



# OSAKA JALT JOURNAL

Volume 2, June 2015

ISSN: 2188-5192



---

# Osaka JALT Journal

Volume 2

June 2015

## Chief Editor

Ryan W. Smithers, *Kwansei Gakuin University*

## Associate Editors

Robert Sheridan

Laura Markslag

Michael Herke

Jennifer Jordan

Steve Porritt

## Osaka JALT Executive Board

Bob Sanderson, President

Ryan W. Smithers, Treasurer & Publications Co-Chair

Julia Kimura, Membership Chair

Mark Pileggi, Publicity Chair

Robert Sheridan, Programs Chair & Publications Co-Chair

Sean Gay, Website Editor

## Disclaimer

The views expressed in the articles in *Osaka JALT Journal* are not necessarily those of the Editors, Osaka JALT, or the Osaka JALT Executive Board.

## Editorial Office

Correspondence relating to *Osaka JALT Journal* should be directed to Ryan W. Smithers or Robert Sheridan, Publications Co-Chairs at [publications@osakajalt.org](mailto:publications@osakajalt.org).

## Submissions

*Osaka JALT Journal* is a peer reviewed journal that is devoted to outstanding research across the spectrum of language teaching and learning. For information regarding submission guidelines and how to submit an article, please see our website <http://www.osakajalt.org/journal/> for more information.

Copyright© Contributors 2015. All articles in *Osaka JALT Journal* are copyrighted by the respective authors and cannot be reproduced or distributed without express permission.

# Contents

- 3 **From the Editor**
- 5 **A Comparison of Textbook and Self-selected Topics in a Japanese EFL Context**  
*By Tsui-Ping Cheng*
- 28 **Improving Motivations through Autonomous Projects**  
*By Ryan W. Smithers*
- 44 **The Potential Impact of Guessing on the Monolingual and Bilingual Versions of the Vocabulary Size Test**  
*By Kurtis McDonald*
- 62 **Using L2 Blogs to Promote Student Participation and Critical Thinking**  
*By Gilbert Dizon and Chris Edelman*
- 78 **Standardized Testing in University**  
*By Damien Healy*
- 98 **Using Focus Groups for ELT Research**  
*By David Bramley*
- 110 **Factors influencing the Adoption of Communicative Textbooks**  
*By Simon Humphries*
- 116 **Wikis Revisited: A Single Software Tool for a Variety of Courses**  
*By Harry Carley*
- 124 **Choose Two, Pick One**  
*By Sean Gay*

## From the Editor

### Welcome to the 2015 Osaka JALT Journal!

**T**he *Osaka JALT Journal*, published by the Osaka Chapter of the Japan Association of Language Teaching (JALT), is a journal that is committed to publishing innovative articles on foreign and second language teaching and learning. It provides authors, most of whom are practitioner-researchers, an opportunity to share with our readers how teaching is improved by incorporating their pragmatic and theoretically based findings into pedagogy.

This issue of the journal opens with *‘A Comparison of Textbook and Self-selected Topics in a Japanese EFL Context’* by Tsui-Ping Cheng, which provides a wealth of information on what classroom topics are best at promoting student engagement—a paper that anyone involved in materials development will not want to miss. In *‘Improving Motivations through Autonomous Projects’*, Ryan Smithers explains how teachers, students, and project work can be molded into a tripartite relationship that results in autonomous learning and improved classroom motivation. Kurtis McDonald’s article, *‘The Potential Impact of Guessing on Monolingual and Bilingual versions of the Vocabulary Size Test’* is a must read for teachers who are engaged in or thinking about conducting vocabulary size testing. The next article, *‘Using L2 Blogs to Promote Student Participation and Critical Thinking’* by Gilbert Dizon and Chris Edelman presents a strong case for getting students involved in blogging. Damien Healy’s paper, *‘Standardized Testing in University’* will be of interest to all English language teachers. *‘Using Focus Groups for ELT Research’* by David Bramley provides excellent insights for practitioner-researchers on how to make the most of focus group sessions.

In this volume of the *Osaka JALT Journal*, we also welcome three reports from authors that presented papers at a JALT conference in the past year. Simon Humphries' report, '**Factors Influencing the Adoption of Communicative Textbooks**', is based on a presentation that he made at the JALT International Conference in Tsukuba last year. Harry Carley's informative report titled, '*Wikis Revisited: A Single Software Tool for a Variety of Courses*' explains the benefits of wikis for foreign language classrooms. Finally, this volume ends with insights for teachers on how to maintain a balanced curriculum by Sean Gay in '*Choose Two, Pick One*'.

Thank you to all of the educators who have shared their research papers and reports with us in this volume of the *Osaka JALT Journal*. Not only has it has been a pleasure to work with all of you, but we have been privileged to have gleaned some nuggets that are bound to make our classrooms better and more informed learning environments.

A final word of thanks to the *Osaka JALT Journal* editorial team. They have devoted much time and effort into this publications, even in the face of pressing work and social commitments. If it were not for their hard work and commitment to editorial excellence, this publication would never have come together.

Enjoy the articles,

Ryan

# A Comparison of Textbook and Self-selected Topics in a Japanese EFL Context

**Tsui-Ping Cheng**

Hitotsubashi University

## **Abstract**

Topics used in language classrooms are vital to promoting student engagement and language learning. Topic selection is often influenced by textbooks. However, while researchers frequently evaluate the authenticity, description, and presentation of the language in textbooks, little research has investigated textbook topics in relation to student interests. To address this gap in the literature, this study compares the topics presented in English language textbooks in Japan with the topics selected by 48 Japanese university students enrolled in compulsory English courses. The textbook topics were compiled from 21 intermediate-level textbooks that are available on the Japanese EFL market, while the self-selected topics were collected from student-led discussions during one academic year. The two sets of topics were categorized and analyzed. The analysis revealed vast discrepancies in terms of the frequency, diversity, and treatment of certain topics. The analysis also showed that the textbooks favored topics related to self and social issues, whereas the students overwhelmingly chose topics that dealt with their school life, local practices, and immediate needs. The findings provide an empirical basis for improving textbook and discussion topics used in the Japanese EFL context.

*Keywords:* textbook topics, self-selected topics, group discussions, willingness to communicate

**E**ncouraging students to speak English poses a real challenge for language teachers in many EFL contexts. The challenge is even greater in English classes in Japan because of the prevalence of teacher-centered methodologies in which students are expected to be silent in the classroom; giving them few opportunities for conversational practice (Talandis & Stout, 2015). This often inhibits the willingness of students to speak English and, consequently, creates a demotivating

environment for language learning. When students finally break their silence and began actively engaging in English conversations, teachers may wonder what factors prompted this increased willingness to use the target language.

Every semester during course evaluations, my students repeatedly select “free discussion” as their favorite and most helpful classroom activity. For the free discussion activity, students choose their own topic. One student commented, “Thanks to this activity, I could discuss many topics with my friends in English. Now I am not hesitant to speak English.” In addition to positive feedback about the activity, I found that the students participated actively and used English consistently in free discussions. Their surprisingly high motivation to speak in English persuaded me to investigate what they talked about and how their self-selected topics differed from those in the textbook. As a result, this study aims to: (1) examine the topics selected by students in a Japanese EFL classroom; (2) compare these self-selected topics with the topics presented in English language textbooks that are available in Japan; and (3) explore the differences, if any, between the textbook-assigned and self-selected topics. This examination will provide language-teaching professionals with valuable insights into instructional practices and topic choices for their local teaching contexts.

## **Background**

The data for this study come from student discussions. To understand the importance of discussion activities, the background section will first outline the rationales for using them in language classrooms. Following this outline, there will be a review of studies that have investigated the impact of discussion topic choice on students’ willingness to communicate in the target language. Finally, the background section will end with a

discussion on how self-selected topics contribute to learner autonomy.

### *Group Work in Language Classrooms*

Discussion activity is a type of group work frequently used in language classrooms to engage learners in meaningful interaction. The importance of group work in language classrooms arises from the role of comprehensible input and interaction in the learning of additional languages (Long & Porter, 1985). During the 1980s and 1990s, language researchers were interested in examining how human cognition and linguistic environment reciprocally interact to affect the processes of second language (L2) learning.<sup>1</sup> From this cognitive-interactionist perspective, Krashen (1985) proposed comprehensible input hypothesis, which argues that the more comprehensible input that learners receive, the more they learn. Long (1996) later formulated the interaction hypothesis, proposing that it is not merely input per se that is essential to the acquisition of a second language, but input that is made comprehensible through negotiation of meaning during interaction. Research on the role of interaction in L2 learning, therefore, examines how input is shaped, modified, and received as learners jointly strive for comprehensibility (e.g., Pica, 1987; Pica & Doughty, 1985a, 1985b; Varonis & Gass, 1985). This line of inquiry highlights the appeal of group work in expanding the amount and variety of language practice for negotiation of meaning during interaction.

Another advantage of group work in language classrooms derives from Swain's (1996) output hypothesis, which states that input is necessary but not sufficient to ensure L2 development. Also essential is output that forces learners to experiment with new linguistic forms and extend their linguistic repertoire. The role that is played by the

---

<sup>1</sup> The term "L2" here is used to mean both second and foreign language.



combination of input, interaction, and output in L2 learning establishes a firm theoretical rationale for using discussion activities in language classrooms.

The social context in which input and output are negotiated is better understood through sociocultural theory (Vygotsky, 1978), which supplies a powerful explanatory framework to describe the collaborative nature of learner–learner interaction (Donato, 1994, 2004). This line of research shifts the focus from the amount of input to “the opportunities for meaningful action that the situation affords” (van Lier, 2000, p. 252) and pulls into view the dynamics of collaboration involved in classroom group work and the value of engaging students in collaborative work as a vehicle for optimal L2 development. Therefore, from both the cognitive-interactionist and sociocultural perspectives, creating a linguistic and social environment favorable to L2 development requires providing learners with ample opportunities to engage in meaningful interaction in language classrooms.

### *Topic Choice for Discussion*

Topics are a vital component of any meaningful interaction in language classrooms because of their direct impact on how language is used (Todd, 1998). Topic selection for classroom activities is often influenced by textbooks. Regarding English as an international language, McKay (2003a) emphasizes the advantages of using topics concerning local and international cultures rather than topics about the culture of native speakers. In so doing, students can learn how to communicate their culture to others in English and experience how English is used for intercultural communication for a variety of purposes. McKay (2003b) further provides an example of Chilean English teachers to help educators recognize the value of including local culture topics in textbooks as a way

of meeting their students' specific needs and purposes for learning English.

From the student's perspective, however, textbook topics may not necessarily be the best choice for classroom discussions. Green, Christopher, and Lam (1997) argued that students are less motivated to engage in a discussion when the discussion topic is imposed. Conversely, students are more personally involved when they have control over the selection of the topic. Green et al. (1997) further stated that giving students the choice of topics for discussion is particularly important in monolingual classrooms, because these students share the same cultural backgrounds and are likely to select similar topics for discussion. Recently, Wolf (2013) reported similar findings in his survey study of Japanese EFL student perceptions of textbook-assigned and self-selected discussion topics. Drawing on data from three monolingual classrooms, Wolf (2013) found that students not only vastly prefer topics limited to a national context, but also demonstrate significantly greater interest, knowledge, and confidence in discussing their own topics.

### *Topic Familiarity and Interest*

The importance of topic familiarity and interest has also been recognized by research on the willingness to communicate (WTC). According to MacIntyre, Dörnyei, Clément, and Noels (1998), WTC in the L2 refers to "a readiness to enter into discourse at a particular time with a specific person or persons, using [an] L2" (p. 547). Several researchers have identified topic familiarity and interest as factors that affect learners' WTC in class (Cao & Philip, 2006; Kang, 2005; MacIntyre et al., 1998; Zarrinabadi, 2014). MacIntyre et al. (1998) argued that, while topic familiarity increases one's linguistic self-confidence, the lack of it might hinder one's readiness to engage in communication with others. Their argument is supported by Zuengler (1993), who found that greater content knowledge can

enable speakers to talk more and, as a result, overcome their limitations in oral proficiency. Likewise, Yu (2009) reported a positive correlation between lexical diversity and topic familiarity in writing and speaking performances on language tests. Kang (2005) also observed that learners feel more comfortable and secure when discussing topics about which they have some background knowledge. On the contrary, unfamiliar topics can impose a linguistic burden on learners that can increase their feelings of insecurity and reduce their willingness to communicate. Regarding topic interest, Kang (2005) found that learners tend to feel excited when conversing about topics that appeal to them, which ultimately enhances the learners' WTC. Kang concluded that, in order to facilitate learner WTC in the L2, teachers should offer topics that the learners are familiar with and interested in. Having learners select topics and present them for discussion, for example, is one way to identify topics that match their interests and generate learner WTC (Kang, 2005). Zarrinabadi (2014) also described the positive effect of this pedagogic practice, noting greater WTC and classroom participation when the students chose the discussion topic.

### *Learner Autonomy*

Allowing learners to choose the topic also enables them to experience a degree of autonomy. The most widely cited definition of learner autonomy refers to learners' abilities to take charge of their learning (Holec, 1981). Although this definition has been debated, the benefits of learner autonomy in language education are increasingly recognized (Benson, 2007; Smith, 2008). The ultimate goal of language teaching is to promote autonomous language-learning behavior and help students become independent learners outside the classroom. Autonomous practices in language learning include

allowing learners to set goals, determine learning methods, decide topics, and evaluate progress (Illés, 2012). Learner autonomy is fostered by giving the learners some control over their learning and involving them in the selection of the topics, materials, and activities used in the classroom.

## **The Current Study**

### ***Participants***

The participants were 46 Japanese students and two exchange students (one Chinese and one Korean), comprising 35 males and 13 females majoring in commerce and management at a university in Tokyo. All the 48 participants were enrolled in a compulsory English communication course for first year undergraduate students of commerce and management. They were placed into three classes based on their TOFEL scores.<sup>2</sup> Each class consisted of 16 students who had scored between 500 and 525 on the paper-based TOFEL test. The course is designed to help the students to communicate with confidence and to become effective communicators in the global business community. Particular attention is given to developing their skills in discussing, presenting, and debating in English.

### ***Research Questions***

In order to examine the disparity between textbook topics and student interests, this study will address the following questions:

- (1) What topics are covered in English language textbooks in Japan?
- (2) What topics did the students choose to talk about in the aforementioned EFL classes?

---

<sup>2</sup> The author was the instructor of the three classes.

(3) What are the differences between the textbook and self-selected topics?

### **Data Collection and Analysis**

The data used in this study was collected from two sources: commercial English textbooks and discussion activities in EFL classes in Japan. This section describes how the topics from these two data sources were collected, categorized, and distributed.

#### ***Textbook Topics***

Based on the proficiency levels of the participating students, twenty-one intermediate-level textbooks that are available in Japan were examined. Specifically, the collection included seven textbooks on integrated skills, seven on discussion and debate, and seven on everyday conversation (see Table 1 and Appendix A). All of the sampled textbooks have a speaking component that corresponds to the objectives of the English communication courses that the participants were taking.

Table 1  
*Textbook Type and Title*

Textbook type	Textbook title	Number of topics
Integrated skills	<i>Breakthrough plus 3</i>	88
	<i>Four corners 3</i>	
	<i>Four corners 4</i>	
	<i>Touchstone 3</i>	
	<i>Interchange 3</i>	
	<i>New language leader intermediate</i>	
	<i>Life intermediate</i>	
Discussion and debate	<i>Debating current issues</i>	107
	<i>For and against</i>	
	<i>Which side are you on?</i>	
	<i>My opinion, your opinion</i>	
	<i>Ideas and issues</i>	
	<i>In focus 2</i>	
	<i>Impact issues 3</i>	
Everyday conversation	<i>Face to face</i>	83
	<i>Communicate 1</i>	
	<i>Topic talk issues</i>	
	<i>Nice talking with you 2</i>	
	<i>English firsthand 2</i>	
	<i>Top notch 2</i>	
	<i>Communication strategies 2</i>	

Table 2 provides a summary of the categories and example topics and their total number listed from the greatest to the least. In total, 278 topics were compiled from the 21 textbooks. The topics were categorized by adapting and synthesizing the thematic categories created by Wolf (2013) and Siegel (2014). The initial categories are self, relationships, leisure, school life, money and jobs, entertainment and fashion, animals,

food and health, places and travel, social issues, communication technology, and culture. After the initial sorting, I added one more category—nature and climate—in order to group topics, such as seasons and weather, that did not really fit into any of the other categories.

Table 2

*Textbook Topics*

Categories	Example topics	Number
Self	Life events, dreams, goals, likes, dislikes, experiences	61
Social issues	Politics, the oil crisis, war, violence, gun control, depopulation, the gender gap	55
Money/jobs	Shopping, career, business ideas	29
Entertainment/fashion	Music, movie, beauty, appearance	22
Food/health	Health care, food choices, restaurants	20
Relationships	Love, friends, marriage	17
School life	Study abroad, school subjects, university life, plagiarism	16
Places/travel	Tourism, travel, transportation	14
Culture	Language, behavior, history, art, ethics and values	14
Communication technology	The internet, cell phones, computers	12
Nature/climate	The environment, global warming, winter	11
Leisure	Free time, leisure activities	3
Animals	Animals, animal rights, endangered species	4

While Siegel (2014) also found topics about “self” to be the most frequent, the textbooks reviewed in her study included only a small number of social topics. The

textbooks sampled in the present study were more advanced, which could explain the higher proportion of social topics. This points to a possible connection between the amount of coverage on social issues and the textbook level.

Upon closer inspection, both the integrated skills and everyday conversation textbooks contained a considerably higher frequency of “self” related topics, with 35 and 20 instances, respectively, whereas the frequency of other categories were much lower and relatively equal. On the other hand, the discussion and debate textbooks preferred topics relating to “social issues,” with 32 instances. The frequency of topics across the textbook types reflected their different themes and priorities of the discourse.

### *Self-selected Discussion Topics*

The self-selected discussion topics were collected from a weekly free-discussion activity in the aforementioned classes. The activity required the students to take turns leading a small group discussion on a topic of their choice. The leader would arbitrarily select a topic that interested the group, prepare three to five relevant questions, lead the discussion, ensure that the members participated in English, and report the results of the discussion to the class. The following is an example of a topic and relevant questions:

Topic: Golden Week

1. What did you do during Golden Week?
2. Did you travel during Golden Week? Where did you go?
3. What was your most memorable event during Golden Week?

Both leaders and members were encouraged to express and exchange their opinions



freely in English. The 16 students in each class were split into four groups of four students. The groups stayed together with the same members for four consecutive free-discussion activities, then new groups were formed for the next round of free discussions. In the end, each student had led a total of eight free discussions across one academic year: four in the summer semester and four in the winter semester.

In total, 367 student-initiated discussion topics were collected over one academic year. The same 13 categories used for the textbook topics were applied to distribute the topics. The categories, example topics, and number of occurrences are listed in Table 3.

Table 3  
*Self-selected Discussion Topics*

Categories	Example topics	Number
School life	School festival, high school, teachers, class schedule, absences, grades, exams, club activities, study abroad, university courses	84
Self	Birthdays, childhood, family, future, dreams, goals, daily routine	63
Places/travel	Golden Week, summer/winter/spring vacation, hometown, travel, commuting, prefectures in Japan	62
Entertainment/fashion	Amusement parks, TV programs, movies, celebrities, hairstyles, fashion	27
Food/health	Breakfast, diet, snacks, sweets, cooking, school cafeteria, restaurants around school	25
Culture	Halloween, Christmas, Japanese New Year, Japanese anime/comics/mascot characters	25
Money/jobs	Part-time jobs, internships, shopping	22
Nature/climate	Seasons, rainy season, typhoons, autumn, winter	19
Leisure	Hobbies, free time, playing sports	13
Social issues	World Cup, Tokyo Olympics, gender roles	11
Communication technology	Smartphones, SNS (social network services)	7
Relationships	Dating, marriage	5
Animals	Animals	4

Siegel (2014) also found topics related to academia and places to be the most frequent in natural conversations outside the classroom between Japanese and foreign university students. Thus, in both monolingual and ELF (English as a lingua franca)

interactions inside and outside the classroom, Japanese students find topics about school life and places to be the most appealing and relevant to their everyday life.

However, one noticeable difference exists between these two different interactions. While the Japanese students in Siegel (2014) frequently took up topics of culture with their international friends and rarely conversed about self-related topics, the Japanese students in this study seldom chose cultural topics for discussion, with only 25 instances, and strongly favored topics about “self.” This indicates that students from predominantly monocultural backgrounds are more willing to discuss personal information and less willing to initiate discussions about cultural topics. Whether students are, as a result, less prepared to successfully communicate with others about their cultures in English is a worthwhile task for future research. In addition, even when the students brought up cultural issues, they showed a clear tendency to linguistically safeguard them in Japanese. For example, when discussing activities for the Japanese New Year, the most popular cultural topic, the students would switch from English to Japanese for cultural customs such as *hatsumode* (first shrine visit of the year), *osechi* (traditional new year foods), *nengajo* (new year greeting postcards), and *otoshitama* (money given by parents and grandparents). Constant efforts to engage students in discussions of local cultural topics are required in order to prepare them for the increasingly frequent ELF interactions that occur around the globe, in which cultural topics frequently arise.

Interestingly, since the data in Siegel (2014) was collected at a university dormitory, topics related to “living situation,” such as Wi-Fi problems and moving out regularly occurred in conversations. Yet these topics were entirely absent in the classroom discussions under investigation. This underscores the context-bound nature of topic selection, as the local context, including the physical environment (dormitory versus

classroom) and the relationship between the participants (dorm mates versus classmates), fundamentally shapes what participants talk about.

Also meriting attention is the fact that the students enjoyed talking about their favorite season, and its associated food, festivals, and natural phenomenon, such as cherry blossoms in the spring and firework displays in the summer. Even though the category of “nature and climate” did not produce a large number of student discussions, the category did identify unique topics, such as typhoons and the rainy season, that were locally oriented and closely tied to the students’ way of life but were almost nonexistent in textbooks.

