A Cross-Sectional Survey on Japanese English-Major University Students' Confidence in the TOEIC Can-Do List

Shinya Ozawa
Hiroshima Shudo University, Japan
<ozawa@shudo-u.ac.jp>

Abstract

This survey was conducted to investigate how university students gain or lose confidence in English communicative domains over four years at university. Self-assessment has been a useful instrument for measuring learners' English proficiency, and the students in this study were required to self-assess their confidence levels in TOEIC Can-Do list items. The results revealed a difference between first- and third-year students' levels of confidence in the Reading domain, but the effect size was small. In the other communicative domains, there was no significant gain in confidence over the four years. Frustratingly, there was a decline in confidence from the third to the fourth year. Even when the limitations of this survey are considered, the results clearly demonstrate that there is a need to improve the curriculum in order to develop student confidence over their four years at university.

Introduction

Following the introduction of the Common European Framework of Reference for Languages (CEFR) and CEFR-J, which was developed for use specifically in Japan, much attention has been paid to their use in language education in Japan. Among the research initiatives on the CEFR, CEFR-J, and other Can-Do lists, several studies have been conducted on the reliability of learners' self-assessments on these lists, most of which confirmed learners' reliability (e.g., Ross, 2006). However, there has not been much consensus or evidence on how a learner's self-assessment can change over time.

In this study, first- to fourth-year English-major students at a local university in Japan were asked to self-assess their confidence in English Can-Do items, and their responses were analyzed. A longitudinal study would have been preferred, but a cross-sectional approach was adopted because of time constraints. The survey was part of a series of attempts that have been made to verify the effectiveness of the curriculum in the department of English Language and

Literature at the university. The researchers hope that the results of this survey can serve as a basis for useful discussions about the future introduction of a new curriculum.

Background of the Study

Since the introduction of the current curriculum in the department, no special attempts have been made to measure its effectiveness thus far. Recognizing this problem, we began our project in 2011 to better understand the general nature of the curriculum, and to suggest ways to improve it. An outline of the curriculum and issues in the department are summarized in the following section.

Outline of the Curriculum

The current English department curriculum was implemented in 2007. It places emphasis on providing abundant English reading inputs for students. There are five compulsory English-skill courses in each semester of the first year and four in each semester of the second year, as shown in Table 1.

In the Progress in English I/II course, the same textbook is used in all the classes. Japanese teachers of English provide instruction for the first-year classes, in which students are required to read the textbook intensively. Native-English speakers teach the second-year Progress in English III/IV classes, in which students are required to deliver oral presentations in English based on reading inputs. As indicated by their titles, the other writing and grammar courses underscore the importance of reading inputs as well. The English-skills courses in the first and second years provide the foundation for more academic-oriented English courses in the third and fourth years, during which students write a graduate thesis on a topic of their choosing.

Table 1. *First- and Second-Year Compulsory English-Skill Courses*

1st year		2nd year	
1st semester	2nd semester	1st semester	2nd semester
Progress in English I	Progress in English II	Progress in English III	Progress in English IV
Reading & Writing I	**Reading** & Writing II	**Reading** & Writing III	**Reading** & Writing IV
Reading & Grammar I	**Reading** & Grammar II	**Reading** & Grammar III	**Reading** & Grammar IV
Speaking I	Speaking II	Speaking III	Speaking IV
Listening I	Listening II	-	-

One concern in the department is that there are no compulsory English-skill courses for the third- and fourth-year students. Only those students willing to continue working to improve their English skills choose to take elective courses such as Presentation and Discussion I/II and Project Work I/II. We, the teachers, believe that the students' English abilities continue to develop in their third and fourth years, since we require them to read academic textbooks in small-

enrollment seminar courses. There is no reliable evidence, however, that students improve their English skills in their last two years.

Investigation of Curriculum Effectiveness

All first-year students at the university are required to take the TOEIC Bridge test three times a year: in April as a placement test and in July and January as achievement tests. Since the proficiency levels of the English-majors are generally higher than those of students in other departments, the placement test was changed from TOEIC Bridge to TOEIC in 2011. Except for these tests, we have not implemented any other means of tracking students' development of their English skills. We commenced our project in 2011, with the purpose of investigating English-major students' English skills using three primary methods: semi-structured interviews with the students, the development of a corpus of English textbooks that we recommend our students to read, and a survey in which students self-assess their English skills by using a Can-Do list. The Can-Do list survey is the focus of this paper.

Literature Review

This section summarizes some of the narrative and systematic reviews of the existing literature on self-assessments. Various previous studies have investigated the reliability of learners' self-assessment of Can-Do items. Before discussing the survey data, it is important to review this particular line of research.

Blanche and Merino (1989) reviewed and summarized all previous studies on self-assessment in foreign language skills. Their extensive literature review identified 16 studies published between 1979 and 1986. Overall, these studies found that self-assessments of proficiency correctly reflected external criteria, such as standardized test scores. In addition, it became clear that when self-assessment items are based on concrete situations, they correlate with learners providing more accurate assessments, with better learners tending to underestimate their proficiencies. The study was a seminal work that synthesized the previous literatures; however, as they pointed out, their study remained as a kind of "prose-based" (Blanche & Merino, 1989, p. 2) subjective narrative review.

In an attempt to draw a more comprehensive and objective picture of self-assessment, Ross (1998) adopted meta-analysis as his procedure. He limited his search to studies that had focused on the correlations between self-assessment and other measures in four language skills, collecting 60 correlations in 10 studies. He found that the overall average correlation r was .63 and effect size g was 1.64; for reading, .61 and 1.56, listening, .65 and 1.71, speaking .55 and 1.33, and writing .52 and 1.23. Overall, correlations between self-assessments and other measures were quite high, with the productive skills of speaking and writing showing lower correlations than the other two skills. He claimed that one of the factors that affected the self-assessments was whether the learners had actually experienced the specific examples that they were being asked about.

Studies comparing self-assessments with other criteria, such as test scores, have also been conducted in Japan. Saida (2008), for example, adopted the DIALANG test (Online Language Diagnosis System, n.d.) to investigate the correlation between the self-assessed English proficiency levels and actual performances of 130 first-year university students in Japan. In

listening, writing, and reading skills, the rates of correspondence between the two were about 60% for each skill. These rates were lower than those found in Alderson's (2005) DIALANG pilot study, in which the rate for each skill was over 80%. Several studies that focused on the correlations between TOEIC scores and Can-Do self-assessments have also been conducted, and they showed the same tendency as above (Ito, Kawaguchi, & Ohta, 2005; Powers, Kim, & Weng, 2008; Powers, Yu, & Yan, 2013).

