

Product Evaluation on the Chinese Language Program for International Students in a Chinese University

Xianjun Tan*¹ , Vincent Pang² , Denis Andrew D. Lajium²

¹ Department of International Cooperation and Exchange, Sichuan University of Science and Engineering, Xueyuan Street, Zigong City, China & Faculty of Psychology and Education, University Malaysia Sabah, Jalan UMS, Kota Kinabalu, Malaysia

² Faculty of Psychology and Education, university Malaysia Sabah, Jalan UMS, Kota Kinabalu, Malaysia

ABSTRACT

Hanyu Shuiping Kaoshi (HSK) is an international standardized test, which means Chinese Proficiency Test in English. A minimum of HSK level 4 is needed for an international student to apply for an academic programme taught in Chinese language in the universities of China. Guided by product evaluation of Context, Input, Process, Product (CIPP) model, the test scores of HSK level 4 of the international students in a Chinese university are analyzed using effect size. Feedbacks from the international students who failed in the test of HSK level 4 are collected based on a questionnaire formulated by the researchers. The strengths, weaknesses, causes of failure and improving direction of the one academic year's Chinese language program in this Chinese university are identified through the evaluation. The research aims to provide references for the improvement of Chinese language teaching in this Chinese university and universities with the same context. The research enriches the literature concerning the evaluation on Chinese language teaching and learning for international students in China.

Keywords: product evaluation, HSK, Chinese language, international students.

1. Introduction and Background

Hanyu Shuiping Kaoshi (HSK) means Chinese Proficiency Test in English. HSK is an international standardized test which is used to assess non-native Chinese speakers' abilities in using the Chinese language. Non-native Chinese speakers include foreign citizens, overseas Chinese, foreign citizens of Chinese origin and China's ethnic minorities. The abilities in using the Chinese language in the daily, academic and professional lives of the non-native speakers are assessed in HSK. HSK level ascends from HSK level 1 to HSK level 6. When the test takers pass level 1, they can meet basic needs for communication and have the ability to further their Chinese language learning.

When the test takers pass level 6, they can understand spoken and written information in Chinese easily and express themselves in Chinese effectively, both in oral and in writing (Chinese Test). Until the end of 2018, there are 349 HSK test centers (Hanban) in mainland China.

According to the *Quality Standards for Higher Education of International Students in China* issued by Ministry of Education of China (2018), for international students applying for an academic programme which is taught in Chinese, the language requirement for these students a Level 4 pass as prescribed by *Chinese Language Proficiency Scales for Speakers of Other Languages*. The *Chinese Language Proficiency Scales for Speakers of Other Languages*, which guides the teaching of Chinese as a foreign language was formulated by the Office of Chinese Language Council International (2007). The scale consist of five levels and Level 4 is equivalent to HSK level 4. A pass of HSK level 4 indicates the ability to converse in Chinese with regard to a wide range of topics and the ability to communicate with native Chinese speakers in a fluent manner. HSK level 4 test consists of three parts, namely, listening comprehension (45%), reading comprehension (40%), and writing (15%). The total score for HSK level 4 is 300, and 100 for each part.

Sichuan University of Science and Engineering (SUSE) is a public university in the southwest of China with two campuses located in Zigong city and Yibin city respectively. It enrolls international students from various countries of the world. International students, who haven not passed HSK level 4 before they enter SUSE are required to take up Chinese language courses in SUSE for an academic year. After the courses, the students must attend the test of HSK level 4 and pass it in order to be qualified for bachelor programs or master programs which are taught in Chinese in SUSE.

The Context, Input, Process and Product (CIPP) evaluation model was developed by Daniel L. Stufflebeam in the late 1960s and it is a comprehensive framework used to guide evaluations (Kellaghan, Stufflebeam, & Wingate, 2003). Over the years, this model has been further developed and CIPP evaluation model checklist was provided by Stufflebeam (2001). The formal definition of evaluation summarizing the key ideas of CIPP evaluation model is as follows: Evaluation is the process of describing, acquiring, supplying, and using depictive and judgmental information about the merit and worth of certain object's goals, plan, execution, and outcomes to guide improvement decisions, offer accountability reports, impart institutionalization decisions, and enhance understanding of the examined phenomena (Kellaghan, Stufflebeam, & Wingate, 2003). Guided by product evaluation in CIPP evaluation model, the researchers intended to analyse the test scores of HSK level 4 of international students after one academic year of Chinese language learning in SUSE. Barriers to meet the language requirement for international students will be diagnosed. Teaching priorities in the Chinese language program will be judged. The goals for improvement of the Chinese language program will be set. Evaluation findings are to be used by the relevant policy makers of SUSE who are in charge of the management of international students. Universities such as SUSE can use the evaluation findings as reference in their teaching practice for international students.

