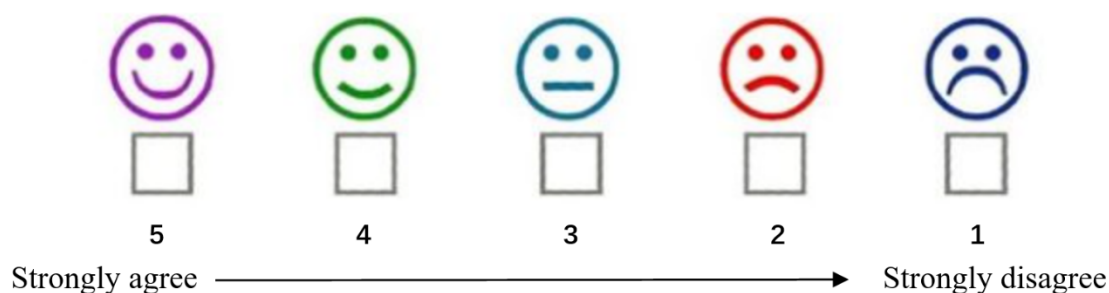
Figure 6.4. The revised instruction included in the Online Survey of Student Perception on Learning for amateur students

Nevertheless, after the pilot implementation, we found that these younger amateur students still had difficulties understanding the simplified questions, which were still too abstract. Younger students were asked to note how much they agreed with statements like, "I put a lot of effort into this course". After discussion with the teacher, it was decided that, for students under 13, "a lot of effort" might not be a concrete enough term. To make their parents or teachers happy, they might be inclined to choose favorable options. As a result, the questions could be set more specifically with concrete data. For example, "I practice an hour or so every day averagely".

6.2.4.6 Employing the proposed assessment framework to assessment systems

Although the proposed assessment framework helped overcome the disadvantages inherent in the traditional assessment of Cantonese operatic singing and yielded assessment results alongside various advantages, how it is used will necessarily affect its efficacy. One-off or disconnected application of the assessment framework can only generate limited effects. As a result, applying the proposed assessment framework into an assessment system to enhance its efficacy is highly recommended for obtaining a sustainable and efficient assessment.

An assessment system is defined as a collection of measurement approaches used to provide a variety of evidence to support educational decision-making (National Research Council, 2001). The most critical distinction between an assessment system and one-off assessments is that the former "allows for the possibility of monitoring trends in achievement and learning levels over time (more like a series of photos) and a better understanding of the relative contribution of various inputs and educational practices to changes in those trends". In contrast, the latter "only provides a snapshot of students achievement levels" (Clarke, 2012, p. 7). Therefore, increasing numbers of international organizations, governments, and other stakeholders are perceiving the significance of developing a robust system for student assessment (UNESCO, 2007; McKinsey and Company, 2007). An effective assessment system can provide sufficient information on quality and quantity, which can be used by stakeholders to make decisions that support the improvement of education quality and student learning (Ravela et al., 2009). With the application of an assessment system, disparate assessments can conceptually and operationally work together.

However, traditional assessments of Cantonese operatic singing are focused on performance. To be specific, both professional and amateur students were provided with concrete feedback in class that was based on the student performance and overall scores at mid-term and at the end of the term, respectively. The overall scores for the students were impression-based rather than evidence-based. The simple utilization of one assessment mode can be argued to restrain the credibility of the assessment. With the application of the proposed assessment framework, a balanced assessment system can be a solution to further optimizing the traditional assessment of Cantonese operatic singing. This might contribute to providing a

more accurate and equitable measure of student learning (see Figure 6.5).
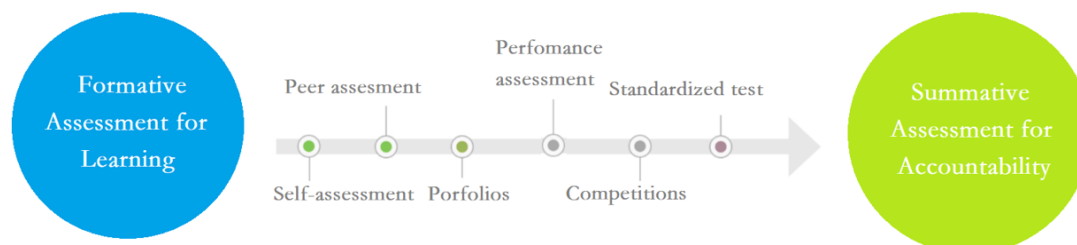
Figure 6.5. Balanced assessment system and its approaches (Andrade, n.d., p. 7)

Assessments can be categorized into formative assessments and summative assessments according to their instructional purposes (Burke, 2010). Regarded as assessment *for* learning, formative assessment aims to provide ongoing feedback to improve learning. By contrast, regarded as assessment *of* learning, the purpose of summative assessment lies in evaluating a final effort to prove learning. The former is usually implemented during a learning segment and uses informal teacher questions, conversations with the student, and so on. The latter is always conducted at the end of a learning segment and employs formal oral interviews, conferences with students, and the like (see Figure 6.6).

Formative Assessment Process	Summative Assessment Process
Assessment *for* Learning	Assessment *of* Learning
Purpose: Provide ongoing feedback to *improve* learning	Purpose: Evaluate final efforts to *prove* learning
Timing: During the learning segment	Timing: At the end of the learning segment
Informal teacher questions	Formal oral interview
Conversation with student	Conference with student
Informal observation	Formal observation
Rough drafts of written work	Final copy of written work
Learning log (in progress)	Final learning log entries
Reflective journal (multiple drafts)	Final journal entries
Mathematics problem solving steps	Mathematics final solution
Practice science experiment	Final science experiment
Rehearsal of presentation	Final presentation
Working portfolio	Showcase portfolio
Practice checklist for do-overs	Final checklist
Practice rubrics (analytical)	Final rubrics (analytical and holistic)
Homework, quizzes	Teacher-made tests
Benchmark/interim tests	High-stakes standardized tests

Figure 6.6. The Balanced Assessment Model (Burke, 2010, p. 25)

The integration of both formative assessment (assessment *for* learning) and summative assessment (assessment *of* learning) is essential and a balanced assessment system helps to integrate the two kinds of assessments seamlessly (Burke, 2010). A variety of approaches can be employed, such as self-assessment, peer assessment, portfolios, performance assessments, standardized tests, and so forth (Andrade and Cizek, 2010). These approaches can form an assessment system that contributes to forming a sustainable assessment. The proposed assessment framework integrates these approaches and constitutes a form of multidimensional assessment. Huebner (2009) stated that "a well-planned approach to balanced assessment will offer teachers, principals, and superintendents with different kinds of data they need to be well-

informed decision-makers" (p. 87). Consequently, further exploration and development of applications of the proposed assessment framework within an assessment system can yield a synergistic effect.

In sum, drawing on the feedback of teachers, students and parents after the pilot implementation, the proposed assessment framework greatly facilitates the teaching and learning of Cantonese operatic singing. Further exploration of its application in practice under diverse contexts is necessary, within which details can be polished and further application exploration can be conducted, so that assessment can play a prominent part in the education and transmission of Cantonese operatic singing.

6.3 Significance of the study

As a traditional artistic genre, the education and transmission of Cantonese operatic singing evolved from being centered on apprenticeships into a modern conservatory tradition. however, according to the content analysis of two typical assessments in *Xiqu* and semi-structured interviews with experts, teachers, students, and parents, considerable room for the improvement of the assessment of Cantonese operatic singing was noted. This was pursuant to increasing the effectiveness, equity, and quality of education and transmission of this traditional artistic genre. The current study can contribute to multiple aspects, including:

A practical application of Bloom's Taxonomy in music education and the Chinese traditional genre. Despite Bloom's Taxonomy having come into being decades ago, its application to Chinese music education, especially within the context of a traditional artistic

genre, has been rare. Introducing Bloom's Taxonomy into Cantonese operatic singing is therefore a bold attempt and innovative educational guidance within this field.

An improvement of the assessment in Cantonese operatic singing. Although the conservatory tradition replaced apprenticeships as the main mode of education and transmission within this field, assessment remains underappreciated and unsystematic. Arbitrariness and subjectivity are commonly witnessed. A criterion- and standard-based assessment framework should facilitate changes to conventions and improve the assessment of Cantonese operatic singing. The full potential of assessment will not be realized until the frameworks and rubrics are fully integrated coherently.

Enhancing synergies for better education in Cantonese operatic singing. Results of assessments are critical indicators of how well school systems are performing and are essential to providing feedback that helps students to do better. When compared to master-centered traditional Chinese educational ideology and ethics, the current study moved learners and the learning process to the center. This helps engage learners and empower them to progress.

The generalization and transmission of Cantonese operatic singing. To generate and transmit a traditional artistic genre, cultivating professional and amateur practitioners is critical, as are knowledgeable audiences. The current study aimed to improve the education of Cantonese operatic singing by formulating a criterion- and standard-based assessment framework that might facilitate the formation of a sustainable and multidimensional assessment system and thereby form a virtuous circle of education and transmission.

An example of assessment formulation for other Chinese music traditions or other disciplines in music education at home and abroad. The modality of the criterion- and standard-

based assessment framework proposed in the current study provides both formulation frames for comprehensive and reliable assessment and flexibility to fit in specific assessment disciplines.

6.4 Limitations of the study

While several useful findings have been obtained, there are two major limitations of this study:

The sample size of the interviews was limited. The interview data of Phase I were collected from six experts in Hong Kong and Guangzhou. The data from Phase II were collected from four teachers working in the conservatory mode and community training, as well as 18 of their students and some of the parents. Although purposive sampling was adopted to select information-rich and representative interviewees, including teachers from different transmission modes and students of various levels, the sample size was still small. Since this is a qualitative study, the findings should not be generalized. A larger-scale study is expected in the future.

To ensure credibility, the triangulation of sources was adopted during qualitative analysis. After utilizing content analysis to examine the existing graded examinations of *Xiqu*, the opinions of experts and instructors in Cantonese operatic singing from different cities were collected and analyzed. Robust, rich, and comprehensive information was the aim. In addition, member-checking was employed in the thematic analysis. Transcriptions were sent to interviewees for confirmation. The interpretations and conclusions were shared with them to clarify their authentic intentions. But as qualitative research, the three major indicators of

assessment quality (validity, reliability, and fairness), cannot be verified in the current study. Adopting a user-centered approach for the formulation of a new assessment framework, the current study focuses on the three major phases of this approach, including the content analysis of two graded examinations in *Xiqu* and semi-structured interviews with experts and teachers of Cantonese operatic singing as a pre-design analysis. The pilot implementation was centered on 1) identifying potential deficiencies, 2) testing whether the introduction of the study, the instruction of the procedure, and the description of the assessment rubrics could be clearly understood by the participants, 3) evaluating the assessment framework application process, and 4) testing the assessment instruments. Since the main purpose of the current study is to formulate a new assessment framework for Cantonese operatic singing, the validation of its quality might be the focus of a future quantitative study.