Generally, we can assume that self-assessment is a reliable measure for predicting learners' proficiencies. However, it also became clear that various factors, such as learners' proficiency levels and actual experiences of the Can-Do items, could also influence responses. Brantmeier, Vanderplank, and Strube (2012), for example, invented self-assessment reading question items adopted from various course materials, so that their items reflected the objectives of a Spanish language course. The participants of the study were enrolled in beginner-, intermediate-, and advanced-level Spanish language courses at a private university in the United States. They completed the self-assessment survey during class, and advanced-level students were asked to take the online DIALANG test to investigate the correspondence between their self-assessment and their actual scores. The results showed that (a) beginner learners rated their speaking skills the lowest, (b) many groups rated reading ability the highest among their language skills, and (c) the advanced students were able to judge their levels appropriately. Result (c) appears to contradict previous findings that advanced students tended to underestimate their proficiencies. As Brantmeier et al. (2012) outlined, this could be because the self-assessment items in the study were criterion-referenced and corresponded with the course objectives. In other words, the learners could easily imagine the situations where their English skills were used (Little, 2014), and were therefore able to provide more accurate answers.

There is also a need to investigate whether or not learners' self-assessments of their own proficiency change in accordance with their English language development. Léger (2009) performed a qualitative investigation of how self-assessments in speaking, conducted by university learners studying French, changed over a 12-week period. She concluded that their self-assessments changed in a positive way; their confidence in self-assessment items increased by the end of the course. Nishida's research (2012) on elementary school students is one of only a few studies to investigate how students' self-assessments develop longitudinally. One hundred and six students aged 10 and 11 participated in the study and the data were collected four times over a one-year period. It became evident that most of the pupils' self-assessments declined after the second term of the year (cf., Chen, 2008). There are, however, few longitudinal studies that focus on how higher-education students rate their own skills. This gap in the existing research was a major factor in our decision to conduct the present survey.

Survey

Our main aim in this survey was to investigate how university students self-assess their English skills over the course of four years of study at a university. It would have been ideal to adopt a longitudinal approach to determine the trajectory of the learners' developments. Due to time constraints, however, we adopted a cross-sectional approach for this study. The research question was how the university students develop their self-assessments of their English skills over the course of four years of study. The study's three hypotheses are as follows: (1) students gain

confidence over four years of study at the university, (2) students rate their reading skills the highest, and (3) students rate items experienced in concrete situations higher than others.

Procedure

The participants were English-major students at a private university in Japan[1]. All students from the first to fourth years were asked to answer the self-assessment questionnaire. The Can-Do statements developed by the Educational Testing Service (ETS; 2000) were used in the survey (Appendix A). There were 75 Can-Do statements covering five communicative domains, including both the business and social aspects of work: listening, speaking, interaction, reading, and writing. First, the Can-Do statements were translated from English to Japanese and tested by 10 postgraduate students at the university, and any statements the participants found confusing were revised. Next, the questionnaire was administered to 348 students, who rated their levels of confidence on a 5-point Likert scale: (5) Can do easily, (4) Can do with very little difficulty, (3) Can do with some difficulty, (2) Can do with great difficulty, and (1) Cannot do at all. Participants were also asked to confirm whether they had actually experienced the situations described in the items. Data from 257 students who answered all of the statements, and their responses were used for the analysis. The statistical analysis was conducted using RStudio 0.99 and HAD software (Shimizu, 2016).

Results

The Kaiser-Meyer-Olkin measure of sampling adequacy was .95 and Barlett's test of sphericity was significant (χ^2 (2775) = 17656.38, p=.0). Therefore, the data was considered to be appropriate for using factor analysis. The number of factors was set at five as suggested in the original version of the questionnaire, and exploratory factor analysis was implemented. Maximum likelihood was adopted as an extraction method because the result can be applied to other samples. The rotation method was Promax because each factor was assumed to correlate with each other. As a result, five factors were identified and labeled as Listening, Speaking, Interaction, Reading, and Writing. Those items with factor loadings below .4 and those that did not correspond with the assumed factors were eliminated from the analysis (Appendix B). Accordingly, 42 out of 75 items were selected. Cronbach's alpha coefficients for each factor were as follows; Listening = .91 (skewness = 0.64, kurtosis = 0.91), Speaking = .91 (skewness = 0.28, kurtosis = -0.27), Interaction = .94 (skewness = 0.71, kurtosis = -0.20), .Reading = .92 (skewness = 0.20, kurtosis = -0.09), and Writing = .94 (skewness = 0.25, kurtosis = -0.10), and the internal consistencies were confirmed as being high enough to be acceptable.

The descriptive statistics of the ratings are shown in Appendix C. As can be seen from the table, the ratings in Speaking and Reading were higher than in the other domains, but the ratings as a whole were quite low and did not substantially improve over the years. The boxplots, with beeswarms, are provided in Appendix D in order to illustrate the individual participants' responses. The differences between each year in each domain were then compared using one-way analysis of variance (ANOVA). The results are shown in Table 2.

No significant differences were found in Listening, Speaking, Interaction, and Writing; the only exception was in Reading. Following this, we adopted a Holm method and conducted post hoc

analysis to make multiple comparisons in Reading. A significant difference was found only between the first- and third-year students. The effect size, however, was very small.

As implied in the previous research, the difficulty and familiarity of the questionnaire – in other words, the participants' experiences of each item - may have had an impact on their responses to them. The boxplots in Appendix E illustrate two groups: 0 is "never experienced" and 1 is "experienced." Although the number in each group was unbalanced and could not provide a precise answer to the question of respondents' relationships to them, there was an overall tendency in each item that already-experienced items were rated higher.