### *Textbook-assigned vs. Self-selected Topics*

The differences and similarities between textbook-assigned and self-selected topics are shown in Figure 1. This Figure compares the frequency of occurrence of all topics in these two collections. A number of points can be drawn from Figure 1.

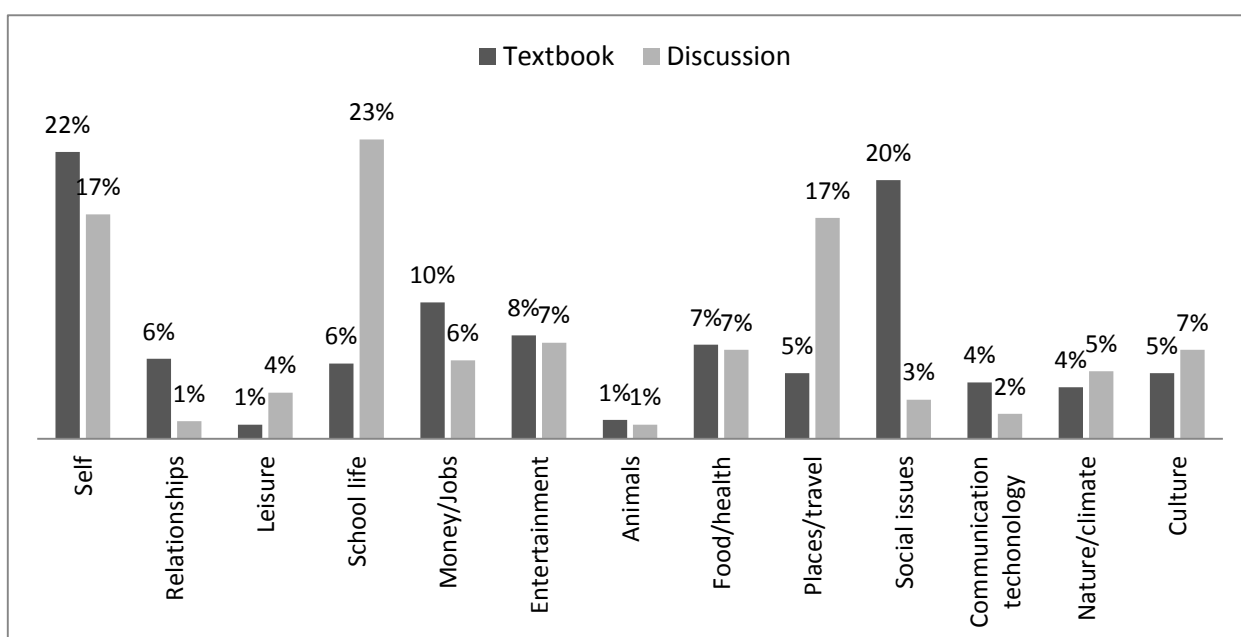


Figure 1. Comparison of textbook and discussion topics.

The most striking difference between the textbook topics and the self-selected topics is the ratio of topics related to school life in the two collections. As seen in Figure 1, while “school life” accounted for only 6% of all textbook topics, it accounted for 23% of the self-selected topics. In other words, topics related to life on campus were ubiquitous in student discussions but relatively scarce in textbooks. In addition, school life produced a rich and varied number of topics ranging from the school festivals to club activities, whereas the school-life topics in the textbooks were limited and superficial. In this regard, the school life presented in the textbooks shows a clear mismatch with the variety of school activities in which the students were engaged. To bridge this gap, it is necessary for textbook writers to give a more accurate and detailed presentation of students’ academic life. Siegel (2014) makes the same observation and suggests that textbooks incorporate more topics, vocabulary, and pragmatic rules that relate to student life on campus. As with the Japanese participants in Siegel (2014), the students in this study also struggled with certain academia-related vocabulary in English, such as course subjects and content, class schedule and credit, and test preparation and items. Increasing the amount and variety of materials on academic life in textbooks is desirable as it emphasizes the relevance of language use in the students’ everyday lives on campus.

Figure 1 also illustrates a marked difference between the textbook and self-selected topics in terms of the ratios of social topics. While social topics comprised 20% of textbook topics, the percentage dropped to a mere 3% in self-selected topics. Moreover, in contrast to the wide range of social issues presented in textbooks, the students’ choice of social topics seemed to suffer from repetition and a lack of depth— exactly opposite the pattern in the “school life” category. Considering the results from both categories, the students seem to strongly favor topics framed in the immediate university context and to

refrain from discussing topics beyond campus life. However, if the course objective is to build critical thinking skills, it is pedagogically beneficial to present current and controversial topics from textbooks, especially from discussion and debate textbooks, in order to intellectually stimulate students and diversify the selection of social issues.

Another obvious discrepancy between the textbook topics and the self-selected topics can be observed in the “places and travel” category. The student discussions contained a considerably higher proportion (17% versus 5%) and a wider variety of topics on places and travel, with most of the instances framed in a local context, such as hometown, prefectures and world heritage sites in Japan, commuting routes to school, and train lines in Tokyo. Although the students were extremely familiar with these topics, they sometimes demonstrated a limited range of vocabulary when attempting to explain them in detail. For instance, while discussing their commute to school, the participating students had difficulty explaining in English different kinds of train tickets and receipts provided by train companies, and, consequently, fell back on Japanese terms—such as *teiki* (commuter pass) and *chien shoumei* (train delay certificate)—as a quick solution to restoring mutual understanding during interaction. Vocabulary items such as these are immediately relevant to their local needs and practices but largely overlooked in textbooks.

The same observation can be made about the “money and jobs” category, where the topic of part-time jobs dominated student discussions. Since many of the students engaged in a variety of part-time jobs after school (some worked as cram school teachers, others train staff, restaurant waiters, etc.), the students expressed great interest in discussing their hourly pay, working hours, job responsibilities, problems at work, and relationships with coworkers and bosses. Given the extensive range of part-time jobs, the

discussions demanded a diverse vocabulary to accurately describe the distinctive responsibilities and interactions in the workplace. Yet the sampled textbooks covered relatively general and superficial career-related topics and showed a lack of depth and comprehensiveness in vocabulary and content. Therefore, it is necessary for teachers and textbook writers to explore the students' personal experiences outside of campus, identify relevant topics, lexical items, and pragmatic conventions, and use the information to shape classroom discussions and textbook discourse. This will not only boost the students' WTC, but also maximize the connection between English learning and learner experiences in local contexts.

In terms of their similarities, self-related topics and universal topics, such as entertainment, food and health, and communication technology, were equally frequent in both datasets. Yet even for these universal topics, the students tended to give them a local twist, with discussions about amusement parks in Japan, favorite foods in the school cafeteria, restaurants around the school, and local cell phone carriers. In contrast, the textbook topics were more general and less locally contextualized.

## **Conclusion**

This study compared the topics presented in English language textbooks in Japan with topics selected by Japanese university students. The results reveal wide discrepancies in terms of the frequency, diversity, and treatment of certain topics. The most dramatic difference between textbooks and topics chosen by students lies in the category of school life. Compared to the broad range of academia-related topics found in student discussions, the dearth of school-life topics in textbooks gives the impression that textbook content is largely detached from the rich student life on campus. To bridge the gap, textbook writers

must present a more comprehensive and detailed description of language use on campus.

Additionally, the analysis indicates that students in the monolingual classroom overwhelmingly choose topics that reflect their local practices and needs across all topic categories. Given the importance of topic familiarity and interest in WTC, local topics effectively facilitate the students' WTC in the L2. The students' tendency to select local topics corresponds to the argument made by Green et al. (1997) and the findings in Wolf (2013). It also supports McKay's (2003a, 2003b) argument for localizing language pedagogy and incorporating locally oriented topics in textbooks in order to meet the students' needs and interests. With locally informed teaching, students will become more capable of expressing their identities and Japanese values when communicating in English with non-Japanese (Hino, 2009). To identify local topics, teachers must become more sensitive to student experiences on and off campus. To improve textbook topics and content in the long run, teachers must avoid becoming passive end-users. Rather teachers should be critical textbook users, critically evaluating textbook discourse and continually providing publishers with feedback (Yuen, 2011).

The analysis also shows that the self-selected topics exhibited a restricted range. Three categories—school life, places and travel, and self—clearly dominated, making up more than half (57%) of the total. The preference for these three categories suggests that topics outside of the students' immediate personal and school circles, such as cultural and social topics, were less frequently selected and quite limited in range. For classes that aim to develop critical thinking or intercultural communication skills, teachers may need to include social topics from textbooks, present supplemental materials, or model example cultural topics in class to redress the imbalance in self-selected topics.

Nevertheless, the self-selected topics identified in this study offer EFL teachers not



only a closer look at the interests and topic preferences shared by students in a predominantly monolingual classroom, but also a repertoire of discussion topics that can be presented in class to promote students' WTC in the L2. The comparison between the textbook and self-selected topics also increases the EFL teachers' sensitivity toward topic choice in relation to their local teaching contexts. Such sensitivity is essential to creating an environment in which students recognize the relevance of language learning and feel secure and motivated to engage in classroom discussions in English.

## References

- Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching*, 40, 21–40.
- Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System*, 34, 480–493.
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.
- Donato, R. (2004). Aspects of collaboration in pedagogical discourse. *Annual Review of Applied Linguistics*, 24, 284–302.
- Green, C. F., Christopher, E.R., & Lam, J. (1997). Developing discussion skills in the ESL classroom. *ELT Journal*, 51(2), 135–143.
- Hino, N. (2009). The teaching of English as an international language in Japan: An answer to the dilemma of indigenous values and global needs in the Expanding Circle. *AILA Review*, 22, 103–119.
- Holec, H. (1981). *Autonomy in foreign language learning* (first published 1979, Strasbourg: Council of Europe). Oxford: Pergamon Press.
- Illés, É. (2012). Learner autonomy revisited. *ELT Journal*, 66(4), 505–513.
- Kang, S. J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System*, 33, 277–292.
- Krashen, S. (1985). *The input hypothesis*. London: Longman.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp.

- 413–468). New York: Academic Press.
- Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207–228.
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K.A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82, 545–562.
- McKay, S. L. (2003a). Toward an appropriate EIL pedagogy: Re-examining common ELT assumptions. *International Journal of Applied Linguistics*, 13(1), 1–22.
- McKay, S. (2003b). Teaching English as an international language: the Chilean context. *ELT Journal*, 57(2), 139–148.
- Pica, T. (1987). Second language acquisition, social interaction, and the classroom. *Applied Linguistics*, 8(1), 3–21.
- Pica, T., & Doughty, C. (1985a). Input and interaction in the communicative language classroom: A comparison of teacher-fronted and group activities. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 115–132). Rowley, MA: Newbury House.
- Pica, T., & Doughty, C. (1985b). The role of group work in classroom second language acquisition. *Studies in Second Language Acquisition*, 7, 223–248.
- Siegel, A. (2014). What should we talk about? The authenticity of textbook topics. *ELT journal*, 68(4), 363–375.
- Smith, R. C. (2008). Learner autonomy. *ELT Journal*, 62(4), 395–397.
- Swain, M. (1996). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 245–256). Oxford: Oxford University Press.
- Talandis G. Jr., & Stout, M. (2015). Getting EFL students to speak: An action research approach. *ELT Journal*, 69(1), 11–25.
- Todd, R. W. (1998). Topic-based analysis of classroom discourse. *System*, 26, 303–318.
- van Lier, L. (2000). From input to affordance: Socio-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245–259). Oxford: Oxford University Press.
- Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, 6(1), 71–90.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wolf, J. P. (2013). Exploring and contrasting EFL learners' perceptions of textbook-assigned and self-selected discussion topics. *Language Teaching Research*, 17(1), 49–66.

- Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31(2), 236–259.
- Yuen, K-M. (2011). The representation of foreign cultures in English textbooks. *ELT Journal*, 65(4), 458–466.
- Zarrinabadi, N. (2014). Communicating in a second language: Investigating the effect of teacher on learners' willingness to communicate. *System*, 42, 288–295.
- Zuengler, J. (1993). Encouraging learners' conversational participation: The effect of content knowledge. *Language Learning*, 43, 403–432.

## Appendix A

### *Sampled Textbooks*

- Browne, C., Culligan, B., & Phillips, J. (2014). *In focus 2*. Singapore: Cambridge University Press.
- Cotton, D., Falvey, D., & Kent, S. (2013). *New language leader intermediate*. Essex, England: Pearson Education.
- Craven, M. (2012). *Breakthrough plus 3*. London: Macmillan Education.
- Day, R. D., Shaules, J., & Yamanaka, J. (2009). *Impact issues 3*. Hong Kong: Pearson Education.
- Dummett, P., Hughes, J., & Stephenson, H. (2013). *Life intermediate*. Singapore: Cengage Learning.
- Flaherty, G. (2002). *Which side are you on?* Tokyo: Seibido.
- Flaherty, G. (2008). *For and against*. Tokyo: Seibido.
- Fuller, D., & Fuller, C. (2011). *Face to face* (second edition). Tokyo: Macmillan Language House.
- Helgesen, M., Brown, S., & Wiltshier, J. (2010). *English firsthand 2*. Hong Kong: Pearson Longman Asia ELT.
- Kenny, T. (2012). *Nice talking with you 2*. Singapore: Cambridge University Press.
- Nishimoto, T., & Hawkins, B. (2009). *Debating current issues*. Tokyo: Seibido.
- McCarthy, M., McCarten, J., & Sandiford, H. (2014). *Touchstone 3* (second edition). Hong Kong: Cambridge University Press.
- McLean, K. (2009). *Topic talk issues* (second edition). Saitama, Japan: EFL Press.
- McLean, P. (2005). *My opinion, your opinion*. Tokyo: Macmillan Language House.
- Paul, D. (2008). *Communication strategies 2*. Singapore: Cengage Learning Asia Pte Ltd.
- Pickering, K. (2012). *Communicate 1*. Oxford: Macmillan Publishers.
- Richards, J. C., & Bohlke, D. (2012). *Four corners 3*. Hong Kong: Cambridge University Press.
- Richards, J. C., & Bohlke, D. (2012). *Four corners 4*. Hong Kong: Cambridge University Press.
- Richards, J. C., Hull, J., & Proctor, S. (2005) *Interchange 3* (third edition). Cambridge: Cambridge University Press.
- Saslow, J., & Ascher, A. (2011). *Top notch 2* (second edition). New York: Pearson Education.
- Sweeney, G. (1999). *Ideas & issues*. Hong Kong: Macmillan Language House.

# Improving Motivations through Autonomous Projects

**Ryan W. Smithers**

Kwansei Gakuin University

## **Abstract**

English as a lingua franca has created motivational dissonances that have resulted in motivation problems with teachers as well as students because English is being taught not as a communication tool, but as a subject that needs validation through student performance on tests. Unfortunately for many teachers, removing standardized testing from the classroom is not an option. This qualitative study on EFL students studying at a tertiary setting demonstrates that in spite of limited flexibility for change, project work allows teachers to make minor changes to their syllabi and yet reap some of the many benefits that are often associated with autonomous learning. The findings here suggest that when given the opportunity to do project work as homework for assessment, students approach their out-of-class learning activities with enthusiasm and derive great pleasure from their chosen tasks. In addition, they become proactive learners who are prepared to invest at least twice as much time studying compared to the amount of time they will spend preparing for teacher-centered assessment via typical standardized testing.

*Keywords:* project-based learning, autonomous learning, motivation, EFL

**T**eachers need to be in the business of empowering students to take control of their own learning because once empowered, students will be better able to: one, harness the power of their curiosity; two, plan for new knowledge; and three, critically evaluate how and what was learnt. This ideology that is synonymous with autonomy theory illustrates why this paradigm is so prevalent in the literature. Unfortunately, research to date has yet to convince educators and institutions to aggressively promote autonomous learning practices like project-based learning (PBL) in the classroom. This study on the effectiveness of PBL at a tertiary setting demonstrates

that for those who are apprehensive about change or need to have standardized testing as part of their curriculum, classroom motivation to learn a foreign language (FL) can be raised and autonomous learning can take place when students have the flexibility to choose their own homework assignments. The results from this study suggest that student homework projects (SHPs) need not have much class time allocated to them to have students and teachers reap the benefits of autonomous learning. Furthermore, SHPs have the potential to re-motivate amotivated students and propel intrinsically motivated students into a flow state—something that is bound to make teaching a lot more rewarding and enjoyable.

### **Research Questions**

While autonomous learning and the kinds of activities that promote it are valued among FL educators today, finding a way to motivate students to take responsibility for their own learning still remains a Pandora's Box. Accordingly, this study seeks determine the efficacy of SHPs to promote autonomous learning as a teaching strategy by addressing the following research questions:

1. Are SHPs autonomy promoting activities that have the ability to empower students to take responsibility for their own learning?
2. Do SHPs have educational merit as a substitute for standardized testing?

### **Background to the Study**

Learning a foreign language is a process that takes much time and effort, with the prerequisite being that learners need to be motivationally equipped to overcome the difficulties that arise during the learning process. Subsequently, the learners most capable

of overcoming the difficulties that come up during the learning of an FL are autonomous learners because they are not only the most motivationally equipped to overcome any and all motivational setbacks that arise during language learning, they are also self-regulated in that they know how to manage their own learning and are effective at utilizing metacognitive strategies (Oxford & Lee, 2008).

Further to this, Boud (1988) sardonically adds that students who do not know how to act independently from their teachers will not likely be any good at learning, nor will they be very employable if they cannot function autonomously. That is to say, it is imperative that students be given the opportunity to learn autonomously, which brings us to the learning paradox: “How do we affect a person by outside influences so that he will not permit himself to be affected by outside influences” (Nelson, 1949, p. 19)? Specifically, how do we create autonomous learners and at the same time have them be autonomous? The answer is the paradox; becoming autonomous and being autonomous cannot be separated.

As it is the purpose of this paper to provide an outline of some simple steps that can be taken so that sound theory can become practice, the background to the study begins with a brief examination of autonomy theory. This will hopefully bring clarity to a confusing concept so that when autonomy is discussed in this paper, the reader will have a clear understanding of what is meant by this term. This paper will then demonstrate why autonomous learning needs to be promoted in the classroom so that the autonomy paradigm can gain greater acceptance in academia, and, hopefully, develop into a universally accepted teaching practice. Finally, the background to the study will introduce the concept of project-based learning as an autonomy-promoting teaching strategy.

### *Autonomy Theory*

#### *Explication*

Ever since Yves Châlon came up with the concept of autonomy in the field of language teaching in 1971, there has been much confusion and criticism about what constitutes autonomous learning (Benson, 2011; Hafner & Miller, 2011). To make it plain what definition of autonomy this paper embraces, and to bring clarity to a confusing concept, this review of the literature examines what exactly autonomous learning for foreign language students is, and it looks at what specifically learner autonomy is not.

Firstly, learner autonomy is not teacher-centered learning. That is, classrooms are not forums for teachers to transmit knowledge to passive vessels. Also, autonomous learning is not self-instruction or distance learning, nor is it a specific methodology or new methodology. So what is learner autonomy? To understand what exactly autonomy is, it is prudent to look at the first universally accepted definition of autonomous learning as a point of reference. The first definition to gain universal acceptance is Holec's (1981) claim that autonomy "is the ability to take charge of one's own learning" (p. 3). Little et al. (2002) have succinctly expounded on this concept by declaring that autonomy involves learner empowerment, learner reflection and appropriate target language use. In other words, learners need to be involved in setting classroom-learning goals, and they need to participate in choosing the contents, rate of progression, and method or technique of study. Furthermore, autonomous learning involves learning from authentic materials because, as McGarry (1995) points out, they have been purposely designed to convey factual information that elicits an affective response from students. Smithers (2013) demonstrated the importance of this in his quasi-longitudinal study of adult EFL learners. His study revealed that cultural interest and proficiency in an FL grow simultaneously.



That is, learners' appetites for movies, print media, and TV programs enjoy a kind of symbiotic relationship with developments in competency.

The concept of autonomy, therefore, necessitates that pedagogical practices become learner-centered, with the goal being to empower students. The focus of classroom learning therefore becomes the active production of a body of knowledge, not the transmission of a body of knowledge. This does not mean that the teacher becomes redundant, but it does create a new role for the teacher. In the autonomous classroom, the teacher becomes the collaborator, with the success or failure of autonomy ultimately linked to the *initiatives* of the teacher. Unfortunately, the dearth of autonomy promoting initiatives found in the literature for inexperienced or pedagogically constrained teachers is likely inhibiting the universal acceptance of autonomy theory.

### *Progression*

Although autonomy theory in language education literature is relatively young, its renown more than makes up for what it lacks in age. In fact, there seems to be widespread agreement on its importance, and its prominence can be seen in second and foreign language education literature, but, unfortunately, pedagogical practices that promote autonomy in the classroom are being opposed by tradition. Subsequently, the universal acceptance of the autonomy paradigm faces many challenges. Firstly, naysayers need to be appeased. These critics of autonomy theory see the promotion of autonomy as impractical and contrary to sensible education (Benson, 2011). This is most likely because they believe education needs to be validated quantifiably through standardized testing. Secondly, Esch (2009), highlighting a study conducted in the UK that examined 558 teachers' pedagogical values and actual assessment practices, stressed that although

teachers value practices that promote autonomy, their practices are to the contrary. Thirdly, Rubin (2008) adds that “although many teachers and texts give a strong nod in the direction of learner-centered learning, changing the paradigm and providing the necessary knowledge and skills for teachers has proven to be quite daunting” (p. 13). Therefore, to help teachers who do not know what initial steps to take to create a learner-centered classroom, or for teachers who cannot abandon standardized testing, this paper proposes project-based learning as a way to revitalize the FL classroom and improve motivations.

### ***Project-based Learning***

Learner autonomy is the result of the interaction that occurs between the teacher and student(s), whereby the teacher’s pedagogical practices promote interdependence, not independence. It follows that the focus has to be on learners and their needs during the learning process. To be precise, learning a foreign language involves not only learning a foreign language, but also learning how to learn, which for many could mean the need for a paradigm shift.

