Investigating Japanese EFL University Students' Needs, Motivations and Perceptions of the New EAP Program

Mikako Nishikawa
Carey Benom
David Cupchak
Pino Cutrone
William Collins
Brien Datzman
Akira Hiroe
May Kyaw Oo
Scott Lucas
Kazuhiko Moji
Toshiaki Nishihara
Sean Albert Sullivan
Akihide Tada
Shin-ichi Tanigawa
Tianqi Zhang

長崎大学大学院多文化社会学研究科・多文化社会学部『多文化社会研究』 2023年第9号　抜刷

Investigating Japanese EFL University Students' Needs, Motivations and Perceptions of the New EAP Program

Mikako Nishikawa[1]
Carey Benom[2]
David Cupchak[2]
Pino Cutrone[1]
William Collins[2]
Brien Datzman[2]
Akira Hiroe[2]
May Kyaw Oo[2]
Scott Lucas[1]
Kazuhiko Moji[1]
Toshiaki Nishihara[2]
Sean Albert Sullivan[1]
Akihide Tada[3]
Shin-ichi Tanigawa[1]
Tianqi Zhang[1]

Abstract

The School of Global Humanities and Social Sciences (SGHSS) at Nagasaki University introduced a new curriculum for English for Academic Purposes (EAP) courses in 2022. "The KEY Program: Academic English for Unlocking Futures" breaks with tradition; unlike the General English Courses at Nagasaki University, the courses in the KEY program explore innovative ways to build communicative competence, especially academic English competence in the target language. In addition to English classes that meet once a week, some of the core classes of the newly introduced KEY program meet twice a week for 90 minutes. The program is designed to empower SGHSS students so that they

[1] School of Global Humanities and Social Sciences, Nagasaki University
[2] Center for Language Studies, Nagasaki University
[3] Center for Japanese Language and Student Exchange, Nagasaki University

progress in their English to meet the language requirement to register for upper-level course subjects taught in English (TOEFL ITP 500 or equivalent). This paper, based on students' progress and program experience, reports on the implemented KEY program, students' needs, motivations, and perceptions of their first year in the Program. Pre- (n=101) and post-program surveys (n=84) have been conducted to understand students' learning experiences, as well as to design, analyze, and assess the overall curriculum. The primary findings provide fresh insight into the current state of EAP courses in higher education. The results of the first-year program help identify areas for improvement so that SGHSS students' learning experience in the KEY program is enriched as we take a several steps toward our goal of "unlocking futures."

Keywords: EAP, Students' Needs, Students' Perception, Integrated-Language Skills

1. Introduction

The KEY program was developed from the Special Course in Academic Skills Program (SCAS), which began in 2015 as a part of the "Global Plus" project at Nagasaki University and ended in 2021. In-service TESOL instructors from Northern Arizona University co-developed the SCAS curriculum under the "Global Plus Course" project initiatives by the Center for Japanese Language and Student Exchange in 2015. The primary purpose of the English course was to (i) raise students' TOEFL ITP scores, (ii) increase the number of students who go abroad, and (iii) increase students' academic English skills (Conway, 2020). From its completion in 2015 to its closure in 2021, five hundred and twenty students have participated in the SCAS Program. The SCAS program provided an English-speaking environment where students had the opportunity to improve their English communicative skills through interactive activities both in class and off-campus (e.g., Summer Camp, KAKEHASHI Project, and JENESYS Project). It has positively impacted English teaching and learning communities at Nagasaki University, even after the Global Plus project had ended its term. In 2021, in deference to the joint missions of three departments, namely the Center for Japanese Language and Student Exchange, the Center for Language Studies, and the School of Global Humanities and Social Sciences, the SCAS program underwent a transformation and was reborn as the "KEY program: Unlocking futures."

The KEY program is an exclusive English program offered to all undergraduate students in the School of Global Humanities and Social Sciences (SGHSS) at Nagasaki Univer-

sity. These English courses are offered as accredited General English Courses for SGHSS students. SGHSS students are now exempt from taking regular General Education English courses. The SGHSS requires that students reach a minimum threshold (≧ 500 on TOEFL ITP or ≧ 5.5 on IELTS) to be able to take academic courses in English offered for their major. To help students achieve such goals, the KEY program emphasizes teaching integrated skill-based academic English courses to build the competency and skills necessary for the students to study the subjects in English. The KEY program puts an emphasis on providing additional language support to meet individual needs. Coaching Fellows (CFs), who act as teaching assistants and tutors, are appointed to schedule meetings with all students throughout the semester. CFs provides individual feedback on class assignments and offer tutorial sessions in person to help students stay motivated and find the strategies that best fit their learning styles. Unlike the SCAS program, the KEY program English courses are not optional, and all SGHSS students must pass the courses as part of the graduation requirement.

2. Review of the Literature

2.1. Designs for Developing EAP courses

English for Academic Purposes (EAP) courses prepare students for tasks required to successfully complete academic subject courses. Past studies on the needs analysis of the EAP students suggest the importance of closing gaps between the level of L2 learners and the high demands of EAP tasks (e.g., Berman & Chang, 2001; Chan, 2001; Hosogoshi & Takahashi, 2015). Therefore, in designing an EAP curriculum in a university, identifying students' views of language needs is essential. Large-scale surveys conducted by Evans & Green (2007), which revisited Hyland's (1997) work on why EAP is necessary for tertiary students in Hong Kong, underscored the importance of identifying the areas of need for students. Hyland's work was based on survey responses from 1,600 undergraduates at five Hong Kong universities. It asked about their attitudes toward EAP classes. According to the study, the majority of students in this study often struggled more with academic speaking and writing (i.e., productive skills) than academic listening and reading (i.e., receptive skills). The findings of this study are consistent with the many calls in the Asian EFL context for more instruction targeting productive skills. However, the findings cannot be generalized without scrutinizing the target population. Since the needs of students vary according to their background and experience with English education, the KEY

program seeks to identify students' needs by asking the reasons and motivations for learning English.