Table 2. *One-Way Analyses of Variance for the Ratings on Five Communicative Domains*

Variable and source	SS	MS	F	p	$\eta 2$
Listening					
Between	0.92	0.31	0.65	0.59	0.01
Error	120.49	0.48			
Speaking					
Between	1.03	0.34	0.67	0.57	0.01
Error	130.29	0.52			
Interaction					
Between	0.79	0.26	0.72	0.54	0.01
Error	92.44	0.37			
Reading					
Between	5.54	1.85	3.63	0.01	0.04
Error	128.62	0.51			
Writing					
Between	3.08	1.03	2.3	0.08	0.03
Error	112.72	0.45			

Discussion

Overall, through first to fourth year, our participants' levels of confidence did not increase as expected, and their confidence either declined or did not change between the third and fourth year. In many Japanese universities, fourth-year students are not required to take many courses, and this situation may partially explain their low self-assessments. Among the five communicative domains, Reading was rated second highest (2.97, $SD = 0.76$) in the first year and highest in the fourth year (3.25, $SD = 0.72$). This tendency corroborates the findings of Brantmeier et al. (2012), and might be due in part to the emphasis our curriculum places on

reading skills. Even though there was a significant difference between first and third year, the effect size was still very small.

Surprisingly, first-year students rated their speaking skills the highest (3.13, $SD = 0.68$), which is contradictory to Ross' findings (1998). This may be partly because the questionnaire items addressed speaking scenarios that were familiar to the participants. For example, speaking contained items such as "introduce myself in social situations and use appropriate greeting and leave-taking expressions (Speaking, 1)," and "describe my daily routine (e.g., when I get up, what time I eat lunch) (Speaking, 8)." It can be assumed that they rated this skill higher compared to the other, seemingly more difficult skills. On the other hand, the Interaction domain, which is similar to Speaking, was rated the lowest. This might have been due to the nature of the questionnaire items. In the Interaction domain, we included items such as "conduct simple business transactions at places such as the post office, bank, drugstore (Interaction, 1)" and "explain to a repairman what is wrong with an appliance that I want fixed (Interaction, 8)." These items required that the participants not only had the experience of interacting with others but also that they could negotiate with specific knowledge. If they had not had the experience mentioned in the items, they might have thought that they were unable to perform those tasks. This again indicates that experience of or familiarity with the questionnaire items is a very important factor in self-assessment of language skills, which supports the findings of Ross (1998).

Conclusion

This study explored how university students gained or lost confidence over a four-year period. The students demonstrated a tendency to rate reading skills higher than other ones, and whether they had experience of the items in question was found to be an important factor affecting ratings. Generally, the third-year students demonstrated more confidence than the first-year ones in the Reading domain, but not in the other domains. In addition, as many of the university teachers could readily guess, the fourth-year students seemed to have lost confidence in each domain, and that the ratings declined from the third to fourth year. We hope that these results will lead to further discussion on how we can ensure that we provide courses that deliver balanced instruction of the five English communicative domains, and how, under the new curriculum, we can enhance and maintain student confidence in the fourth year.

There are two issues to be considered for future studies. First, because of time constraints, this study adopted a cross-sectional approach to investigate the university students' self-assessments on Can-Do items. A longitudinal study is also needed to reinforce this study's findings. Second, the Can-Do list in this study was designed for TOEIC test-takers, and some question items were unfamiliar to the students. Since familiarity and experience affect self-assessments, similar studies, which adopt criterion-referenced and concrete Can-Do items, might affect the accuracy of self-assessments.

About the Author

Shinya Ozawa is a professor in the Department of English at Hiroshima Shudo University in Japan. His research interest lies in the integration of ICT in the university language classrooms.

Note

1. The homogeneity of each group cannot be guaranteed because of the placement test change from TOEIC Bridge to TOEIC in 2011. As a reference, a t-test was conducted to compare the first- and second-year students who took TOEIC (t = -0.473, df = 240, p = .636), and third- and fourth-year students who took TOEIC Bridge (t = 0.160, df = 252, p = .873). It can therefore be assumed that the students were at similar English proficiency levels when they entered the university.

References

Alderson, C. A. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. New York: Continuum.

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language Learning, 39*, 313–338. doi:10.1111/j.1467-1770.1989.tb00595.x

Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me? Individual self-assessment by skill and level of language instruction. *System, 40*(1), 144–160. doi:10.1016/j.system.2012.01.003

Chen, Y. M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research, 12*(2), 235–262. doi: 10.1177/1362168807086293

DIALANG [Online language diagnosis system] (n.d.). Lancaster, UK: Lancaster University. Retrieved from https://dialangweb.lancaster.ac.uk

Educational Testing Service. (2000). *TOEIC Can-do guide: Linking TOEIC scores to activities performed using English*. Retrieved from https://www.ets.org/Media/Research/pdf/TOEIC_CAN_DO.pdf

Ito, T., Kawaguchi, K., & Ohta, R. (2005). *A study of the relationship between TOEIC scores and functional job performance: Self-assessment of foreign language proficiency*. Retrieved from http://www.toeic.or.jp/library/toeic_data/toeic_en/pdf/newsletter/1_E.pdf

Léger, D. S. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals, 42*(1), 158–178. doi:10.1111/j.1944-9720.2009.01013.x

Little, D. (2014). The Common European Framework and the European Language Portfolio: involving learners and their judgements in the assessment process. *Language Testing, 22*(3), 321–336. doi:10.1191/0265532205lt311oa

Nishida, R. (2012). A longitudinal study of motivation, interest, can-do and willingness to communicate in foreign language activities among Japanese fifth-grade students. *Language Education & Technology, 49*, 23–45. Retrieved from http://ci.nii.ac.jp/naid/110009470694

Powers, D. E., Kim, H-J., & Weng, V. Z. (2008). *The redesigned TOEIC (listening and reading) test: Relations to test-taker perceptions of proficiency in English* (ETS Research Report No. 08-56). Retrieved from https://www.ets.org/Media/Research/pdf/RR-08-56.pdf

Powers, D. E., Yu, F., & Yan, F. (2013). *The TOEIC listening, reading, speaking, and writing tests: Evaluating their unique contribution to assessing English-language proficiency*. Retrieved from https://www.ets.org/Media/Research/pdf/TC2-03.pdf

Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Language Testing, 15*(1), 1–20.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Education, 11*(10), 1–13. Retrieved from http://pareonline.net/pdf/v11n10.pdf

Saida, C. (2008). The use of the common European framework of reference levels for measuring Japanese university students' English, *JACET Journal, 47*, 127–140. Retrieved from http://ci.nii.ac.jp/naid/110007467221/en

Shimizu, H. (2016). Free-soft no tokei-bunseki soft HAD: Kino no syokai to tokei gakusyu kyoiku kenkyu-jissen ni okeru riyo no teian [The development of free statistical analysis software HAD: Introducing its functions and suggesting how to use for statistical learning, teaching/researching practices], *Journal of Media, Information and Communication, 1*, 59–73. Retrieved from http://jmic-weblab.org/ojs/index.php/jmic/article/view/6/5