6.5 Implications for future study

This study has developed an assessment framework for Cantonese operatic singing, which fills a gap in the related literature. Based on the statement of limitations of this research and previous discussions, future research might concentrate on the following three issues: 1) in-depth educational research into Cantonese operatic singing assessments as breakthrough points, 2) research pertaining to administration or policy strategies surrounding the education and transmission of Cantonese operatic singing based on the data derived from assessment, 3) the transmission and development of the traditional genre through educational influence within contemporary society, and 4) further research in music education assessment.

Treating the proposed assessment framework as a breakthrough point in the exploration of the assessment of Cantonese operatic singing might stimulate more research into related educational issues. Up to time of writing, history and musical ontology have been the mainstreams of research into Cantonese operatic singing. Research pertaining to its education has been limited (Leung, 2021) and research into educational assessment has been nonexistent. The current study has filled this gap and opened a door to further research, thereby improving and developing the education of Cantonese operatic singing. As one of the three important components of education, assessment significantly influences instruction and curriculum (Lebler, 2015). When these components are well aligned and calibrated, student learning can be maximized (Rhode Island Department of Education, n.d.). Therefore, the assessment framework proposed in this study is a foundation for a string of future studies. These might include its application within an assessment system and thereby extend to other aspects of education, such as the synergies found amongst assessment, instruction, and curriculum.

Secondly, more studies about educational administration and policy might be inspired by the generation of assessment data. Unlike mathematics or languages, a lack of precision in the existing assessment of music education stems from its informality and subjectivity. Accurate measurements to evaluate quality are lacking (Hanna, 2007). Thus, although there is a consensus surrounding the intrinsic value of music within the curriculum, the academic legitimacy of music programs will inevitably be questioned if standardized means of assessment are lacking (Asmus, 1999). This affects administrative decision making in a policy environment that values accountability. Reliable assessment data might facilitate discussions or proposals of more robust strategies in administration and policy.

How assessment impacts education and thereby affects the transmission and development of traditional forms within contemporary society is a rewarding research orientation. The ultimate purpose of assessment is not only to provide invaluable information to students, teachers, schools, and parents but also to ensure that instruction is as effective as possible in boosting student learning in music (Asmus, 1999). A high-quality conservatory mode education facilitates the cultivation of excellent artists; community training fosters future professional students and knowledgeable audiences. Both are vital to the transmission and development of this traditional genre within contemporary society. This makes them worthy of exploration.

6.6 Recommendations for further research

Based on the findings and limitations of this study, further research can be suggested. A few are proposed below.

Validity and reliability need to be ensured in future quantitative research. These are the two most essential and critical characteristics in the evaluation of any measurement tools or instruments of convincing research (Mohajan, 2017) and contribute to enhancing the accuracy of assessment (Tavakol and Dennick, 2011). An ideal assessment is expected to have "high reliability and generalizability coefficients, low standard errors, and small confidence intervals" (Rhode Island Department of Education, n.d., p.17). Although frameworks can improve the validity and reliability of assessment (Pearce, 2018), the validity and reliability of the assessment must be verified and reported when measuring study outcomes with the application of an assessment framework. Content validity and criterion validity should be valued in any

future study. "An assessment that is highly reliable is not necessarily valid. However, for an assessment to be valid, it must also be reliable" (Rhode Island Department of Education, n.d., p. 17). A reliable assessment should generate similar results for the same student, even it is conducted more than once. Thus, test-retest reliability should be the focus of future studies, within which external consistency can be verified. Inter-rater reliability should also be given attention. It is the extent to which equivalence of ratings can be acquired by an assessment method when different assessors use it.

More methods can be adopted in future studies to enrich the data collection. For example, observations, focus groups, and questionnaires and surveys are excellent data collection methods that I plan to employ in the subsequent quantitative research.

Broader applications of the proposed assessment framework can be explored. No single mode of assessment can provide a well-rounded and representative measurement of student learning in the arts (Gruber, 2008). Gruber (2008) noted that "testing, observation, finished product, and portfolios, when used individually, do not provide a completely accurate representation of learning in art. However, like the four legs of a table, the four assessment strategies form balanced support for a comprehensive assessment plan" (p. 43). An assessment system that features the application of the proposed assessment framework can be established. Utilizing various modes of assessment, such an assessment system could provide convincing assessment results, so that sustainability, multidimensionality, and effectiveness of assessment might be guaranteed. This would boost the development of the education and transmission of Cantonese operatic singing. Besides, since the proposed criterion- and standard-based

assessment framework is compatible in multiple circumstances, further application in different settings, contexts, and cultures can be implemented.

Strategies of administration or policy may guarantee the effective application of the assessment framework and its assessment results. The current study centers on the formulation of an assessment framework for Cantonese operatic singing. Still, no matter how good the assessment framework might be, validity and reliability cannot be ensured if it is used arbitrarily. Therefore, matched administration and policy are imperative. On the other hand, results from the assessments are critical to establishing how well school systems are performing and providing feedback that enhances the quality of education and education programs. Thus, how to use results to fit original purposes marks the efficiency and success of assessment.

There is a possibility of extending the application of the proposed assessment framework to other fields like acting and acrobatic fighting, which constitute the four major skills of Cantonese opera together with singing and reciting. This might include other areas in the artistic genre, such as singing, instrument performance, music drama, or even dancing, which consists of musical and kinesthetic elements. Furthermore, the proposed assessment framework can be popularized to assess different levels in music education, such as, primary and secondary levels as well as the community sector. After sufficient applications and further refinement, this assessment framework can be generalized to different educational settings, contexts, and even cultures.

There is also the promise of alignment with other educational components to advance the synergies of the education of Cantonese opera. The kernel of Bloom's Taxonomy is that the components of education (assessment, instructions, and curriculum) should serve and advance

educational goals and learning objectives. This involves not only designing fit-for-purpose assessments but also matching instructions and curriculum. With all three components integrated, the synergies of education in Cantonese opera might be remarkable.

6.7 Conclusions

As the traditional mode of transmission and education of Cantonese operatic singing, apprenticeship featured in a quasi-parental relationship, informal learning, and oral approach. Its assessment is inclined to be master-centered and subjective. But it had taken the place of the apprenticeship by the conservatory tradition and the community training in the mid-twentieth century. Consequently, significant changes have taken place in the transmission and education of this traditional genre. Such as teacher-student relationship, institutionalized approaches, formal/non-formal learning. But the traditional master-centered conception in apprenticeship still influences the design and practice of assessment in the conservatory tradition and the community training, which might not meet the need of contemporary education. For better education and transmission of Cantonese operatic singing, the current study strives to formulate a criterion- and standard-based assessment framework concerning a revised theoretical framework derived from Bloom's Taxonomy to ensure quality and promote sustainability and multidimensionality.

In response to Research Question 1, the design of the existing graded examinations in *Xiqu* is still developing. No ground can be detected in the grade division of these two examinations. From the perspective of the construct, COSGEHK published a unified and normative design of test items, whereas no regularity could be perceived in the test items of POGE. COSGEHK

proposed a series of rubrics based on the established graded examination framework of UWL, whereas POGE failed to publish any complete rubrics. The design of the criteria, standards, weightings, and grading and evaluation of COSGEHK is more normative than those of POGE. Common practice in the existing graded examinations of Chinese opera features 1) dependence on subjective perception, rather than criterion- and standard-based assessment, 2) a focus on the assessment of performance skills and a lack of multidimensionality, and 3) an inclination to momentary judgment and a lack of sustainability.

Compared to POGE, the relatively successful design of COSGEHK rests on the cooperation between Cantonese operatic artists and the educational experts at UWL, as well as on UWL's well established standardized assessment. This shows the value of educational expertise and frameworks in formulating assessments. Thus, for the sake of transmission and education, educational ideology ought to be valued alongside the characteristics of Cantonese operatic singing when formulating Cantonese operatic singing assessments. To restrain the subjectivity of assessment, a criterion- and standard-based assessment was introduced into the framework formulation of the current study. To ensure the multidimensionality of assessment, the current study employed a revised theoretical framework that was derived from Bloom's Taxonomy, which included four domains: cognitive, psychomotor, affective, and behavioral. These contribute to guiding comprehensive assessment. To facilitate sustainability of assessment, a balanced assessment was suggested that was based on the proposed framework. Since experts, teachers, and even some of the students believed that *Xiqu* was hard to be assess but assessing student learning was less controversial, the current study focused on the assessment of student learning in art education, rather than on artistic achievement. It was

recognized that a user-friendly design would affect the popularity of a newly proposed assessment framework. Consequently, the proposed assessment framework is computer-aided, within which algorithms are preset in alongside assessment rubrics. Calculations are triggered automatically by the assessor judgments of student performance for each criterion. The above answer the second research question, according to which the assessment framework was formulated.

Using thematic analysis, the characteristics of traditional assessments made during daily teaching and learning in Cantonese operatic singing are marked by 1) opaqueness, 2) factitious scoring, 3) oral assessment, 4) incompleteness and non-normativity 5) generalization and implicitness, 6) a lack of objectivity and fairness, and 7) a lack of instructiveness. These fully respond to the third research question and constitute problems that should be mitigated when formulating a new assessment framework.

After the pilot implementation, it was concluded that the new framework contributed to better assessments, which were 1) instructive and efficient 2) easy to start and user-friendly, 3) detailed and well-rounded, 4) normative and systematic, 5) clear, 6) documentable, 7) objective and fair, 8) transparent, 8) focused on process as well as scores, and 9) convincing. These respond to the fourth research question.

Although scarce research has focused on the assessment of Cantonese operatic singing, this study strives to draw new attention to the subject and provide a basis for in-depth exploration from a new angle. The findings and outcomes of this study improve the antiquated ideology and unsystematic and non-normative practice of existing assessment in Cantonese operatic singing. It is hoped this will contribute to improving the education and transmission

of this traditional artistic genre. To further improve the proposed assessment framework, a large-scale study could be conducted to ensure its reliability and validity. Future studies could extend its application to the other major skills of Cantonese opera, such as acting movement and acrobatic fighting, or to more diversified modes, such as balanced assessment. More research and innovation are needed to delve into the curriculum and instruction of Cantonese operatic singing.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. In W. S. Brainbridge (Ed.), *Berkshire Encyclopedia of human-computer interaction,* (Vol. 2, pp. 763-768). Berkshire. https://doi.org/10.7551/mitpress/6918.003.0015

ABRSM. (n.d.). *Graded music exam marking criteria.* The Associated Board of the Royal Schools of Music. Retrieved from https://gb.abrsm.org/en/our-exams/information-and-regulations/graded-music-exam-marking-criteria/

Ainley, J., & Schulz, W. (2020). Framework development in international large-scale assessment studies. In H. Wagemaker (Ed.), *Reliability and validity of international large-scale assessment* (pp. 23–36). Springer. https://doi.org/10.1007/978-3-030-53081-5_3

Airasian, P. W. (1994). The impact of the taxonomy on testing and evaluation: Bloom's Taxonomy: A forty-year retrospective. In L. W. Anderson, & L. A. Sosniak (Eds.), *Bloom's taxonomy: A forty-year retrospective* (pp. 82-102). University of Chicago Press.