2.2. Developing the KEY Program

Past studies suggest academic success of L2 learners depends on their skill integration abilities. They need to apply integrated language skills (i.e., reading, listening, speaking and writing) to incorporate ideas from other sources (Leiki & Carson, 1994; 1997; Grabe & Zhang, 2013). However, L2 learners often face challenges in synthesizing information from other sources in academic writing (e.g., Cumming, 2013; Cumming, Lai, & Cho, 2016; Rea-Dickins, Kiely, & Yu, 2007). For example, students are expected to write a report based on what they read (reading-to-write), take notes during a lecture (listening-to-write), or present their work by summarizing the main points of the lecture (reading-to-speak) in English. All these integrated tasks involve high cognitive demands since they require the combination of a different set of skills (i.e., reading, listening, speaking, and writing). Ideally, EAP courses should provide a wide range of topics (e.g., vocabulary and genre) and tasks that stimulate the knowledge that is typically required of them to perform well in their academic studies.

Before the launch of the program, monthly face-to-face meetings were held for all team members of the KEY program. The meetings aim to gain a common understanding of the integrated skill-based course. "Active Learning" and "Critical Thinking" were determined as key elements of teaching in the program. Active Learning is a method of encouraging students to engage in class materials such as discussions, problem-solving tasks, case studies, role-plays, peer reviews, and flipped learning. According to previous studies, Active Learning promotes higher-order thinking as shown in Bloom's Taxonomy (e.g., analysis, synthesis, evaluation, and reflection) (Tabrizi & Rideout, 2017). One of the crucial components of Active Learning is Critical Thinking, which is also known as Critical Thinking Motivator (Faust & Paulson, 1998). In the KEY program, a wide range of topics that promote problem-solving (i.e., critical thinking) were selected that are meaningful to the program participants.

The KEY program also offers individual support both in and outside of the classroom hours so the program can keep the participants motivated to learn. The need for students to have substantial and extensive support is well established by previous studies (Cohen,

2003: Gardner, 1998). These studies highlight that language aptitudes, motivations, and learning styles affect L2 learners. It is essential to provide English education tailored to students' needs to keep their motivations high. Therefore, our textbooks and other materials were chosen by taking into account learners' motivations (e.g., topics that fulfill learners' needs) and language proficiency levels.

Although not all SGHSS students go abroad to study, all students in SGHSS are expected to acquire fundamental academic English skills (e.g., reading-to-write, listening-to-speak, reading-to-discuss) in these EAP courses before they complete the second year of their university life. In 2022, the following English courses (as shown in Table 1) were offered to first-year students in the KEY program.

Table 1. KEY Program (2022)

First-semester	Listening & Speaking I (L&S I) (twice/week)	Reading & Writing I (R&W I) (twice/week)	Reading & Discussion I (R&D I) (once/week)
Second-semester	Listening & Speaking II (L&S I) (twice/week)	Reading & Writing II (R&W II) (twice/week)	Reading & Discussion II (R&D II) (once/week)
			IELTS (once/week)

Some of the core classes of the newly introduced KEY program meet twice a week for 90 minutes to increase their L2 learning opportunities. We divided classes (n=101) into three cohorts, each of which comprised mixed levels of proficiency. In the second semester, students were divided into one advanced and two standard groups according to their English language proficiency. Proficiency was based on TOEFL ITP scores; students took the TOEFL ITP in the middle of the first semester. In addition, all students are expected to take the IELTS Academic Practice Test at the end of the second semester of the first year. The KEY program is unconventional in many ways from traditional General English courses offered in other faculties.

Since this is the first attempt to implement such an intensive integrated skill-based academic English course for all first-year students in the SGHSS, it is essential to gain insights into the program's outcomes. If it is successful, it might serve as a role model for building Academic English competence among students in other faculties.

3. Research Questions

This study aimed to evaluate the extent to which the KEY program fulfilled the needs of the SGHSS students in the first semester of their university studies. The following research questions (RQs) were posed in this study:

(1) To what extent did students' self-perception of their language skills change after completing the first semester of the KEY program?

(2) Which classes, materials, and activities in the KEY program did students find helpful in improving their English skills?

4. Methodology

4.1. Data Collection

The current study combined quantitative and qualitative analyses to deepen our understanding of the students' perceptions of the KEY program. The pre- and post-program surveys had both open and closed questions: 5-point Likert Scale items provided data for the quantitative analysis and the responses to opened questions for the qualitative data. The qualitative data complemented quantitative data and gave further insight into the students' needs and satisfaction with the program.

4.2. Sampling

The participants in this study consisted of 101 first-year university students (female, N = 70; male, N = 31) who were enrolled in the SGHSS in April 2022. Due to ethical concerns, we distributed a letter explaining that the survey responses would only be used for research purposes. The letter also stated that students' responses would not affect their academic grades. They responded to the survey only when they had agreed to participate in the study. In the pre-survey, we collected 101 responses and in the post-survey, we collected 84 responses.

4.3. Procedures

In this study, we adopted questionnaires from Evans and Green (2007) on difficulties in academic skills. The questionnaires in our study were composed of 58 items for the pre-program survey and 54 items for the post-program survey. The first four items and their sub-questions were related to students' international experiences, including their

previous experience of English education outside of school curriculums, as well as their English proficiency test scores. Some parts of the survey items were the same for both pre- and post-program surveys, and they were mainly composed of three sections: (1) reasons for studying English (motivations), (2) English skills which they find (more/less) difficult in academic settings (needs), and (3) students' satisfaction with the KEY program (only in the post-survey).

Two online surveys (pre and post) were conducted by having participants scan a QR Code; the pre-program survey was conducted during the program orientation in April, and the post-program survey was given during the last day of classes in August. The questionnaires were designed to deepen our understanding of students' needs and perceptions of the KEY program. Responses were mostly recorded using a 5-point Likert Scale ranging from 1 (strongly disagree) to 5 (strongly agree); participants also had an opportunity to fill out some open-ended questions to elaborate on the reasons for their ratings.