2. The Chinese Language Program

The one-year Chinese language program for international students was conducted on Yibin campus of SUSE. The learning outcomes of the programs were measured by the test scores of HSK level 4 of the international students. The courses provided for the international students are shown in Table 1.

Table 1 Courses for the International Students

	Courses	Class Hours of Each Week	Duration of Weeks
First Semester	Elementary Comprehensive Chinese	6	14
	Elementary Chinese Practice	6	14
	Oral Chinese	2	14
	Chinese character	2	14
Second Semester	Elementary Comprehensive Chinese	8	16
	Elementary Chinese Practice	4	16
	Chinese listening	4	16
	Chinese character	2	16

The second semester was from March to June of 2019. In the second semester, there were special courses for international students of SUSE, namely Chinese culture courses. SUSE organized cultural experience activities for the international students. In the last segment of the program, the international students visited Xingwen Global Geopark of Sichuan Province, Modern Agricultural Park of Xingwen county, and Yongshou Village of Xingwen county, to experience Chinese culture and actual development of China. Some thematic cultural courses, such as Chinese tea, Chinese liquor, etc were also provided to the students. Besides the Chinese courses in the classroom and Chinese culture courses in and outside classroom, there were various cultural activities on campus or outside campus of SUSE in which the international students could experience and learn about the actual development of China.

The research questions concerning the Chinese language program are as follows:

1. What are the differences in the attainment of listening, reading and writing?
2. What are the differences in attainment according to gender?
3. What are the differences in attainment according to region of origin?

Some researchers in China analyzed the scores of HSK level 4 in their universities before. Based on the research of Xu Xinyuan (2013) from Tongji University, Yu and Jiang (2014) from Xinjiang University and Xu Lin (2016) from Kaili University, the performance in the listening comprehension part of the international students in their

universities is better than the performance in the reading comprehension part, and the worst performance of the international students is in the writing part. Women may be better at second language learning than men and female learners generally do better than male learners (Ellis, 1994). As stated by Ellis and believed by many people, female language learners seem to have inborn advantage over male language learners. According to Xu Lin (2016), the pass rate of female international students in Kaili University was better than male international students in HSK level 4 test. According to Xu Xinyuan (2013), international students from Kazakhstan, Turkmenistan, Equatorial Guinea, Antigua and Barbuda had a bad performance in the HSK level 4 test,

3. Method

CIPP evaluation is a decision and accountability-oriented evaluation approach. The model was selected by the researchers because it could promote the ongoing improvement efforts and overcome deficiency. The needs and problems within a specific environment are assessed by context evaluation, which helps to assess goals and identify what needs to be done. Work plans and budgets are assessed by input evaluation, which helps to design improvement efforts. Activities are monitored and assessed by process evaluation, which helps to implement improvement efforts. Intended and unintended outcomes are assessed by product evaluation, which helps to make informed decision to go on, halt or improve the efforts. This research only involves product evaluation as it makes emphasis on the attainment of outcomes which are measured by test results.

In this research, the target population is the cohort of international students who finished one-year Chinese language program in SUSE. The international students who took HSK level 4 in the HSK test center of SUSE were chosen for this evaluation research.

One of the objectives of product evaluation is to collect descriptions and judgments of outcomes. In this research, test scores of HSK level 4 of international students in SUSE were used to measure program outcomes. HSK test center of SUSE has the archive of test scores of international students who attended the HSK test in SUSE. All the test scores were provided in Excel format by the Chinese Testing International Co., Ltd in Beijing.

To compare outcomes across groups, the magnitude of effect or effect size is used. Effect size is the difference between the means in the form of standard score (Hunter, Schmidt & Jackson, 1982). Effect size (ES) can be calculated mathematically

using the formula $ES = \frac{x_2 - x_1}{s_p}$ In the formula, x_1 and x_2 are the means of the estimates of

the respective groups and s_p is the pooled standard deviation of the groups which is

computed from the following formula: $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ In the formula,

n_1 and n_2 are the sample sizes and s_1 and s_2 their corresponding standard deviations (Glass, McGaw & Smith, 1981). The use of categories for the interpretation of effect size are summarized in Table 2 (Cohen, 1988, in Pang, 2014).