Airasian, W., & Miranda, H. (2002). The role of assessment in the revised taxonomy. *Theory into Practice, 41*(4), 249–254. https://doi.org/10.1207/s15430421tip4104_8

Amer, A. (2006). Reflections on Bloom's revised taxonomy. *Electronic Journal of Research in Educational Psychology, 4*(1), 213–230. https://www.redalyc.org/articulo.oa?id=293123488010

Anderson, L. W., & Krathwohl, D. R. (Ed.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Anderson, L. W., & Sosniak, L. A. (1994). *Bloom's taxonomy: A forty-year retrospective. Ninety-third yearbook of the National Society for the Study of Education, Part II.* University of Chicago Press.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational leadership*, *57*(5), 13-19. https://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=3270122&site=ehost-live&scope=site

Andrade, H., & Cizek, G. J. (Eds.). (2010). *Handbook of formative assessment*. Routledge. https://doi.org/10.4324/9780203874851

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. Jossey-Bass.

Armstrong, S., Chan, S., Malfroy, J., Thomson, R., & Hendry, G. (2008). *Assessment guide: Implementing criteria and standards-based assessment.* The University of Western Sydney. https://www.westernsydney.edu.au/__data/assets/pdf_file/0004/449860/PVC5557_Assessment_Guide_LR3.pdf

Arnab, R. (2017). *Survey sampling theory and applications*. Elsevier Science & Technology.

Asmus, E. P. (1999). Music assessment concepts: A discussion of assessment concepts

and models for student assessment introduces this special focus issue. *Music*

Educators Journal, 86(2), 19–24. https://doi.org/10.2307/3399585

Astin, W. A., & Antonio, A. L. (2012). *Assessment for excellence: The philosophy and*

practice of assessment and evaluation in higher education. Rowman & Littlefield

Publishers.

Ball, S. J. (1990). *Politics and policy making in education*. Routledge.

Bansler, J. P., & Havn, E. (2010). Pilot implementation of health information systems: Issues

and challenges. *International Journal of Medical Informatics*, *79*(9), 637-648.

https://doi.org/10.1016/j.ijmedinf.2010.05.004

Barber, M., & Mourshed, M. (2007). *How the world's best-performing school systems come*

out on top. McKinsey & Co. https://www.mckinsey.com/industries/education/our-

insights/how-the-worlds-best-performing-school-systems-come-out-on-top

Bhat, B. A., & Bhat, G. J. (2019). Formative and summative evaluation techniques for

improvement of the learning process. *European Journal of Business & Social*

Sciences, *7*(5), 776-785.

https://www.researchgate.net/publication/333633265

Biggs, J., & Tang, C. (2007). *Teaching for quality learning at university*. McGraw-Hill

Education.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American*

Psychologist, *61*(1), 27. https://doi.org/10.1037/0003-066X.61.1.27

Bloom, B.S. (1956). *Taxonomy of educational objectives: The classification of educational goals.* McKay.

Boyatzis, R.E. (1998). *Transforming qualitative information: Thematic analysis and code development.* Sage Publications.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*, 77–101. https://doi.org/10.1191/1478088706qp063oa

Brophy, T. S., & Albert, K. (Eds.). (2008). *Assessment in music education: Integrating curriculum, theory, and practice*. GIA Publications.

Brown, S. (2005). Assessment for learning. *Learning and Teaching in Higher Education,* (1), 81-89.

https://www.qub.ac.uk/directorates/AcademicStudentAffairs/CentreforEducationalDevelopment/FilestoreDONOTDELETE/Filetoupload,120807,en.pdf

Brown, S., & Glasner, A. (1999). *Assessment matters in higher education: Choosing and using diverse approaches*. McGraw-Hill Education.

Burke, K. (2010). *Balanced assessment: From formative to summative*. Solution Tree Press.

Cai, X.X. (2010). *Xiyi Renchuan-Taiwan Dui Chuantong Xiqu Yishi De Jiyi Chuangcheng Yue Jilu Baocun.* [Opera is inherited by people-The technique inheritance and the record conservation of the traditional opera master in Taiwan]. *Wenhua Yichan, [Cultural Heritage] 2010*(1), 32–50.

Cate, T. J. T., & De Haes, J. C. J. M. (2000). Summative assessment of medical students in

the affective domain. *Medical Teacher*, *22*(1), 40-43.

https://doi.org/10.1080/01421590078805

Chan, F. N., Wu, R. Z., & Chan, Z. L. (2007). 粵劇六十年 *[60 years within Cantonese*

opera]. The Chinese University Press.

Chan, S.Y. (1997). 實地考查與戲曲研究. [Ethnography and Xiqu research].香港中文大學

粵劇研究計劃 [Cantonese Opera Research Programme, CUHK].

Chan, S. Y. (1999). 香港粵劇導論 [Introduction and discussion of Cantonese opera].The

Chinese University Press.

Clarke, M. (2012). *What matters most for student assessment systems: A framework paper.*

Open Knowledge Repository. Retrieved from

https://openknowledge.worldbank.org/handle/10986/17471

Clarke, V., & Braun, V. (2014). Thematic analysis. In T. Teo (Ed.), *Encyclopedia of*

critical psychology (pp. 1947-1952). Springer.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Routledge.

Colwell, R. (2003). The status of arts assessment: Examples from music. *Arts Education*

Policy Review, *105*(2), 11-18. https://doi.org/10.1080/10632910309603457

Crooks, T. (2001). The validity of formative assessments. In P. A. Howard-Jones., L. Franey.,

R. Mashmoushi., & Y. C. Liao (Eds.), *Proceedings of British educational research*

association annual conference (pp. 13 - 15). University of Leeds.

Dalto, J. (2014). *Teaching skills: The psychomotor domain of learning and learning*

objectives. Vector Solutions. Retrieved from

https://www.vectorsolutions.com/resources/blogs/teaching-skills-the-psychomotor-domain-of-learning-and-learning-objectives/

Davies, R. S. (2020). *Designing surveys for evaluations and research.* EdTech Books. https://edtechbooks.org/designing_surveys

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education, 42*(3), 347–360. https://doi.org/10.1080/02602938.2015.1111294

DeLuca, C., & Bolden, B. (2014). Music performance assessment: Exploring three approaches for quality rubric construction. *Music Educators Journal*, *101*(1), 70-76. https://doi.org/10.1177/0027432114540336

Denscombe, M. (2014). *The good research guide* (4th ed.). Open University Press.

Derrell, T. (2015). *5 common types of educational assessment*. American Intercontinental University. Retrieved from https://www.aiuniv.edu/Blog/2015/July/Educational-Assessment?cv=1

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., & Messer, B. L. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR), and the Internet. *Social Science*, *38*(1), 1-18. https://doi.org/10.1016/j.ssresearch.2008.03.007

Drisko, J. W., & Maschi, T. (2016). *Content analysis*. Oxford University Press.

Duke, R. A. (1999). Measures of instructional effectiveness in music research. *Bulletin*

of the *Council for Research in Music Education, 143,* 1–48.

http://www.jstor.org/stable/40319011

Evans, M. (2002). *Peter Merholz and Nathan Shedroff on user-centered design.* Retrieved

from https://www.digital-web.com/articles/peter_merholz_and_nathan_shedroff/

Gardner, J. (2012). *Assessment and learning* (2nd ed.). SAGE Publications.

Gareis, C. R., & Grant, L. W. (2015). *Teacher-made assessments: How to connect*

curriculum, instruction, and student learning (2nd ed.). Routledge.

https://doi.org/10.4324/9781315773414

Gareth, T., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. In C. Willig, &

W. Stainton-Rogers (Eds.), *The Sage handbook of qualitative research in psychology* (2nd

ed., pp. 17-36). SAGE Publications. https://doi.org/10.4135/9781526405555.n2

Geisert, P. (1972). *The dimensions of measurement of the affective domain* (ED069663).

ERIC. https://files.eric.ed.gov/fulltext/ED069663.pdf

Gibbs, G., & Rowntree, D. (1998). *Teaching in higher education: Theory and evidence.* The

Open University.

Gläser, J., & Laudel, G. (2013). Life with and without coding: Two methods for early-

stage data analysis in qualitative research aiming at causal explanations. *Forum,*

Qualitative Social Research, 14(2). https://doi.org/10.17169/fqs-14.2.1886

Great Schools Partnership (n.d.). *The glossary of education reform.* Retrieved from

https://www.edglossary.org/assessment/

Hauenstein, A. D. (1998). *A conceptual framework for educational objectives: A holistic approach to traditional taxonomies*. University Press of America.

Hendry, G. D., Armstrong, S., & Bromberger, N. (2012). Implementing standards-based assessment effectively: Incorporating discussion of exemplars into classroom teaching. *Assessment & Evaluation in Higher Education, 37*(2), 149-161. https://doi.org/10.1080/02602938.2010.515014

Hochschild, J. L. (2009). Conducting intensive interviews and elite interviews. In M. Lamont & P. White (Eds.), *Workshop on interdisciplinary standards for systematic qualitative research* (pp. 124-127). National Science Foundation.

Hong Kong Association of Cantonese Opera Scholars. (n.d.). *粵曲考級試（演唱）考試大綱*. [Syllabus of Cantonese Opera Graded Examination (Singing)]. Retrieved from http://www.hkacos.com/doc/02.pdf

Hong Kong Association of Cantonese Opera Scholars. (n.d.). *《粵曲考級試》（演唱）內容設計和考試範疇的常見問題* [Frequently Asked Questions in the Content Design and Examination of the Cantonese Opera Test (Singing)]. Retrieved from http://www.hkacos.com/doc/01B.pdf

Huang, Y. (2019). *浅论粤港两地"粤剧进校园"* [Brief Discussion of 'Integration of Cantonese Opera and Campus' in Guangdong Province and Hong Kong]. *Home Drama.* (2), 124-125

Hughes, C. (2007). *Practical guidelines for writing assessment criteria & standards.*

Retrieved from

http://www.tedi.uq.edu.au/downloads/assessment/WritingCriteriaandStandards.pdf

Iaquinta, R., & Impedovo, M. A. (2013). Expanding the boundaries of learning: The role of

vocational orientation. In P. M. Pumilia-Gnarini (Ed.), *Handbook of research on didactic*

strategies and technologies for education: Incorporating advancements (pp. 289-297).