Table 2. Students' Self-perception of Language Skills (Questionnaire Items)

Questions	Areas of skill
Q17. Understanding main ideas of lectures	Listening
Q18. Understanding questions	
Q19. Identifying different views and ideas	
Q20. Taking notes	
Q21. Following a discussion and understanding classmates	Speaking
Q22. Speaking from notes	
Q23. Asking questions	
Q24. Answering questions	
Q25. Participating actively in discussions	
Q26. Presenting information/ideas	Reading
Q27. Reading quickly to get overall meanings (Skimming)	
Q28. Reading quickly to find information (Scanning)	
Q29. Understanding the organization of a text	
Q30. Identifying supporting ideas and examples	
Q31. Understanding meanings of difficult words	
Q32. Using own words in notetaking	
Q33. Planning and organizing the structure of essays	Writing
Q34. Expressing ideas clearly and logically	
Q35. Synthesizing information and ideas by summarizing and paraphrasing	
Q36. Writing coherent paragraphs by linking sentences smoothly	
Q37. Using appropriate academic styles by referencing sources	
Q38. Proofreading a written essay	

4.4. Analysis

To address RQ1, data analysis was carried out using SPSS (ver. 25). Descriptive statistics summarized the means and standard deviations of students' responses according to the 5-point Likert Scale. Wilcoxon matched-pairs signed-rank tests were conducted to compare the means of pre-program and post-program survey responses in four L2 skills (i.e., listening, speaking, reading, and writing) to identify which of these skills the students had gained confidence in. Then we calculated the mean for each category. Another Wilcoxon matched-pairs signed-rank test was performed to determine whether the differences in responses to each item were significant between the pre-program and post-program surveys.

To answer RQ2, first, the data analysis was carried out to report the percentages of their item responses on class satisfaction in the 5-point Likert Scale, means, and standard deviations of the survey. Next, Cronbach's alpha was also reported for reliability. Then, the AMOS (ver. 27), Structural Equation Modelling (SEM), was applied to investigate causal relationships among instruction materials, class activities, and CF support for R&W I, L&S I, and R&D I classes. Finally, the open-ended responses were analyzed to gain insights into students' perspectives of the KEY program for qualitative analysis They were coded until the major themes of their responses emerged. Their comments were then compared to illustrate their differences according to their language proficiency.

5. Findings

5.1. Students' Language Proficiency Levels

Assessing and understanding students' current level of English has some practicality issues. Ideally, we wanted to assess their English academic skills in all four areas (i.e., listening, reading, writing, and speaking); however, the cost and time required to administer English proficiency tests for all 101 students before and after the semester at SGHSS were prohibitive. The SGHSS administers the TOEFL ITP test. It was department-funded, and all first-year students took it in mid-semester. Table 2 refers to the scores of the TOEFL ITP test conducted on June 28, 2022, whereby the score for Section I averaged 51.2 (SD = 4.4); the score for Section II averaged 48.2 (SD = 5.0); the score for Section III averaged 51.2 (SD = 4.0) and the Total Score averaged 502.0 (SD = 36.1). Students were divided into two groups using their scores with a total score of 500 as the cut-off as shown in Table 3,

in which 42 students were classified as an upper group (test score ≧ 500) and 57 as a lower group (test score < 500). This study used these data for further analysis.

Table 3. TOEFL ITP Scores (June 28, 2022)

Items	Minimum	Maximum	Median	Mean	SD
Section I (Listening)	45	64	50	51.2	4.4
Section II (Grammar & Structure)	35	68	47	48.2	5.0
Section III (Reading)	39	63	51	51.2	4.0
Total Score	427	630	497	502.0	36.1

N=99.

Table 4. TOEFL ITP Score-Based Groups

Group	Number	Percentage
Upper	42	42.4
Lower	57	57.6

N=99. Note: Upper (TOEFL ≧ 500) Lower (TOEFL <500)

5.2. Students' Self-Rating of Language Difficulties (Pre- and Post-Program Surveys)

We divided 22 survey items (from Q17 to Q38) into four categories based on four skills (as presented in Table 4), Listening (4 items, from Q17 to Q21), Speaking (5 items, from Q22 to Q26), Reading (7 items, from Q27 to Q32), and Writing (6 items, from Q33 to Q38). Then we calculated the mean for each category. The results (as shown in Table 5) indicated that the KEY program had contributed to the increase in students' self-perception ratings in all four English skills (Listening, Speaking, Reading, and Writing) over the semester. The Wilcoxon signed-rank test was used to determine if the changes in the ratings were significant. The results of the test (as shown in Table 6) indicated that the differences in students' self-perceptions between the pre-program and the post-program survey were statistically significant: Listening ($z = -3.225$, $p < .005$), Speaking ($z = -4.286$, $p < .001$), Reading ($z = -2.632$, $p < .01$), and Writing ($z = -3.233$, $p < .005$).

Table 5. Descriptive Data of Pre- and Post-Surveys

Measure		Mean	Median	SD
Pair 1	Pre-Listening	2.88	2.80	0.80
	Post-Listening	3.16	3.30	0.76
Pair 2	Pre-Speaking	2.50	2.60	0.93
	Post-Speaking	2.90	2.90	0.84
Pair 3	Pre-Reading	3.06	3.17	0.78
	Post-Reading	3.26	3.17	0.72
Pair 4	Pre-Writing	2.44	2.33	0.82
	Post-Writing	2.76	2.75	0.72

N=84.