Table 2 Interpretation of effect size

Value of effect size	Effect
$ES < 0.2$	Nil
$0.2 \leq ES < 0.5$	Small
$0.5 \leq ES < 0.8$	Medium
$ES \geq 0.8$	Large

4. Findings

4.1 Analysis of Scores of HSK Level 4

After one academic year's Chinese language learning, the international students of SUSE took HSK level 4 in the HSK test center of SUSE in May, June and July of 2019 respectively. Due to the different levels of Chinese learning, the international students chose to attend the HSK level 4 at different months. Based on the data of HSK test center of SUSE, Table 3 was formulated by the researcher.

Table 3 Students attended the HSK level 4 in HSK test center of SUSE

Date of Test	Number of Students	Percentage
May 11 of 2019	4	3
May 26 of 2019	8	7
June 16 of 2019	14	12
June 29 of 2019	37	32
July 13 of 2019	53	46
	116	100

Source: HSK test center of SUSE & own compilation

It can be seen that only 10% of students attended the test in May. The test scores in the three months were collected by the researchers for evaluation. For students who attended the test twice, the latter test score was chosen by the for evaluation. Demographic information of the HSK level 4 test takers is formulated by the researchers in Table 4.

Table 4 Demographic information of HSK level 4 test takers

Nationality	Region	Number of students		
		Male	Female	Total
Azerbaijan	Asia	1		1
Cambodia	Asia	7	4	11
Kazakhstan	Asia	7	4	11
Laos	Asia	1	1	2
Mongolia	Asia	1		1
Nepal	Asia	9		9
Pakistan	Asia		1	1

Tajikistan	Asia	7	1	8
Thailand	Asia	1	3	4
Uzbekistan	Asia	2	1	3
Yemen	Asia		1	1
Angola	Africa	1		1
Benin	Africa	1		1
Botswana	Africa		2	2
Burundi	Africa	5	1	6
Chad	Africa	1		1
Cote d'Ivoire	Africa	2		2
D.R.Congo	Africa	1		1
Equatorial Guinea	Africa	3	5	8
Ethiopia	Africa	1		1
Gabon	Africa	1		1
Ghana	Africa	2	1	3
Guinea	Africa	1		1
Liberia	Africa	1		1
Morocco	Africa	1		1
Republic of Congo	Africa		3	3
Tanzania	Africa	5		5
Togo	Africa	1		1
Uganda	Africa		1	1
Zambia	Africa		3	3
Zimbabwe	Africa	8	5	13
France	Europe	1		1
Russia	Europe	1	3	4
Costa Rica	North America	1		1
Papua New Guinea	Oceania	2		2
		76	40	116

Source: HSK test center of SUSE & own compilation

From Table 4, it can be seen that 76 male international students took the test, that is 66% of the total target population and 40 female international students took the test, that is 34% of the total target population. All together, these international students come from 35 countries. There are 56 students from 20 African countries, including Angola, Benin, Botswana, Burundi, Chad, Cote d'Ivoire, D.R. Congo, Equatorial Guinea, Ethiopia, Gabon, Ghana, Guinea, Liberia, Morocco, Republic of Congo, Tanzania, Togo, Uganda, Zambia and Zimbabwe. There are 52 students from 11 Asian countries, including Azerbaijan, Cambodia, Kazakhstan, Laos, Mongolia, Nepal, Pakistan, Tajikistan, Thailand, Uzbekistan and Yemen. There are five students from two European countries, namely Russia and France. There are two students from Oceania, the country is Papua New Guinea. There is only one student from Costa Rica, North America.

4.2 Comparison of Scores of Listening, Reading and Writing Skills

Table 5 Comparison of scores of each skill of HSK level 4

Source: HSK test center of SUSE & own compilation

	Mean	SD	Effect Size	Effect
Listening	65.28	16.79	-0.043	NIL
Reading	64.48	19.99		
Listening	65.28	16.79	0.255	SMALL
Writing	69.64	17.39		
Reading	64.48	19.99	0.275	SMALL
Writing	69.64	17.39		

From Table 5, it can be seen that there is no difference between the listening skill and the reading skill, the performance of students in the writing skill is slightly better than in the listening skill and the reading skill.

4.3 Comparison of Scores Based on Gender

Table 6 Comparison based on gender

		Mean	SD	Effect Size	Effect
Listening Scores	Male	65.88	16.78	-0.103	NIL
	Female	64.15	16.95		
Reading Scores	Male	65.62	20.29	-0.165	SMALL
	Female	62.33	19.49		
Writing Scores	Male	70.11	17.07	-0.077	NIL
	Female	68.75	18.17		

Source: HSK test center of SUSE & own compilation

From Table 6, it can be seen that there is no difference between males and females in the listening skill and the writing skill, the performance of male students is slightly better than females in the reading skill.