Unfortunately, many educators feel that change and the unknown are best avoided if possible, but taking the first step towards creating an autonomous-learning environment need not be a conundrum. In fact, there are many language-learning activities that can help teachers transition to a learner-centered classroom (see, for example, lists of activities in Benson, 2011; Little et al., 2002). One such activity, “learning by doing” has been an approach to education since the idea was first promoted by John Dewey, the philosopher, psychologist, and educational reformer, one hundred years ago (“Project-based learning”, 2013). Today, Dewey’s approach to pedagogy has been reborn in the

concept of project-based learning. In simplest terms, this pedagogy requires that students research something of interest to them and share what they have learned with their classmates and teacher—usually by way of a presentation.

Learning through projects not only challenges students to represent their newly acquired knowledge, but it also necessitates that they apply what they learn (Smithers, 2014). Furthermore, students engage in the study of authentic material, either collaboratively or individually and learn to excel at four aspects of foreign language learning that are embodied in the concept of self-regulated learning. That is, they develop a strong sense of self-efficacy and are good at using metacognitive, cognitive, and social strategies; the four best predictors of successful language learning (Takeuchi, 2013). Finally, in line with good autonomous learning practices, projects promote learner reflection by concluding with self- and group evaluations. Thus, this paper proposes that teachers start by giving students greater say in what is learned out of class by giving students the opportunity to study autonomously via SHPs.

## **The Study**

### *Subjects*

This study qualitatively examines the effect that SHPs had on students studying Basic English at a four-year engineering university in Japan. Two atypical fourth-year students were selected for interviews. One intrinsically motivated participant (Student A) was atypical in that he already had enough credits to graduate, but felt that his English skills were insufficient for what lay ahead of him once he graduated. The other, an extrinsically motivated student (Student B), was at the other end of the spectrum. He, as a result of a lack of motivation and effort during his freshman, sophomore, and junior years, found

that he was lacking the required credits to graduate. Specifically, he found that he needed to pass an English language course.

### *Procedures*

As mentioned above, this study involves an atypical sampling of this demographic because atypical samplings, also known as critical case samplings, are known for their comprehensive and scrupulous representations of phenomena being investigated (Dörnyei, 2007). This means, in addition to being wide-ranging and thorough, the data will be, most importantly, relevant.

The interview, conducted jointly with both students, sought to discover the merit of implementing a PBL activity like an SHP in lieu of the kind of mid-term exam that is typical of a teacher-centered learning pedagogy—specifically, an activity that would not take up too much class time, but would allow for autonomous learning.

At the start of the year, during class orientation, the students were approached about the idea of substituting their mid-term exam for a PBL activity. At that time, the students enthusiastically requested that they be permitted to do SHPs and evidence their learning by making individual presentations in front of the class on topics of their choosing. Thus, the students and the teacher mutually agreed on a rubric (see Appendix A) that was utilized by the teacher and fellow students so that the students could be brought into the evaluation process as stakeholders. As well as assessing their peers' presentations, they also conducted self-evaluations in order to engage in self-reflection because these two activities have been proven to promote autonomy and improve learner motivation (Brown, 2004), in addition to promoting higher order thinking (Cheng & Warren, 2005). The final score the students received was based on an aggregation of the student's self-evaluation,

the teacher's evaluation, and peer evaluations.

## **Results and Discussions**

### *Evaluations*

The evaluations, especially the self-evaluations, revealed some interesting results. There was a noticeable trend for students to evaluate themselves more critically than their peers. For example, both Student A and B evaluated themselves at 64 per cent, while their peers scored them at 92 and 76 per cent respectively. In the literature, low self-confidence and being overly critical of self has been viewed as a trait common to the Japanese (Heine et al., 1999; Kitayama & Markus, 2000), but this does not mean that this trend is immutable. In hindsight, the students likely could have benefited from a lesson on assessment, which may have helped to lessen the disparity between self- and peer evaluations. Accordingly, it is recommended that assessment training be given sufficient consideration during project orientation to help reduce scoring disparities.

In addition, the students were all influenced by the perceived complexity of the presentations, which also stresses the importance of assessment training. Student A's presentation on 'How to Appreciate White Wine' was very complicated and overly technical. In fact, much of the terminology used to explain about wine and wine making (i.e. astringency, maceration, tannins, etc.) would have likely been new to anyone lacking exposure to such terminology. As a result, Student A's presentation left his peers dumbfounded. Not only was this presentation perceived as having a high level of English fluency in the face of structural, grammatical, and pronunciation errors, but it was the first presentation and subsequently became the standard by which all other presentations were judged. This is especially apparent in light of Student B's presentation on 'Tips for Tennis.'

His presentation was grammatically and structurally much more sound and easier to understand, with his presentation receiving the highest mark in the class by the teacher, but a mean of only 76 per cent from the group, while his self-evaluation was 64 per cent.

Of note, one student in the class who clearly did not make a passing effort was not afraid to let his self-evaluation reflect his lack of effort. He rated his performance 40 per cent and was scored 46 per cent by his peers, which demonstrates the effectiveness of learner empowerment. This student who made a poor effort was able to reflect critically on his learning, and he accordingly accepted responsibility for the outcome.

### *Interview*

In regards to the interview, Students A and B were subject to various questions about the PBL assignment. The interview took about 25 minutes to complete, with the students being given the freedom to answer in their mother tongue or in English. The following is the portion of the interview that is most pertinent to this study:

- Teacher: How did you feel about doing this presentation?  
Student A: Very tough!  
Student B: Very tough!  
Teacher: How much time did you spend preparing?  
Student A: Two weeks.  
Student B: I first wrote my speech in Japanese, then translated it to English... ten hours.  
Student A: I spent about twenty or thirty hours. I had too much information to go through.  
Teacher: Did you study more for the presentation than a mid-term test?  
Student A: YES!  
Teacher: How much time do you usually spend preparing for a mid-term test?  
Student A: For a class in Basic English... less than 5 hours.  
Teacher: How about for other subjects?  
Student A: Five minutes.

- Teacher: You don't study very much.
- Student A: No, but I get perfect marks.
- Student B: For a mid-term test I usually spend about two or three hours in preparation.
- Teacher: Which is more beneficial to learning, a typical mid-term test or a presentation like this?
- Student A: Absolutely this presentation!
- Student B: Absolutely! This was fun!
- Teacher: Any other comments about this project?
- Student A: It's good.
- Student B: It was fun. I have retained what I studied.

Some important information can be gleaned from these questions and answers. Firstly, looking at the students' responses to the first question reveals that they both possess a strong sense of self-efficacy. They both felt that the SHPs they had chosen were very challenging, yet because they both took ownership of their respective projects, they formed a strong sense of commitment to their projects and conjured up the appropriate self images to see themselves through to the successful completion of the task, demonstrating that recent findings into the interrelation of motivation, identity and autonomy are valid (see, for example, Gao & Lamb, 2011; Ushioda, 2013). In addition, due to the fact that self-efficacy continues to develop throughout an individual's life and is greatly enhanced when a difficult task is completed (Bandura, 1977), the successful completion of a SHP can therefore be seen as an effective way of strengthening one's self-efficacy so that future challenges will be met with a higher probability of success. This is especially important in light of Smithers' (2013) recent findings that drew attention to motivational rollercoaster ride that L2 learners experience during their initial years of learning.

From a more pragmatic angle, the data also reveals that SHPs create in students a desire to learn. Student A revealed that in regards to typical English mid-term exam

preparations he was prepared to spend up to five hours preparing, whereas with this project he was willing to invest between twenty to thirty hours over the course of two weeks to prepare for his presentation. In fact, his deep enjoyment and total commitment to his SHP could be likened to what Csikszentmihalyi (1990) calls a flow state; a state in which emotions and intrinsic motivation are combined to produce the optimal learning experience for a student. Student B, who started the class unmotivated, also demonstrated a similar pattern of motivated behavior. He declared that because he enjoyed what he was studying, he was willing to spend three times as much time acquiring the knowledge necessary to complete this project than he would have if he had been preparing for a typical mid-term test, and he felt that he retained what he had studied. These findings validate similar findings from another study on PBL conducted in a tertiary setting with EFL students by Matsubayashi and Kawamura (2013). Their findings revealed that students enjoyed working on projects of their choosing because they were motivated to learn, and, as a result, the students also discovered the cliché that the more you learn, the more you realize how much you don't know; and the more you realize how much you don't know, the more you will want to learn.

### **Limitations and Future Research**

The results from this study were limited in that conclusions could not be made about why the students tended to be overly critical of themselves and their abilities during self-evaluations. Is this truly a culturally specific phenomena? Also, is this something that can be remedied by assessment training so that self-assessment scores can be a better indicator of true abilities? Furthermore, research is needed to determine if assessment training can prevent student bias towards perceived language proficiency.



## Conclusion

Teachers who are in agreement with the theory of autonomy in principle, but are not sure what initiatives to take, or are lacking confidence to transition out of a traditionally expository style of teaching may find that SHPs are a viable first step. As examined above, SHPs are a fruitful teaching strategy that can re-motivate amotivated students and propel motivated students into a flow state.

Furthermore, the results above concur with a growing trend in the literature that shows how project work, like SHPs, allows students to be empowered in the learning process so that students: one, commit more effort to learning; two, take responsibility for the ultimate success or failure for their learning; and three, owing to a strengthening of students' senses of self-efficacy, become more motivationally secure to tackle increasingly complex learning tasks.

## References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215. doi:10.1037/0033-295X.84.2.191
- Benson, P. (2011). *Teaching and researching autonomy* (2nd ed.). England: Pearson.
- Boud, D. (1998). *Developing student autonomy in learning* (2nd ed.). Oxon: Taylor & Francis.
- Brown, D. (2004). *Language assessment: Principles and classroom practice*. New York, NY: Longman.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93-121. doi: 10.1191/0265532205lt298oa
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: HarperCollins.
- David, J. L. (2008). What research says about project based learning. *Educational Leadership*, 65(5), 80-82. Retrieved from [http://www.ascd.org/publications/educational\\_leadership/feb08/vol65/num05/Project-Based\\_Learning.aspx](http://www.ascd.org/publications/educational_leadership/feb08/vol65/num05/Project-Based_Learning.aspx)
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Qualitative, qualitative*

- and mixed methods*. Oxford: Oxford University Press.
- Esch, E. (2009). Crash or clash? Autonomy 10 years on. In R. Pemberton, S. Toogood, & A. Barfield (Eds.), *Maintaining Control: Autonomy and language learning* (pp. 27-4). Hong Kong: Hong Kong University Press.
- Fenner, A. (2000). Learner autonomy. In A. Fenner & D. Newby (Eds.), *Approaches to materials design in European textbooks: Implementing principles of authenticity, learner autonomy, cultural awareness* (pp. 78-139). Austria: Council of Europe. Retrieved from <http://archive.ecml.at/documents/materials.pdf#search='annebrit+Fenner+learner+autonomy'>
- Gao, X., & Lamb, T. (2011). Exploring links between identity, motivation and autonomy. In G. M. Xuesong & T. Lamb (Eds.), *Identity, motivation and autonomy in language learning* (pp. 1-8). Bristol: Multilingual Matters.
- Heine, S. J., Markus, H. R., Lehman, D. R., & Kitayama S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106(4), 766-94.
- Holec, H. (1981). *Autonomy and foreign language learning*. Oxford: Pergamon Press.
- Kitayama, S., & Markus, H. R. (2000). The pursuit of happiness and the realization of sympathy: Cultural patterns of self, social relations, and well-being. In E. Diener & E. M. Suh (Eds). *Culture and Subjective Well-being* (pp. 112-161). Cambridge: The MIT Press.
- Little, D., Ridley, J., & Ushioda, E. (2002). *Towards greater learner autonomy in the foreign language classroom*. Dublin: Authentik.
- Hafner, C. A., & Miller, L. (2011). Fostering learner autonomy in English for science: A collaborative digital video project in a technological learning environment. *Language Learning & Technology*, 15(3), 68-86.
- Matsubayashi, Y., & Kawamura, A. (2013). *Project-based learning as a source of motivation*. Paper presented at the 40th (2013) JACET Summer Seminar, Kusatsu, Gunma.
- McGarry, D. (1995). *Learner autonomy: The role of authentic texts*. Dublin: Authentik.
- Nelson, L. (1949). *Socratic method and critical philosophy: Selected essays*. New York, NY: Dover Publications.
- Oxford, R. L., & Lee, K. R. (2008). The learners' landscape and journey: A summary. In C. Griffiths (Ed.). *Lessons from good language learners* (pp. 306-317). Cambridge: Cambridge University Press.
- Rubin, J. (2008). Reflections. In C. Griffiths (Ed.). *Lessons from good language learners* (pp. 10-15). Cambridge: Cambridge University Press.
- Project-based learning. (2013). Wikipedia, the free encyclopedia. Retrieved August 30,

2013, from [http://en.wikipedia.org/wiki/Project-based\\_learning](http://en.wikipedia.org/wiki/Project-based_learning)

Smithers, R. W. (2013, August). *The ebb and flow of a language learner's motivational profile: A quasi-longitudinal study*. Paper presented at the 40th (2013) JACET Summer Seminar, Kusatsu, Gunma.

Smithers, R. (2014). Using projects to increase autonomy, motivation, and language learning. *Osaka JALT Journal*, 1, 19-32.

Takeuchi, O. (2013). *Motivation, autonomy and metacognition: A self-regulatory approach to L2 learning*. Paper presented at the 40th (2013) JACET Summer Seminar, Kusatsu, Gunma.

Ushioda, E. (Ed.). *International perspectives on motivation: Language learning and professional challenges*. New York, NY: Palgrave Macmillan.

## Appendix A

### *Presentation Rubric*

	Name: _____
<b>Student spoke only in English</b>	<b>/10</b>
<b>Pronunciation</b>	<b>/10</b>
<b>Voice and eye contact</b>	<b>/10</b>
<b>Presentation was interesting</b>	<b>/10</b>
<b>Presentation was easy to understand</b>	<b>/10</b>
<b>Total</b>	<b>/50</b>

# The Potential Impact of Guessing on Monolingual and Bilingual Versions of the Vocabulary Size Test

**Kurtis McDonald**  
Kobe College

## **Abstract**

This study explores the potential impact of guessing on monolingual English and bilingual Japanese versions of the Vocabulary Size Test (VST; Nation & Beglar, 2007) in order to better understand the estimates that these tests provide. Eight first-year Japanese university students took 140-item versions of the VST modified to include an ‘I don’t know’ option and, immediately after completion, participated in brief semi-structured retrospective interviews about the reasoning behind the answers they selected on a sample of self-identified unknown items. Findings indicate these learners employed a large amount of both informed and uninformed guesses and that their guesses contributed to higher vocabulary size estimates, with estimates garnered from the Japanese version of the test showing a greater range of difference. The findings also suggest that informed guesses are guided by both partial knowledge and test strategy use, with potentially useful finer distinctions able to be identified.

*Keywords:* guessing, partial knowledge, test language, test strategy use, vocabulary knowledge, Vocabulary Size Test

**G**iven the key role that second language (L2) vocabulary knowledge plays in L2 language proficiency (Nation & Webb, 2011), the ability to accurately measure this knowledge is clearly of importance to both learners and their instructors. The importance of vocabulary size, in particular, can be seen in the research into vocabulary load and coverage requirements for unassisted reading comprehension. Following Hu and Nation’s (2000) determination that knowledge of at least 98% of the

words in an authentic written text is likely required for unassisted reading, Nation's (2006) analysis of the vocabulary loads of authentic written texts demonstrated that knowledge of 8,000 to 9,000 of the most frequently used English word families would be required to read newspapers and novels without support. For instructors, understanding how L2 learners' vocabulary sizes compare to such benchmarks can be used to inform decisions on the suitability of texts for language study as well as frequency levels of words to target for study (Nation, 2013; Nguyen & Nation, 2011).

The desire to adequately measure vocabulary size has spurred the development of many different tests to this end over the years, each with its own advantages and disadvantages (for a more detailed discussion, see Milton, 2009). The Vocabulary Size Test (VST; Nation & Beglar, 2007) is one test that has received a great deal of attention since its introduction. Intended to efficiently measure written receptive knowledge of the most frequently used 14,000 word families of English of the spoken section of the British National Corpus (BNC), the original version of the VST tests knowledge of 140 randomly sampled words, 10 from each 1000-word level. Each word is presented in a simple, non-defining sentence along with four possible definitions as multiple-choice options. Vocabulary size estimates are produced by multiplying the total number of correct answers by 100. Although Beglar (2010) presented convincing Rasch-based validity evidence for the original VST and Nation (2013) considered it to be "proven, reliable, and very practical" (p. 37), recent literature has identified a number of factors that users of the test should consider for the test's validity, reliability, and practicality to be maximized in any particular testing context. Among the most discussed factors have been the language in which the answer options should be presented (see Elgort, 2013; Karami, 2012; McDonald, 2014; Nguyen & Nation, 2011; Stewart, 2009) and the degree to which

guessing is likely to be involved (see Stewart, 2014; Zhang, 2013).

The potential value of bilingual versions of the VST is one area that has rightly received a great deal of attention in recent years. Bilingual versions present the answer options in the first language (L1) of the test takers which, it is thought, allows for learners of lower proficiency levels to better demonstrate any knowledge they have of an L2 word without calling on the reading and grammatical skills required when the definitions are presented only in the L2 (Nguyen & Nation, 2011). Although there are many factors to consider, studies by Nguyen and Nation (2011), Karami (2012), and Elgort (2013) have outlined many of the key characteristics of well-constructed bilingual versions of the VST and Stewart (2009), Elgort (2013), and McDonald (2014) have shown that lower proficiency learners do generally receive higher scores on bilingual versions of the VST than on monolingual versions.

As a four-option, multiple-choice test, the potential influence of guessing is another area of research that has been pursued in the literature. Although the VST was intentionally created as a measure sensitive enough to detect partial and intuitive knowledge (Nation, 2012; Nation & Webb, 2011), Stewart (2014) has questioned the degree to which random, uninformed guessing might contribute to overestimations of vocabulary size. To address such a concern, Zhang (2013) suggested that the use of an ‘I don’t know’ option in concert with announced penalties for wrong guesses can improve reliability, though he also advised that the testing purposes should ultimately inform the degree of sensitivity sought.

Although valid, reliable, and practical measures of vocabulary size are needed, previous research has offered more questions than answers in terms of these concerns with the VST. By employing a mixed methods approach, this study seeks to provide initial

answers to the following research questions as they relate to a comparison of the bilingual Japanese VST with the original monolingual VST with a small sample of Japanese university EFL learners:

1. How much do informed and uninformed guesses contribute to learners' vocabulary size estimates on each version of the test?
2. What kind of information is used when learners make informed guesses?

### **Method**

This small-scale, mixed methods study was carried out in order to follow up on previous research by the author, outlined in McDonald (2014), which quantitatively compared the responses of 133 Japanese first-year students from various departments at a small, private women's college in western Japan on full versions of both a bilingual Japanese VST (administered in April) and the original monolingual English VST (administered in September). In the current study, eight volunteers out of the original 133 participants agreed to take the VST again in late January and to submit to semi-structured retrospective interviews about the reasoning behind the answers they selected on a sample of self-identified unknown items afterward. All of the participants had recently completed one academic year of coursework with the author as their instructor at the time this research was conducted.

### ***Participants***

Table 1 lists the participants' pseudonyms, majors, average TOEIC Institutional Program (IP) scores, April bilingual Japanese VST scores, September monolingual VST scores, and the version of the test each was administered for this follow-up study conducted in



January.

Table 1

*Summary of Participant Information*

Participant	Major	TOEIC (IP) Average	Japanese VST (April)	English VST (Sept.)	Assigned VST (Jan.)
Sayo	English	625	83	71	Japanese
Shoko	English	598	81	78	English
Chika	English	595	83	76	English
Rio	Bioscience	443	83	69	Japanese
Ayami	English	440	74	65	Japanese
Yuri	Intercultural Studies	353	77	79	English
Reina	Bioscience	325	77	62	English
Natsuko	Music	308	65	55	Japanese

*Note.* The TOEIC (IP) scores listed represent the average from two administrations of the test conducted in April and November. VST = Vocabulary Size Test,  $k = 140$ .

***Instruments***

Paper versions of the original monolingual English VST and the parallel bilingual Japanese version of the same 140-item test (translated by Sasao and Nakata in 2010) downloaded from Paul Nation’s website were used in this study. However, both versions were modified to include a fifth answer option of ‘I don’t know’/‘わかりません’ [*wakarimasen*] on all items. The following example illustrates the same item from the 14,000-word-family frequency level as expressed on each version of the test.

Monolingual English version

LIMPID: He looked into her **limpid** eyes.

- a. clear
- b. tearful
- c. deep brown
- d. beautiful
- e. I don’t know

Bilingual Japanese version

LIMPID: He looked into her **limpid** eyes.

- a. 澄んた [*sunda*]
- b. 涙でいっぱいの [*namidade ippai no*]
- c. こけ茶色の [*kogecha iro no*]
- d. 美しい [*utsukushii*]
- e. わかりません [*wakarimasen*]

### *Procedure*

Besides the addition of the 'I don't know' option, important changes were also made to the testing procedure for the January administration of the test, which was conducted in three passes. Unlike the previous test administrations in April and September, which encouraged students to guess on items that they were unsure of, the participants of the current study were explicitly instructed to select the 'I don't know' option for any items they did not know confidently, thus, constituting their first pass through the test. Following completion of the test under these conditions, participants were then instructed to return to all of the items which they had answered with 'I don't know' and to select what they felt to be the best answer from the original four options by circling it in red, constituting a second pass through the test. After completing the test again under these conditions, the participants were asked to return to all of the guessed items again and to circle the item numbers for all of the items which they had chosen answers through completely uninformed, random guessing for a third and final pass. All participants were given as much time as needed to complete the test under these conditions.