Table 6. Wilcoxon Matched-Pairs Signed-Rank Tests (Four Skills)

Pre-survey		Post-survey	W	z	p	Effect Size (r)
Pre-Listening	–	Post-Listening	766.00	−3.225	0.001	−0.249
Pre-Speaking	–	Post-Speaking	683.00	−4.286	0.000	−0.331
Pre-Reading	–	Post-Reading	873.50	−2.632	0.008	−0.203
Pre-Writing	–	Post-Writing	866.50	−3.233	0.001	−0.249

Another Wilcoxon signed-rank test was performed to further explore which items in the questionnaire had greater impacts on the changes in students' self-perception of their English skills in the pre- and post-program surveys. Table 7 presents the results of the test showing that in the post-program survey, students' ratings of their perceptions of their English skills had increased in all 22 items compared to the pre-program survey. This suggests that the respondents had gained more confidence after studying in the KEY program for one semester.

Table 7. Wilcoxon Matched-Pairs Signed-Rank Test (Individual Items)

Pre-survey		Post-survey	W	z	p	Effect Size (r)
Pre Q17. Understanding main ideas of lectures	–	Post Q17. Understanding main ideas of lectures	439.50	−1.617	0.106	−0.125
Pre Q18. Understanding questions	–	Post Q18. Understanding questions	441.00	−3.012	0.003	−0.232
Pre Q19. Identifying different views and ideas	–	Post Q19. Identifying different views and ideas	298.00	−2.808	0.005	−0.217

Pre-survey		Post-survey	W	z	p	Effect Size (r)
Pre Q20. Taking notes	–	Post Q20. Taking notes	436.00	−2.236	0.025	−0.173
Pre Q21. Following a discussion and understanding classmates	–	Post Q21. Following a discussion and understanding classmates	362.00	−2.606	0.009	−0.201
Pre Q22. Speaking from notes	–	Post Q22. Speaking from notes	540.00	−2.368	0.018	−0.183
Pre Q23. Asking questions	–	Post Q23. Asking questions	232.50	−3.549	0.000	−0.274
Pre Q24. Answering questions	–	Post Q24. Answering questions	334.50	−2.000	0.045	−0.154
Pre Q25. Participating actively in discussions	–	Post Q25. Participating actively in discussions	215.50	−4.906	0.000	−0.379
Pre Q26. Presenting information/ideas	–	Post Q26. Presenting information/ideas	192.00	−3.940	0.000	−0.304
Pre Q27. Reading quickly to get overall meanings (Skimming)	–	Post Q27. Reading quickly to get overall meanings (Skimming)	386.00	−1.587	0.112	−0.122
Pre Q28. Reading quickly to find information (Scanning)	–	Post Q28. Reading quickly to find information (Scanning)	386.50	−2.013	0.044	−0.155
Pre Q29. Understanding the organization of a text	–	Post Q29. Understanding the organization of a text	439.50	−2.221	0.026	−0.171
Pre Q30. Identifying supporting ideas and examples	–	Post Q30. Identifying supporting ideas and examples	446.00	−2.692	0.007	−0.208
Pre Q31. Understanding meanings of difficult words	–	Post Q31. Understanding meanings of difficult words	292.00	−1.899	0.058	−0.147
Pre Q32. Using own words in notetaking	–	Post Q32. Using own words in notetaking	723.00	−.421	0.674	−0.032
Pre Q33. Planning and organizing the structure of essays	–	Post Q33. Planning and organizing the structure of essays	716.00	−1.554	0.120	−0.120
Pre Q34. Expressing ideas clearly and logically	–	Post Q34. Expressing ideas clearly and logically	454.50	−2.267	0.023	−0.175
Pre Q35. Synthesizing information and ideas by summarizing and paraphrasing	–	Post Q35. Synthesizing information and ideas by summarizing and paraphrasing	247.50	−3.065	0.002	−0.236
Pre Q36. Writing coherent paragraphs by linking sentences smoothly	–	Post Q36. Writing coherent paragraphs by linking sentences smoothly	386.50	−2.890	0.004	−0.223
Pre Q37. Using appropriate academic styles by referring sources	–	Post Q37. Using appropriate academic styles by referring sources	307.50	−3.337	0.001	−0.257
Pre Q38. Proofreading a written essay	–	Post Q38. Proofreading a written essay	281.00	−2.614	0.009	−0.202

Furthermore, 17 out of 22 questionnaires responses showed statistically significant differences between students' self-perceptions in the pre- and post-program survey. These items relate to three types of skills that are essential in academic settings: basic L2 learning skills, class participation and communication skills, and writing skills. Q18. Understanding questions ($z = -3.012$, $p < .005$) is related to basic L2 learning skills; Q23. Asking questions ($z = -3.549$, $p < .001$), Q25. Participating actively in discussions ($z = -4.906$, $p < .001$), and Q26. Presenting information/ideas ($z = -3.940$, $p < .001$) are related to class participation and communication skills; Q35. Synthesizing information and ideas by summarizing and paraphrasing ($z = -3.065$, $p < .005$), Q36. Writing coherent paragraphs by linking sentences smoothly ($z = -2.890$, $p < .005$) and Q37. Using appropriate academic styles by referencing sources ($z = -3.337$, $p < .005$) are related to writing skills.

In contrast, there were no statistically significant differences in the self-perceptions of students for the other five items that require more advanced and comprehensive skills: Q17. Understanding main ideas of lectures ($z = -1.617$, $p > 0.1$), Q27. Reading quickly to get overall meanings (Skimming) ($z = -1.587$, $p > .1$), Q31. Understanding the meanings of difficult words ($z = 0.058$, $p > .05$), Q32. Using own words in notetaking ($z = -.421$, $p > .5$), and Q33. Planning and organizing the structure of essays ($z = -1.554$, $p > .1$).

5.3. Students' Views of the KEY Program: Quantitative Analysis

RQ2 focused on investigating students' perceptions of KEY program classes, class activities, instructional materials, and help with assignments from CFs on a 5-point Likert Scale ranging from 1-5, whereby 1 indicated "definitely not helpful," and 5 indicated "definitely helpful".

Overall, students ($N = 84$) found all activities and assignments helpful (Mean = 4.51), particularly class activities, instruction materials, and assignments that involved communication and discussion as well as one-on-one input involving essays and advice from CFs.