4.4 Comparison of Scores Based on Continents

Scores of students from Asia, Africa and Europe are chosen for analysis. Only one student from Costa Rica and two students from Papua New Guinea, who are excluded from the analysis due to the small number.

Table 7 Comparison based on continents

	Number	Mean	SD	Effect Size	Effect
Listening Scores	56 African Students	63.16	15.81	0.249	SMALL
	52 Asian Students	67.15	16.27		
Listening Scores	56 African Students	63.16	15.81	0.201	SMALL
	5 European Students	67.4	25.22		

Listening Scores	52 Asian Students	67.15	16.27	0.012	NIL
	5 European Students	67.4	25.22		
Reading Scores	56 African Students	60.91	17.69	0.366	SMALL
	52 Asian Students	67.96	20.75		
Reading Scores	56 African Students	60.91	17.69	0.298	SMALL
	5 European Students	66.8	21.63		
Reading Scores	52 Asian Students	67.96	20.75	-0.055	NIL
	5 European Students	66.8	21.63		
Writing Scores	56 African Students	63.88	16.46	0.725	MEDIUM
	52 Asian Students	75.69	16.14		
Writing Scores	56 African Students	63.88	16.46	0.652	MEDIUM
	5 European Students	73.8	13.88		
Writing Scores	52 Asian Students	75.69	16.14	-0.126	NIL
	5 European Students	73.8	13.88		

Source: HSK test center of SUSE & own compilation

From Table 7, it can be seen that there is no difference between Asian students and European students in the listening skill, the performance of Asian students and European students is slightly better than African students in the listening skill.

There is no difference between Asian students and European students in the reading skill, the performance of Asian students and European students is slightly better than African students in reading skill.

There is no difference between Asian students and European students in the writing skill, the performance of Asian students and European students is better than African students in the writing skill.

4.5 Feedback from the International Students Who Failed in the Test

The researcher formulated a questionnaire to get feedback from the international students who failed in the test of HSK level 4. The target population are the students who failed in the test and chose to learn Chinese for another academic year in SUSE. The researcher sent the questionnaire in Word format to one student in this population and asked the student to gather the feedback from this population. The student sent back all the feedback to the researcher online. The questionnaire collected the gender, age and nationality and focused on four questions: "What are the strengths of the Chinese language program in SUSE?", "What are the weaknesses of the Chinese language program in SUSE?", "What caused your failure in the test of HSK level 4?", and "In order to help the international students pass HSK level 4, how to improve the Chinese language program in SUSE?".

The feedback of two male students are excluded in the research because they did not attend the test of HSK 4 after one academic year's Chinese learning. The feedback of the students who attended the test and failed are chosen by the researcher. Five male students gave their feedback, three from African countries and two from Asian countries, their ages range from 22 to 25. Three female students gave their feedback, they are from African countries and their ages range from 19 to 21. In their original words, the strengths of the Chinese language program in SUSE include:

"Teachers are very dedicated and Chinese language is taken as seriously as any other major at SUSE"

"patience and dedication of teachers"

"separate classes for reading, listening and writing"

"good explanations and PowerPoint presentations"

"The teachers take classes very seriously"

"Reading, listening and practice are in separate classes"

"The teachers are very kind and sincere, they really help us learn the language"

"Good learning environment and classes are taken seriously"

The weaknesses include:

"inability to finish HSK book 4 during the end of academic year"

"they barely finish HSK books"

"I think there is no weakness, everything is organized"

"Teachers spend more time on HSK 1, 2 and 3 and less time on 4 which is the most difficult"

"The time spent on HSK 4 is not enough, students are forced to learn the last book on their own"

Causes of failure include:

"I wasn't able to finish HSK books in time"

"inadequate preparation"

"I wasn't 100% ready"

"there was not enough practice"

"laziness"

"We didn't finish last HSK 4 book"

"I wasn't responsible and I often went on trips to other cities"

"One year was not enough for me"

In order to improve the Chinese language program, their suggestions include:

"more time on teaching HSK book 4 and revise it"

"give PowerPoint of each lesson ahead"

"The school can provide more activities that help Chinese language practice"

"cruise through HSK 1, 2, 3 and spend more time on HSK 4 and plan school activities specifically to practice Chinese language"

"provide more exercises for students; Teachers should make more HSK 4 practice tests"

"allocate more time in studying HSK 4 books"

“Students should be reminded frequently that if they don’t pass HSK 4, they will repeat another year of Chinese, so that they should work harder”

“more explanations and exercises on HSK 4 specifically”

4.6 Discussions on Findings

For the differences in the attainment of listening, reading and writing, based on the analysis of the scores of HSK level 4 in SUSE, the situation is different from the research before (discussed in Section 2), the best performance is not in the listening skill. As analyzed in the section 4.2, there is no difference between listening skill and reading skill, the best performance of the students is in the writing skill, which is slightly better than listening skill and reading skill.