Appendix A: TOEIC Can-Do Questionnaire

(Adapted from "TOEIC Can-do guide: Linking TOEIC scores to activities performed using English," by ETS, 2000, Appendix A)

IN LISTENING, I CAN

1. understand simple questions in social situations such as "How are you?" "Where do you live?" and "How do you feel?"
2. understand a salesperson when she or he tells me prices of various items
3. understand someone speaking slowly and deliberately, who is giving me directions on
4. understand explanations about how to perform a routine task related to my job
5. understand a co-worker discussing a simple problem that arose at work
6. understand announcements at a railway station indicating the track my train is on and the time it is scheduled to leave
7. understand headline news broadcasts on the radio
8. understand a client's request made on the telephone for one of my company's major products or services
9. understand a person's name when she or he gives it to me over the telephone
10. understand play-by-play descriptions on the radio of sports events that I like (e.g., soccer, baseball)
11. understand an explanation given over the radio of why a road has been temporarily closed
12. understand someone who is speaking slowly and deliberately about his or her hobbies, interests, and plans for the weekend
13. understand directions about what time to come to a meeting and the room in which it will be held
14. understand an explanation of why one restaurant is better than another
15. understand a discussion of current events taking place among a group of persons speaking English

IN SPEAKING, I CAN

1. introduce myself in social situations and use appropriate greeting and leave-taking expressions
2. state simple biographical information about myself (e.g., place of birth, composition of family)
3. describe the plot of a movie or television program that I have seen
4. describe a friend in detail, including physical and personality characteristics

5. describe my academic training or my present job responsibilities in detail

6. order food at a restaurant

7. talk about topics of general interest (e.g., current events, the weather)

8. describe my daily routine (e.g., when I get up, what time I eat lunch)

9. talk about my future professional goals and intentions (e.g., what I plan to be doing next year)

10. tell a co-worker how to perform a routine job task

11. telephone the airline to change my flight reservations to a different time and day

12. tell a colleague at work about a humorous event that recently happened to me

13. give a prepared half-hour formal presentation on a topic of interest

14. adjust my speaking to address a variety of listeners (e.g., professional staff, a friend, children)

15. tell someone directions on how to get to my house or apartment

IN INTERACTIVE SKILLS, I CAN

1. conduct simple business transactions at places such as the post office, bank, drugstore

2. telephone a restaurant to make dinner reservations for a party of three

3. give and take messages over the telephone

4. explain written company policies to a new employee

5. discuss with a co-worker the best way to accomplish a job task

6. discuss with an electronics salesperson the features I want on a new videocassette recorder

7. meet with a doctor and explain the physical symptoms of my illness

8. explain to a repairman what is wrong with an appliance that I want fixed

9. request information over the telephone (e.g., check airline schedules with a travel agent)

10. meet with a real-estate agent to discuss the type of house I would like to buy

11. talk to an elementary school class about what I do for a living

12. discuss world events with an English-speaking guest

13. discuss with my boss ways to improve customer service or product quality

14. telephone a department store and find out if a certain item is currently in stock

15. conduct an interview with an applicant for a job in my area of expertise

IN READING, I CAN

1. read, on storefronts, the type of store or services provided (e.g., "dry cleaning," "book store")
2. read and understand a train or bus schedule
3. read and understand a restaurant menu
4. find information that I need in a telephone directory
5. read office memoranda written to me in which the writer has used simple words or sentences
6. read and understand traffic signs
7. read and understand simple, step-by-step instructions
8. read and understand an agenda for a meeting
9. read and understand a travel brochure
10. read and understand magazine articles like those found in Time or Newsweek, without using a dictionary
11. read and understand directions and explanations presented in computer manuals written for beginning users
12. identify inconsistencies or differences in points of view in two newspaper interviews with politicians of opposing parties
13. read highly technical material in my field or area of expertise with no use or only infrequent use of a dictionary
14. read and understand a popular novel
15. read and understand a letter of thanks from a client or customer

IN WRITING, I CAN

1. write a list for items to take on a weekend trip
2. write a one- or two-sentence thank-you note for a gift a friend sent to me
3. write a brief note to a co-worker explaining why I will not be able to attend the scheduled meeting
4. write a postcard to a friend describing what I have been doing on my vacation
5. fill out an application form for a class at night school
6. write clear directions on how to get to my house or apartment
7. write a letter requesting information about hotel accommodations for a future vacation
8. write a short note to a co-worker describing how to operate a standard piece of office equipment (e.g., photocopier, fax machine)
9. write a memorandum to my supervisor explaining why I need a new time off from work
10. write a letter introducing myself and describing my qualifications to accompany an employment application
11. write a memorandum to my supervisor describing the progress being made on a current project or assignment

12. write a complaint to a store manager about my dissatisfaction with an appliance I recently purchased

13. write a letter to a potential client describing the services and/or products of my company

14. write a 5-page formal report on a project in which I participated

15. write a memorandum summarizing the main points of a meeting I recently attended

Appendix B: Factor Loadings for TOEIC Can-Do Questionnaire

Item	1	2	3	4	5	Communality
		Factor loading				
I6	**.95**	-.11	-.08	.01	.06	.92
I10	**.84**	-.18	-.06	.01	.23	.80
I14	**.78**	-.15	-.04	.07	.21	.69
I13	**.71**	.22	.07	-.15	-.17	.62
I11	**.71**	-.14	.09	.13	.04	.55
I4	**.70**	.18	-.04	-.02	-.05	.52
I5	**.68**	.21	-.01	.01	-.10	.52
I15	**.67**	.16	.13	-.15	-.09	.52
I12	**.62**	.26	-.01	.06	-.18	.49
I8	**.61**	-.10	.10	.14	.08	.42
I9	**.53**	.14	.02	.15	.10	.33
S15	**.40**	.10	.01	.35	.07	.30
I7	**.37**	-.09	.18	.30	-.02	.26
I2	**.31**	.29	-.07	.27	.15	.28
L7	-.05	**.76**	-.15	.20	.06	.65
L8	.11	**.70**	-.05	.01	.07	.51
L5	.04	**.69**	-.11	.19	.08	.54
L11	.11	**.68**	-.15	.10	.14	.52
L4	-.03	**.64**	-.05	.28	.05	.50
L6	-.24	**.64**	-.10	.31	.22	.62
L15	.19	**.63**	.05	.03	-.10	.45
L10	-.01	**.62**	-.02	.04	.06	.38
L13	-.13	**.58**	-.12	.31	.21	.50
R10	.01	**.51**	.24	-.29	.24	.46
R8	-.04	**.51**	.21	-.25	.46	.58
R12	.06	**.50**	.35	-.32	.08	.49
S11	.16	**.46**	.11	.24	-.09	.31
R13	.11	**.45**	.35	-.36	.10	.48
I1	.10	**.43**	-.01	.34	.08	.32
S10	.28	**.41**	.10	.31	-.12	.37
R14	.10	**.40**	.11	-.26	.32	.35
S14	.21	**.35**	.14	.22	-.13	.26
S13	.09	**.32**	.22	.30	-.16	.27
I3	.27	**.31**	-.05	.26	.12	.25
W9	.21	-.24	**.70**	.06	.07	.61