Table 8. Did You Find These Classes Helpful?

Class activity/Assignment	5 Definitely	4 Probably	3 Possibly	2 Probably not	1 Definitely not	Mean	SD
R&W I	31.7%	46.5%	5%	0%	0%	4.32	0.584
L&S I	48.5%	31.7%	3%	0%	0%	4.46	0.630
R&D I	35.6%	41.6%	5%	1%	0%	4.55	0.568
R&W I (Instructional Materials)	43.6%	35.6%	3%	1%	0%	4.49	0.611
L&S I (Instructional Materials)	45.5%	32.7%	5%	0%	0%	4.35	0.649
R&D I (Instructional Materials)	38.6%	31.7%	10.9%	2%	0%	4.29	0.785
Oral activities	61.4%	19.8%	1%	1%	0%	4.70	0.555
Report	44.6%	30.7%	5.9%	2%	0%	4.42	0.732
Discussions	58.4%	18.8%	4%	2%	0%	4.61	0.695
Essays	64.4%	16.8%	1%	1%	0%	4.74	0.540
CFs	67.3%	12.9%	3%	0%	0%	4.77	0.499

N=84.

Construct reliability was assessed using Cronbach's Alpha. The result revealed that the KEY Satisfaction survey with 11 items was found reliable (x = .838). The table below shows correlation coefficient between these classes, instruction materials, activities, and CF support.

Table 9. Correlations

	1	2	3	4	5	6	7	8	9	10	11
1. R&W I	1										
2. R&W I Materials	0.44	1									
3. L&S I	0.62	0.32	1								
4. L&S I Materials	0.26	0.56	0.40	1							
5. R&D I	0.40	0.31	0.39	0.29	1						
6. R&D I Materials	0.32	0.46	0.34	0.45	0.39	1					
7. Oral activities	0.07	0.15	-0.01	0.14	0.35	0.22	1				
8. Report	0.41	0.43	0.25	0.37	0.45	0.29	0.45	1			
9. Discussion	0.07	0.34	0.21	0.31	0.33	0.40	0.47	0.51	1		
10. Essays	0.19	0.22	0.08	0.28	0.39	0.32	0.5	0.52	0.33	1	
11. CFs	0.29	0.18	0.27	0.09	0.20	0.32	0.18	0.26	0.12	0.22	1

The findings revealed their survey responses were not so strongly correlated as we initially thought. Using the same data set, Structural Equation Modelling (SEM) was applied to investigate how instruction materials, class activities, and CF Support impacted class ratings with the IBM software AMOS (ver. 27). The schematic representation was drawn based on our hypothesis that the overall rating of each class (i.e., whether students

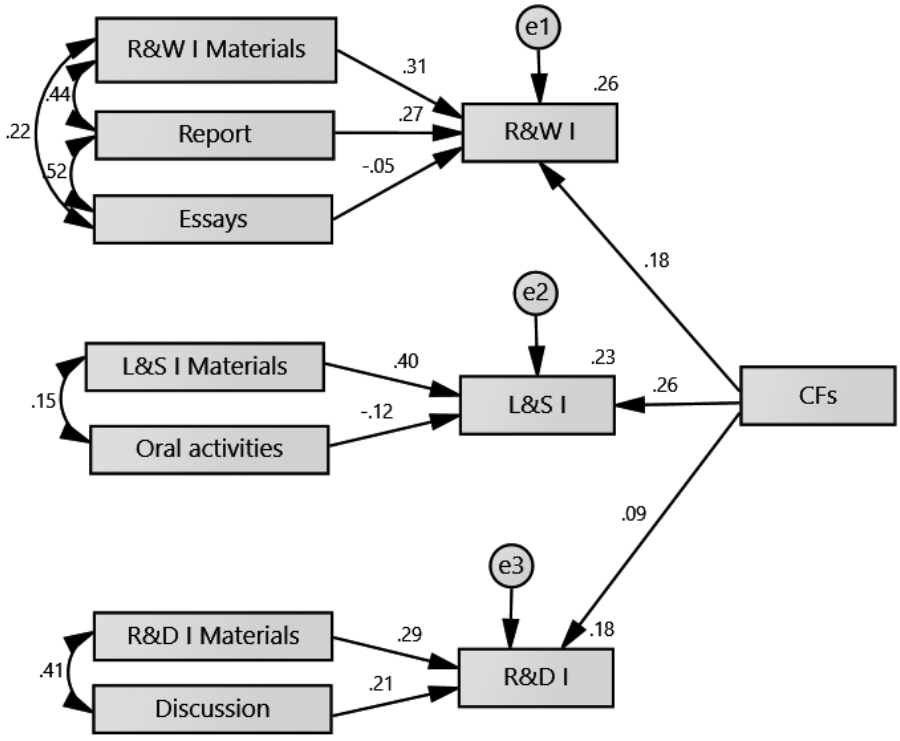
found the class helpful on the 5 point-Likert Scale) depended on how they felt about three factors, namely instructional materials, class activities, and CF support. Figure 1 shows the results of the linear regression analysis with standardized estimates. The number in the diagram is an R square (i.e., coefficient), which represents the proportion of the variance for a dependent variable (i.e., R&W I, L&S I, R&D I) explained by an independent variable (i.e., Instruction Materials, Essays, Reports, Oral Activities, Discussions, CF Support.) For goodness of fit, Chi-square values are used for testing the model as shown in Table 10. It was used to evaluate whether the model differs significantly from the data (Kline, 2016). The hypothetical model provided an acceptable model of fit (CMIN = 206.566 and Chi square/df = 40, $p \leq .05$).

Table 10. Model Fit Summary

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	37	206.566	40	.000	5.164
Saturated model	77	.000	0		
Independence model	22	335.793	55	.000	6.105

N = 84.