IGI Global.

Jago, C. (2009). *A history of NAEP assessment frameworks* (ED ED509382). ERIC.

https://eric.ed.gov/?id=ED509382

Joffe, H., & Yardley, L. (2004). Content and thematic analysis. In D. F. Marks & L. Yardley

(Eds.), *Research methods for clinical and health psychology* (pp. 56-68). SAGE Publications.

Johnson, R. R. (1998). *User-centered technology, a rhetorical theory for computers and*

other mundane artifacts. State University of New York Press.

Kahraman, Z. (2010). Using user-centered design approach in course design. *Procedia,*

Social and Behavioral Sciences, 2(2), 2071–2076.

https://doi.org/10.1016/j.sbspro.2010.03.283

Keshavarz-Ghorabaee, M., Amiri, M., Zavadskas, E. K., Turskis, Z., & Antucheviciene, J.

(2021). Determination of objective weights using a new method based on the removal

effects of criteria (MEREC). *Symmetry (Basel), 13*(4), 525.

https://doi.org/10.3390/sym13040525

Kizlik, B. (2012). *Measurement, assessment, and evaluation in education.* Retrieved from

https://online.stu.edu/articles/education/educational-measurement-assessment-

evaluation.aspx.

Krahenbuhl, K. S. (2016). Student-centered education and constructivism: Challenges,

concerns, and clarity for teachers. *The Clearing House: A Journal of Educational

Strategies, Issues and Ideas*, *89*(3), 97-105.

https://doi.org/10.1080/00098655.2016.1191311

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into

Practice*, *41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Krathwohl, D. R., & Anderson, L. W. (2009). *A taxonomy for learning, teaching, and

assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Kreitzer, A., & Madaus, G. (1994). Empirical Investigations of the hierarchical structure of

the taxonomy. In L. Anderson & L. Sosniak (Eds.), *Bloom's Taxonomy: A forty-year

retrospective* (pp. 64–81). The National Society for the Study of Education..

Krippendorff, K. (2013). *Content analysis: an introduction to its methodology* (3rd ed.).

SAGE Publications.

Lai, B.J. (2001). *廣東戲曲簡史*. [*A brief history of Chinese opera in Guangdong*].

Guangdong People's Publishing House.

Lai, B. J. & Huang, J. M. (1988). *粵劇史* [Cantonese opera history]. China Theater Press.

Lai, K. (2010). *香港粵劇敘論* [Description and discussion of Hong Kong Cantonese

opera]. Joint Pub. Co.

Lam, T. C. M. (1995). *Fairness in performance assessment* (ED391982). ERIC.

   http://ericae.net/db/edo/ED391982.htm

Lambert, N. M., & McCombs, B. L. (1998). *How students learn: Reforming schools through*

   learner-centered education. American Psychological Association.

Lau, C.C. & Xin, Y. X. (1995). *粵劇研討會論文集* [Papers and Proceedings of the

   International Seminar on Cantonese Opera Joint Publishing (H.K) & Centre of Asian

   Study, University of Hong Kong.

Lebler, D., Carey, G., & Harrison, S. D. (2015). *Assessment in music education: From policy*

   to practice. Springer International Publishing.

Lehman, P. (2008). A vision for the future: Looking at the standards. *Music Educators*

   Journal, *94*(4), 28-32. https://doi.org/10.1177/00274321080940040103

Leung, B. W. (2014). Teachers' transformation as learning: Teaching Cantonese opera in

   Hong Kong schools with a teacher–artist partnership. *International Journal of Music*

   Education, *32*(1), 119-131. https://doi.org/10.1177/0255761413491174

Leung, B. W. (2015). Transmission of Cantonese opera in the conservatory tradition:

   two case studies in South China and Hong Kong. *Music Education Research, 17*(4), 480–

   498. https://doi.org/10.1080/14613808.2014.986081

Leung, B. W. (2018). A proposed model of transmission of Cantonese opera in Hong

   Kong higher education: From oral tradition to conservatoire. *Arts and Humanities in*

   Higher Education, 19(2), 144–166. https://doi.org/10.1177/1474022218791465

Leung, B. W. (2021). Informal learning of Cantonese operatic singing in Hong Kong: An autoethnographic study. *Pedagogy, Culture & Society*, 1-17. https://doi.org/10.1080/14681366.2021.1934090

Leung, B. W. (Ed.) (2018). *Traditional music in the modern world: Transmission, evolution, and challenges* (Vol.24*)*. Springer.

Leung, P. K. (1982). 粵劇研究通論 [General studies of Cantonese opera]. Cantonese Opera Research Project Press.

Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher, 32*(7), 3–13. https://doi.org/10.3102/0013189X032007003

Luo, D. (2018). *香港八和会馆八和粤剧学院粤剧教育现状研究* [A study on the status quo of Cantonese Opera education in Bahe Cantonese Opera Academy of Chinese Artists Association of Hong Kong]. (Unpublished master's thesis, Guangzhou University, China). Retrieved from https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201901&filename=1018242346.nh&uniplatform=NZKPT&v=bzb2g0aDTbWpkuf8l85Yibstzf_pE4esPOAZfbuUS64IxBIegcWcSY7mhw_Zszks

Lynch, D., Laws, K. R., & McKenna, P. J. (2010). Cognitive behavioral therapy for major psychiatric disorder: does it really work? A meta-analytical review of well-controlled trials. *Psychological Medicine*, *40*(1), 9-24. https://doi.org/10.1017/S003329170900590X

Lyu, Z. Z. (2018) 粵劇在高校的傳承現狀與思考——以廣州中心的考察 [The transmission status and reflection of Cantonese opera in tertiary schools: Guangzhou-

centered Investigation]. *廣東開放大學學報 [Journal of the Open University of Guangdong]*. (6), 7

Mantie, R. (2019). The philosophy of assessment in music education. In T. S. Brophy (Ed.), *The Oxford handbook of assessment policy and practice in music education* (1st ed., Vol 1, pp. 33-56). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190248093.013.2

Marichal, J. L., & Roubens, M. (2000). Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, *124*(3), 641-650. https://doi.org/10.1016/S0377-2217(99)00182-4

Mark, M. L., & Madura, P. (2013). *Contemporary music education*. Cengage Learning.

Mark, M., & Gary, C. (2007). *A history of American music education* (3rd ed.). Rowman & Littlefield Education.

Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives*. Corwin Press.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). SAGE Publications.

McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). *Instrument development in the affective domain* (3rd ed.). Springer New York. https://doi.org/10.1007/978-1-4614-7135-6

McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative

assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation, 18*(2), 1-15. https://doi.org/10.7275/tmwm-7792

Mertler, C. A. (2000). Designing scoring rubrics for your classroom. *Practical Assessment, Research, and Evaluation*, *7*(1), 25. https://doi.org/10.7275/gcy8-0w24

Mishook, J.J., & Kornhaber, M. L. (2006). Arts integration in an era of accountability. *Arts Education Policy Review*, *107*(4), 3–11. https://doi.org/10.3200/AEPR.107.4.3-11

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of "Spiru Haret" University*, *17*(4), 59-82. https://doi.org/10.26458/1746

Mok, M. M. C., Leung, B. W., Kuo, B. C., Liu, Z. Y., Lam, S. M., Ng, C. W., ... & Choi, T. W. (2017). Kinetic chain assessment and learning computerized system for Cantonese opera: The development, validation and application. In *2017 International Conference on Applied System Innovation (ICASI)* (pp. 762-765). IEEE. https://doi.org/10.1109/ICASI.2017.7988542.

Morgan, C., Dunn, L., O'Reilly, M., & Parry, S. (2004). *The student assessment handbook: New directions in traditional and online assessment*. Routledge Falmer

Mueller, J. (2006). *Authentic assessment toolbox.* North Central College. Retrieved from http://jonathan.mueller.faculty.noctrl.edu/toolbox/index.htm

National Assessment of Educational Progress. (n.d.) *Assessment frameworks*. National Assessment of Educational Progress. Retrieved from https://nces.ed.gov/nationsreportcard/assessments/frameworks.aspx

Ng, W. C. (2015). *The rise of Cantonese opera*. University of Illinois Press.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291-310. https://doi.org/10.1037//0033-295X.108.2.291

O'Neill, G., & Murphy, F. (2010). *Assessment: Guide to taxonomies of learning.* University College of Dublin Teaching and Learning Resources. Retrieved from https://www.ucd.ie/t4cms/ucdtla0034.pdf

Oakland, T. (1997). Affective assessment. *Psicologia Escolar e Educacional*, *1* (2-3), 11-21. https://doi.org/10.1590/S1413-85571997000100002

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50006-X

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology (1st ed.,* pp. 224-239). Guilford Publications.

Pearce, J., Edwards, D., Fraillon, J., Coates, H., Canny, B. J., & Wilkinson, D. (2015). The rationale for and use of assessment frameworks: Improving assessment and reporting quality in medical education. *Perspectives on Medical Education*, *4*(3), 110-118. https://doi.org/10.1007/s40037-015-0182-z

Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music

cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences*, *1423* (1), 378-395. https://doi.org/10.1111/nyas.13654

Pellegrino, J.W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment.* National Academy Press.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3–14. https://doi.org/10.3102/0013189X031007003

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human-computer interaction.* John Wiley & Sons.

Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed.). Routledge Falmer.

Reid, N. (2006). Thoughts on attitude measurement. *Research in Science & Technological Education*, *24*(1), 3-27. https://doi.org/10.1080/02635140500485332

Rhode Island Department of Education. (n.d.). *Comprehensive assessment system: Rhode Island criteria & Guidance.* Rhode Island Department of Education. Retrieved. Retrieved from https://www.ride.ri.gov

Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. SAGE Publications. https://doi.org/10.4135/9781446288009

Royal Roads University. (n.d.). *Writing effective assessment criteria*. Royal Roads University. Retrieved from http://ctet.royalroads.ca/writing-effective-assessment-criteria

Sadler, R. D. (2005). Interpretations of criteria-based assessment and grading in higher

education. *Assessment & Evaluation in Higher Education, 30*(2), 175-194.

https://doi.org/10.1080/0260293042000264262

San, K. S. J. (1993). *Hongchuan" Guizhi Ji Qi Jinji*. [The regulations and taboos of the "Red

Boats"]. In J. Li (Ed.), *Xianggang Yueju Koushushi* [Oral History of Hong Kong

Cantonese Opera] (pp. 12–20). Joint Pub.