After completing the modified test in three passes, each participant was interviewed individually about the reasoning and thought processes that influenced their second pass selection of guesses on items that they had originally identified as unknown. Due to time constraints, each interview was limited to a selection of the highest frequency guessed items, ranging from 20-40 items per individual. The questions were asked in English though participants were informed that they could respond in English or Japanese.

### **Results**

The responses on the modified versions of the tests administered in January were analyzed

to take into account the three levels of self-perceived knowledge that the testing procedure sought to distinguish. First, all of the participants' responses (from both the first and second passes through the test) were checked to establish a *score with all guesses*. This score represents the total score that each participant would likely have gotten if they had been initially encouraged to answer all items as they had done in previous administrations of the tests in April and September. Next, all of the responses to items identified as completely uninformed, random guesses during the participants' third pass through the test were removed so that scores based only on responses which the participants used self-perceived knowledge or informed guessing could be tabulated as each individual's *score with only informed guesses*. Finally, all guesses were removed from scoring, leaving only the participants' initial first-pass responses (based only on what was thought to be known information), to tabulate each learner's *score without any guesses*. The scores garnered from each method of analysis are presented in Tables 2 and 3 below along with the participants' scores from the two previous times they took the VST in the same academic year for the sake of comparison.

Table 2

*Scores for Each Participant on Previous Administrations of the VST and the Modified Japanese VST*

Participant	Japanese VST (April)	English VST (Sept.)	Modified Japanese VST (Jan.)		
			Score With All Guesses	Score With Only Informed Guesses	Score Without Any Guesses
Sayo	82	72	90	73	56
Rio	82	70	84	62	56
Ayami	75	64	89	63	49
Natsuko	65	55	73	61	36
Average	76	65.25	84	64.75	49.25

*Note.* VST = Vocabulary Size Test,  $k = 140$ .

Table 2 shows each participant’s scores from earlier administrations of the tests as well as the three scores able to be tabulated from the modified ‘I don’t know’ version of the test administered in January. As you can see, the first student listed, Sayo, received an 82 on the Japanese version of the test in April, a 72 on the English version administered in September, and a 90 on the modified Japanese version in January when all her responses were included. When all of her self-identified random guesses were removed, her score was 73. After all guesses were removed from scoring, her score fell to 56. Similar trends can be seen in the scores of the three other participants who took the modified Japanese version.

Table 3  
*Scores for Each Participant on Previous Administrations of the VST and the Modified English VST*

Participant	Japanese VST (April)	English VST (Sept.)	Modified English VST (Jan.)		
			Score With All Guesses	Score With Only Informed Guesses	Score Without Any Guesses
Shoko	80	77	75	62	53
Chika	84	75	73	62	49
Yuri	77	79	70	56	44
Reina	76	61	63	57	37
Average	79.25	73.00	70.25	59.25	45.75

*Note.* VST = Vocabulary Size Test,  $k = 140$ .

Table 3 presents the same matrix of information for the four participants who took the modified English version of the VST in January. The first student listed, Shoko, received an 80 on the Japanese version of the test in April, a 77 on the English version administered in September, and a 75 on the modified English version in January when all her responses were included. When all of her self-identified random guesses were removed, her score was 62. After both her informed and uninformed guesses were

completely removed from scoring, her score fell to 53. Again, comparable trends can be seen in the scores of the other participants listed.

A simple comparison of the different scores able to be compiled when all responses are counted, when only informed guesses are counted, and when no guesses of any kind are counted demonstrates the great disparities among the participants' scores on the modified versions of both tests administered in January. These differences are particularly striking when the vocabulary size estimates are calculated by multiplying each score by 100 as the design of the VST stipulates. For the four participants who took the modified bilingual version of the test, the vocabulary size estimates garnered range from an average of 8,400 word families when all responses are counted, to an average of 6,475 when answers based on self-professed knowledge and informed guesses were added together, to an average of 4,925 when only answers based on self-professed knowledge were scored, thereby eliminating all self-identified guesses of any kind. Expressed in a different way, the participants who took the modified bilingual version of the test could increase their scores by an average total of 41% through guessing, with roughly 23% attributable to informed guessing of some kind and 18% due to random chance through completely uninformed guessing. The average scores from the four participants who took the modified English version of the test followed a similar trend, though the overall range was not as large. When all correct answers on the modified English version were counted, the average vocabulary size estimate was 7,025 word families for these participants, compared to 5,925 when limited to items answered based on self-professed knowledge and informed guessing and 4,575 when all guesses were excluded. In terms of percentages, the participants who took the modified English version of the test could increase their scores by an average total of 35% through guessing, with

16% of that attributable to some kind of strategic guessing and 19% due to uninformed guessing.

Figures 1 and 2 present visual representations of the trends for each of the three scoring methods for both versions of the modified tests administered in January. As with the overall mean scores, the mean scores at each frequency level vary quite substantially and this disparity generally gets larger as word family frequency decreases, especially for the modified version of the bilingual Japanese test.

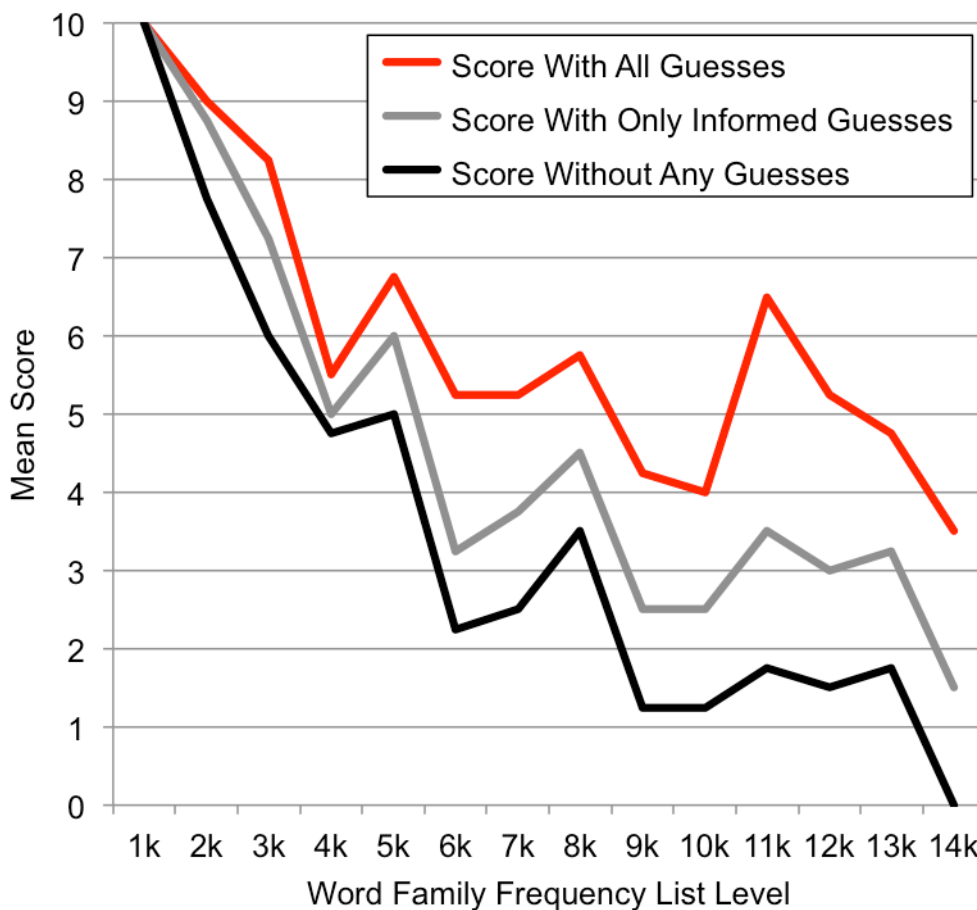


Figure 1. Mean number of correct responses at each frequency list level for each method of scoring the modified bilingual Japanese Vocabulary Size Test ( $n = 4, k = 140$ ).

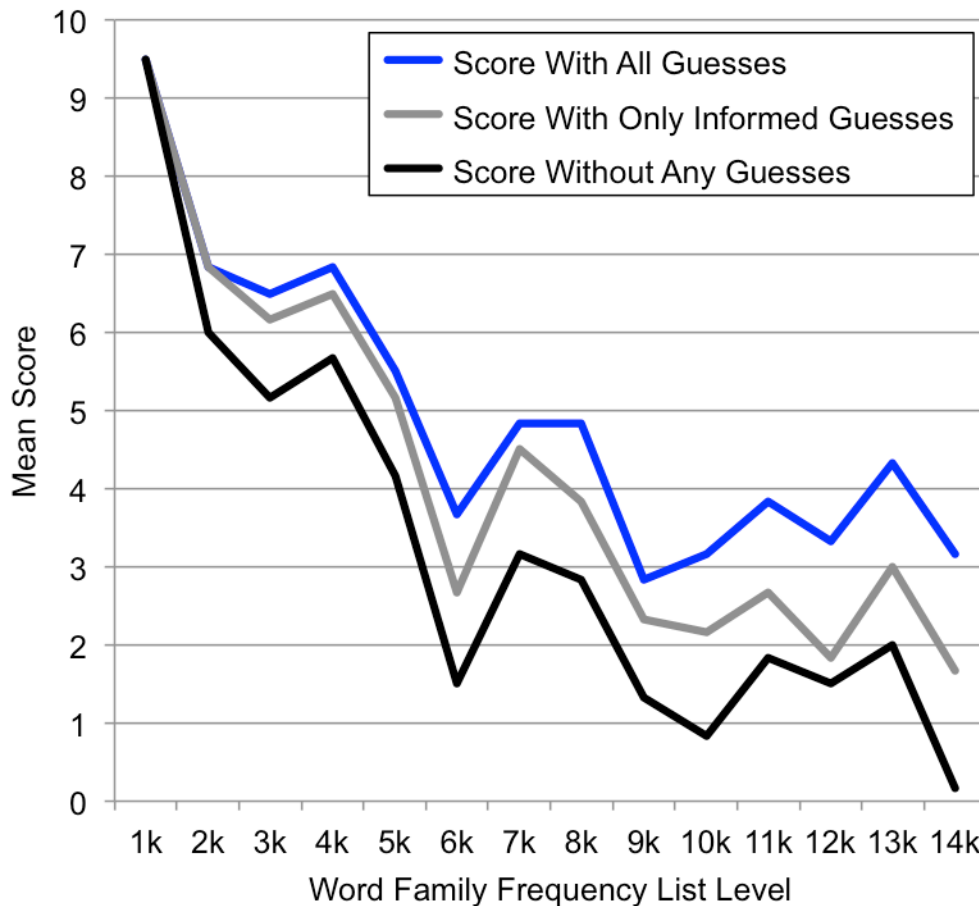


Figure 2. Mean number of correct responses for each frequency list level for each method of scoring the modified monolingual English Vocabulary Size Test ( $n = 4$ ,  $k = 140$ ).

The follow-up interviews conducted with the participants immediately after completing the modified versions of the VST administered in January revealed that a wide range of information was called upon to help learners make informed guesses. In particular, information that can be attributed to partial knowledge and test strategy use was clearly identified in these interviews. In what follows, a few of the finer distinctions within each of these categories are illustrated with quotations from the learners; the words in brackets have been added for clarity.

### *Partial knowledge-informed guesses*

Several degrees of partial knowledge can be identified in the interview data. One interesting distinction that can be drawn here is between true and false partial knowledge. Two excerpts from Sayo's interview might help illustrate the chief distinction here. When asked to explain the reasoning behind her correct guess of the word 'lonesome,' Sayo's reply shows that *true partial knowledge* of a word partly guided her response.

Interviewer: ...Why did you pick C [孤独な (*kodokuna*)]?

Sayo: Umm, lonely, so, from lonely, I guess C.

Interviewer: Mm-hmm. You know the Japanese word means lonely but you don't know this word?

Sayo: Yeah.

However, with her informed guess of the word 'jug,' she employed *false partial knowledge* that, nonetheless, still directed her to select the correct answer.

Interviewer: ...Why did you pick A [水入れ (*mizu-ire*) ]?

Sayo: ...so, uh, 'jug', uh, 'jug-cuzzi' ...uhh, so *ofuro/jug-cuzzi* [Jacuzzi].

Interviewer: I see.

Sayo: Yeah.

Interviewer: So you had a different word...

Sayo: Yes, I thought water so I choose, chose A.

Another key distinction able to be identified among all informed guesses based on partial knowledge concerns responses which learners made based on "a feeling," "an image," or "an impression;" what could all be classified as *intuition*. An excerpt from Chika's interview illustrates how learners' intuitions can help narrow answer options in a way that might be helpful, even if it doesn't always lead to the correct answer. This can be seen with reasoning she employed with 'nil.'



Interviewer: ...Why did you pick A [very bad]?

Chika: *Eto*, uh, in my opinion, [the letter] 'N' have bad image, so I chose "very bad".

However, faulty intuitions can just as easily lead learners astray, as can be seen in the explanation Chika offers for her incorrect answer for 'jug.'

Interviewer: ...Why did you pick D [a weapon that explodes] for 'jug'?

Chika: *Eto*, huh, mmm...uh, my image, 'jug' is weapon.

### *Test strategy-informed guesses*

The data uncovered through the interviews also makes it clear that conscientious learners utilize a number of test strategies as they strive to identify what they consider to be the best answers to items with words that are completely unknown or only partially known. An example from Yuri's interview shows how even the ostensibly decontextualized sentences in which target words are presented can yield clues which learners can use to guide them to certain answers. Here, Yuri explains how she used the item sentence stem provided for the word 'drawer,' "The drawer was empty," to inform her correct answer selection.

Interviewer: ...Why did you pick A [sliding box]?

Yuri: Uh, 'empty' is '*karappo*'...so... 'box'.

Other instances show how learners replace the target word with the answer options in order to judge their potential suitability and eliminate distractors from the selection process, as the following excerpt from Shoko's interview illustrates.

Interviewer: The next guess was [Third 1000] #9 [rove] and you picked A [getting drunk]...

Shoko: Uh, from the sentence.

Interviewer: He couldn't stop roving?

Shoko: I cut B and D.

Interviewer: I see. Why did you cut those?

Shoko: Mmm. I don't think he couldn't stop "working hard"...haha. I think 'roving' is negative.

## Discussion

### *The contribution of informed and uninformed guesses to vocabulary size estimates on each version of the test*

Examining the three very different vocabulary size estimates garnered from the administration of the modified versions of VST illustrates just how much of an impact guessing can have on the scores that result. Indeed, the estimates varied dramatically for the modified bilingual version of the test, from an average high of 8,400 word families assumed to be known when all correct guesses were included to an average low of 4,925 word families when all correct guesses were excluded, a difference of 3,475 word families. Viewed in terms of Nation's (2006) vocabulary load and coverage figures, the high end of these estimates suggests vocabulary sizes sufficient for reading authentic newspaper articles at 98% coverage while the low end implies that these learners would require perhaps several more years of study before being able to read authentic newspaper articles without assistance. The estimates from the modified English version also varied substantially from an average high of 7,025 word families assumed to be known when all correct guesses were counted to an average low of 4,575 word families known when they were not, a difference of 2,450 word families.

Of more particular interest regarding the construct of written receptive vocabulary knowledge are the differences in scores when items answered correctly due to informed guessing of some kind are distinguished from those answered correctly due to random, uninformed guessing on both versions of the test. Of the average word family increases

in the estimates garnered from the bilingual Japanese version when all correct guesses are counted, 1,925 word families can be attributed to some form of informed guessing on the Japanese version, compared to the 1,100-word-family increase found with the English version. These figures can also be compared to the degree of successful self-professed uniformed guessing determined to be roughly 1,550 word families on the Japanese version and 1,350 on the English version.

The moderate differences exhibited between the two versions suggest that the participants were able to employ more partial and strategic knowledge to respond correctly to items on the bilingual version of the VST than on the monolingual version. This finding would seem to support the view that the bilingual Japanese version likely does measure vocabulary size with a finer degree of sensitivity than the monolingual English version, at least with learners of certain proficiencies. However, further analysis of the bilingual Japanese version is needed before its validity can be assumed.

### *The information used when learners make informed guesses*

The interview data collected in this study suggests that a wide range of reasoning, based on both accurate and inaccurate assumptions related to vocabulary knowledge, intuition, and testwiseness, underpinned the informed guessing employed by the learners interviewed. Closer analysis of the interview data revealed that the participants utilized both *partial knowledge* and *test strategies* extensively when making informed guesses on both versions of the VST. Within both of these general categories, several finer distinctions could be identified. *True partial knowledge*, *false partial knowledge*, and *intuition* were found to inform guesses based on *partial knowledge* while context clues interpreted from the item sentence stems, answer option substitution, and distractor

evaluation/elimination were seen as evidence of the use of *test strategies*.

### **Limitations**

Several limitations of this study are worth noting. First, the eight participants who volunteered to participate in this study might not be representative of the broader population, especially of students who strongly dislike English or lengthy English tests. Next, the bilingual Japanese version of the VST used in this study has not undergone the same degree of rigorous piloting and validation that the monolingual English version has. Finally, there were several shortcomings in the way the interviews were conducted. These shortcomings include general problems in communicating questions and answers due to language barriers between myself and the learners, a lack of follow-up questions into how well items and answer options were understood and what combination of information may have been used to narrow choices before guesses were ultimately made, and the incompleteness in the interview data due to the decision to limit the discussion of guessed items by time rather than ensuring that all guessed items were covered.

### **Conclusion**

This study found evidence which suggests that a large amount of guessing can be expected by Japanese university EFL students of lower to intermediate English proficiency levels taking full, 140-item bilingual Japanese and monolingual English versions of the VST when they are compelled to answer all items. A comparative analysis of the test and interview data collected in this study further intimates that guessing may have more of an impact on increasing scores on the bilingual Japanese version of the test, presumably due to the learners' abilities to better understand and evaluate the answer options in their L1.

While this suggests the bilingual Japanese version may indeed be a more sensitive measure of vocabulary size, more in-depth research with a wider range of learner proficiency levels represented is needed.

The preliminary investigation into the kind of information used when making informed guesses conducted in this study allowed for finer distinctions of informed guessing to be identified as true partial knowledge, false partial knowledge, and intuition. More robust research in this area, as well as that of test strategy use, is also called for. Although there are several limitations to this study, it is believed that the findings it suggests may help spur on further research into the range of factors that may affect the administration of measures of English vocabulary size with native speakers of Japanese. In the meantime, it is hoped that this study, along with the growing body of research on the VST, can at least make users of this test more aware of the range of issues and options involved in administering it and in analyzing and interpreting the vocabulary size estimates it can produce.

### **Acknowledgements**

This study was supported by a research grant from the Kobe College Research Institute. I would like to thank the learners who volunteered their time as well as Paul Nation and Eton Churchill for their insightful comments on earlier drafts of this paper.

### **References**

- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27(1), 101–118. doi:10.1177/0265532209340194
- Elgort, I. (2013). Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing*, 30(2), 253–272. doi:10.1177/0265532212459028

- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Karami, H. (2012). The development and validation of a bilingual version of the Vocabulary Size Test. *RELC Journal*, 43(1), 53–67. doi:10.1177/0033688212439359
- McDonald, K. (2014, September). *How do monolingual and bilingual versions of the Vocabulary Size Test compare?* Paper presented at the 18th Annual Conference of the Japan Language Testing Association, Ritsumeikan University, Shiga, Japan.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol, UK: Multilingual Matters.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82. doi:10.1353/cml.2006.0049
- Nation, I. S. P. (2012). *Vocabulary Size Test Information and Specifications*. Retrieved from <<http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf>>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31, 9–13.
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.
- Nguyen, L. T. C., & Nation, I. S. P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal*, 42(1), 86–99. doi:10.1177/0033688210390264
- Stewart, J. (2009). A comparison of test scores between monolingual and bilingual versions of the Vocabulary Size Test: A pilot study. In A. M. Stoke (Ed.), *JALT2008 conference proceedings*. Tokyo: JALT.
- Stewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly*, 11(3), 271–282. doi:10.1080/15434303.2014.922977
- Zhang, X. (2013). The I don't know option in the Vocabulary Size Test. *TESOL Quarterly*, 47(4), 790–811. doi:10.1002/tesq.98

# Using L2 Blogs to Promote Student Participation and Critical Thinking

**Gilbert Dizon**

Himeji Dokkyo University

**Chris Edelman**

Ritsumeikan University

## **Abstract**

The authors studied the effects of student blogging on participation in two English for specific purposes (ESP) courses designed for first-year university medical students. The students were divided into two online groups: a class blog (CB) group where learners were encouraged to read and respond to others' posts, and a personal blog (PB) group in which students wrote and reflected individually. Both groups followed a similar procedure of writing a timed blog post in class and then were assigned to either write a response to another blog post or write a follow-up to their own entry from class as a homework assignment. Participation, or the number and frequency of homework entries written, was found to be significantly higher in the CB group, with 75 percent of the students participating actively. Whereas in the PB group, participation was marginal, with 80 percent of the students either not participating or sporadically participating. Furthermore, the authors found that online discussions via blogs fostered critical thinking by giving the learners opportunities to express and support their opinions as well as read and understand multiple views. The results of the study highlight the importance of social interaction in L2 writing and reinforce existing research on the effect that it has on L2 computer-mediated communication (Kitade, 2000; Ortega, 1997; Wu, Yen, & Marek, 2011).

*Keywords:* L2 writing, computer-assisted language learning, blogging

**I**t is important for L2 instructors to incorporate activities that promote participation while directing students toward becoming more autonomous learners. Given this, more and more language teachers are integrating technology into their curricula in

an attempt to reach this goal. However, technology alone is not the answer. How technology is implemented plays an even greater role, as Stockwell (2012) declares, “Technology does not inherently facilitate language learning, but rather, it is how technology is used that dictates whether or not language learning occurs through its use” (p. 87). With that in mind, this study sought to compare the effectiveness of class and personal blogs.