Figure 1. The Model for Causal Relationship with Standardized Estimates



The standardized estimate shows correlations between the indicator variables in boxes (Figure 1). For example, the R&W I instructional materials explained 31% of the variance associated with the R&W I class rating. Likewise, the reports accounted for 21 %, and essays for -5% of the R&W I class ratings. We also added an error (e1) to include any unknown factors that might be influencing how students rated the R&W I. In the case of L&S I, the instructional materials explained 40% while oral activities negatively correlated at 12%. The instructional materials and discussions were associated with the R&D I at 29% and 21 % respectively. The factor, "CF support" had the closest correlation with the L&S I class at 26%, followed by the R&W I at 18% and the R&D I at 9%. The findings may have been reflected the hours CFs supported KEY program activities. Overall, instructional materials (i.e., textbooks and handout materials) were more closely related to the ratings of all three classes.

Not all path coefficients were significant. The Covariances table below represents an estimate of the covariance between factors and measurement errors for the negatively worded items. The Covariances between L&S I Materials and Oral Activities (p = .178), and between R&W I Materials and Essays (p = .05) were not found significant at the .05 level.

Table 11. Covariance: Default Model

			Estimate	S.E.	C.R.	P	Label
R&D I Materials	<-->	Discussion	.219	.064	3.436	***	
L&S I Materials	<-->	Oral Activities	.050	.037	1.345	.178	
Report	<-->	Essays	.204	.048	4.223	***	
R&W I Materials	<-->	Report	.199	.055	3.657	***	
R&W I Materials	<-->	Essays	.074	.038	1.958	.050	

There are some possible reasons why these covariances were not significant. In L&S I, students were awarded points (10% of their grades) for participating in oral activities outside their classroom hours to meet with CFs individually and practice the IELTS Speaking Task 2. Although students found the activities meaningful and valuable, they may not have seen a strong link between the class instruction and the activities. Likewise, students in R&W I were also awarded points (10% of their grades) for writing up their essays on selected topics to get some feedback from CFs. Although they found the activities meaningful and valuable, again, they may not have seen a strong link between the class instruction and the activities.

Overall, the result from SEM gave a glimpse of the relationships among the classes, materials, activities, and support provided by CFs. However, the quantitative analyses do not provide a full picture of students' perceptions of the classes. It is thus important to carry out qualitative analyses.

5.4. Students' Views of the KEY Program: Qualitative Analysis

The questionnaires also included some open-ended questions regarding class activities. A grounded approach was used to classify the comments for each class until major themes emerged using the software NVivo (Ver. 12.) There was a total of 214 comments: 76 comments for R&W I, 76 comments for L&S I and 62 comments for R&D I. Four major themes emerged from the open coding. They were a) teaching quality, b) content, c) suitability, and d) accomplishment.

Table 12. Themes Emerged from Students' Comments

Major Themes	Examples	R&W I (76)	L&S I (76)	R&D I (62)
Teaching Quality	Teacher's personality	11	11	3
	Easy-to-understand instruction	14	20	10
	Activities (e.g., discussion, presentation)	3	13	8
	Atmosphere	0	1	0
Content	Enjoyable	0	7	5
	Easy to understand	5	1	0
	Comprehensible to some degree	8	4	8
	Initially difficult, but eventually understand	8	9	7
	Difficult to understand	7	3	13
	Textbook	0	0	2
Suitability	Level	8	2	0
	Pace	4	2	4
	Amount of work	1	2	0
	Support (peer support, CF support)	0	2	2
Accomplishment	Gained essential knowledge and skills	12	19	15
	Struggled due to the lack of L2 skills	15	7	10
	Not enough confidence to carry out the tasks	5	3	2
Total Number of Coding References		101	106	89

The highest number of coding reference was a) teaching quality (N = 94), followed by d) accomplishment (N = 88), c) content (N =87) and Suitability (N = 8). In other words, the findings suggest students cared slightly more about teaching quality and accomplishment than the content they learned. Regarding students' comments on accomplishment, we

witnessed mixed messages being displayed. The following section illustrates some examples by comparisons.

5.5. Variation in Students' Views by Proficiency Level

Students' comments in the open response segment of surveys provided students' views towards class instructions and activities and clarified there were some gaps between higher (test score ≧ 500) and lower proficiency (test score < 500) students. Several students in the higher-proficiency group commented that they were mostly satisfied with the contents, levels, and pace of instruction in the KEY courses. Examples of such comments include:

R&W I (higher)

- *The instructor carefully taught me the vocabulary necessary for writing essays, as well as how essays should be organized.*
- *The instructions were easy to understand, and the pace of the classes was at an appropriate speed.*
- *I realized that the more I read and wrote in class, the more knowledge I gained in various aspects, such as the organization of an essay and grammar.*

L&S I (higher)

- *The instructions and explanations were easy to understand.*
- *The instructor gave easy-to-understand explanations.*
- *I was able to gain applicable skills in listening and presentation.*

R&D I (higher)

- *I always had discussions with my classmates, and I was able to exchange various opinions, which broadened my horizons.*
- *I didn't have any trouble understanding the class itself because it was conducted in relatively easy English.*
- *I think that my understanding has progressed by obtaining various information through discussions.*

Students in the lower proficiency group were also satisfied with the course contents. However, their comments showed a lack of confidence. Some comments below showed

how they lacked the knowledge and skills required to perform well in class. Examples of such comments include:

R&W I (lower)

- *I felt my lack of basic knowledge, and missed, though not all, class instructions. It is my fault.*
- *I lack basic knowledge as a student.*
- *The instructors taught me something which I could not follow, and finally I came to understand it.*

L&S I (lower)

- *I sometimes could not understand the intentions of the instructors.*
- *I sometimes could not follow the instructor's English, but I grasped their intentions.*
- *I did not understand the instructors' English perfectly, but I did understand what to do next.*

R&D I (lower)

- *Some words and sentences were difficult to understand.*
- *It was quite difficult to have an impromptu discussion.*
- *Because my listening skills are weak, there were times when I couldn't catch what was being said.*

The result indicates a gap between the knowledge and skills gained in high-school English classes, primarily General English, and Academic English skills required in higher education. Such findings suggest some students probably needed more language scaffolding in a classroom where English is used as a medium of instruction to perform the academic tasks required in higher education. The result might imply the need for streaming students according to their proficiency levels to put them into different classes in the future.