Saxon, D. P., Levine-Brown, P., & Boylan, H. R. (2008). Affective assessment for

developmental students, part 1. *Research in Developmental Education*, *22*(1), 1-4.

https://www.academia.edu/1813234/Affective_assessment_for_developmental_students_

part_1

Schulte-Tenckhoff, I., & Nettl, B. (1989). The Western impact on world music: Change,

adaptation, and survival. *Cahiers de Musiques Traditionnelles, 2*, 274.

https://doi.org/10.2307/40240051

Scott, S. J. (2012). Rethinking the roles of assessment in music education. *Music Educators

Journal*, *98*(3), 31-35. https://doi.org/10.1177/0027432111434742

Stemler, S. (2001). *An introduction to content analysis* (ED458218). ERIC.

https://files.eric.ed.gov/fulltext/ED458218.pdf

Stephens, M., & Ormandy, P. (2019). An evidence-based approach to measuring affective

domain development. *Journal of Professional Nursing*, *35*(3), 216-223.

https://doi.org/10.1016/j.profnurs.2018.12.004

Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, *20*(3), 5-15. https://doi.org/10.1111/j.1745-3992.2001.tb00065.x

Stronge, J. H. (2018). *Qualities of effective teachers* (3rd ed.). ASCD.

Suskie, L. (2018). *Assessing student learning: A common sense guide*. John Wiley & Sons.

Swedberg, R. (2020). Exploratory research. In C. Elman, J. Gerring, & J. Mahoney (Eds.), *The production of knowledge: Enhancing progress in social science* (pp. 17-41). Cambridge University Press.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's Alpha. *International Journal of Medical Education*, *2*(1), 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Teddlie, C., & Yu, F. (2007). Mixed Methods Sampling. *Journal of Mixed Methods Research*, *1*(1), 77–100. https://doi.org/10.1177/2345678906292430

Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. In C. Willig & R. W. Stainton (Eds.), *The SAGE handbook of qualitative research in psychology* (2nd ed., pp. 17-37). SAGE Publications.

The Hong Kong Academy of Performing Arts (n.d.). *Chinese Opera.* The Hong Kong Academy of Performing Arts. Retrieved from https://www.hkapa.edu/co#

Thorpe, M. (2000). Encouraging students to reflect as part of the assignment process: Student responses and tutor feedback. *Active learning in higher education*, *1*(1), 79-92. https://doi.org/10.1177/1469787400001001006

Treagust, D. F. (1995). Diagnostic assessment of students' science knowledge. In D. F.

Treagust (Ed.), *Learning science in the schools* (pp. 339-358). Routledge.

United Nations Educational, Scientific, and Cultural Organization (UNESCO). 2007. *Education for all global monitoring report 2008: Education for all by 2015. Will we make it?* UNESCO/Oxford University Press.

Wagemaker, H. (2020). *Reliability and validity of international large-scale assessment: Understanding IEA's comparative studies achievement*. Springer. https://doi.org/10.1007/978-3-030-53081-5

Wang, P.C. (2019). Assessment in music classes in China. In T. S. Brophy (Ed), *The Oxford Handbook of Assessment Policy and Practice in Music Education* (1st ed., Vol.1, pp. 339-358). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190248093.013.15

Wang, Y. J. (2019). *初探音乐师范专业粤剧鉴赏课程建设——以华南师范大学音乐学院为例* [Preliminary study on the construction of Cantonese opera appreciation course for normal school students - A case study of Music College of South China Normal University]. *Song of the Yellow River, 13,* 87. https://doi.org/10.19340/j.cnki.hhzs.2019.13.068

Wang, Z. D., Wang, Y. M., Li, K., Shi, J., & Wang, F. Y. (2021). The comparison of the wisdom view in Chinese and Western cultures. *Current Psychology*, 1-12. https://doi.org/10.1007/s12144-020-01226-w

Waugh, C. K., & Gronlund, N. E. (2013). *Assessment of student achievement* (10th ed.). Pearson.

Weinstein, A., Feldtkeller, B., Malizia, A., Wilson, S., Bailey, J., & Nutt, D. J. (1998). Integrating the cognitive and physiological aspects of craving. *Journal of Psychopharmacology*, *12*(1), 31-38. https://doi.org/10.1177/026988119801200105

Weisberg, H., Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3rd ed.). SAGE Publications.

Whiting, L. S. (2008). Semi-structured interviews: Guidance for novice researchers. *Nursing Standard*, *22*(23), 35-40. https://doi.org/10.7748/ns2008.02.22.23.35.c6420. PMID: 18323051

Wilson, T., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126. https://doi.org/10.1037/0033-295X.107.1.101

Wormald, B. W., Schoeman, S., Somasunderam, A., & Penn, M. (2009). Assessment drives learning: an unavoidable truth? *Anatomical sciences education*, *2*(5), 199-204. https://doi.org/10.1002/ase.102

Yuen, S. F. (1993). *Duo Hangdang De Yishu Shijian.* [The artistic practice of multiple roles]. In J. Li (Ed.), *Xianggang Yueju* Koushushi [Oral History of Hong Kong Cantonese Opera] (pp. 66–74). Joint Pub.

Yuen, S. F. (2008). *Yueju De Tuibian Yu Weiji*. [The transformation and crisis of Cantonese opera]. In S. S. Chow & L. Y. Cheng (Eds.), *Part II of Yueji Guoji Yantaohui Lunwenji [Proceedings of International Cantonese Opera Symposium],* (pp. 7–14). Cantonese Opera Research Project, The Chinese University of Hong Kong.

Yung, B. (1989). *Cantonese opera: Performance as creative process*. Cambridge University

Press.

Zhang, J. J., & Ma, D. (2018). *廣東省佛山市"粵劇進校園"的思考. [The reflection of "Cantonese Opera in Schools" in the context of Foshan, Guangdong]*. Chinese Music Education.

Appendix A

<div align="center">Semi-structured Interview Protocol in Phase I</div>

Script before the interview: I'd like to thank you once again for being willing to participate in the interview aspect of my study. Our interview today will last approximately one hour during which I will be asking you about your opinions on Cantonese operatic singing.

Here is a consent form indicating that I have your permission (or not) to audio record our conversation.

If yes: Thank you! Please let me know if at any point you want me to turn off the recorder or keep something you said off the record.

If no: Thank you for letting me know. I will only take notes of our conversation.

Before we begin the interview, do you have any questions? [Discuss questions]

If any questions (or other questions) arise at any point in this study, you can feel free to ask them at any time. I would be more than happy to answer your questions.

Procedure:

Step 1: Explain the purpose and procedures of the interview for the interviewee.

Step 2: Building rapport. Ask the interviewee about cued questions about their retelling and reflecting on the experience of learning and teaching in Cantonese operatic singing, as well as their macroscopic opinions on its assessment and teaching.

Step 3: Apprehension phase. Provide participants with observational fieldnote and relevant documents to help the interviewee to go over what s/he has just experienced.

Step 4: Exploration phase. The questions should address two of the three research question: 1) What elements that can be measured and assessed in Cantonese operatic singing? (e.g. finding out the measurable elements of Cantonese operatic singing, like intonation, rhythm, meter, pronunciation, and so forth.) 2) How to establish criteria and standards-based assessment for Cantonese operatic singing on the purpose of its transmission and education?

Basic interview questions:

Section 1: Introductory questions

1. Can you tell me about your teaching/learning/practicing experience in Cantonese opera?

Section 2: Transition questions.

2. Cantonese operatic singing graded examination made its debut this year. Here is its syllabus. What is your opinion about this graded examination and its syllabus? Any advantages and disadvantages from your perspective?

3. What do you think about the teaching & learning of Cantonese operatic singing? You can talk about its status in the past, present, even your expectations in the future.

4. What do you think about the assessment of Cantonese operatic singing? You can talk about

its status in the past, present, even your expectations in the future.

5. Is the assessment in aligned with the teaching? Or is there a gap between them? Is an assessment system necessary for Cantonese operatic singing?

Section 3: Key questions

6. What elements are measurable and assessable?

7. Do you think a criteria- and standard-based assessment is necessary for the education and transmission of Cantonese operatic singing?

8. How to ensure diversified skills and knowledge be tested in the examination? (Show the "Figure 4.3. The comparison diagram of components of learning domains involved in POGE and COSGEHK" to the interviewees and explain the implication thereby leading them to tease out the answer to this question)

9. What classification should be employed when formulating the graded examination? For example, should it be classified into roles, categories of operatic singing manners, or something else?

10. Given consider to the construction and diversity, what test items should be included in the examination?

11. Professional and traditional terms or plain language, which should be adopted in writing the description of criteria?

12. How many levels should be adopted to be the assessment standards?

13. Is it necessary to introduce weightings into the assessment? If so, how to set up weightings to meet the requirements of specific assessments?

14. Which type of rubrics should be employed, holistic or analytic rubrics?

Section 4: Closing questions

15. Before we conclude this interview, do you have some suggestions for the assessment criteria of Cantonese operatic singing?

Thank you for your time. That's all the questions I have for you. Do you have any questions for me?

I will send you a copy of your mapped responses so you can identify, and I can correct any mistakes. I will also be in touch with information about the results and outcomes of this work.

Thank you again.

Appendix B

**Consent Form & Information Sheet for Participants Aged 9 to 15 and Participants
Aged 16 to 17 (Non-minimal Risk Studies)**

THE EDUCATION UNIVERSITY OF HONG KONG
Department of Cultural and Creative Arts

CONSENT TO PARTICIPATE IN RESEARCH

**Proposing a Criterion- and standard-based Assessment Framework for Cantonese
Operatic Singing with the Application of the New Bloom's Taxonomy**

I hereby consent my child, _____, to participate in the captioned project
supervised by Prof. Leung Bo Wah and conducted by Luo Yue, who are staff and student of
the Department of Cultural and Creative Arts at The Education University of Hong Kong.

I understand that information obtained from this research may be used in future research and
may be published. However, my right to privacy will be retained, i.e., the personal details of
my child will not be revealed.

The procedure as set out in the **<u>attached</u>** information sheet has been fully explained.   I
understand the benefits and risks involved. My child's participation in the project is voluntary.

I acknowledge that we have the right to question any part of the procedure and can withdraw
at any time without negative consequences.

Signature: _____

Name of Participant: _____

Date: _____

Signature: _____

Name of Parent / Guardian*: _____

Date: _____

Please delete as appropriate

INFORMATION SHEET

Proposing a Criterion- and standard-based Assessment Framework for Cantonese Operatic Singing with the Application of the New Bloom's Taxonomy

You are invited to participate in a project supervised by Prof. Leung Bo Wah and conducted by Luo Yue, who are staff and student of the Department of Cultural and Creative Arts in The Education University of Hong Kong.