## **Background**

### ***L2 Blogging***

Research on L2 blogging has shown that it can have a positive impact on L2 writing development (Armstrong & Retterer, 2008; Vurdien, 2012; Wu, 2006). However, rather than providing a comprehensive overview of the advantages of blogging over traditional writing activities, the authors have determined four specific benefits which are relevant to this study: 1) decreased anxiety, 2) improved motivation, 3) increased learner autonomy, and 4) enhanced critical thinking.

Many studies have demonstrated the influence that learner anxiety and perceived L2 competence can have on students’ willingness to communicate (Baker & MacIntyre, 2000; MacIntyre, 1994; Yashima, 2002). Therefore, it is critical for L2 learners to be comfortable whenever engaging in a communicative task. According to Miura and Yamashita (2004), blogging provides students with a safe outlet to express themselves, thereby leading to empowerment and learning. Citing a study by Nagel and Anthony (2009), Nepomuceno (2011) reinforces this claim, “blogs can potentially boost the affective domain of learners, making them perceive writing not as a dreaded class activity but as a means to help express emotions and ideas,” (p. 97).



In a study involving six L2 French learners, Kissau, McCullough, and Pyke (2010) found that online second language learning could decrease learner anxiety as well as promote willingness to communicate, especially among weaker students. Giving students time to prepare their responses, lower-level learners were able to understand the topic of discussion and contribute more than in face-to-face settings where interactions take place synchronously. In other words, the online interactions “leveled the playing field” (Kissau et al., 2010, p. 286) by enhancing the confidence of the less proficient students and giving them equal opportunities to communicate.

At the motivational level, Lee (2010, 2011) notes that blogs help foster positive attitudes by encouraging learners to read and respond to other student entries. In Lee’s (2011) study, 16 American university students participated in Spanish language blogs while studying abroad for a semester in Spain. The participants of the study contributed to three types of blogs: personal blogs, a class blog, and a project blog which incorporated both teacher-assigned and free-topic discussions.

Lee (2011) found that many of the learners worked together to construct knowledge and critically reflect upon their experiences abroad. Therefore, not only were they able to develop their writing skills, but they also were able to gain a sense of community through blogging. Similarly, Dizon (2013) illustrated that a class blog can develop into an online community, thus supporting L2 writing through collaborative dialogue.

Increased learner autonomy is another oft-cited benefit of blogging. Iida (2009) explains that blogging is an effective task for L2 students, “not simply because it provides additional opportunities to study the language, but because it encourages them to be autonomous” (p. 6). Likewise, Lee (2011) found that blogging fostered learner autonomy through self-regulation and self-management. However, Little (1996) suggests that social

interaction is critical in order for learner autonomy to be developed. As a result, it is important to place learners in contexts where they can meaningfully communicate with others.

Mynard (2007) looked at the effect that blogging had on learner autonomy in a study involving 26 female Japanese EFL students studying abroad in the UK. Although participation in the personal blogs were completely voluntary and no guidelines were given in terms of frequency or content, a majority of the students (85%) regularly wrote on their blogs over a 12-week period. According to the results of her study, Mynard concluded that the blogs helped the learners reflect on their own learning and provided opportunities for them to become autonomous learners.

According to Zhang (2009), blogs “foster critical thinking by encouraging students to evaluate what they read and write” (p. 68). In a study involving 41 L2 English learners, Baker and Ismail (2009) found that the participants viewed blogging as a useful and interesting way to improve their critical thinking skills. Fola-Adebayo (2014) also examined students’ perceptions on the impact that blogging can have on the development of critical thinking skills. One hundred questionnaires were randomly administered to students who were enrolled in an English for academic purposes (EAP) course that incorporated blogging at a university in Nigeria. Seventy-six students responded to the questionnaire, with over three-fourths of them (76%) indicating that blogging helped them reflect and think critically about that they had learned in class. Based on the results of the study, Fola-Adebayo (2014) determined that the learners in the study were able to improve their thinking skills through a process of reflection, analyzing, and bonding.

### ***Learner Participation and Behavior***

The degree of participation students exhibit when using CALL is dependent on a variety

of factors. As Desmarais, Duquette, Renié, and Laurier (1998) discovered, language proficiency may play a large role in learner participation and behavior. In their study, they tracked the behavior of students as they progressed through a multimedia program the researchers developed called Vi-Conte. Rather than proceeding through the program as the developers intended, Desmarais et al. (1998) determined that low-proficiency learners tend to engage in disorganized behavior, i.e., non-linear progress, compared to learners with higher levels of proficiency. Moreover, the authors found that college-age students, such as the ones involved in the present study, were more likely to display disorganized behavior as opposed to adult students. As a result, younger, less proficient students may experience more difficulty using technology for language learning purposes, thus resulting in lower levels of participation.

Technical capability also affects how students behave in CALL environments. While today's students may be proficient in daily computer tasks such as emailing, using search engines, as well as interacting via social networks, they may not be able to apply the same skills during CALL tasks. In a survey involving 911 university students in basic foreign language programs, Winke and Goertler (2008) concluded that many students lack the proficiency with or access to the specialized tools that are essential for CALL, including working with advanced features in word-processing programs and maintaining a website. They referred to this disparity as the personal versus academic/professional computer literacy divide. Therefore, sufficient learner training is vital to successful implementation of CALL in the classroom.

### **Research Question**

Before deciding to implement CALL in the L2 classroom, teachers must determine whether or not the activity is appropriate for the learners involved. In other words, would

using CALL foster language learning in a meaningful way or would a traditional activity suffice? Therefore, the authors' goal in this study was to discover the impact that class blogging had on the students' participation, as defined by the number and frequency of homework posts completed, compared to reflective writing via individual student blogs.

### **Participants**

The participants were composed of two intact ESP classes of first-year medical students at a university in western Japan. Both the class blog (CB) and personal blog (PB) groups met once a week at different times on different days using the same room. Both classes were taught by the same instructor using the same syllabus, materials and activities, apart from the types of blogs, which was the only differentiation between the two classes. The study included a total of 67 participants, 29 in the CB group and 36 in the PB group. While the learners were able to grasp the main ideas from most reading activities in the textbook, e.g. medical charts and nurse-patient conversations, the teacher/researcher observed that proficiency levels of both groups were quite low in regards to oral production, with most students struggling to speak in the L2 when role-playing various hospital scenarios or discussing medical issues. As a result, blogging was an effective task because it gave the students opportunities to meaningfully participate in the classroom. This is significant because some learners may not feel comfortable actively participating in a traditional classroom environment. As Ortega (1997) argues, “students that do not normally participate much in traditional classroom discussion seem to dramatically increase their participation in the electronic mode” (p. 85).

### **Methodology**

Students were given the same initial blog activity in each class: writing their opinion

about a topic after watching a video and/or reading a short article on an individual blog page. The topics that were chosen were the ones that ran parallel with course material. For example, if students were studying cancer as part of their curriculum, the topic of smoking was used as a blog activity. This was done to provide students with the language terms and ideas necessary to reflect upon and respond to the topic in the English language.

Blogger was chosen as the website to be used for this task. It is a free blog hosting service offered by Google. Students were introduced to the website from the first class and made clearly aware of the outlines of the coursework. Students used randomly assigned numbers instead of names in order to maintain anonymity. This design was chosen to help lower student anxiety by “enabling learners to take risks in communication that they may be less willing to take if their actual identity is known” (Stockwell & Tanaka-Ellis, 2012, p. 77).

During the first three weeks, students were given explicit training and provided with opportunities to familiarize themselves with the functions and features of the site because, as Hubbard and Romeo (2012) state, learner training is key to student success with CALL tools. “For students to be effective users of a computer tool or learning application, they must first understand how to operate it and then become comfortable with its operation” (Hubbard & Romeo, 2012, p. 45).

Student participation was tracked in both groups, which helped to identify the students who were not actively participating in the blogs. These students were approached about their non-participation. Specifically, the students from the PB group were reminded to post comments about the class video or reading, while the students from the CB were encouraged to post responses to their classmates' comments. After the third week, no feedback or reminders on participation were given to the students. It was after these three

weeks of training and practice that the researchers started collecting data from the blogs for this study. Therefore, based on the definition put forth by Burchfield and Sappington (1999) who describe participation as “the number of unsolicited responses volunteered” (p. 290), any contributions made to the blogs after this point can be deemed as a form of participation. As Rocca (2010) notes, participation may come in different forms, such as student questions or responses, and may “take a few seconds or an extended period of time” (p. 187). While quality is important, Rocca (2010) further states that it is much more subjective, and, as a result, researchers and instructors are provided with quantitative measures that can be more clearly defined.

Only 10 minutes were assigned for in-class writing on the blogs as the classes had to maintain a strict schedule in regards to advancing through the curriculum. Students were instructed that upon entering the classroom, they should turn on their computers and create a new blog page for posting. After this, monitors were turned off and students were given time for the topic activities. Upon finishing the activities, the instructor asked students to turn on their computers and type their reactions/opinions to the topic. The use of electronic dictionaries was discouraged, but allowed, with an emphasis placed on writing fluency and output. The amount of time remaining was called out at every minute mark so that students were aware of the time constraint to finish the activity.

The blog procedure was exactly the same in both classes. The difference was that the students in the PB group were unable to view any post other than their own and were given the homework assignment of continuing, expanding upon, or revising their original blog post because the time constraints made it difficult for them to write detailed posts in class. Therefore, the activity allowed the students to reflect upon their writing in order to correct any errors and/or write any additional opinions on the topic that they were could

not complete in class. Below are examples of an unedited in-class post and homework post by Student 46. As the excerpts illustrate, there was a marked improvement in word count and depth between the in-class post with the homework post.

- *In-class post*: I think that Nuclear Energy Radiation is dangerous. Because Nuclear Energy Radiation give us healthy damage. So, We should stop Nuclear Energy.
- *Homework post*: Radioactivity is terrible. It affect to people, plant and animals. People will be sick on radioactivity. Plant is wither and vegetables don't grow normally. Animals will be sick and people can't eat them. I think that must be take measure against radioactivity of Fukushima nuclear power plant quickly.

The CB group had three subgroups, two consisting of ten students and one consisting of nine. Students were able to view and comment on the posts of the other students they shared a subgroup with. Students from all subgroups were given a homework assignment of reading and responding to the post of at least one other student. It was posited by the authors that the CB group would show the highest level of participation due to the socially interactive nature of the homework assignment. Hester and Francis (2004) define social interaction as “any situation in which a person produces an action addressed or directed towards another and/or which invites or makes possible a response from another” (pp. 3-4). Based on this definition, the class blogs were an effective medium for social interaction because they encouraged the students to share their opinions and respond to each other’s posts.

## Results and Discussion

Table 1 below shows the participation patterns (i.e., the number and frequency of

homework posts written by the students in the class blogs). Four types of participation were identified: non-participation, sporadic participation, active participation, and full participation. Non-participation is defined as a complete lack of follow-up comments or replies during the nine-week study. Sporadic participation is characterized by more than one follow-up post but less than seven. Active participation is defined as a total of seven or eight follow-up posts, with one entry posted weekly at least seven out of the nine weeks. Full participation is defined as a minimum of one entry each week over the course of the study.

Table 1

*Class Blog Participation*

	Students	%	# of posts	%
Non-Participation	2	6.9	0	0
Sporadic Participation	6	20.7	26	11.2
Active Participation	13	44.8	123	52.8
Full Participation	8	27.6	84	36.0
Total	29	100	233	100

As Table 1 illustrates, active participation or full participation was demonstrated by nearly three-fourths (72%) of the students in the class blogs. Non-participation was fairly low, with only two students (7%) not posting in any of the nine weeks. Six students fell into the sporadic participation category, accounting for approximately 20% of the total number of students doing the class blogs. Also notable was the fact that the CB group wrote more than twice the number of posts as the PB group (233 posts vs. 88 posts), even though the CB group had fewer students (29 students vs. 36 students).

The class blogs encouraged not only participation, but also critical thinking.



Although there is no authoritative definition of critical thinking, Long (2003) created a list of abilities that are important to the process (based on Mayfield, 2001). The following unedited exchange represents some of those skills, such as the capacity to “prepare persuasive arguments using evidence” (p. 231), as illustrated by Student seven’s ability to support an argument for euthanasia using the learner’s own personal experience. Moreover, Student six sympathized with and ultimately accepted an opposing view. In other words, the writer was able to “understand multiple perspectives” (p. 231).

S7: Assisted suicide is OK. Because I think that it is help peole by the pain. My grandfather is died by the lung cancer. He is very painful. I did not see him. I think that people want to no pain died. So I agree this thinking.

S2: I agree. I think this idea is Ok. Your grandfather was very hard... Suffer from pain, is unpleasant.

S6: I thought not good Assited Sucide. Becouse Assisted Sucide is help by the pain but Assited Sucide is kill human. But I read your opinion and my opinion was changed.

The anonymity provided by the blogs also allowed the students to share strong opinions with one another, even if those views were in disagreement with others. As Horwitz, Horwitz, and Cope (1986) assert, negative evaluation from peers is a major source of anxiety among L2 learners. However, according to Pasfield-Neofitou (2012), the anonymity often provided in computer-mediated communication (CMC) can encourage students to exchange their thoughts with their peers. They posit that “anonymity in CMC can have positive effects, such as giving learners more confidence to participate than they may have in face-to-face communication” (p. 101). As the unedited example below illustrates, this was evident in the class blogs where the learners were able to openly disagree with each other without fear of negative or upsetting

responses from the other blog participants:

S28: I think abortion is ok. Because bring up child is not easy. If women is poor, they may not enough child care. That is not happy for child.

S26: I don't think so. Because if have an abortion, we lose one life. So, we have to give this problem more careful consideration.

Table 2

*Personal Blog Participation*

	Students	%	# of posts	%
Non-Participation	21	58.4	0	0.0
Sporadic Participation	8	22.2	29	33.0
Active Participation	4	11.1	31	35.2
Full Participation	3	8.3	28	31.8
Total	36	100	88	100

Despite having more participants, active participation was significantly lower in the PB group (see Table 2), with only four students posting between seven to eight follow-up comments. Similarly, considerably fewer students (three) participated every week. As a result, non-participation and sporadic participation made up for 80 percent of the participation in the PB group. These results correspond with Lee's (2010, 2011) findings that social interaction promotes participation by encouraging students to respond to other's blog posts.

These results show that the class blogs promoted student participation by allowing the learners to communicate in a meaningful setting with their peers. While participation in the PB group was quite low, participation in the CB group was high, with the majority of the students either actively or fully participating in the weekly homework assignments.

In addition, due to the anonymity provided by the blogs, the students in the CB group were able to improve their critical thinking skills through online discussions without having to worry about negative evaluations from their peers.

### **Conclusion**

Class blogging is clearly more effective than a traditional task such as individual reflective writing. Furthermore, the class blogs helped support critical thinking by giving the students opportunities to share their experiences, support their views, and critically reflect on a variety of pertinent topics with others. However, as previously mentioned, the students experienced difficulties using the blogging software for language learning purposes. Even simple tasks such as logging in proved problematic for some, which led to the training period being extended two additional weeks beyond the initial orientation class for the blog. Therefore, discovering how learner training might affect student participation in a CALL environment would be a worthwhile avenue of study.

This study has several limitations. It is based solely on a group of medical students at one university in western Japan. Also, questionnaires were not administered to gain a more complete picture of how social interaction via the class blogs affected participation. Moreover, additional factors such as technical capability, Internet access, and gender may have influenced the results of the study. Therefore, future research could employ surveys along with other methods to better measure the effect of social interaction on student participatory patterns.

This study underscores the importance of social interaction on student participation and contributes to previous findings on the positive impact that it can have in the context of L2 CMC (Kitade, 2000; Ortega, 1997; Wu, Yen, & Marek, 2011). Although the learners

received instruction from the same teacher with the same materials and activities, it was the collaborative element of the blogs that fostered participation in one group, while a lack of social interaction influenced poor participation in the other.

## References

- Armstrong, K., & Retterer, O. (2008). Blogging as L2 writing: A case study. *AACE Journal*, 16(3), 233-251.
- Baker, N. A., & Ismail, K. (2009). Using blogs to encourage ESL students to write constructively in English. *ASEAN Journal of Teaching and Learning in Higher Education*, 1, 45-57.
- Baker, S., & MacIntyre, P. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning*, 50, 311-341.
- Burchfield, C. M., & Sappington, J. (1999). Participation in classroom discussion. *Teaching of Psychology*, 26, 290-91.
- Desmarais, L., Duquette, L., Renié, D., and Laurier, M. (1998). Evaluating learning and interactions in a multimedia environment. *Computers and the Humanities*, 31(3), 327-349.
- Dizon, G. (2013). Class weblog as an online community: Collaborative learning through blogging. *The Kyoto JALT Review*, 1, 45-63.
- Fola-Adebayo, T. (2014). Promoting metacognitive learning using blogs in an L2 learning community. *Journal of Language and Literature*, 24, 64-80.
- Hester, S., & Francis, D. (2004). *An Invitation to Ethnomethodology Language, Society, and Interaction*. London, UK: SAGE Publications.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70, 125-132.
- Hubbard, P., & Romeo, K. (2012). Diversity in learner Training. In G. Stockwell (Eds.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 33-48). New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139060981
- Iida, A. (2009). Research in weblog pedagogy: Blogging and developing learner autonomy in a JFL context. *The Language Teacher*, 33(2), 3-7.
- Kissau, S., McCullough, H., & Pyke, J. G. (2007). "Leveling the playing field" The effects of online second language instruction on student willingness to communicate. *CALICO Journal*, 27(2), 277-297. Retrieved from

[https://calico.org/html/article\\_793.pdf](https://calico.org/html/article_793.pdf)

- Kitade, Keiko. (2000). L2 learners' discourse and SLA theories in CMC: Collaborative interaction in Internet chat. *Computer Assisted Language Learning*, 13(2), 143-166.
- Lee, L. (2010). Fostering reflective writing and interactive exchange through blogging in an advanced language course. *ReCALL*, 22(2), 212-222.
- Lee, L. (2011). Blogging: Promoting learner autonomy and intercultural competence through study abroad. *Language Learning & Technology*, 15(3), 87-109.
- Little, D. (1996). Freedom to learn and compulsion to interact: promoting learner autonomy through the use of information systems and information technologies. In R. Pemberton, S. L. Edward, W. W. F. Or, & Pierson, H. D. (Eds.), *Taking control: Autonomy in language learning* (pp. 203–219). Hong Kong: Hong Kong University Press.
- Long, C. (2003). Teaching critical thinking in Asian EFL contexts: Theoretical issues and practical applications. *Proceedings of the 8<sup>th</sup> Conference of Pan-Pacidific Association of Applied Linguistics*. Retrieved from <http://www.paaljapan.org/resources/proceedings/PAAL8/pdf/pdf022.pdf>
- MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, 12, 241-247.
- Mayfield, M. (2001). *Thinking for Yourself: Developing Critical Thinking Skills through Reading and Writing* (5th ed.). United States: Thomas Learning.
- Miura, A., & Yamashita, K. (2004). Why do people publish Weblogs? An online survey of Weblog authors in Japan. In K. Morgan, C. A. Brebbia, J. Sanchez, & A. Voiskounsky (Eds.), *Human perspectives in the Internet society: Culture, psychology and gender* (pp. 43–50). Southampton, England: WIT Press.
- Mynard, J. (2007). A blog as a tool for reflection for English language learners. *The Asian EFL Journal*, 24, 1-10, Retrieved from <http://asian-efl-journal.com/teaching-articles/2007/11/05/a-blog-as-a-tool-for-reflection-for-english-language-learners/>
- Nagel, D. M., & Anthony, K. (2009). Writing therapy using new technologies—the art of blogging. *Journal of Poetry Therapy*, 22(1), 41-45.
- Nepomuceno, M. M. (2011). Writing online: Using blogs as an alternative writing activity in tertiary ESL classes. *TESOL Journal*, 5, 92-105.
- Ortega, L. (1997). Processes and outcomes in networked classroom interaction: Defining the research agenda for L2 computer-assisted classroom discussion. *Language Learning & Technology*, 1(1), 82-93.

- Pasfield-Neofitou, S. E. (2012). *Online Communication in a Second Language: Social Interaction, Language Use, and Learning Japanese*. Bristol, UK: Multilingual Matters.
- Rocca, K. A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education*, 59(2), 185-213.
- Stockwell, G. (2012). Introduction. In G. Stockwell (Ed.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 1-13). New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139060981
- Stockwell, G., & Tanaka-Ellis, N. (2012). Diversity in environments. In G. Stockwell (Ed.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 71-89). New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139060981
- Vurdien, R. (2012). Enhancing writing skills through blogs in an EFL class. *The EUROCALL Review*, 20, 155-158.
- Winke, P., & Goertler, S. (2008) Did we forget Someone? Students' computer access and literacy for CALL. *CALICO Journal*, 25(3), 482-509.
- Wu, W.S. (2006). The effect of blog peer review and teacher feedback on the revisions of EFL writers. *J. Education and Foreign Languages and Literature*, 3(2), 125-138.
- Wu, W.-C. V., Yen, L. L., & Marek, M. (2011). Using online EFL interaction to increase confidence, motivation, and ability. *Educational Technology & Society*, 14(3), 118-129.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86(1), 54-66.
- Zhang, D. (2009). The application of blog in English writing. *Journal of Cambridge Studies*, 4(1), 64-72.

# Standardized Testing in University

**Damien Healy**

Osaka Institute of Technology

## Abstract

This paper will look at a growing trend in Japanese universities where first and second year students are required to study TOEIC listening and reading skills in compulsory oral English classes. With globalization a key word for Japanese universities nowadays, the idea of teaching first and second year students an internationally recognized standardized test will seem appealing to more Japanese universities. Therefore, a critical look at standardized testing, which is what the TOEIC test is, will be given. It will be shown that both validity and reliability are essential if a test is to be considered accurate in what it is meant to do. The TOEIC test has a degree of reliability, but the test lacks validity. Other problems associated with this test will be described such as teachers being forced to teach to the test and it will be posited that such passive skills used in the test do not help students improve their active skills. The paper will also show that TOEIC literature has overestimated the tests abilities, showing that TOEIC scores are not as accurate as they appear to be. It is suggested that if Japanese students are to become truly global citizens, universities need to allow students to spend more time learning and practicing active English activities in compulsory oral English classes.