(table continues)

| Item | | Factor loading | | | | Communality |
	1	2	3	4	5	
W7	-.05	.13	**.68**	.12	.04	.50
W4	-.21	-.14	**.63**	.50	.05	.71
W3	-.03	-.31	**.61**	.41	.09	.64
W14	.10	.28	**.60**	-.12	-.21	.51
W10	.09	.06	**.59**	.15	-.05	.38
W13	.29	.13	**.58**	-.16	.01	.46
W11	.37	-.06	**.58**	-.08	.06	.49
W12	.42	-.14	**.56**	-.05	.08	.53
W8	.12	.16	**.55**	-.01	.11	.36
W5	.04	.01	**.55**	.10	.20	.35
W15	.27	.24	**.51**	-.21	-.03	.44
W6	-.02	.04	**.46**	.32	.11	.33
S8	.07	-.01	.11	**.74**	-.07	.57
S2	.07	-.09	-.04	**.73**	.02	.55
L12	-.01	.14	-.12	**.69**	.14	.53
S7	.13	.11	.07	**.64**	-.13	.47
S1	.03	.05	-.05	**.62**	.05	.39
S4	.17	.08	.10	**.60**	-.15	.43
L1	-.13	.00	-.01	**.60**	.03	.38
S6	.09	.15	-.10	**.54**	.22	.38
L2	-.24	.46	-.07	**.54**	.03	.56
W2	-.17	-.33	.45	**.53**	.25	.67
S9	.08	.26	.18	**.51**	-.16	.40
L3	-.07	.30	.00	**.51**	.13	.38
S3	.01	.37	.15	**.49**	-.27	.48
S12	.20	.41	-.01	**.44**	-.12	.42
L14	.19	.19	-.06	**.44**	.12	.28
W1	-.07	-.20	.41	**.43**	.26	.47
S5	.10	.33	.22	**.37**	-.15	.32
L9	.02	.19	-.12	**.36**	.30	.27
R7	-.02	.18	-.01	-.05	**.76**	.61
R4	.12	.04	.04	.03	**.72**	.53
R6	-.04	.20	-.02	-.04	**.70**	.53
R9	.06	.19	.01	-.03	**.68**	.51
R3	-.06	-.02	-.02	.23	**.67**	.51
R5	.05	.02	.09	.11	**.66**	.46
R2	-.06	-.03	.00	.29	**.65**	.51
R1	-.03	-.18	.00	.43	**.54**	.51
R11	.06	.26	.15	-.16	**.50**	.37
R15	.20	.16	.20	-.06	**.45**	.31

Appendix C: Comparisons of University Students' Ratings of Confidence in Five Communicative Domains

Factor	1st year[a] M (SD)	1st year[a] 95% CI	2nd year[b] M (SD)	2nd year[b] 95% CI	3rd year[c] M (SD)	3rd year[c] 95% CI	4th year[d] M (SD)	4th year[d] 95% CI
Listening	2.32 (0.68)	[2.16, 2.48]	2.32 (0.71)	[2.15, 2.48]	2.44 (0.63)	[2.24, 2.64]	2.44 (0.72)	[2.27, 2.60]
Speaking	3.13 (0.68)	[2.97, 3.30]	3.20 (0.65)	[3.03, 3.38]	3.11 (0.73)	[2.90, 3.32]	3.03 (0.80)	[2.86, 3.20]
Interaction	1.67 (0.52)	[1.53, 1.81]	1.80 (0.64)	[1.65, 1.95]	1.82 (0.66)	[1.64, 2.00]	1.75 (0.62)	[1.60, 1.89]
Reading	2.97 (0.76)	[2.80, 3.13]	3.05 (0.65)	[2.88, 3.22]	3.34 (0.72)	[3.13, 3.54]	3.25 (0.72)	[3.09, 3.42]
Writing	2.18 (0.72)	[2.03, 2.34]	2.43 (0.59)	[2.27, 2.59]	2.45 (0.63)	[2.26, 2.64]	2.40 (0.70)	[2.24, 2.55]