The introduction of the research

By establishing a criterion- and standard-based assessment system, this study aims to enhance the assessment in the education of Cantonese operatic singing, which may further facilitate the inheritance, popularization, and development of this traditional art form.

The current study needs information and feedback from knowledgeable people by virtue of their expertise, professional roles, experience in Cantonese operatic singing. You are considered to be the suitable person.

The methodology of the research

The study will consist of two phases. In Phase I, after a content analysis of two *Xiqu* graded examinations, six eminent Cantonese opera experts from Guangdong Province and Hong Kong will be interviewed to solicit their views about the assessment in Cantonese operatic singing. While in Phase II, teachers and students from conservatory tradition or community training in Guangdong Province will be invited to participate in a series of pilot implementations of the assessment rubrics that were established based on the proposed assessment framework. After each class, the teacher and a number of voluntary students will be invited to participate in a semi-structured interview to provide opinions and suggestions.

The semi-structured interview will last for no longer than one hour (for the teacher) or half an hour (for the students).

No financial benefits will be involved except for some souvenirs to express feelings of gratitude.

The potential risks of the research (State explicitly if none)

The current study adopts semi-structured interviews, which will cause no harm and no potential risks to the participants.

Your participation in the project is voluntary. You have every right to withdraw from the study at any time without negative consequences. All information related to you will remain confidential and will be identifiable by codes known only to the researcher.

Your child's participation in the project is voluntary. You and your child have every right to withdraw from the study at any time without negative consequences. All information related to your child will remain confidential and will be identifiable by codes known only to the researcher.

Describe how results will be potentially disseminated

The result of the semi-structured interview will be used as the information and data in my thesis and the publication of journal articles, conference presentations, and the like.

If you would like to obtain more information about this study, please contact me by email at   or telephone number   or my supervisor Prof. Leung Bo Wah by email at

If you have any concerns about the conduct of this research study, please do not hesitate to contact the Human Research Ethics Committee by email at hrec@eduhk.hk or by mail to the Research and Development Office, The Education University of Hong Kong.

Thank you for your interest in participating in this study.

Luo Yue

香港教育大學
文化及創意藝術學系

參與研究同意書

應用新布魯姆分類法為粵曲演唱倡議一個評價体系

茲同意敝子弟＿＿＿＿＿＿＿＿＿＿參加由梁寶華教授負責監督，駱悅負責執行的研究計劃。她／他們是香港教育大學的教員和學生。

本人理解此研究所獲得的資料可用於未來的研究和學術發表。然而本人有權保護敝子弟的隱私，其個人資料將不能洩漏。

研究員已將所附資料的有關步驟向本人作了充分的解釋。本人理解可能會出現的風險。本人是自願讓敝子弟參與這項研究。

本人理解本人及敝子弟皆有權在研究過程中提出問題,並在任何時候決定退出研究，更不會因此而對研究工作產生的影響負有任何責任。

簽署：

參加者姓名：

日期：

簽署：

父母／監護人*姓名：

日期：

*請刪去不適用者

有關資料

應用新布魯姆分類法為粵曲演唱倡議一個提出的評價体系

誠邀閣下及貴子女參加梁寶華教授負責監督，駱悅負責執行的研究計劃。她／他們是香港教育大學的教員和學生。

研究計劃簡介

本研究旨在通過建立基於規範與標準的評價體系，加強粵劇演唱的表演教學和以教育為中心的教學方法，進一步促進粵劇演唱這一傳統藝術形式的傳承、普及和發展。本研究需要粵曲和粵劇相關的專業知識和經驗，貴子女是合適的人選，故誠邀請其參與本研究。

研究方法

本研究分為兩階段。在研究的第一階段，我們會邀請至少六位來自廣東省和香港的著名粵劇專家，聽取他們对于粵曲演唱的意見。在研究的第二階段，我們會從兩間香港和廣東的有關粵劇的學院的選取至少兩個班級進行課堂觀察，觀察與記錄參加者的日常教與學。每節課結束後，老師和一些志願學生將被邀請參加一個半結構性的面談，進一步瞭解與與課堂及粵曲演唱相關的資訊。

課堂觀察將持續一節課的時間，而半結構性訪談將不超過一小時(對教師)或半小時(對學生)。

除贈送紀念品以表達謝意外，本研究不會涉及任何經濟資助。

說明任何風險 (若無，請明確指出)

本研究採用觀察法和訪談法，不會對參與者造成傷害，也沒有潛在的風險。

閣下及貴子女的參與純屬自願性質。閣下及貴子女皆享有充分的權利在任何時候決定退出這項研究，更不會因此引致任何不良後果。凡有關貴子女的資料將會保密，一切資料的編碼只有研究人員得悉。

描述將如何發佈研究結果

課堂觀察和半結構化訪談的結果將被用作資料和資料發表在我的論文和將來的期刊文章、教育相關陳述及展示等形式中。

如閣下想獲得更多有關這項研究的資料,請以電郵

或 電 話   與 本 人 或 本 人 的 導 師 梁 寶 華 教 授

聯 絡 。

如 閣 下 對 這 項 研 究 的 操 守 有 任 何 意 見,可 隨 時 與 香 港 教 育 大 學 人 類 實 驗 對 象 操 守 委 員 會 聯 絡 (電 郵 : hrec@eduhk.hk; 地 址 : 香 港 教 育 大 學 研 究 與 發 展 事 務 處)。

謝 謝 閣 下 有 興 趣 參 與 這 項 研 究 。

駱悅

Appendix C

**Consent Form & Information Sheet for Participants Aged 16 to 17 (Minimal Risk Studies)
and Participants Aged 18 or Above**

THE EDUCATION UNIVERSITY OF HONG KONG
Department of Cultural and Creative Arts

CONSENT TO PARTICIPATE IN RESEARCH

Proposing a System of Assessment for Cantonese Operatic Singing with the application of the new Bloom's Taxonomy

I, _____, hereby consent to participate in the captioned project supervised by Prof. Leung Bo Wah and conducted by Luo Yue who are staff and student of the Department of Cultural and Creative Arts in The Education University of Hong Kong.

I understand that information obtained from this research may be used in future research and may be published. However, my right to privacy will be retained, i.e., my personal details will not be revealed.

The procedure as set out in the **attached** information sheet has been fully explained. I understand the benefits and risks involved. My participation in the project is voluntary.

I acknowledge that I have the right to question any part of the procedure and can withdraw at any time without negative consequences.

Signature:

Name of Participant:

Date:

INFORMATION SHEET

Proposing a System of Assessment for Cantonese Operatic Singing with the application of the new Bloom's Taxonomy

You are invited to participate in a project supervised by Prof. Leung Bo Wah and conducted by Luo Yue, who are staff and student of the Department of Cultural and Creative Arts in The Education University of Hong Kong.

The introduction of the research

By establishing a criterion- and standard-based assessment system, this study aims to enhance the assessment in the education of Cantonese operatic singing, which may further facilitate the inheritance, popularization, and development of this traditional art form.

The current study needs information and feedback from knowledgeable people by virtue of their expertise, professional roles, experience in Cantonese operatic singing. You are considered to be the suitable person.

The methodology of the research

The study will consist of two phases. In Phase I, after a content analysis of two *Xiqu* graded examinations, six eminent Cantonese opera experts from Guangdong Province and Hong Kong will be interviewed to solicit their views about the assessment in Cantonese operatic singing. While in Phase II, teachers and students from conservatory tradition or community training in Guangdong Province will be invited to participate in a series of pilot implementations of the assessment rubrics that were established based on the proposed assessment framework. After each class, the teacher and a number of voluntary students will be invited to participate in a semi-structured interview to provide opinions and suggestions.

The classroom observation will last as long as a class, whereas the semi-structured interview will last for no longer than one hour (for teacher) or half an hour (for students).

No financial benefits will be involved except for some souvenir to express feelings of gratitude.

The potential risks of the research

The current study adopts observation and interview, which will cause no harm and no potential risks to the participants.

Your participation in the project is voluntary. You have every right to withdraw from the study at any time without negative consequences. All information related to you will remain confidential and will be identifiable by codes known only to the researcher.

Describe how results will be potentially disseminated

The result of the classroom observation and the semi-structured interview will be used as the information and data in my thesis and the publication of journal articles, conference presentation

and the like.

If you would like to obtain more information about this study, please contact me by email at   or telephone number   or my supervisor Prof. Leung Bo Wah by email at

If you have any concerns about the conduct of this research study, please do not hesitate to contact the Human Research Ethics Committee by email at hrec@eduhk.hk or by mail to Research and Development Office, The Education University of Hong Kong.

Thank you for your interest in participating in this study.

Luo Yue

香港教育大學
文化及創意藝術學系

參與研究同意書

應用新布魯姆分類法為粵曲演唱倡議一個評價体系

本人同意參加由梁寶華教授負責監督，駱悅負責執行的研究計劃。她／他們是香港教育大學的教員和學生。

本人理解此研究所獲得的資料可用於未來的研究和學術發表。然而本人有權保護本人的隱私，本人的個人資料將不能洩漏。

研究員已將所附資料的有關步驟向本人作了充分的解釋。本人理解可能會出現的風險。本人是自願參與這項研究。

本人理解本人有權在研究過程中提出問題,並在任何時候決定退出研究，更不會因此而對研究工作產生的影響負有任何責任。

簽署:
_____

參加者姓名：
_____

日期：
_____

有關資料

應用新布魯姆分類法為粵曲演唱倡議一個提出的評價体系

誠邀閣下參加梁寶華教授負責監督，駱悅負責執行的研究計劃。她／他們是香港教育大學的教員和學生。

研究計劃簡介

本研究旨在通過建立基於規範與標準的評價體系，加強粵劇演唱的表演教學和以教育為中心的教學方法，進一步促進粵劇演唱這一傳統藝術形式的傳承、普及和發展。本研究需要粵曲和粵劇相關的專業知識和經驗，您是合適的人選，故誠邀請您參與本研究。

研究方法

本研究分為兩階段。在研究的第一階段，我們會邀請至少六位來自廣東省和香港的著名粵劇專家，聽取他們对于粵曲演唱的意見。在研究的第二階段，我們會從兩間香港和廣東的有關粵劇的學院的選取至少兩個班級進行課堂觀察，觀察與記錄參加者的日常教與學。每節課結束後，老師和一些志願學生將被邀請參加一個半結構性的面談，進一步瞭解與與課堂及粵曲演唱相關的資訊。課堂觀察將持續一節課的時間，而半結構性訪談將不超過一小時(對教師)或半小時(對學生)。

除贈送紀念品以表達謝意外，本研究不會涉及任何經濟資助。

說明任何風險 (若無，請明確指出)

本研究採用觀察法和訪談法，不會對參與者造成傷害，也沒有潛在的風險。

閣下的參與純屬自願性質。閣下享有充分的權利在任何時候決定退出這項研究，更不會因此引致任何不良後果。凡有關閣下的資料將會保密，一切資料的編碼只有研究人員得悉。

描述將如何發佈研究結果

課堂觀察和半結構化訪談的結果將被用作資料和資料發表在我的論文和將來的期刊文章、教育相關陳述及展示等形式中。

如閣下想獲得更多有關這項研究的資料,請以電郵

或電話　　　　　　　　與本人或本人的導師梁寶華教授

聯絡。

如閣下對這項研究的操守有任何意見,可隨時與香港教育大學人類實驗對象操守委員會聯絡(電郵:hrec@eduhk.hk; 地址:香港教育大學研究與發展事務處)。

謝謝閣下有興趣參與這項研究。

駱悅

Appendix D

The Introduction of the Assessment System Frameworks

Based on the background of Cantonese opera singing, a new assessment system is formulated. Now we are seeking expert opinions to further improve it. Before an interview, please refer to this introduction for a preliminary understanding of the design of the new assessment system. Thank you very much for your participation!