*Keywords:* TOEIC, standardized testing, validity, reliability

**T**ests are of the utmost importance to help make rational educational decisions because they not only provide us with information about students' abilities, but they also allow us to measure students' proficiencies in a subject, discover how much of a courses' objectives students have mastered, and see students' strengths and weaknesses, which allow us to work with students on areas that they need improvement on. Consequently, this begs the question of which test to use in compulsory EFL classes at a university. Obviously, the type of test that is chosen will depend on the kind of information being sought.

Powers (2010), suggests that the *Test of English for International Communication* (TOEIC) is a valid test in that it “measures a person’s ability to communicate in English in the context of daily life and the global workplace environment using key expressions and common, everyday vocabulary” (pp. 1-2). Based on this statement by Powers, it can be inferred that this test is designed for both business and lay people, and that it will help all people communicate through English. However, Wilson (1993), informs us that the TOEIC test was designed to measure workers’ English language skills used in international corporations, not lay people. Bresnihan (2013) posits that the majority of students who have not worked in international business will find the contexts and contents difficult to understand, and this will lead to students feeling demotivated in their oral English classes. Long (2014), concludes that even students with high TOEIC scores are not ready to communicate orally in rapid and complicated interactions and negotiations, suggesting that the TOEIC Listening and Reading test does not help students improve their active communicative abilities, which an oral English class should be helping students acquire and improve.

This paper concurs with Wilson, Bresnihan, and Long’s findings which suggest that the TOEIC Listening and Reading test might not be the best use of time for a compulsory first year oral English university class. Furthermore, this paper demonstrates that the TOEIC test lacks validity. It is not a test of communicative ability, only one of passive comprehension, which means that for teachers who are forced to teach to the test, the amount of time given to useful communicative exercises in the classroom is reduced. In addition, Cihl and Culligan (2011) tell us that the TOEIC test uses low frequency vocabulary which students will probably not encounter during their four years at university, and, therefore, Nation and Waring (1997) suggest students will not be able to



retain them by the time they graduate. Furthermore, by the time first grade students have graduated, the test scores will no longer be valid as TOEIC scores are only considered valid for two years.

So as to bring clarity to the complexity of standardized testing, this paper begins with a brief history of testing. This is followed by an examination of standardized testing and the importance of test validity and reliability. Due to the high acceptance TOEIC enjoys as a standardized test in Japan, this paper continues its examination of standardized testing by scrutinizing the TOEIC test in relation to its validity and reliability to determine the impact of TOEIC testing at Japanese universities.

### **History of Testing**

The earliest known assessment program, according to Odendahl (2011), was carried out in 210 B.C.E. by the Han dynasty for “civil service examination that assessed both scholarship and military skills” (p. 2). These tests continued with many changes until 1905. Over time, the difficulties of testing—such as cheating or gaming the system—were discovered. In Europe, testing was often done orally. Then, at the University of Cambridge in the mid-1700s, a written examination was undertaken, and in the late 1790s a numerical score was introduced as opposed to the older system of either pass or fail. In 1845, Horace Mann, suggested introducing written examinations with an allotted time, making the examination more objective and helping to make teachers accountable for their students’ results. In the mid-1890s, Dr. Rice introduced statistical methods into the scrutiny of curricula and teaching methods.

According to Odendahl (2011), Galton and Pearson were interested in “using quantitative techniques to classify observations of traits and investigate human

development systematically” (p. 5). They used distribution curves, percentiles, the computation of correlations, and the concept of statistical significance; all for the quantitative study of material.

Between 1900 and 1917, schools in the United States embraced the idea of measuring students’ intelligence because many students were failing existing school tests. With this tool, students could be streamed into different classes according to their IQ test scores, making it more efficient for teachers and students to succeed. In 1914, Frederick J. Kelly introduced a method to make reading tests easier to score, less time-consuming, and more precise; these were multiple choice and short-answer items. These methods improved efficiency and reduced costs and subjectivity in scoring. However, some opponents to these standardized tests suggested that they, “restricted assessment to factual information and thus led to curricular emphasis on superficial memorization” (Odendahl, 2011, p. 13). Nowadays, this occurs in Japanese junior and senior high schools, as standardized tests are widely used in English lessons, thus restricting students’ abilities to communicate in English. In fact, standardized tests are used in most educational systems worldwide today and can be referred to as norm-referenced tests. Some of these tests, such as the TOEIC test, do not help with communicative language acquisition because, as Canale and Swain (1980) inform us, there needs to more to a test than simply measuring grammatical competence; there also needs to be a sociolinguistic and strategic competence component.

### **Standardized Tests**

Standardized tests are the quintessential norm-referenced tests, and, according to Bachman (1990), are comprised of three characteristics:

First, standardized tests are based on a fixed, or standard content, which does not vary from one form of the test to another. Second, there are standard procedures for administering and scoring the test, which do not vary from one administration of the test to the next. Finally, standardized tests have been thoroughly tried out, and through a process of empirical research and development, their characteristics are well known (p. 74).

We can see how students perform on these types of test by contrasting their score with the scores of a given group, or norm. The norm group (NR) is a large group of individuals who are similar to the individuals for whom the test is designed. The NR is given the test and then the characteristics or norms of this group's performance are used as reference points for interpreting the performance of other students who take the test.

Dörnyei (2007) tells us that “if we have a sufficiently big sample size, the characteristics of the people in the group will approach a very special pattern termed ‘normal distribution’” (p. 27). Within the sample, a few people will attain a very high or low value with most of the people clustered around the middle or average range. This gives us a bell-shaped curve, and it has been shown that the greater the sample, the more normal the distribution and regular the curve becomes. Dörnyei (2007) goes on to tell us that with the bell-shaped curve “it has unique properties upon which it is possible to build a whole range of mathematical procedures that have led to the development of statistics” (p. 28).

Standardized tests measure students' general educational skill development in learning outcomes on a national basis, and they highlight the strengths and weaknesses of students and facilitate comparisons between students' general level of achievement and ability. However, they do not measure broader analytical, evaluative, communicative or persuasive skills. These are the skills that should be taught at university, as students have very little experience using these skills in junior and senior high school. Also,

standardized test results can be misused, for example, by branding a university or teacher as below standard (Shortt, 2014). Thus, it is important to carefully consider to what concepts are key for choosing a test.

### **Key Concepts of Testing**

Hughes (1989) explains that tests should be created which “consistently provide accurate measures of precisely the abilities in which we are interested” (p. 6). To achieve this goal, a test has to have validity—construct validity, criterion-related validity, and content validity—and reliability.

#### ***Construct Validity***

Bachman & Palmer (1996) explain that this “pertains to the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores” (p. 21). We must be able to justify the interpretations we make about test scores; not simply argue that they are valid. To reach this goal, we must be able to demonstrate that the test scores relate to the area that we intend on testing and very little else. So, to show this evidence, we must define the construct we want to measure. According to Hughes (1989), a construct is an ability or a trait which is hypothesized in a theory. It is a specific definition of an ability that provides the basis for a given test or test task and for interpreting scores derived from this task. Therefore, construct validity refers to the extent to which we can interpret a given test score as an indicator of the abilities we want to measure. Mackey and Gass (2005) suggest “construct validity can be enhanced when multiple estimates of a construct are used” (p. 108). So, to have only one construct in a test can possibly skew the test score interpretations. Also, construct validity test score interpretations should be

able to be generalized so that not only interpretation of the present cohort is possible, but the results can be generalized for other cohorts too.

### *Criterion-Related Validity*

Criterion-related validity is when we compare the test scores in our test with a well-established test which is measuring the same construct. If there is a good correlation between the two sets of scores, we can then say that our test has been demonstrated to have criterion-related validity. There are two kinds of criterion-related validity: concurrent validity and predictive validity.

#### *Concurrent Validity*

Concurrent validity is when the test and the criterion are administered at about the same time. For example, when an oral interview of at least thirty minutes is required in order to test all of the grammar studied through a semester and time constraints limit the time available for oral examinations to ten minutes, a random selection of students should take the longer thirty-minute test followed by the shorter ten-minute test. If there is a high level of agreement between the scores, then the shorter test can be considered valid and subsequently implemented to test the remaining students.

#### *Predictive Validity*

This looks at the degree to which a test can predict a test-takers future performance. Placement tests have predictive validity. A student is placed in a certain class because it is presumed that the score attained on the placement test shows the abilities of that student and how they will perform in the future. The sole purpose of placement tests are to make

the teaching environment better for all concerned, as it is thought best to keep students with similar abilities together in one class.

### ***Content Validity***

A test has content validity if the test has a representative sample of the structures with which it is meant to be testing. Tests which contain a proper sample of the relevant structures should correspond to the test specifications which will be drawn up at the beginning of test construction. The test specifications should relate to what was actually taught in the classes. Sometimes test specifications are written which only cover the easy parts to test. However, it is important to cover all the areas of study; otherwise a negative backwash effect will occur where students ignore the areas not tested. Content validity is very important. The greater the content validity, the more likely the test will accurately measure what it is intended to measure.

### ***Reliability***

A test is reliable if it measures consistently. If a group takes an identical or near identical test and consistently achieves similar results, then the test can be considered reliable (Huerta-Macías, 2002; Mackey & Gass, 2005; Richards & Renandya, 2002). Reliability is an essential quality of test scores. Without it, the scores cannot provide us with information about the ability we want to measure. It is possible to quantify the reliability of a test in the form of a reliability coefficient. This allows us to compare the reliability of different tests. The ideal reliability coefficient is one. This means that the exact same score will be achieved by a group regardless of when it happens to be administered. This is impossible because three factors affect test reliability: 1) test method facets; 2) personal

attributes; and 3) random factors. On the other hand, a coefficient of zero would mean there was no reliability to the test. The most important factor in setting a minimum acceptable level of reliability is the purposes for which the test is intended. If the test is a high-stakes one, then the reliability should be very high; if a low-stakes test, then the reliability can be lower. However, the test designer should always try to achieve as high of a coefficient as possible. Bachman and Palmer (1996) suggest if the test construct is relatively narrow then higher levels of reliability can be expected. Also, if the test tasks are uniform in their characteristics, then higher levels of reliability can be expected. Lado posits that “good vocabulary, structure and reading tests are usually in the .90 to .99 range, while auditory comprehension tests are more often in the .80 to .89 range. Oral production tests may be in the .70 to .79 range” (as cited in Hughes, 1989, p. 32). Bachman (1990) argues that reliability and validity are not independent entities, but are complimentary, and that “in order for a test score to be valid, it must be reliable” (p. 160).

### **The TOEIC Test**

In an effort to produce a test that would reliably measure students’ language proficiency, ETS (Educational Testing Services) introduced the TOEIC test in 1979. According to ETS, “it directly measures the ability of non-native speakers of English to listen and read in English in the global workplace” (2014b, para. 1), and has found popularity in Japan due to the Ministry of Education, Culture, Sports, Science and Technology’s (MEXT) declaration that English levels in Japan will improve if tests like the TOEIC test are administered in secondary and tertiary education (MEXT, 2013). ETS (2014a) tell us that 52% of test takers for TOEIC listening and reading were full-time students in 2013 and 49% of Japanese test-takers were taking it for learning purposes, not for job applications

or graduation, showing us that it is a popular test in schools and universities.

The TOEIC test taken by the majority of students is the Listening and Reading test because, according to the Ministry of Economy, Trade and Industry (METI; 2013), this test is effective for developing business communication skills. The need for globalization, and the fact that Japan's English level is ranked 135th in the world and 27th in Asia, have compelled the Japanese government to establish *The Council on Promotion of Human Resource for Globalization Development*, which warns that it is imperative to raise English standards if Japan wants to stay competitive with its rivals ("The Council on Promotion," 2011). As a result, the TOEIC test has become the means through which universities are seeking to prepare graduates to become globally competitive.

### ***Objectives of the TOEIC test***

This test assesses English listening and reading skills, which are deemed important in the workplace. In fact, the test creators, ETS (2014c), allude that listening skills developed as a result of this test will help with "face-to-face communication, meetings, videoconferencing, podcasts and telephone conversations" (para. 1), whereas the reading skills developed will help with "emails, reports, newsletters and other forms of business correspondence" (para. 1). Although the skills mentioned above suggest that communicative abilities can be developed as a result of this test, only passive skills are being tested, not active ones. ETS (2014c) also suggest that this test helps with globalization development, which is what METI wants to promote, by informing us that nearly 14,000 organizations in 150 countries throughout the world "trust the TOEIC test to determine who has the English language skills to succeed in the global workplace" (para. 3). Universities want to promote their institute and students as globalized and this



is why the TOEIC test is becoming more popular. In fiscal 2013, ETS (2014d) tell us that about 7 million people worldwide took TOEIC programs. Wilson (1989) suggests 80% of TOEIC examinees are in Japan, and the 2013 report on test takers worldwide admit that the vast majority of test takers are from Asia. Students who want to be able to function in the global workplace are more likely to choose a university where such tests are available. Also, Sarich (2014), informs us that Japanese companies are increasingly using TOEIC as a measure of English ability in the consideration of hiring and advancement, and, due to this, many universities have introduced TOEIC preparation courses into their English curricula.

### *Uses of the test*

This test is very popular in universities since the TOEIC test can be taken as an Institutional Program, which means that the test can be taken at a time and location of the university's choosing. Add to this the fact that the test is reasonable and comes in three different versions: one, the TOEIC Listening and Reading test for beginner to advanced learners; two, the TOEIC Bridge test for beginner to intermediate level learners; and three, the TOEIC Speaking and Writing test for intermediate to advanced students.

Some universities test their students twice to find out how the students have progressed; once in their first year and again in their second year. However, this is a norm-referenced test not a criterion-referenced test. Thus, it shows the test-takers abilities in relation to other test-takers. It does not measure progress in instructional objectives, which a criterion-referenced test would do. Therefore, it can be argued that such tests should not be used as achievement tests (Brown, 1995; Gorsuch, 1997). Hughes (1989) informs us that "criterion-referenced tests are often to be preferred, not least for the

beneficial backwash effect they are likely to have” (p. 19). A university that relies on these tests to measure their students’ progress may be mistaken. Some universities might be persuaded to regularly implement the TOEIC test because ETS imply that it can measure progress; as ETS data documents by showing how average scores increase with every sitting of the test, thereby suggesting progress in English ability. Specifically, ETS (2014a) tell us that in Japan 79% of test takers had previously taken the test and 49% had taken it three or more times, so there is an unwritten rule in which students infer they can get a higher score by simply taking the test more than once. Thus it seems that universities are justified by having the students repeatedly take the TOEIC test. Unfortunately, these results do not guarantee the reliability or validity of the test.

### **Validity in the TOEIC test**

Powers (2010) suggests that when test takers are asked how well they can perform language tasks, they give reasonably accurate self-assessments, which ETS refers to as *can-do* reports, such as when using a Likert scale. When these *can-do* reports agree with the test takers’ TOEIC score, it shows evidence for the validity of TOEIC scores. This report is published by ETS leaving the results open to criticisms of bias. Bachman & Palmer, 1988 (cited in Bachman, 1990, p. 148), discovered that asking test takers how difficult different aspects of language use are for them is a much better indicator than asking how well they can perform language tasks. Coombe (2002) brings up many issues related to self-assessment in language such as an individual’s judgment being subject to variables whose influence is difficult to establish. That is, extraneous factors affect the accuracy of self-assessment. Furthermore, because validity, like reliability, depends on systematic analysis, the question is raised as to whether short term self-assessments lend

themselves to consistency. Self-assessment is great for developing autonomy and promoting learning (Harris, 1997; Oscarson, 1989). However, it cannot be used to validate a TOEIC score and therefore the TOEIC test. Blue (1988) found that self-assessment does not measure language proficiency accurately, and Boud and Falchikov (1989) note that over or under-assessment can happen. Matsuno (2009) also found this in a study conducted in Japan where more able students assessed themselves at a lower level, and low-level students assessed themselves above their real level. Mistar (2010) maintains that to increase the validity and reliability of self-assessment, students need to be trained in how to self-assess. Therefore, the claims that test takers self-assessment show evidence for the validity of TOEIC scores is questionable.

Woodford (1982) informs us that 2,710 test takers were involved in the first TOEIC test, of which 500 were selected to validate the test. They decided to use concurrent validity by testing these 500 with the TOEFL test, which was already well-established. Of these 500, only 100 participated in the validity test. This seems to be a rather small number to validate a new test, considering that it is only 3.7% of the test takers. Woodford goes on to tell us that the input for the validity tests, both in listening and reading, were in English. However, questions were then asked in Japanese by a Japanese examiner and the test takers were encouraged to respond in Japanese. Cunningham (2002) tells us that this makes interpretation of the results and the construct validity of the test questionable, since this is meant to be a test of communicative ability, not the ability to comprehend. It should be noted that even though there is a Speaking and Writing TOEIC test, the majority of students in Japan take the Listening and Reading test. ETS (2014d) inform us that 769,220 tertiary level students took the Listening and Reading test in 2013, and only 1,981 took the Speaking and Writing test.

Woodford (1982) explains that with only passive skills being tested, we can also interpret a test taker's active skills through score interpretation. Also, ETS (2012b) states that "TOEIC Listening and Reading tests are indirect measures of speaking and writing proficiency" (p. 4). As was already stated, Long (2014) concluded that passive skills do not correlate with active skills as well as ETS is suggesting. TOEIC can assess receptive competency but it lacks validity in suggesting that it can assess communicative competency.

Content validity, as mentioned earlier, is when a test accurately measures what it is intended to measure. ETS's claim that this test measures English communication skills in a business context do not seem to be supported. It is highly unlikely that a reading test can be classed as a genuine communicative skills test. It is definitely an important skill, but does not support the content validity suggested by ETS, and, therefore, the test results and validity should be considered suspect.

### **Reliability in the TOEIC test**

ETS (2012a) suggest that the TOEIC test is a highly reliable test shown to be within 25 points for one Standard Error of Measurement. However, the report also mentions that this is true only two thirds of the time. Cunningham (2002) informs us that 99% confidence in the score will increase the original 25 points to 63.5 points for one standard error of measurement, lessening the test's reliability.

TOEIC reliability is achieved by having a clearly organized test with clear instructions and non-culturally based inputs. Also, because this is a multiple-choice test which is scored using a computer, it shows high reliability. In Japan, there is a glut of books on the market, and there are many language schools that help students increase

their scores by becoming test wise. Test takers are taught how to improve their test scores by learning techniques that allow them to intelligently guess the correct answers to unknown questions. This suggests an avenue for further study would be to see if test takers that focus solely on learning test taking techniques to improve their overall score can accomplish this without actually increasing their language ability.

### **Impact on Education**

For teachers who find themselves forced to teach to the test, Shepard (2000) warns that standardized testing can lead to limiting the breadth and depth of material covered (as cited in Odendahl, 2011). Also, the use of multiple-choice formats and short reading passages constrict the range of skills assessed. Therefore, students have fewer opportunities to explore a variety of ways to communicate in English. This runs counter to the Ministry of Education, Culture, Sports, Science and Technology's plans to develop students' communicative abilities (MEXT, 2013).

Students are given snippets of grammar, or low frequency vocabulary which they should memorize so that they can guess which answer is probably correct. This is similar to a Japanese *Juku*, where students learn only what is needed to pass entrance exams by rote memorization. Such exercises help students pass entrance tests and gain high scores on TOEIC tests, but their English ability outside the test is very poor. As mentioned earlier, Japan's English ability is 135th in the world and 27th in Asia; this result alone suggests that teaching to the test, which is how English is taught in junior and senior high schools, is not effective for learning a foreign language.

Students' interest has also waned since the test's introduction. This can be seen through end of term questionnaires. Comments such as "I wanted more speaking practice"

and “I thought university English class would be more interesting than high school English” show us that students do not want a repeat of what they have gone through in high school. With so much time taken up studying how to improve their TOEIC scores, there is little time left for practicing communicative interactions which Hedge (2000) informs us have “value both for negotiating comprehensible input from other students and for opportunities to practice using the language they have learned until it is automatized” (p. 262). Due to a lack of communicative interaction in the class when these students graduate and start working in companies, they will be unprepared to deal with foreign customers who cannot communicate in Japanese, which runs counter to *The Council on Promotion of Human Resource for Globalization Development’s* goal to promote global competitiveness. Besides, even if graduates never have dealings with foreign customers while working at their respective vocations, a simple holiday overseas could become a stressful or troubling experience if TOEIC testing results in compromised English communication abilities. The students who increase their vocabulary as a result of earnestly studying for the TOEIC test, unfortunately, learn mainly low-frequency technical words. After first year compulsory courses are over, students are very unlikely to come across these words over the next three years of their studies in university, and, therefore, the lack of exposure will result in the words not being retained. Also, if students take the TOEIC test in the first and second year, two years will have passed before they graduate and start working in industry. ETS (2012a) inform us that TOEIC scores are only considered valid for two years as students’ proficiency can improve or decline over time. Companies which use TOEIC scores as a prerequisite for employment understand this and will likely require job seekers to have test results no older than two years. Students need to practice and improve their active English skills to have a chance at becoming

global citizens, and the best place to do this is in an oral English class. Therefore, it would be better for the students if they were offered elective TOEIC classes in their third or fourth year of studies.

### **Conclusion**

Tests are very important to gauge the abilities of students. They help teachers see how their students are progressing, they set goals for students to aim for, and they let society judge their strengths and weaknesses. It was shown that statistical analysis and standardized tests were introduced around the start of the last century, both of which are widely used nowadays. Standardized tests were looked at, and it was explained that to accurately measure a student's ability, the test taker must ensure that the correct test type and construct are chosen and that the test's validity and reliability are accurate.