For the differences in attainment according to gender, based on the analysis in the section 4.3, there is no difference between males and females in the listening skill and the writing skill, the performance of male students is slightly better than females in the reading skill. Hence, in the whole, in this research, the performance of male language learners is slightly better than female language learners. This finding is different from the result obtained in Kaili University before.

For the differences in attainment according to region of origin, there is no difference between students from Asia and students from Europe in the listening skill, reading skill and writing skill of HSK level 4. The performance of students from Asia and Europe is better or slightly better than students from Africa in the three skills of HSK level 4.

Based on the feedback from the students who failed in the test and chose to learn Chinese for another academic year in SUSE, the relevant findings include: the strengths of the Chinese language program focus on the dedication of teachers and serious Chinese teaching; the weaknesses lie in the unfinished books, especially book of HSK 4; the causes of failure in the test are due to personal reasons of the test takers; in order to help international students pass HSK 4, the students hope that the teachers spend more time on book of HSK 4.

5. Conclusions

Based on the analysis above, barrier to meet the language requirement for international students of SUSE is not in the writing comprehension part but in the listening comprehension and reading comprehension parts, which indicates that teaching priorities in the one academic year’s Chinese language program of SUSE should be given to the listening comprehension and reading comprehension parts. The Chinese language teachers should use various methods to help the international students overcome the difficulties in Chinese listening and reading. The policy makers of SUSE and the involved Chinese language teachers should set goals to improve the listening and reading abilities of the international students while keeping the strengths in the writing part. Female international students could be outperformed by their male counterparts in learning Chinese language, as shown in SUSE, which reminds the relevant Chinese language teachers that they should pay special attention to the performance of the female international students in the process of language teaching. In terms of country of origin,

the relevant Chinese language teachers should pay special attention to the international students who come from African countries, as their performance in HSK level 4 is relatively worse. This research is conducted from the perspective of international students, more researches could be done to enrich the literature of evaluation on the Chinese language program for international students, such as, from the perspective of teachers, from the diachronic perspective, etc.

References

- Chinese Test*. (n.d.). About the HSK. Retrieved June 13, 2020 from <http://www.chinesetest.cn/gosign.do?id=1&lid=0>
- Ellis, R. 1994. *The Study of Second Language Acquisition*. New York : Oxford University Press. pp. 202
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social science research*. Beverly Hills, CA: Sage.
- Hanban. (n.d.). *About Chinese test*. Retrieved June 13, 2020 from http://www.hanban.org/tests/node_7475.htm
- Hunter, J. E., Schmidt, F. L., & Jackson, G. B. 1982. *Meta-analysis: Cumulating research findings across studies (Vol. 4)*. Sage Publications, Inc.
- Kellaghan, T., Stufflebeam, D. L., & Wingate, L. A. (Eds.). 2003. *International handbook of educational evaluation*. Kluwer International Handbooks of Education (Vol. 9). Kluwer Academic Publisher.
- Ministry of Education of China. 2018. *Quality standards for higher education of international students in China*.
- Pang, V. 2014. *Curriculum evaluation: An application in a smart school curriculum implementation*. Kota Kinabalu, Malaysia: Penerbit Universiti Malaysia Sabah.
- Stufflebeam, D. L. 2001. *CIPP evaluation model checklist: A tool for applying the CIPP model to assess projects and programs*. Retrieved from <https://wmich.edu/evaluation/checklists>.
- The Office of Chinese Language Council International. 2007. *Chinese language proficiency scales for speakers of other languages*. Beijing: Foreign Language Teaching and Research Press.
- Xu Lin & Zhang Yong. 2016. *A multivariate analysis on Lao students' HSK (level IV) achievement*. *Journal of Chuxiong Normal University* 10: 39-42
- Xu Xinyuan. 2013. *The analysis and countermeasures research of the new HSK examination: The case of Tongji University's preparatory foreign students*. *China Examinations* 8: 48-54
- Yu Liang & Jiang Min. 2014. *An analysis of the new HSK test results of central Asian and Russian learners of Chinese in Xinjiang University and pedagogical implications*. *Journal of Xinjiang University (Philosophy, Humanities & Social Science)* 4: 148-151