<u>A new assessment system</u>

Rationale

According to Bloom's Taxonomy, which is firstly proposed by American educational psychologist Benjamin Bloom in 1956, human learning can be divided into the following three domains:
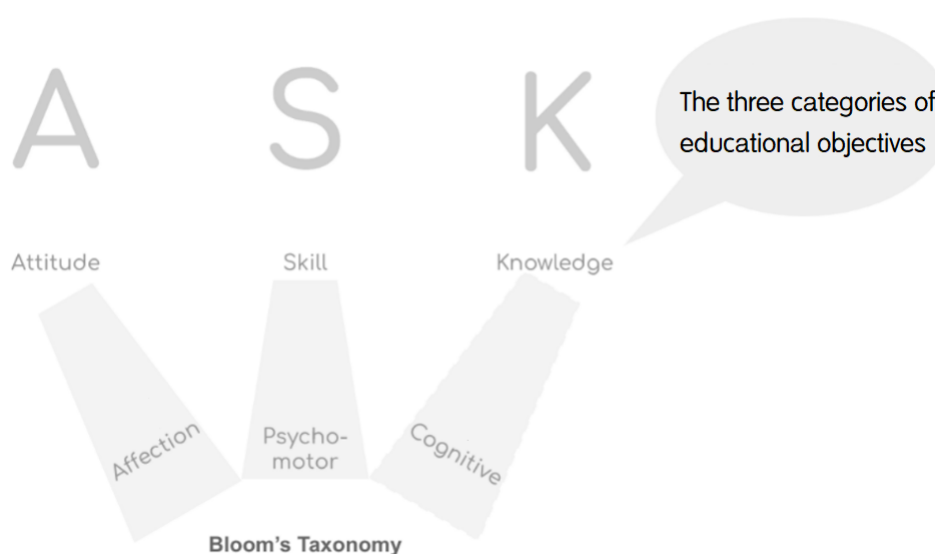
Fig.1 The original Bloom's Taxonomy

"Competence" is a comprehensive development of these three domains and eventually manifests itself.

Fig. 2 The mechanism of all the domains in Bloom's Taxonomy

Therefore, the comprehensive output is produced based on the overall behavior comprising the cognitive, psychomotor, and affective domains.

Based on the above rationale, I attempt to formulate an assessment system according to human being's study progression in every domain for Cantonese operatic singing. The interpretation of every domain and its application in the specific subject in this traditional genre is illustrated as below:

1. Cognitive Domain

This domain involves open-ended questions and answers about relevant knowledge. Teachers can judge students' cognitive level and ability of relevant knowledge points from their answers, which range from low to high hierarchy, including:
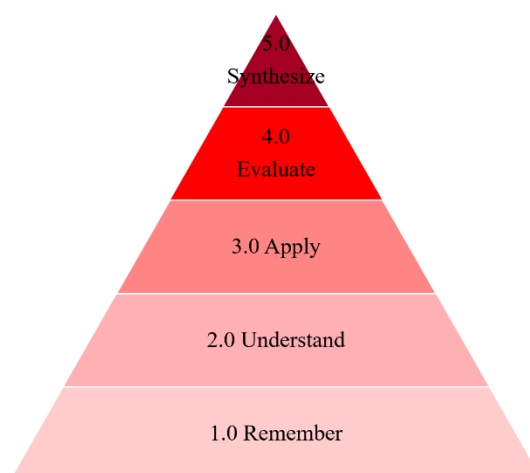
Teachers are requested to use the following table to select the corresponding evaluation for students based on the understanding of the above assessment hierarchies.

Assessment Rubrics in Cognitive Domain						
Name:			Grade:			
Standards Items	Low ————————————————→ High					Weightings
	1.0 Remember	2.0 Understand	3.0 Apply	4.0 Analyze	5.0 Synthesize	
Q1	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/3
Q2	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/3
Q3	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/3
					Total score:	

Notice: For further subdivisions, each level from 1.0 to 5.0 is divided into 4 levels. The more to the right, the higher level it represents. For each line from 1.0 to 5.0, only one grid can be selected with a √. For example, the assessment of a student could be:

Standards Items	Low ————————————————→ High					Weightings
	1.0 Remember	2.0 Understand	3.0 Apply	4.0 Analyze	5.0 Synthesize	
Q1	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ⊙ ○ ○	○ ○ ○ ○	1/3
Q2	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ⊙ ○	○ ○ ○ ○	1/3
Q3	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ⊙	○ ○ ○	1/3
					Score:	78.33

2. Psychomotor Domain

The psychomotor domain includes "the manipulative or motor-skill area", which can be used to assess the singing skill in Cantonese operatic singing. The progressive hierarchy of this domain is as follow:

Assessment Rubrics of Cantonese Operatic Singing Skill (For Junior Students)					
Name:				Grade:	

Standards / Items	Low ⟶ High					Weightings
	1.0 Perceive	2.0 Simulate	3.0 Manipulate	4.0 Articulate	5.0 Master	
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Breath	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Articulation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Timber	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
					Total score:	

Assessment Rubrics of Cantonese Operatic Singing Skill (For Senior Students)					
Name:			Grade:		

Standards \ Items	Low → High					Weightings
	1.0 Perceive	2.0 Simulate	3.0 Manipulate	4.0 Articulate	5.0 Master	
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Breath	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Articulation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Timber	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Flavor	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
					Total score:	

3. Affective Domain

The affective domain contains students' interests, attitudes, and values during learning.

The progressive hierarchy of this domain is as follow:
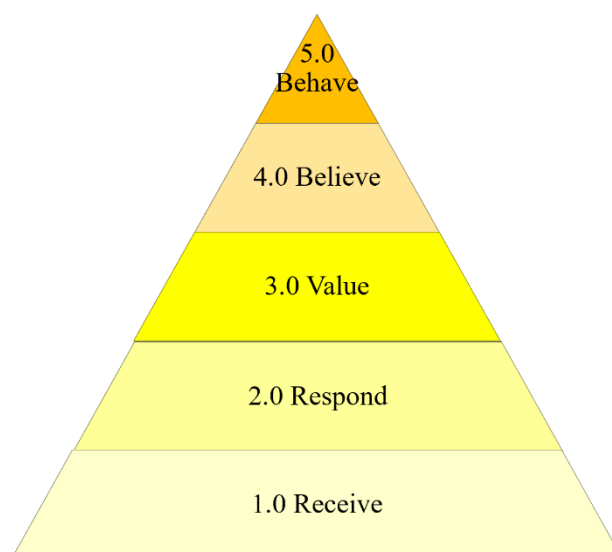
This domain can be assessed via two routes: 1) teachers' observation and evaluation and 2) student questionnaire. Unlike the other domains, Affective Domain will not be counted into the results rightly or wrongly. Its purpose is to collect the information of students' psychology and attitude in learning, so that timely concerns or intervention for academically marginal students can be triggered.

Observational Assessment toward Students' Attitude and Psychology in Learning

Grade:_____   Name:_____

Criteria / Standards	Lower-level			Higher-level	
	1.0 Receive	2.0 Respond	3.0 Value	4.0 Believe	5.0 Behave
11. Regarding the importance of this subject, the student's behavior is:					
12. When confronted with learning difficulties thereby receiving encouragement and instruction of solutions from the teacher, the student's behavior is:					
13. When the teacher stimulates the student's interest in learning in certain approaches, the student's behavior is:					
14. When encouraged to devote more to the course, the student's behavior is:					
15. Regarding the confidence to master the knowledge and skills of this course, the student's behavior is:					
16. Regarding the connection between this course and Cantonese opera, the student's performance is:					
17. Regarding the extra effort that may be required to learn the course well, the student's behavior is:					
18. Regarding the relevance of Cantonese opera to the inheritance of traditional Chinese culture, the student's behavior is:					

Questionnaire on Students' Learning Status and Psychology (For Professional Students)

This questionnaire will not be calculated into the score, nor will it be made public. It is only to understand your psychological condition during the learning process of this course so that the teacher can give you timely care and support. Therefore, please answer this questionnaire truthfully. Please tick the box below the icon between "Strongly Agree" and "Strongly Disagree" to choose your answer.

Standards / Criteria	Strongly agree				Strongly disagree
	1	2	3	4	5
11. I think it is important for my career to learn this course well.					
12. I have a feeling of powerlessness when studying this course					
13. I am very interested in learning this course well.					
14. I would like to spend more time and energy on this course.					
15. I am confident that I can master the knowledge and skills of this course.					
16. I don't think this course is essential for Cantonese opera.					
17. To learn this course well, I have taken the initiative to consult teachers or predecessors.					
18. I think this course is closely related to the transmission of traditional Chinese culture,					

| | which gives me a sense of mission and pride. | | | | | |

Questionnaire on Students' Learning Status and Psychology (For children)

This questionnaire will not be calculated into the score, nor will it be made public. It is only to understand your psychological condition during the learning process of this course so that the teacher can give you timely care and support. Therefore, please answer this questionnaire truthfully. Please tick the box below the icon between "Strongly Agree" and "Strongly Disagree" to choose your answer.