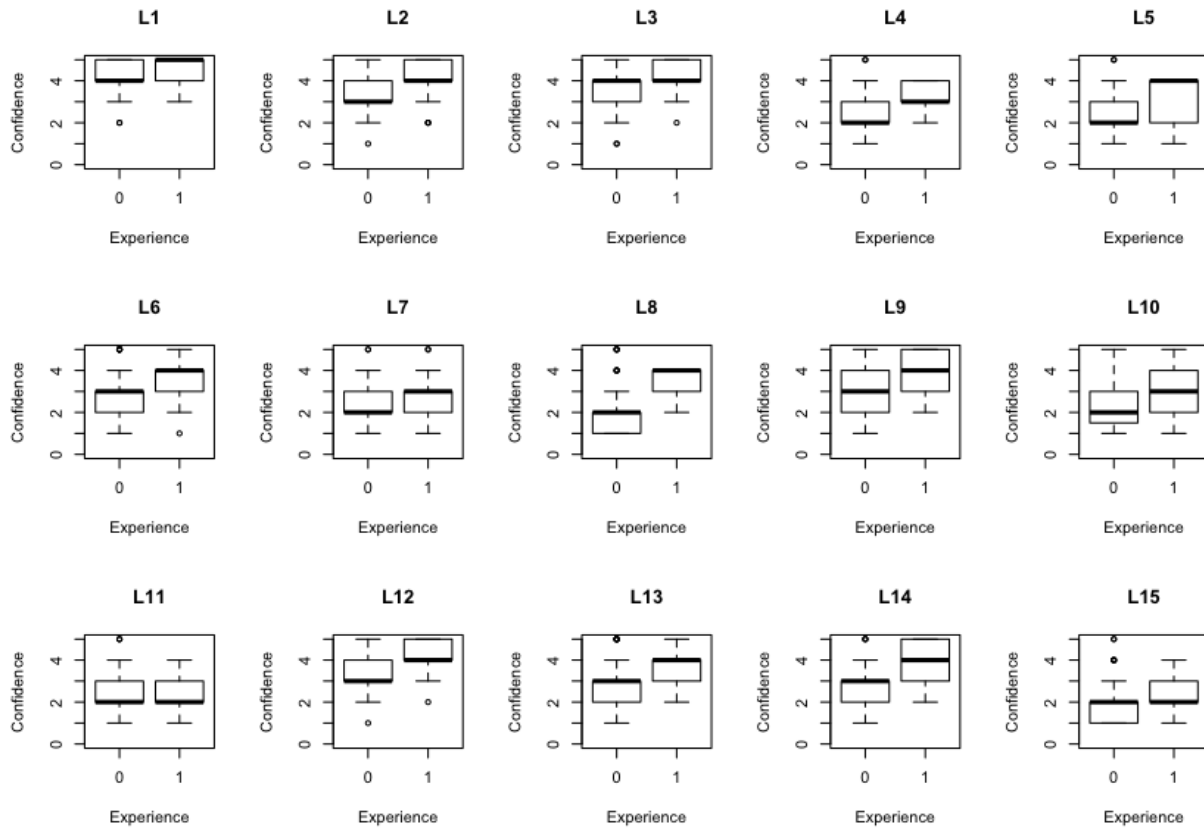*Note.* CI = confidence interval.   [a]$n = 73$. [b]$n = 67$. [c]$n = 47$. [d]$n = 70$.