Teaching circumstances were briefly looked at, and it was explained that TOEIC is now a high-stakes test in some universities. Validity and reliability for the TOEIC test were described, showing that the test has limited reliability, and its validity is questionable. It was explained that the use of the TOEIC test has had a negative impact on oral English classroom practices for first year students. The TOEIC test might be good for an office worker who is dealing with foreigners through email and possibly phone conversations, and it might also be helpful for third and fourth year students who can show a well-recognized qualification to prospective employers, but for first and second year university students, it is something which will be demotivating and will limit opportunities for them to participate in oral English practices which will develop their communicative competence.

## References

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Blue, G. (1988). Self-assessment of listening comprehension. *International Review of Applied Linguistics*, 16, 149-156.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18, 529-549.
- Bresnihan, B. D. (2013). Useful information teachers and administrators should know about the TOEIC. In N. Sonda & A. Krause (Eds.), *JALT 2012 Conference Proceedings*. Tokyo: JALT
- Brown, J. D. (1995). Differences between norm-referenced and criterion-referenced tests. In J. D. Brown & S. O. Yamashita (Eds.), *Language Testing in Japan* (pp. 12-19). Tokyo: JALT.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, (1) 1-47
- Cihi, G. & Culligan, B. (2011). *TOEIC vocabulary secrets revealed*. Retrieved from <http://www.wordengine.jp/research/pdf/TOEIC%20Vocabulary%20-%20Secrets%20Revealed%20Dec%2016.11.2v.pdf>
- Coombe, C. (2002). *Self-assessment in language testing: Reliability and validity issues*. Retrieved from <http://www3.telus.net/linguisticsissues/selfassess.html>
- Cunningham, C. R. (2002). *The TOEIC test and communicative competence: Do test score gains correlate with increased competence?* (Unpublished M.A. Dissertation). University of Birmingham, Retrieved from [http://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/matefltesl\\_dissertations/Cunndiss.pdf](http://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/matefltesl_dissertations/Cunndiss.pdf)
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Educational Testing Services. (2012a). *Examinee handbook listening and reading*. Retrieved from [http://www.ets.org/Media/Tests/TOEIC/pdf/TOEIC\\_LR\\_examinee\\_handbook.pdf](http://www.ets.org/Media/Tests/TOEIC/pdf/TOEIC_LR_examinee_handbook.pdf)
- Educational Testing Services. (2012b). *Examinee handbook speaking and writing*. Retrieved from [http://www.ets.org/Media/Tests/TOEIC/pdf/TOEIC\\_Speaking\\_and\\_Writing\\_Examinee\\_Handbook.pdf](http://www.ets.org/Media/Tests/TOEIC/pdf/TOEIC_Speaking_and_Writing_Examinee_Handbook.pdf)
- Educational Testing Services. (2014a). *2013 report on test takers worldwide: The*



- TOEIC listening and reading test*. Retrieved from [http://www.ets.org/s/toeic/pdf/ww\\_data\\_report\\_unlweb.pdf](http://www.ets.org/s/toeic/pdf/ww_data_report_unlweb.pdf)
- Educational Testing Services. (2014b). *About the TOEIC test*. Retrieved from <http://www.toeic.or.jp/english/toeic/about.html>
- Educational Testing Services. (2014c). *About the TOEIC listening and reading test*. Retrieved from [http://www.ets.org/toeic/listening\\_reading/about?WT.ac=toeichome\\_listeningreading\\_121127](http://www.ets.org/toeic/listening_reading/about?WT.ac=toeichome_listeningreading_121127)
- Educational Testing Services. (2014d). *TOEIC program data and analysis 2013*. Retrieved from [http://www.toeic.or.jp/library/toeic\\_data/toeic\\_en/pdf/data/TOEIC\\_Program\\_DAA.pdf](http://www.toeic.or.jp/library/toeic_data/toeic_en/pdf/data/TOEIC_Program_DAA.pdf)
- Gorsuch, G. J. (1997). Test purposes. *The Language Teacher*. Retrieved from [http://jalt-publications.org/old\\_tlt/files/97/jan/gorsuch.html](http://jalt-publications.org/old_tlt/files/97/jan/gorsuch.html)
- Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, 51 (1), 12-20.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Huerta-Macías, A. (2002). Alternative assessment: Responses to commonly asked questions. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 338-343). Cambridge: Cambridge University Press.
- Long, R. W. (2014). Complexity and fluency indicators of “good” speakers. In N. Sonda & A. Kraus (Eds.), *JALT 2013 Conference Proceedings*. Tokyo: JALT
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Matsuno, S. (2009). Self-, peer-, and teacher assessments in Japanese university EFL writing classroom. *Language Testing*, 26 (1), 75-100.
- Ministry of Economy Trade and Industry. (2013). *Part 2 tackling various subjects to realize the levels of dependability and security of information system software required in a safe and secure advanced information society*. Retrieved from [www.meti.go.jp/english/press/data/pdf/20090528\\_01c.pdf](http://www.meti.go.jp/english/press/data/pdf/20090528_01c.pdf)
- Ministry of Education, Culture, Sports, Science and Technology. (2013). *Five proposals and specific measures for developing proficiency in English for international communication*. (Commission on the Development of Foreign Language Proficiency). Retrieved from <http://www.mext.go.jp/component/english>

/\_\_icsFiles/afieldfile/2012/07/09/1319707\_1.pdf

- Mistar, J. (2010). A study of the validity and reliability of self-assessment. *Universitas Islam Malang*, Indonesia. Retrieved from <http://journal.teflin.org/index.php/teflin/article/viewFile/244/180>
- Nation, P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.). *Vocabulary description, acquisition and pedagogy*. Cambridge: Cambridge University Press
- Odendahl, N. V. (2011). *Testwise: Understanding educational assessment, Volume 1*. Lanham: Rowman and Littlefield Education.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rational and application. *Language Testing*, 6 (1), 1-13.
- Powers, D. E. (2010). *Validity: What does it mean for the TOEIC tests?* Retrieved from <http://www.ets.org/Media/Research/pdf/TC-10-01.pdf>
- Sarich, E. (2014). A guide to planning and executing a TOEIC preparation course. *The Language Teacher*, 38 (1), 17-21.
- Shortt, P. (2014). *Comparing standardized educational and teacher generated classroom achievement tests*. [PowerPoint slides]. Retrieved from Current perspectives on educational testing for masters in teaching and learning: [http://study.hiberniacollege.net/matl/2014/matl\\_cpoet/session2/task1/print/matl\\_cpoet\\_s2\\_t1\\_print.pdf](http://study.hiberniacollege.net/matl/2014/matl_cpoet/session2/task1/print/matl_cpoet_s2_t1_print.pdf)
- The Council on Promotion of Human Resource for Globalization Development. (2011). *An interim report of the council on promotion of human resource for globalization development*. Retrieved from [http://www.kantei.go.jp/jp/singi/global/1206011interim\\_report.pdf](http://www.kantei.go.jp/jp/singi/global/1206011interim_report.pdf)
- Wilson, K. M. (1989). *Enhancing the interpretation of a norm-referenced second language test through criterion referencing: A research assessment of experience in the TOEIC testing context*. Retrieved from <http://www.ets.org/Media/Research/pdf/RR-89-39.pdf>
- Woodford, P. E. (1982). *An introduction to TOEIC: The initial validity study*. TOEIC research summary number 0. Retrieved from [https://www.ets.org/research/policy\\_research\\_reports/publications/publication/1982/cglx](https://www.ets.org/research/policy_research_reports/publications/publication/1982/cglx)

## Using Focus Groups for ELT Research

**David Bramley**

Osaka Jogakuin University

### **Abstract**

A focus group is one way of finding out directly from students through discussion, rather than through the medium of a written survey, what they think about their language programs and materials. Furthermore, as a focus group is conducted when necessary in the students' native language, it gives students a chance to more accurately communicate their opinions. Written questionnaires are a fixed medium. When a written questionnaire is administered, the questions in it cannot be changed or modified to find out individual responses. On the other hand, a focus group interview can be more enquiring, in that researchers may modify questions or ask follow up questions to pursue a line of enquiry not encompassed in a written document. Focus groups can give researchers insights as to how to form questions which are more likely to get the kinds of responses that will help give researchers greater insights and which are therefore more useful to include on a more comprehensive survey. This article details the construction of focus group questions and subsequent focus group interviews on the topic of tablet computers, namely iPads, and the use of eBooks in the students' English study program.

*Keywords:* focus groups, tablet computers, eBooks, survey, interview, questionnaire

**F**ocus groups are a tool that is widely used in marketing in order to determine customer satisfaction to projects, to gain insights into future directions to follow, and as a tool to develop more detailed surveys. There is less data on their applications outside the context of business, and this is especially true for language education. However, focus groups are considered by faculty to be a reliable and important tool for gathering feedback from students by allowing the teacher or program coordinator to determine students' level of satisfaction with a particular program (Braskamp & Ory, 1994; Fowle, 1999). Furthermore, focus groups allow researchers to gain more

information in less time than it would take with individual interviews, thus making them more time efficient (Nasser-McMillan & Borders, 2002). This article examines various aspects of a focus group and then describes a set of focus group interview sessions which were used to develop a questionnaire about students' use of iPads.

### **What is a focus group?**

Focus groups were first designed during World War II to assess the reactions of the public to wartime propaganda. However, after the war the idea was adapted by businesses to do marketing research. A focus group has quite small number of people; usually between four and 15, but the average is eight. In marketing, the group is has a moderator and it focuses on a particular product or topic. The aim of a focus group is to discuss a topic, rather than to ask formal written questions to each individual in the form of a questionnaire. The type of information produced from a focus group is qualitative rather than quantitative, focusing on data such as preferences and beliefs. Focus groups may also be directed at the solution of problems as a way of generating ideas or predicting outcomes. Such focus groups may consist of experts or well-informed individuals who share their point of view on a topic. In an academic setting, focus groups may consist of teachers, students, administrators or researchers.

### **Uses of a focus group**

From an academic perspective, focus groups can be used for a variety of research purposes. One use of a focus group is to gather initial data when there is not enough information on a proposed research topic (Nasser-McMillan & Borders, 2002). Stewart and Shamdasani (1990) suggest a number of uses for focus groups in research, such as

getting data, creating research hypotheses, and gathering information about services and programs. McLeod et al. (2000) talk about using focus groups to design a valid questionnaire. Finally, speaking on the topic of gauging satisfaction levels among learners, Anderson (1996) believes that focus groups “offer a very flexible way of sampling student opinion” (p. 16).

### **Focus group, survey or both?**

Whether a focus group or a survey is used depends on the goals and the type of data sought by the researcher. If qualitative data is sought, then a focus group may be more appropriate than a questionnaire because a focus group can pursue deeper answers to questions and allows participants to elaborate or explain their opinions or experiences. In addition, the act of group participation can help people to build on each other’s ideas and comments. Focus groups are best used when researchers want to hear a range of opinions on a complex issue that they want to understand better. Focus groups help researchers begin to understand the reasons for a situation or problem. They are useful for getting detailed information on a specific issue because the moderator can get people to elaborate on their answers. Finally, focus groups can help researchers to develop a survey because they will help to develop meaningful questions to include (Eliot, 2010). Researchers should keep in mind that a focus group requires a time commitment and needs to be planned in advance; time is needed to plan questions, space has to be reserved, and participants have to be invited (Dzija et al., 2005).

Surveys provide quantitative data that can be used to answer questions such as, “How much?” “How often?” or “How large?” While they require time to plan questions, distribute surveys, and process the answers, they can be administered in a much shorter

time than focus group sessions. Moreover, unlike focus group sessions they can be given to a potentially unlimited number of people. They do not even require participants to be physically present in a particular place at a particular time because they may be administered via, telephone, email or the Internet. Finally, a variety of question types, for example, multiple choice and open-ended questions can be used (Dzija et al., 2005). However unlike focus groups, answers to open-ended questions cannot be elaborated upon. Surveys are best used when researchers want to understand the frequency with which something happens, when decisions will be based on numbers, and when key issues have already been defined through a qualitative approach such as focus groups or interviews and researchers now want to know what the larger population thinks (Eliot, 2010). Thus before researchers embark on either a focus group/interview study or a survey study, it is crucial that they understand exactly what kind of data they need and for what purposes they will be using the results.

### **Focus group questions**

As one of the uses of a focus group is to get subjects to elaborate on the research topic, the types of questions which are asked during the focus group sessions are important. Some advice given by researchers is not to ask leading questions, or questions which would result in simple Yes/No answers (Adams & Cox, 2008). Questions should be open-ended and be arranged in a logical pattern ("Using," n.d). To this end, a four-question sequence is recommended ("Focus," 2012). In this sequence a main, open-ended question, for example, "What do you think about ...?" or "What's the biggest problem with ...?" is asked, then the follow up question asks the respondent to elaborate on the answer using probing questions, such as "What do you mean?" or "Give me an example" are used to

clarify the responses. The fourth part of the sequence is to provide prompts or cues if necessary. A prompt may be, “You mentioned X. What about Y and Z?” To use such a sequence of questions successfully, it is important that focus group moderators discuss and prepare the question types beforehand. If possible, a pilot session should be conducted to check the smoothness of the questioning technique and to see if the questions elicit the kind of response required.

### **Focus group case study**

Osaka Jogakuin University and Junior College in Osaka, Japan gave iPads to all of its incoming students in 2012. The students initially used the iPads for a wide variety of learning tasks such as checking words in the dictionary, translating, and doing research on the Internet. They also used iPads to make presentations to their classmates and to make both video and sound recordings. In addition, teachers used the iPads to present textual, audio, and visual materials. Later in 2012, the students received one of their reading texts in the form of an in-house eBook. This was followed by 12 more in-house eBooks and two commercial eBooks which are now being used exclusively to teach reading and discussion to first year students, as well as listening skills to second and third year students.

### **Construction of a survey using focus groups**

As part of ongoing research into the production and use of eBooks as well as the classroom use of tablet computers by Swenson, Bramley, and Cornwell (2013) and Bramley (2014), a decision was made to construct a survey about students’ iPad use. The main purpose of this survey was to see which classes were having the highest frequency

of use in order to make changes in the courses to improve the students' learning experiences with the technology that was adopted by the school. The researchers were not interested in assessing students' comfort with the technology as the pedagogical decision had been made as an institution to adopt it. However the researchers are interested in improving their use of the technology and hoped the resulting data would provide some direction for where the program needs to be improved or to train students better in the use of the new technology. In order to get data and insights for the construction of a qualitative survey, it was decided that focus group sessions would be held. To this end, a variety of open-ended questions, probing questions and prompts were prepared with the aim of generating student conversation on the topics that the researchers wanted to hear about. Initially it was decided to ask focus group questions on both eBook and iPad use. (See Figure 1).



eBooks

What did you think about the eBooks that you used?

(First year – spring term, Second year the previous autumn term)

Questions about First Year Spring term e-texts / Second year previous autumn eBooks

Which eBook did you like best?

What activities did you like in the eBook? (Examples?)

What did you think about the activities in the eBooks?

Prompts:

pop over

gallery

sound files

review quizzes – vocab - content - dialogue

How much of the book did you use?

iPad Use

What do you like about the iPad?

What do you dislike?

Where do you use the iPad?

How often do you use the iPad in class? (Examples of use?)

How often do you use the iPad at home? (Examples of use?)

How do you use your iPad now? (Second years)

In what other ways do you use the iPad? (First year)

*Figure 1.* Focus group questions about eBook and iPad use.

**The focus group sessions**

The focus group sessions were held in two, twenty-five minute sessions. Students were encouraged to attend the sessions by the promise of book vouchers for those who stayed for a whole session. Students who attended the sessions were divided into three groups. The group size varied from four to six students. The attendees were all first or second year students. The moderators allowed the discussion to develop in Japanese, as this is

the language in which the students were fully able to discuss their true opinions, questions were used by the moderators only to move the conversation in different directions when it appeared to dry up and when it seemed appropriate that students elaborate more upon a topic. In addition, as the purpose of the focus group was to obtain data that could indicate suitable survey question directions, it was necessary to provide some direction to the conversation. This was perhaps a departure from the "traditional" role of the focus group moderator in that the moderators were more active in the conversation. The focus group sessions were all recorded and the results were translated into English and transcribed. (See Figure 2 for excerpts).

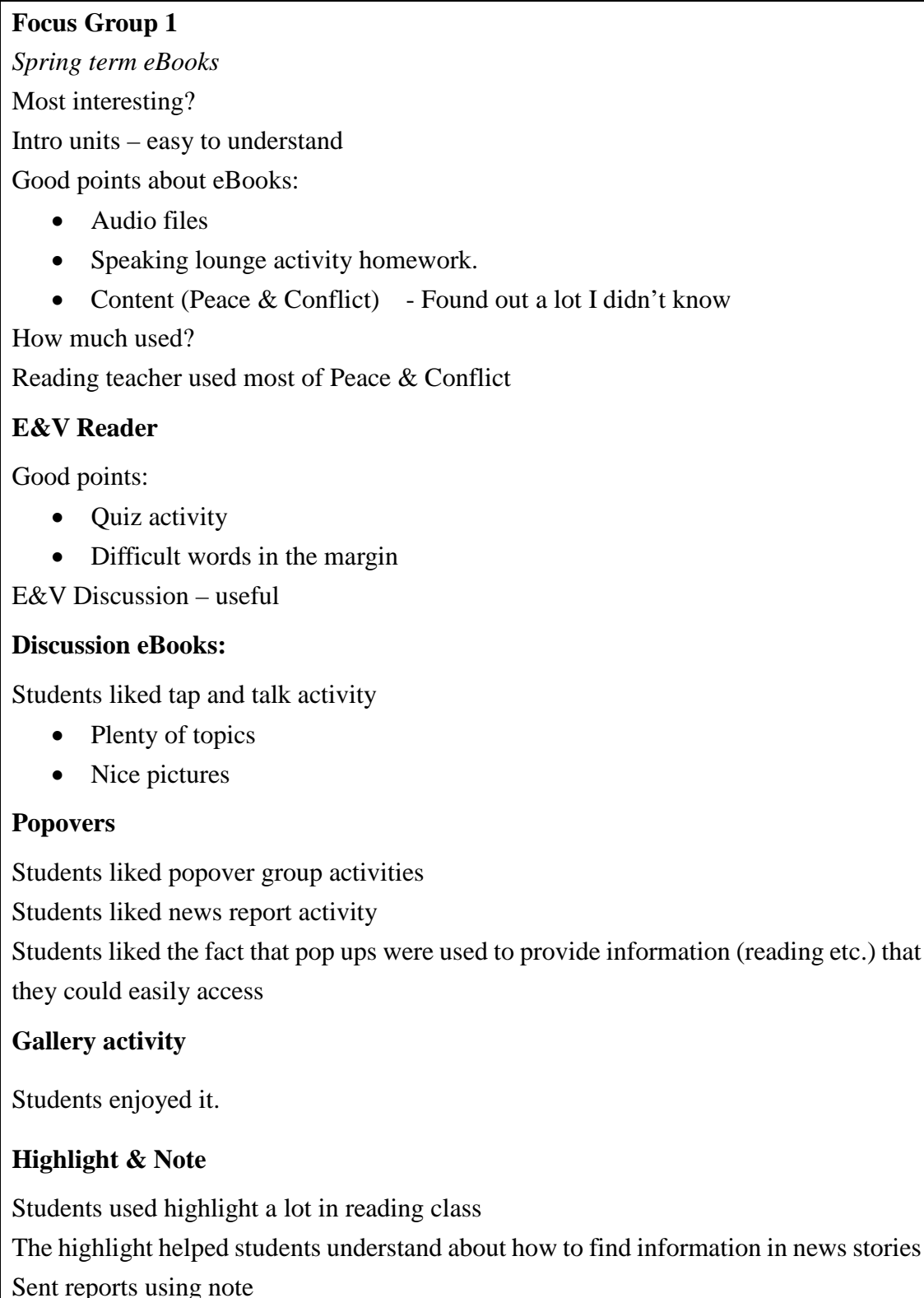


Figure 2. Focus group results.

The focus groups provided a variety of useful data concerning student impressions of both the iPad and the eBooks they had been using. However, although we had originally decided to include both eBooks and iPads in the survey, it was eventually decided that we would focus the survey on iPads. There were a number of reasons for this. First, through the focus groups we discovered that not all students had had the same amount of exposure to the eBooks. For example, second year students had only started to use the eBooks from the autumn term. On the other hand, all first year and second year students had had iPads from the time they entered the school. In addition, including the use of eBooks in the survey would necessitate a very long questionnaire that would be very time consuming to construct, administer, and evaluate. Also, focusing only on iPads would help to narrow down more specific areas for further research. Finally, perceptions of the eBooks were of a qualitative nature (see Figure 2) whereas a survey would produce more quantitative data such as the frequency of downloading apps, frequency of use in different classes, and type of use while at school and at home. Responses from the focus group questions were used as the basis of a 45-item survey on iPads, which was translated into Japanese and administered to more than 350 students. The results of this survey are discussed in the article, "Tablet Computers and eBooks: Student Feedback" (Bramley, 2015).

## **Conclusion**

While the results of the survey showed some differences between first and second year students in their use of iPads, these were not outstanding. The main information gained was from the process of planning and holding of the focus group sessions and their outcomes. From this process we gained more insight into how to plan and hold focus group sessions, what questions to ask in order to gain more ideas about the research issues

that we need to pursue and how to limit the scope of the research to a manageable size.

In this respect, focus groups are a valuable learning tool for researchers.