6. Discussion

The KEY program was originally developed to increase integrated skills (i.e., reading and writing, listening, and speaking, reading and discussion) under the premise that the

academic success of L2 learners depends on their skill integration abilities (e.g., Leiki & Carson, 1994; 1997; Grabe & Zhang, 2013). Hyland (1997) and other scholars called for identifying the areas of students' needs as the key to developing a successful EAP curriculum. Past studies (e.g., Evans & Green, 2007) also suggested that L2 students put more demands on developing productive skills than receptive skills. Therefore, this study focused on surveying students' perceptions of confidence in acquiring specific language skills after their first semester experience of the KEY program.

6.1. RQ1. To What Extent Did Students' Self-perception of Language Skills Change after Completing the First Semester of the KEY Program?

Results from the descriptive data as well as the Wilcoxon signed-rank tests suggest that the KEY program had positively affected students' self-perception ratings in all four English skills (listening, speaking, reading, and writing) over the semester. The analyses of survey responses between pre-and post-questionnaires demonstrated the areas in which students gained significant confidence over the semester of the KEY program (e.g., understanding lectures, asking questions, notetaking, participating actively in discussions, understanding organization of a text, and expressing ideas clearly and logically).

The survey results also helped identify the areas of specific Academic English skills that students need to develop in the KEY program. This study did not conduct post-survey interviews and, thus, we are limited in what we can surmise about our students' psyche. However, the item responses in the survey that did not gain any significant improvement in the post-survey can provide valuable feedback in improving the program as below. We discuss point by point the items in which students did not show improvement in the pre- and post-survey to make improvement.

6.1.1. Listening: Q17 -*Understanding main ideas of lectures*

The materials used in the listening course were authentic materials. Authentic materials are materials that were not necessarily designed for language learners in a classroom. The materials consisted of shortened *TED Talks*, news clips, and podcasts. It is possible that the students were unfamiliar with these kinds of materials in a language class where the main idea of the passage might not be as straightforward as more traditional materials. The instructors should take careful consideration when transitioning the students from more traditional materials designed for language classes to more authentic

materials. For example, the instructors can make the level or topic more appropriate for students. Authentic materials can sometimes be more engaging and motivating, but they can also be more difficult because they were not designed with the learner in mind.

6.1.2. Reading: Q27 -*Reading quickly to get overall meaning (Skimming)*

In the reading course, there was great emphasis put on increasing each student's reading speed, fluency, and comprehension. Despite that emphasis, the balance between those three skills can be lost when there is a large gap in proficiency among the students in a class. Grouping the students according to their reading level would better prepare the instructors on choosing appropriate reading texts.

6.1.3. Q31 -*Understanding meanings of difficult words*

Without exception, all participants in the KEY program need to increase their lexicon. Most of the instructors have made efforts to assess students' background knowledge about new words, and from the students' responses to this item in the survey, they are keenly aware of their low vocabulary. However, the instructors could share more strategies with the students to assist them in understanding difficult words in context.

6.1.4. Q32 -*Using own words in notetaking*

No significant improvement reflects the low vocabulary level of this group of students. Nevertheless, the R & W I class spent a lot of time on paraphrasing and the importance of paraphrasing.

6.1.5. Writing: Q33 -*Planning and organizing the structures of essays*

Not significant improvement could also reflect mixed proficiency levels within the same course. Some of the lower-level students would benefit greatly from more time spent on outlining and planning, whereas some of the higher-level students might not need to spend as much time on this practice.

6.2. RQ2. Which Class, Materials, and Activities in the KEY Program Did Students Find Helpful in Improving their English Skills?

Descriptive data of the survey results found all class activities and assignments helpful (Mean = 4.51), as well as the assistance from the Coaching Fellows. Structural Equation Modelling (SEM) helped understand the relationships of the students' ratings to overall

class satisfaction. The results revealed that instruction materials were closely related to their satisfaction rating. However, the qualitative analysis of students' comments revealed various reasons why they highly rated their classes. Here, we discuss the main themes: a) teaching quality, b) content, c) suitability, and d) accomplishment point by point.

6.2.1. Theme 1: *Teaching Quality*

Students often described the classes in the KEY program as fun and enjoyable. It is not only because the instructors were all efficient in communicating with the students, but they also created a positive learning environment for them. Even when students struggled in class due to the lack of L2 skills, they understood the instruction thanks to the instructor's personality. They also appreciated that the KEY classes offer more opportunities to discuss, present and interact with each other in English in class.

Excerpt

My teacher made a lot of gestures, and he gave us an enjoyable class regardless of whether we were good at English or not. There were times when we did a little complicated activity, but he said, "You don't have to be so nervous," and the atmosphere in the classroom softened, so I think everyone was able to relax and concentrate on the lesson. (L&S I)

6.2.2. Theme 2: *Content*

Most students praised instructors for giving clear and easy-to-understand instructions in class. Even those students who said they struggled to understand the content, many of them said they eventually understood or understood to a certain extent. Since it was a mixed group of students with different levels of proficiency, there were different *needs* displayed in the comments. Some thought the topics were easy, while others claimed they were hard to understand. Instructors tried to adjust the difficulty level so that students could understand the instruction. One of the techniques used in class was pair activities or peer reviews. It helped students to help each other and check whether they misunderstood their tasks. Some appreciated the content from a previous week reviewed before moving on to the new unit. Students also liked the routines of activities so that they knew what to expect rather than figuring out each time what to do next in class. Others mentioned that they were able to keep up with their homework as they received regular notifications on their Learning Management System page.