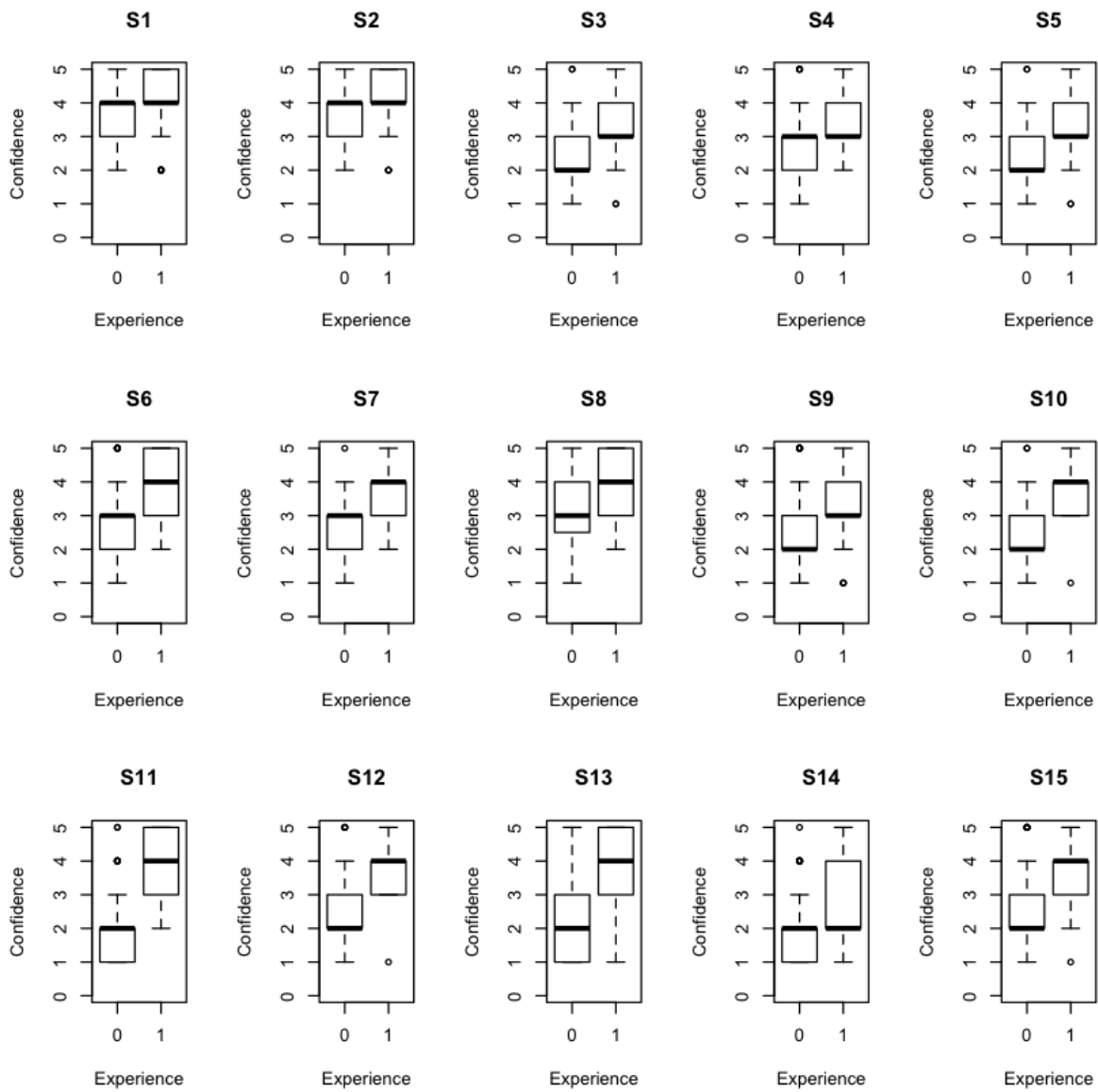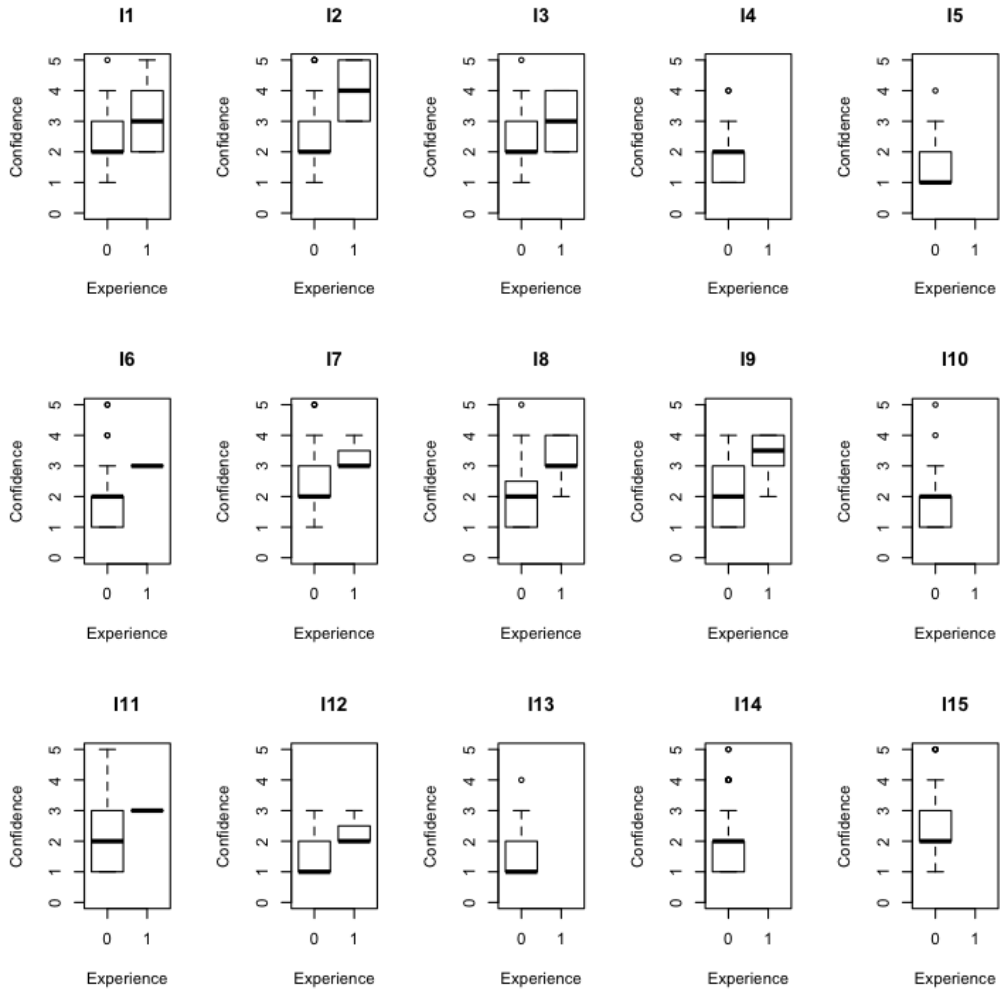# Appendix D: Individual Responses in Each Communicative Domain
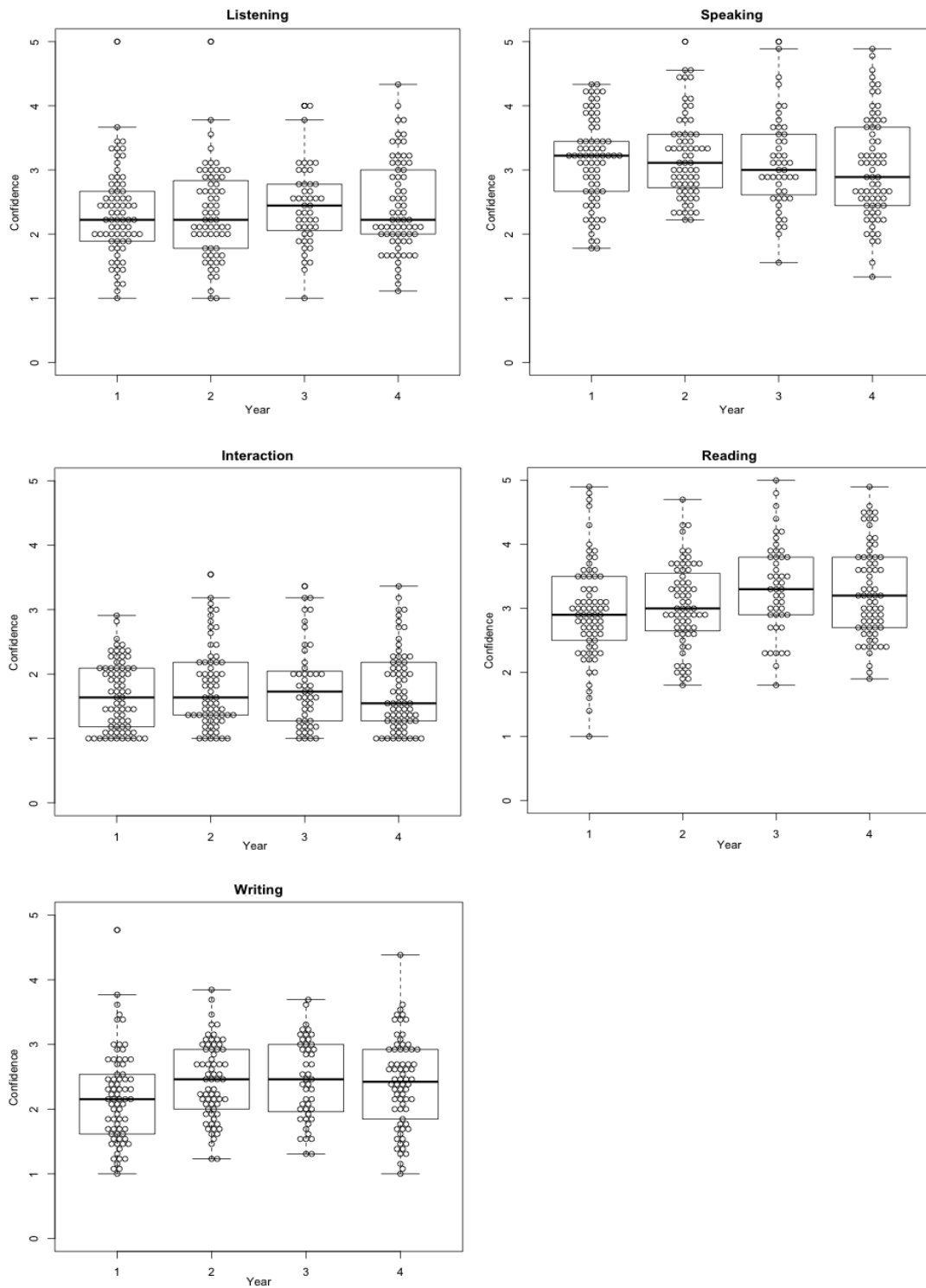
Note. Boxplots show the maximum, third