## References

- Adams, A., & Cox, A. L. (2008). Questionnaires, in-depth interviews and focus groups. In P. Cairns, & A. L. Cox (Eds.), *Research Methods for Human Computer Interaction* (pp. 17-34). Cambridge, UK: Cambridge University Press.
- Anderson, I. (1996). Focus groups as a means of obtaining student feedback. *ELT Management Newsletter*, 20, 16-18.
- Bramley, D. (2014). Tap into the future: A study of iPads and interactive eBooks in an e-learning project. *Osaka JALT Journal*, 1, 89-108. Retrieved from <http://www.osakajalt.org/storage/journal/Osaka%20JALT%20Journal%20March%202014.pdf>
- Bramley, D. (2015). Tablet computers and eBooks: Student feedback. *Osaka Jogakuin University Journal*, 11, 23-43.
- Braskamp, L.A., & Ory, J.C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco, CA: Jossey-Bass.
- Dzija, J., Foster, J., Hernandez S., Nardi, S., Theriault, P., & Wynne, C. (2005). *How to conduct effective focus groups and surveys*. Retrieved from [http://www.ats.msstate.edu/assessment/docs/How\\_to\\_Conduct\\_Effective\\_Focus\\_Groups.pdf](http://www.ats.msstate.edu/assessment/docs/How_to_Conduct_Effective_Focus_Groups.pdf)
- Eliot, S. (2010). *Survey or focus group: Which to use when*. Retrieved from <http://www.qualitative-researcher.com/group-wisdom/survey-or-focus-group-which-to-use-when/>
- Focus group questionnaire fundamentals*. (2012). FocusGroupTips.com. Retrieved from <http://www.focusgrouptips.com/focus-group-questionnaire.html>
- Fowle, C. (1999). Focus Groups: an effective means of gaining course feedback. *ELT Management*, 28, 26-29.
- McLeod, P.J., Haegher, T.W., Steinhert, Y., & Boudreau, D. (2000). Using focus groups to design a valid questionnaire. *Academic Medicine*, 75(6), 671.
- Nasser-McMillan, S.C., & Borders, L.D. (2002). Use of focus groups in survey item development. *The Qualitative Report*, 7(1). Retrieved from <http://www.no.va.edu/ssss/QR/QR7-1/nassar.html>

- Stewart, D.W., & Shamdasani, P.N. (1990). *Focus groups: Theory and practice*. Newbury Park, CA: Sage.
- Swenson, T., Bramley, D., & Cornwell, S. (2013). Making interactive eBooks: More than just cutting and pasting. *Osaka Jogakuin University Journal*, 10, 17-30. Retrieved from [http://ir-lib.wilmina.ac.jp/dspace/bitstream/10775/3127/1/B\\_02Swenson.pdf](http://ir-lib.wilmina.ac.jp/dspace/bitstream/10775/3127/1/B_02Swenson.pdf)
- Using focus groups to get student feedback*. (n.d.). Carnegie Mellon University Eberly Center for Teaching Excellence. Retrieved from <http://www.cmu.edu/teaching/assessment/howto/assessteaching/focusgroups.html>

# Factors Influencing the Adoption of Communicative Textbooks

**Simon Humphries**

Kansai University

**T**extbooks are one of the main tools of the trade in language teaching (Littlejohn, 2011) and it follows that they may facilitate curriculum change because they can provide visible frameworks for teachers and students to follow (Rubdy, 2003). However, if teachers do not understand and take ownership of innovative approaches, they stick to the status quo (Humphries & Burns, 2015). Last November, in my presentation at the JALT Conference in Tsukuba, I outlined factors that can promote or inhibit the appropriate use of an innovative textbook based on a study at a rural technical college. Due to space limitations, this is a short report that summarizes the main points. Please see Humphries (2014) for the complete original chapter that was the source for the JALT presentation.

## **The *Kosen*: A context of change**

The study took place in a rural *kosen*, which is an engineering college for students aged 15-20. The management decided to introduce a communicative curriculum for students in Grades 11 and 12. (Grade 10 students continued to use the old curriculum.) Under the new curriculum, students studied for the TOEIC Bridge test using a textbook from a Western publisher in place of the textbook approved by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). The two textbooks differed substantially (see

Table 1). The MEXT-approved textbook supported the *yakudoku* approach. In general, *yakudoku* is a teacher-centered, form-focused approach that is predominantly conducted in Japanese, and focuses on translating a text from English to Japanese. In contrast, the new textbook was designed to facilitate communicative language teaching (CLT). It was learner-centered, meaning-focused, and written entirely in English (apart from a glossary and a detachable phrase book that contained Japanese translations). Moreover, while the MEXT-approved textbook contained long reading passages, the new textbook emphasized listening and speaking.

Table 1  
*Textbook Comparisons*

<b>MEXT-approved textbook</b>	<b>New textbook</b>
Teacher-centered	Learner-centered
Emphasis on reading comprehension	Emphasis on listening and speaking
Overt grammatical usage instruction	Focus on communicative situations
Japanese publisher, aimed at Japanese high schools	British publisher, aimed at a variety of institutions in East Asia
Instructions and explanations in Japanese	Only glossary and ‘phrase book’ contain Japanese
Low output, highly structured exercises	Meaning-focused exercises, designed to encourage students to creatively share information and opinions

### **The Study**

As indicated above, the new textbook required a fundamental change in teaching



approach compared to the standard MEXT-approved textbook used previously. However, the teachers received no training for the new approach. The opportunity arose, therefore, to explore, in the absence of training, the factors that can promote or inhibit the appropriate use of an innovative textbook.

Four Japanese teachers of English (JTEs) at the *kosen* volunteered to participate. They taught Grade 10 (old curriculum) and Grade 11 (new curriculum), which enabled a comparison of the influence of the two different types of textbooks on their attitudes and practices. The collected data consisted of classroom observations and semi-structured interviews during a six-week period.

### **Factors influencing textbook use**

The following factors influenced the teachers' use of the new textbook: sociocultural factors, uncertainty, limited training, negligible external influences, internal *laissez-faire*, student issues, and unsuitable materials (see Table 2). As a result, rather than adopt the CLT approach, they continued to use *yakudoku* (Humphries & Burns, 2015).

Table 2  
*Factors That Arose during the Study*

Category	Example
Sociocultural evidence	All the teachers predominantly used <i>yakudoku</i>
Uncertainty	The teachers tended to externalize their feelings of uncertainty
Limited training	Only one teacher valued his training but struggled to apply it
Negligible external influences	None of the teachers felt any pressure from external sources
Internal <i>laissez-faire</i>	No pressure to conform but some teachers wanted guidance
Student issues	Even the highest proficiency learners seemed to prefer silent deskwork
Unsuitable materials	The teachers struggled to explain sociocultural differences

(For a description of these factors, please see Humphries, 2014, pp. 260-264.)

### **Implications**

Based on the issues that arose from the attempts at curricular change through the introduction of innovative textbooks in the *kosen*, we can discuss the implications for effective change from two perspectives: policymaking and relevant training.

#### ***Policymaking***

Before introducing new policies, two initial factors should be considered. First, teachers need appropriate working conditions (such as small classes and adequate contact hours).

Second, policymakers must consider the suitability for local contexts. After considering these initial factors, policies will be more successful if they are coordinated with three supporting components: materials, assessment, and training. The change at the *kosen* incorporated the first two components by introducing a new textbook and TOEIC Bridge, but a lack of teacher confidence in and understanding of CLT led the teachers to fall back on *yakudoku*.

### ***Relevant training***

Training needs to be relevant to classroom issues. The study revealed that teachers need support in the areas of methodology, textbook selection, foreign sociocultural topics, student issues, and instructional language proficiency (see Table 3).

Table 3  
*Recommendations for Teacher Training to Support Change*

<b>Category</b>	<b>Example</b>
Methodology	Workshops focusing on the materials that facilitate discussion of methodological problems
Textbook selection	Help teachers to analyze textbooks
Foreign sociocultural topics	Students themselves can research foreign cultures
Student issues	Help teachers develop strategies to develop students' intrinsic motivation
Instructional language proficiency	Train teachers to effectively use the target language and/or vernacular

(These recommendations are described in detail in Humphries, 2014, pp. 265-267.)

## Conclusion

In conclusion, this case study provides insights into the complexity of curriculum innovation. Curriculum planners should look beyond the simplicity of using the textbook as a change agent. Instead, they should consider what ongoing teacher support should accompany the gradual change to new practices.

## References

- Humphries, S. (2014). Factors influencing Japanese teachers adoption of communication-oriented textbooks. In S. Garton & K. Graves (Eds.), *International Perspectives on Materials in ELT* (pp. 253-269). New York: Palgrave Macmillan.
- Humphries, S., & Burns, A. (2015). "In reality it's almost impossible": CLT-oriented curriculum change. *ELT Journal*. Advance online publication. doi:10.1093/elt/ccu081.
- Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials Development in Language Teaching* (2nd ed., pp. 179-211). Cambridge: Cambridge University Press.
- Rubdy, R. (2003). Selection of materials. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 37-57). London: Continuum.

# Wikis Revisited: A Single Software Tool for a Variety of Courses

**Harry Carley**  
Matsuyama University

**T**his paper corresponds to a presentation that was given at Osaka Tech Day 2014 and rationalizes how wikis may be applied in foreign language courses. A wiki is a free and secure software tool that has been available for over 10 years. Wikis offers versatility in teaching for lecturers while at the same time allowing learners hands on experience with educational technology tools. They can allow for project-based work, writing individually or by collaboration, or social communication in a secure environment, such as blogging. For teachers, wikis can assist in class organization and monitoring of student progress. The wiki team of administrators and software engineers are constantly upgrading and adding new features too. Wikis can be used for almost any subject and can be produced in any language. Additionally, they can be utilized in small or large classes. Furthermore, a wiki is a suitable mode of instruction for almost any EFL class, like a writing course that wants to promote a ‘green attitude’ by having a paperless classroom or a presentation course that utilizes videos and research gathering techniques.

Utilizing a wiki in course work is not a recent web tool for educators since the software capabilities have been in existence for many years. From the time of their initial introduction, wikis have been lauded by instructors as a viable teaching tool at all levels of education. The use of wikis is best suited to learning environments that may include information gathering activities, presentations, or writing on an individual or

collaborative level. The implementation of wikis in the instructional process can benefit learner and lecturer alike. From the standpoint of an instructor, the popularity of wikis can be credited to their flexibility in usage. Students, meanwhile, are motivated to use them due to their convenient and attractive features. Additionally, Internet connectivity allows for access through any PC or mobile device in or outside a traditional classroom environment.

### **Registration**

Specific wikis such as Wikispaces.com offer free unlimited creation of wikis for teachers and students upon joining. Thusly, there are no fees to be concerned about and the number of students in a class or the number of classes an instructor wishes to implement wikis into is of no concern.

### **Safety and Security**

As teachers are ultimately in charge of the well-being and safety of their learners, control and security of information displayed on the Internet should be of the utmost concern for teachers. The nature of the Internet and the collaborative-based design of wikis allows anyone utilizing a search engine to find and view wikis, which necessitates that teachers take proactive steps to insure the safety of their students.

Upon signing up for a wiki account, an instructor is asked for a security preference. The settings can later be adjusted by the wiki administrator (instructor) at any time. The permission settings are as follows:

- Public: Anyone can view and edit pages.
- Protected: Anyone can view pages, but only members can edit.
- Private: Only members can view or edit pages.

For most educational environments either the second or third option would be most appropriate. The first option, 'Public,' is probably the least desirable since it would allow anyone to have access to the wikis. In addition to the abovementioned three settings, there is also a Custom setting that allows administrators to customize their wikis by making alterations to background themes or personalizing wiki names. New accounts are automatically given a 'YOURWIKINAME.wikispaces.com' name, but this can be changed to a simpler name like 'YOURWIKINAME.com'. It should be noted that some customizations of wikis may be subject to a nominal fee. A final note about wiki security. To restrict any type of spam attack or postings from outside an intended group, most wikis, and Wikispaces.com in particular, have put safe guards into place.

### **Data Storage**

As far as the actual usage of the wikis, storage restrictions for downloading videos, images, or other items from the Internet seem to be rather limitless. For instructors who have numerous or large classes, there is no concern about students downloading too many pictures or that the class may reach a certain gigabyte limit half way through a course.

### **Simplicity**

Wikis are adaptable to most any course. Whether it is a class that involves writing, presentations, or any kind of task or project-based assignments, the user friendliness of the software makes it a 'do all' tool for educators. An extremely useful feature is its ability to allow two users to be working on the same project at the same time. This is especially beneficial in writing courses where collaborative undertakings may be desired.

The important factor in utilizing a wiki is that students are learning with the tech

tools of today. Gone are the days of pencil and paper, today's educational environment involves computers and creativity. Wikis allow students to become actively involved in the information gathering process by tying the Internet to tasks and projects. In addition, wiki pages offer a variety of add-on features called widgets. Widgets allow students to add video, calendars, spreadsheets, and other options. No one course is likely to utilize all the widgets available, but there is much room for innovation and flexibility.

Wikis that are created through Wikispaces.com are automatically saved by the software as learners construct their work. An icon will appear indicating 'draft saved'. Additionally, although users are requested to do so, all work is routinely saved, should students at the end of a class fail to log out and just shut down the computer. The next time they return to their wiki all of their previous tasks will have been saved so that they can continue working from right where they left off.

Another especially useful feature for instructors is the 'changes made' icon. By clicking this button teachers can effortlessly follow student progress in or out of the classroom. The feature records every change made, along with the time and date the change is made. This is useful when checking whether learners are completing homework assignments or even accessing their wikis outside of regular class time. It is also helpful to discern which members of a group are actually contributing to team work and which are not.

### **Latest Feature**

As wikis are constantly updated and improved with added components, one of the most recent improvements is the ability to comment on work that has already been submitted. Comments or questions can be posted by anyone registered to that wiki. This works



exceptionally well in writing or discussion based courses; large or small. Postings can be made by instructors to individual students to offer advice and comments on their writing or from student to student; whichever the mode, all writings and comments can be made and viewed by students and instructors in a timely manner. This online access can allow real-time assessment and communication to flow freely.

Also, if desired, comments may be posted in a more blog-like fashion with multiple replies and remarks that can be viewed effortlessly by all. This is a very practical feature for generating dialogue between students, which adds to the multiplicity of wikis, thereby demonstrating the value of wikis for use in education.

### **Detriments**

The drawbacks to utilizing wikis within a learning environment seem negligible at best. There are those who will always be apprehensive about Internet security. By utilizing proper precautions, this can become a nonissue. This is especially true when students always remember to apply safety protocol regarding personal and private information, such as not posting their phone number or home address. Aside from Internet security, there are three other areas where issues could arise:

1. With student access to the Internet granted, learners may get easily distracted or engrossed in a certain item. For example, if students are required to describe in English two Japanese festivals that they have downloaded video for, they may become so engrossed in watching videos that they may lack the time to write any English descriptions. Without proper monitoring by the instructor, classes could easily dissolve into nothing the more than playtime on the Internet.
2. An issue of concern for conversation classes is the lack of actual discourse that

might take place in this kind of environment. If tasks are geared only toward information gathering by students, no tangible communication or language skill training may be taking place. To alleviate this concern, the use of presentations could be encompassed into a course. In this fashion, learners not only show what they have accomplished, but they can utilize English to reflect on learning by explain to their peers why and what they have accomplished.

3. Writing and commenting on posts may also become cause for concern if the comments become too casual and lead to only short text type messaging between users. In this case students should be encouraged to write longer posts utilizing complete sentences.

As mentioned above, each of these possible disadvantages to wikis can be easily eliminated with a little careful lesson planning, monitoring, and a thorough explanation of what the instructor expects students to accomplish in classes where wikis are utilised.

### **Added benefits**

With the implementation of a wiki or similar software, the need for physical paper can be all but eliminated, paving the way for a more ecological and greener classroom. While one class may not amount to a noticeable reduction of paper, spread over a number of courses campus wide, the savings to a university or learning institution could become immense. As many campuses now are also Wi-Fi enabled, being restricted to computer rooms is no longer a deterrent. It provides freedom for students as well as instructors to access their wikis at anytime and anywhere. In large writing courses, wikis can make the management of an introductory writing course for non-English majors possible by allowing instructors to simplify the workload and ease the burden of having to haul

around students' writings in countless notebooks.

Although this paper has attempted to justify the use of wikis in a primarily EFL classroom there are no barriers to applying wikis in other languages Any language can be used as long as the necessary language fonts are installed on the computer. This is a simple process that can be effortlessly implemented on any Windows seven computer by going to the Start Menu and selecting Control Panel, and then Regional and Language Options. From this box a user can go to Date, Time, Language, and Regional options. Under Languages, additional languages may be added by checking the desired language box.

If any confusion or uncertainty by the instructor does arise there are also wiki pages that explain the 'how to' of wikis. As wikis have been in existence for a number of years now, there are already many examples of wikis used in various classes posted by teachers on the Internet. Additionally, over the years there have been many presentations and scholarly articles written on their extended applications in various educational environments. In fact, any project or task that any current instructor is considering undertaking has most likely already been executed as a wiki in some shape or form.

## **Conclusion**

As discussed in this article, wikis offer innumerable safe and secure opportunities for learning. Used individually or in groups, wikis allows students to take ownership of their learning tasks or projects. Under the guidance of an instructor, students can expand their knowledge of a particular subject while at the same time utilizing modern technology, which they may be called upon to implement after graduation and ascension into the working world. This connection between class work and the practical learning of life

skills can help stimulate learning and creativity amongst students. Wikis are an engaging educational tool that are procurable at no cost with a wealth of learning benefits that are not confined to any one style of instruction. The possibilities for educational enterprise in or out of the classroom are immeasurable. All that is left is the imagination and resourcefulness of the instructor.

## Choose Two, Pick One

**Sean Gay**

Ritsumeikan University

**A**s an educator, one of the key aspects of our work is preparing a functional and effective syllabus. As a teacher trainer, one issue which is extremely difficult to handle is explaining how to organize and balance a syllabus in an extremely short interval of time. Fortunately, a strong theoretical background can lead to simple and effective answers to complex problems. As such, the author explained the following simple but powerful tool for maintaining a balanced curriculum at the 2014 JALT Osaka Back to School event.

### **Language as Componential**

In understanding language, one of the fundamental questions is the nature of language as a unitary or componential construct. While some research supports connections between the four skills: speaking, listening, writing, and reading; language is best seen as being composed of the four skills rather than the four skills being aspects of a larger singular construct (Sawaki, Stricker, & Oranje, 2008, p. 7). This is important for assessment, but it is also essential in preparing lesson structure. As each of the skills needs to be seen as separate, it is important that all of the skills are addressed through the course of the syllabus.

### ***Integration***

These four skills, as addressed above, are separate and distinct constructs that are connected and complement each other. The visual skills of writing and reading can be

useful counterbalances and, when learned together, increase the effectiveness of both skills (Grabe & Zhang, 2013, p. 10). In fact, any combination of the four skills can increase learning (Movahed & Karkia, 2014, p. 60). What is important was not which skills are utilized, but that there are a minimum of two skills used in each activity.

### *Cognitive engagement*

The theory of cognitive engagement offers a framework for understanding the efficacy of integration. Cognitive engagement may result in a state of flow, or optimal experience, which results from activities centered between being too easy, thus boring, and too difficult, thus stressful (Csikszentmihalyi, 1988). This state of flow can increase motivation and support a potent learning environment. This is consistent with pedagogical literature, which asserts that engagement and processing are key components to learning; this also ties back to the integration of the separate skills through a connectionist framework (McLaughlin et al., 2005, p. 7). Utilizing multiple complementary skills increases the cognitive load which promotes engagement. Without this engagement, student interest in the content becomes irrelevant.

### **Fluency, Accuracy, and Complexity (FAC)**

The four skills are not the only delineating factors for language. Language use and competency can be divided into fluency, accuracy, and complexity (Skehan, 1998). Fluency is the ability to utilize or comprehend language quickly and smoothly. Accuracy is the utilization or comprehension of correct grammar and vocabulary. From a pedagogical perspective, complexity can best be seen as the acquisition of new knowledge, or the combination of old knowledge into larger constructs. Each of these competencies is present in each skill. Unlike the four skills, however, these three competencies often

interfere with each other (Nicholson, 2014, pp. 265-266). The combination of these competencies, therefore, leads to an affective cognitive burden.

### **Choose two, pick one**

The culmination of these theoretical underpinnings leads to a simple concept for the formulation of activities and the balancing of syllabi. In order to promote connections between the different language skills, it is important to integrate at least two of the four skills in each activity. This increases the cognitive load in a manner which promotes engagement. On the other hand, it is essential to focus on only one of the FAC competencies. If the focus is not distinct, or if there are multiple competencies targeted, then the cognitive burden will make the activity difficult and non-productive for the students. A chart, which maps the skills and competencies used in each activity, can be set up to promote adherence to the “choose two pick one” principle. Furthermore, that chart can be used to visualize the balance of skills and competencies in the syllabus. There is always difficulty in addressing how a syllabus will be balanced; difficulty demonstrates reflection and consideration of the pedagogical principles involved in the composition of a syllabus. The “choose two pick one” principle is useful in focusing teachers’ attentions on the goals of activities on the syllabus. This focus can help teachers balance their syllabi to support a strong and well-rounded learning environment.

### **References**

- Csikszentmihalyi, M. (1988). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge: Cambridge University Press.
- 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.

- McLaughlin, M., McGrath, D. J., Burian-Fitzgerald, M. A., Lanahan, L., Scotchmer, M., Enyeart, C., & Salganik, L. (2005). *Student content engagement as a construct for the measurement of effective classroom instruction and teacher knowledge*. Washington, DC: American Institutes for Research.
- Movahed, R., & Karkia, P. (2014). Reading/Listening & the 4/3/2 on EFL Students' Speaking Skills. *International Journal of Linguistics (IJL)*, 6(1), 53-66.
- Nicholson, S. J. (2014). An Analysis of the Task-Based Syllabus. *International Journal of English Language Education*, 2(1), 259-272.
- Sawaki, Y., Stricker, L., & Oranje, A. (2008). Factor structure of the TOEFL Internet-based test (iBT): Exploration in a field trial sample (Research Report No. RR-08-09). Princeton, NJ: Educational Testing Service.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.