6.2.3. Theme 3: *Suitability*

Higher-level students tend to comment that they felt the class was at their level and the pace was right. Those students also claimed that the amount of work was sufficient.

6.2.4. Theme 4: *Accomplishment*

There were many positive responses about the skills and knowledge gained in all three classes. Some students with lower proficiency, however, did not increase their confidence over the semester to comprehend class instruction. Many of these students do not believe they can carry out the tasks for their lack of L2 skills. The KEY program endorsed teaching academic topics that required more complex processing of cognitive learning as the instructors tried to implement active learning and critical thinking into their activities. Since students with lower proficiency did not have the basic L2 skills to comprehend instructions, they were not able to fully engage in the topics and activities. These students might need more language scaffolding, especially considering it was probably their first-time taking EAP classes in English.

7. Limitations and Future Implications

This study was limited only to questionnaires to measure students' confidence and self-perception of language skills after completing the first semester of the KEY Program. Nevertheless, the methods used in our study helped us gain insights that will lead to improvements in the KEY program.

There are still many new components to consider when creating a new English program and trying to make it as effective as possible. In this study, we tried to gain some insight into the students' perceptions of the materials used in the KEY program and their overall cognizance of the program as it pertains to their improvement. In general, the students found the KEY program and its materials to be beneficial in their development of the various skills needed to be an effective user of the English language. However, there are some aspects that the instructors could take into consideration to continually improve the program.

For instance, when choosing a text for both listening or reading the instructor could consider the wide range of proficiency levels and choose texts that better suit the range.

Unfortunately, it is rather complicated to assess students' proficiency levels prior to entering university. Mixed proficiency levels are not unique to this program or to university students in Japan. So, teachers need to find better ways to keep higher-proficiency students motivated while not entirely demotivating students with low proficiency. Another suggestion would be to consider reducing the number of participants in a course to twenty-five students. The program's 101 participants were divided up into three groups, but creating a fourth group could help create an environment that would allow instructors to better assist lower-proficiency students and challenge the higher-proficiency students.

8. Conclusion

In conclusion, this exploratory study contributes to our understanding of the attitudes and experiences of first-year SGHSS students in the KEY program. By soliciting students' opinions, administrators and instructors will be able to better adapt the curriculum to best suit the needs of students in this context moving forward. Thus far, we feel that the KEY program is off to a promising start. The findings of this study suggest that the KEY program has already had a positive effect on students' academic skills and self-confidence. Moreover, the results also indicate that the students were generally satisfied with the instruction they received. Immediate directions for future analysis emerge from the above discussions of limitations and implications. First, in an effort to dig deeper into how we can improve, it would be useful to interview students at the end of the semester. Moreover, while the importance of self-confidence and self-perceptions cannot be understated, the researchers recognize that this is only the first step. Eventually, students will need to be assessed on performance and their actual English skills. To this end, the IELTS Academic Practice Test that students will take at the end of their first year will help us more adequately measure students' English proficiency in each of the four skills (i.e., listening, speaking, reading, and writing). Ideally, our program can develop a system whereby such a test would be administered at the start and again at the end of students' first year of study. Finally, with a year under our belt, instructors in the KEY program are better equipped to deal with the pedagogical challenges of teaching in this context moving forward. Through a process of trial and error, and by working together, our instructors are dedicated to helping our students ultimately reach an advanced level of English proficiency.

Acknowledgement:

We would like to show our gratitude to the dedicated administrative staff [E. Miyawaki, A. Tokushige] for supporting the program to success.

References

Berman, R., & Cheng, L. (2001). English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. *Canadian Journal of Applied Linguistics, 4* (1), 25-40.

Chang, S. Y. (2021). English Medium Instruction, English-enhanced instruction, or English without instruction: The affordances and constraints of linguistically responsive practices in the higher education classroom. *TESOL Quarterly, 55* (4), 1114-1135.

Cohen, A. D. (2003). The learners' side of foreign language learning: Where do styles, strategies, and. tasks meet? IRAL 41, 279-291.

Conway, J. (2020). The effects of a year-long English for academic purposes program: The SCAS program in 2018-19. *Journal of Center for Language Studies Nagasaki University, 8*, 19-35.

Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly, 10* (1), 1-8.

Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes, 23*, 47-58.

Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes, 6* (1), 3-17.

Faust, J. L., & Paulson, D. R. (1998). Active learning in the college classroom. *Journal on Excellence in College Teaching, 9* (2), 3-24.

Grabe, W., & Zhang, C. (2013). Reading and writing together: A critical component of English for academic purposes teaching and learning. *TESOL Journal, 4* (1), 9-24.

Hosogoshi, K., & Takahashi, S. (2015). The use of integrated listening, reading, speaking and writing. tasks on students' productive skills in a university EAP course. *Professional and Academic English, 45*, 22-30.

Hyland, K. (1997). Is EAP necessary? A survey of Hong Kong undergraduates. *Asian Journal of English Language Teaching, 7* (2), 77-99.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York: The Guilford Press.

Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly, 28* (1), 81-101.

Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly, 31* (1), 39-69

Rea-Dickins, PR, Kiely, R, Yu, G (2007). Student identity, learning and progression: the affective and academic impact of IELTS on "successful" candidates. In IELTS research reports (Vol. 7). Manchester: British Council and Canberra: IDP Education Australia.

Tabrizi, S., & Rideout, G. (2017). Active learning: Using Bloom's taxonomy to support critical pedagogy. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE), 8* (3), 3202-3209.

Appendices

Appendix A

Pre-survey:
https://docs.google.com/forms/d/e/1FAIpQLSc737y2ZyVu_5EUyUcPqTcF6MCUHwviAmiGnZwlQ5d5pVMTSw/viewform?usp=sharing

Appendix B

Post-survey:
https://docs.google.com/forms/d/e/1FAIpQLSeDkdSFho5eJ_yGk8I7Jem0hG92Y6H1Oi12_6rI5W94oWORjQ/viewform?usp=sharing