quartile, median, first quartile, and minimum of the data. Beeswarm plots show individual responses.
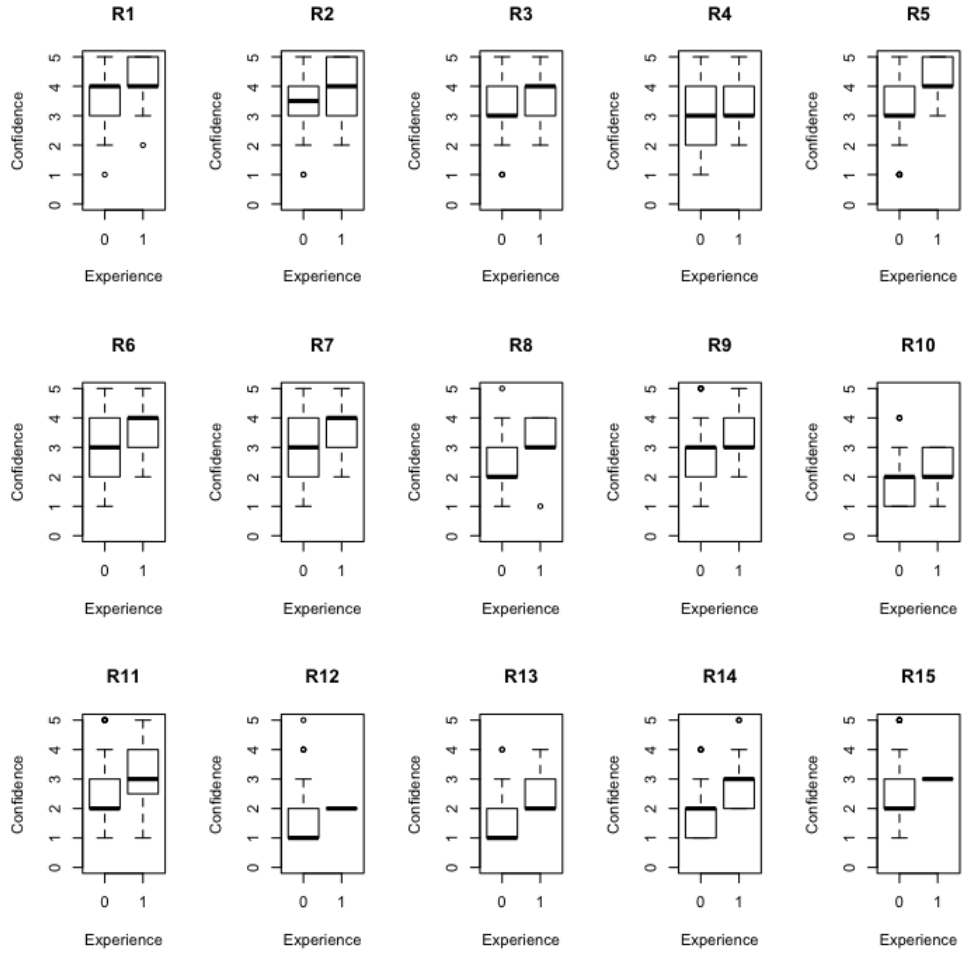
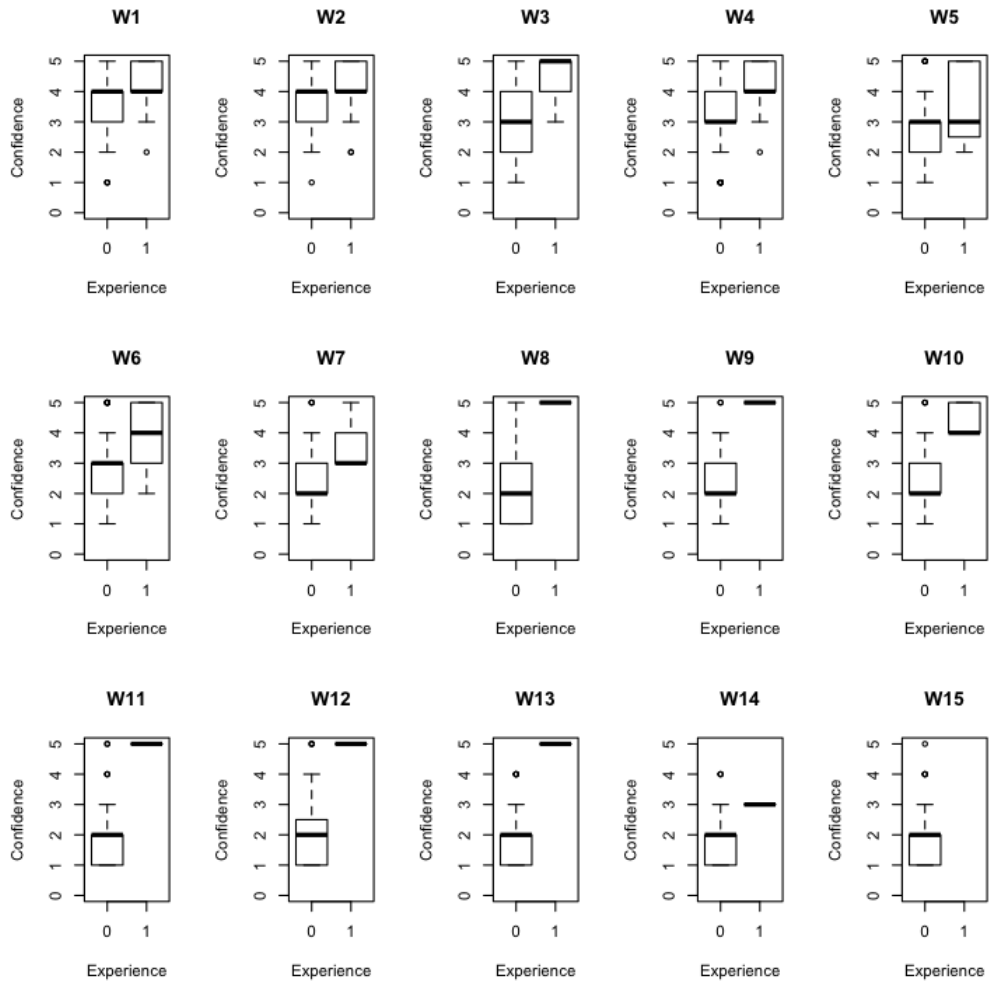# Appendix E: Experience Factors on Self-Assessment