Strongly agree   Strongly disagree

1	I am fond of learning Cantonese operatic singing.	☺ ☐	☺ ☐	😐 ☐	☹ ☐	☹ ☐
2	I have a feeling of powerlessness during learning Cantonese operatic singing.	☺ ☐	☺ ☐	😐 ☐	☹ ☐	☹ ☐
3	I have the confidence to learn Cantonese operatic singing well.	☺ ☐	☺ ☐	😐 ☐	☹ ☐	☹ ☐
4	I am willing to spend considerable time and energy learning Cantonese operatic singing.	☺ ☐	☺ ☐	😐 ☐	☹ ☐	☹ ☐
5	I think this course is important to learning	☺ ☐	☺ ☐	😐 ☐	☹ ☐	☹ ☐

	Cantonese operatic singing.					
6	It is my families' expectation instead of my decision to learn Cantonese operatic singing.	☺ ☐	☺ ☐	☐ ☐	☹ ☐	☹ ☐
7	I feel happy and satisfied during learning this course.	☺ ☐	☺ ☐	☐ ☐	☹ ☐	☹ ☐
8	I intend to keep learning Cantonese opera/Cantonese opera even if I encounter difficulties.	☺ ☐	☺ ☐	☐ ☐	☹ ☐	☹ ☐

4. Behavioral Domain

This domain embraces an integration of cognitive, affective, and psychomotor domains, which can be used to assess the drama class. The progressive hierarchy of this domain is as follow:

5.0 Perfect
4.0 Perform
3.0 Adapt
2.0 Assimilate
1.0 Acquire

Assessment Rubrics of Drama Class in Cantonese Operatic Singing

Standards / Items	Low ⟶ High										Weightings
	1.0 Acquire		2.0 Assimilate		3.0 Adapt		4.0 Perform		5.0 Perfect		
Assessment Rubrics of Drama Class in Cantonese Operatic Singing											
Name: _____ Grade: _____											
Technical Requirements											
Intonation	○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		20%
Pulse & Rhythm	○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		20%
Cantonese pronunc	○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		20%
Changqiang	○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		20%
Artistic Requirements											
Artistic expression in line with the repertoire requirements (including emotion, environment, character)	○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		20%
The appearance in	○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		○ ○ ○ ○		20%
									Total score:		

Appendix E

Assessment toolkits in Excel Format

Assessment Rubrics in Cognitive Domain

Name: _____   Grade: _____

Standards / Items	Low → High					Weightings
	1.0 Remember	2.0 Understand	3.0 Apply	4.0 Analyze	5.0 Synthesize	
Q1	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/3
Q2	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/3
Q3	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/3

Total score: _____

Assessment Rubrics of Cantonese Operatic Singing Skill (For Junior Students)

Name: _____   Grade: _____

Standards / Items	Low → High					Weightings
	1.0 Perceive	2.0 Simulate	3.0 Manipulate	4.0 Articulate	5.0 Master	
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Breath	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Articulation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Timber	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%

Total score: _____

Assessment Rubrics of Cantonese Operatic Singing Skill (For Senior Students)

Name: _____   Grade: _____

Standards / Items	Low → High					Weightings
	1.0 Perceive	2.0 Simulate	3.0 Manipulate	4.0 Articulate	5.0 Master	
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Breath	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Articulation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Timber	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6
Flavor	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	1/6

Total score: _____

Assessment Rubrics of Drama Class in Cantonese Operatic Singing

Name: _____ Grade: _____

Standards \ Items	Low — 1.0 Acquire	2.0 Assimilate	3.0 Adapt	4.0 Perform	5.0 Perfect — High	Weightings
Technical Requirements						
Intonation	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Pulse & Rhythm	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Cantonese pronunc	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Changqiang	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
Artistic Requirements						
Artistic expression in line with the repertoire requirements (including emotion, environment, character)	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%
The appearance in	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	20%

Total score:

Observational Assessment toward Students' Attitude and Psychology in Learning

Name: Grade:

Standards / Items	Low ——————————————→ High				
	1.0 Receive	2.0 Respond	3.0 Value	4.0 Believe	5.0 Behave
1. Regarding the importance of this subject, the student's behavior is:	○	○	○	○	○
2. When confronted with learning difficulties thereby receiving encouragement and instruction of solutions from the teacher, the student's behavior is:	○	○	○	○	○
3. When the teacher stimulates the student's interest in learning in certain approaches, the student's behavior is:	○	○	○	○	○
4. When encouraged to devote more into the course, the student's behavior is:	○	○	○	○	○
5. Regarding the confidence to master the knowledge and skills of this course, the student's behavior is:	○	○	○	○	○
6. Regarding the connection between this course and Cantonese opera, the student's performance is:	○	○	○	○	○
7. Regarding the extra effort that may be required to learn the course well, the student's behavior is:	○	○	○	○	○
8. Regarding the relevance of Cantonese opera to the inheritance of traditional Chinese culture, the student's behavior is:	○	○	○	○	○

Appendix F

Online Students Learning Psychological Survey

Questionnaire on Students' Learning Status and Psychology (For Professional Students)

Name *

Please input your name

Grade *

Please input your grade

This questionnaire will not be calculated into the score, nor will it be made public. It is only to understand your psychological condition during the learning process of this course so that the teacher can give you timely care and support. Therefore, please answer this questionnaire truthfully. Please tick the box below the icon between "Strongly Agree" and "Strongly Disagree" to choose your answer. *

	Strongly agree				Strongly disagree
	1	2	3	4	5
1. I think it is important for my career to learn this course well.	○	○	○	○	○
2. I have a feeling of powerlessness when studying this course.	○	○	○	○	○
3. I am very interested in learning this course well.	○	○	○	○	○
4. I would like to spend more time and energy on this course.	○	○	○	○	○
5. I am confident that I can master the knowledge and skills of this course.	○	○	○	○	○

6. I don't think this course is important for Cantonese opera.	○	○	○	○	○
7. To learn this course well, I have taken the initiative to consult teachers or predecessors.	○	○	○	○	○
8. I think this course is closely related to the transmission of traditional Chinese culture, which gives me a sense of mission and pride.	○	○	○	○	○

Submit

Appendix G

Summary of Test Items of POCG

Grade	Chouhang (丑行)	Daomadan/wudan (刀馬旦/武旦)	Hualian (花臉)	Laodan (老旦)	Laosheng (老生)	Qingyi/huadan (青衣/花旦)	Xiaosheng (小生)	Wusheng (武生)
1	1) singing, 2) reciting.	1) Knowledge of Peking opera, 2) Singing & Reciting, and 3) Acting movement & Acrobatic fighting	1) Singing, 2) Reciting	Singing	1) Knowledge of Peking opera, 2) singing	1) Singing, 2) Reciting.	1) Singing, 2) Reciting	1) Singing, 2) Reciting, 3) The skills of posture, 4) Basic skills.
2	1) singing, 2) reciting.	1) Knowledge of Peking opera, 2) Singing & Reciting, and 3) Acting movement & Acrobatic fighting	1) Singing, 2) Reciting	1) Singing, 2) Reciting	1) Knowledge of Peking opera, 2) singing, 3) Reciting.	1) Singing, 2) Reciting.	1) Singing, 2) Reciting	1) Singing, 2) Reciting, 3) The skills of posture, 4) Basic skills.
3	1) singing, 2) reciting.	1) Knowledge of Peking opera, 2) Singing & Reciting, and 3) Acting movement & Acrobatic fighting	1) Singing, 2) Reciting	1) Singing, 2) Reciting	1) Knowledge of Peking opera, 2) singing, 4) Reciting.	1) Singing, 2) Reciting.	1) Singing, 2) Reciting	1) Knowledge of Peking opera, 2) singing, 3) Reciting, 4) Posture and *bazi*, 5) Basic skills.
4	1) singing, 2) reciting.	1) Singing & Reciting, 2) Acting movement & Acrobatic fighting.	1) Singing, 2) Reciting, 3) Acting movement	1) Singing with posture, 2) Reciting	1) Knowledge of Peking opera, 2) singing, 5) Reciting.	1) Singing, 2) Reciting.	1) Singing, 2) Reciting	1) Knowledge of Peking opera, 2) singing, 3) Reciting, 4) Posture and *bazi*, 5) Basic skills.

Grade	Chouhang (丑行)	Daomadan/wudan (刀馬旦/武旦)	Hualian (花臉)	Laodan (老旦)	Laosheng (老生)	Qingyi/huadan (青衣/花旦)	Xiaosheng (小生)	Wusheng (武生)
5	1) singing, 2) Performance	1) Singing & Reciting, 3) Acting movement & Acrobatic fighting.	1) Singing, 2) Reciting, 3) Acting movement	1) Singing with posture, 2) Reciting, 3) Performance imitation.	1) Knowledge of Peking opera, 2) singing, 3) Performance with reciting.	1) Singing, 2) Reciting.	1) Singing with posture, 2) Reciting, 3) Imitated performance.	1) Knowledge of Peking opera, 2) singing, 3) Reciting, 4) Posture and *bazi*, 5) Basic skills.
6	inclusive performance	1) Singing & Reciting, 4) Acting movement & Acrobatic fighting.	1) Singing & Reciting 2) Play performance, 3) Acting movement imitation.	1) Singing with posture, 2) Reciting, 3) Imitated performance	1) Knowledge of Peking opera, 2) Inclusive performance, 3). The elaboration of the character's costume and makeup.	1) Singing, 2) Reciting.	1) Singing with posture, 2) Reciting, 3) Imitated performance.	1) Knowledge of Peking opera, 2) singing, 3) Reciting, 4) Posture and *bazi*, 5) Basic skills.
7	inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	1) Inclusive performance, 2) The elaboration of the skills that the character used in the play.	Inclusive performance	Inclusive performance	1) Knowledge of Peking opera, 2) Inclusive performance.
8	inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	1) Knowledge of Peking opera, 2) Inclusive performance.
9	inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	Inclusive performance	

The Education University of Hong Kong Library
For private study or research only.
Not for publication or further reproduction.

Appendix H

Semi-structured Interview Protocol in Phase II

I'd like to thank you once again for being willing to participate in the interview aspect of my study. Our interview today will last approximately one hour. I will be asking you about your opinions after conducting a pilot implementation of the proposed assessment framework for Cantonese operatic singing.

Basic interview questions:

Section 1: Introductory questions

1. Can you tell me about your experience of assessment in Cantonese opera?

2. Based on your teaching/learning/practising experience, what do you think of the assessment that you have had in Cantonese operatic singing?

Section 2: Transition questions.

3. Previously we have conducted a pilot implementation of the proposed assessment framework.

    For teachers: What do you think during or/and after applying this assessment framework? Are there any distinct gaps between the assessment results and the student's regular academic performance?

    For students: What do you think after receiving the assessment results with applying this assessment framework? Are there any distinct gaps between the assessment results and your average academic performance?

Section 3: Key questions

4. What are the merit and demerit of this assessment framework?

5. Can this assessment framework help your teaching/learning?

6. Would you like to use it during your teaching/learning in Cantonese operatic singing in the future?

Section 4: Closing questions

7. Before we conclude this interview, do you have some suggestions for improving this newly proposed

   assessment framework?

Thank you for your time. That's all the questions I have for you. Do you have any questions for me?

I will send you a copy of your mapped responses so you can identify and I can correct any mistakes. I will

also be in touch with information about the results and outcomes of this work.

Thank you again.