

How teachers promote young language learners' engagement: Lesson form and lesson quality

Keywords: Engagement; foreign language; mixed-methods; motivation; classroom practice

W. L. Quint Oga-Baldwin
Waseda University

Yoshiyuki Nakata
Doshisha University

Cite as:

Oga-Baldwin, W. L. Q., & Nakata, Y. (2020). How teachers promote young language learners' engagement. *Language Teaching for Young Learners*, 2(1), 101–130.
<http://doi.org/10.1075/ltyl.19009.oga>

Funding information:

This study was funded by a JSPS KAKENHI grant-in-aid for young scientists (B) [24720260]

Abstract

Teachers of young learners often seek guidance on how best to engage and motivate their students. In this study, we aimed to document engaging teacher practices in the context of foreign language classes in Japanese elementary schools. We surveyed 16 public elementary school foreign language classes in western Japan using quantitative (questionnaire; external rating) and qualitative (naturalistic observation) tools grounded in self-determination theory. Classes were sorted into three groups of high, middle, and low teacher support based on student surveys, and observed for practices that influenced student engagement in each tercile. Results indicate that students are most responsive in classrooms involving teacher warmth and strictness, homeroom teacher involvement, appropriate pacing, instructional clarity, and a balance of activities. We offer descriptions of how these practices were employed, with implications for classroom practice and teacher training.

1. Introduction

Instruction in English as a foreign language has seen a sharp increase in elementary schools across the world (Garton, Copland, & Burns, 2011), with notable growth in Asia (Butler, 2015). Japan has been no exception to this trend. At the same time, many elementary school teachers in Japan and around the world are struggling to adapt to these new curricular policies (Butler, 2005; Copland, Garton, & Burns, 2014). Many teachers, especially those in the Japanese context, lack training in foreign language pedagogy (Butler, 2015). Teachers in these settings struggle with issues of classroom management, motivation, and particularly how to get students to speak English (Copland, Garton, & Burns, 2014). On top of this, many teachers in at all levels of Japanese education struggle with implementing communicative teaching methods (Nishino & Watanabe, 2008). Recognizing that these issues are considered of central curricular importance in Japan and many other countries (Butler, 2015; MEXT, 2008), elementary teachers need practical guidance on how to effectively engage and motivate their students in language learning.

As noted by Copland and Garton (2014), many training manuals for teachers of young language learners (YLLs) offer concrete suggestions for instruction, but may still lack empirical classroom-based analysis of the efficacy of these practices and the principles behind them. Studies have indicated the importance of classroom teaching to improving students' motivation to learn the language (Nikolov, 1999), but these have not identified theoretically unified principles for organizing instruction. A key point to note is that ideas such as motivation and engagement may be very difficult to separate from the classroom context in which they occur (Brophy, 2010). With regard to foreign language motivation in elementary schools, Butler (2015) has stated that “we need more contextualized approaches to understand motivation” (p. 319). Thus,

in order to understand the relationship between motivation and elementary foreign language learning, empirical research on effective classroom practices is necessary.

In order to answer some of the challenges of motivating and managing students in these educational settings, we sought to provide teachers of young learners with instructional principles based on robust theory and empirical classroom practices by answering the following research question: *What instructional practices in elementary foreign language classrooms support and engage students?* We used a *quan-QUAL-quan* mixed-method approach (Creswell, 2009) to define instructional principles through empirical observation of elementary school classrooms in Japan. We triangulated students' reports of teacher support with external ratings of students' behavior, then described features common to highly engaged classes to deduce connected practices that promote enjoyment, attention, and cognition. For the purpose of this discussion, we use the term *principle* to mean a fundamental, flexible approach to instruction that underlies and links a number of classroom instructional approaches, while *practices* are the visible, observable actions that teachers take.

2. Literature review

2.1 Engagement

Motivation is one of the major issues faced by teachers of YLLs (Copland, Garton, & Burns, 2014). In class, students' personal and situational motivation manifests itself as engagement in learning tasks (Oga-Baldwin & Nakata, 2017; Philp & Duchesne, 2016; Svalberg, 2009). According to the current conceptualizations, motivation is best understood as the antecedent attitudes that students bring to class (Oga-Baldwin, 2019), or perhaps in lay terms the *why* that students attribute to their language studies (Nikolov, 1999; Noels, Pelletier, Clément, & Vallerand, 2000). Engagement is the momentary state where students' motivation becomes action.

Engagement is represented by students' emotional investment, often called *emotional engagement*; positive classroom behavior, titled *behavioral engagement*; and active thinking and learning, also known as *cognitive engagement* (Fredricks, Blumenfeld, & Paris, 2004). From this perspective, students' active use of the language, including listening and speaking, are included in the concept of engagement (Philp & Duchesne, 2016; Svalberg, 2009). Engagement can be understood to further prime students' future motivation (Oga-Baldwin & Nakata, 2017, Oga-Baldwin et al., 2017; Oga-Baldwin, 2019; Hiver & Al-Hoorie, 2019).

Thus, the noted challenges with motivation, speaking, and classroom management (Copland, Garton, & Burns, 2014) appear to be linked under the question of how to engage students. Currently, foreign language study in Japanese elementary schools is organized around the idea of increasing engagement and experience with English (Butler, 2015; Oga-Baldwin & Nakata, 2014b). At the same time, teachers in this context are often not trained as foreign language teachers (Fennelly & Luxton, 2011). Teachers of young language learners need specific principles based on empirical analysis in order to provide an engaging language learning environment. Prior models of engagement (Oga-Baldwin & Nakata, 2017; Oga-Baldwin, Nakata, Parker, & Ryan, 2017; Jang, Kim, & Reeve, 2012) have made use of the principles employed by self-determination theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2017).

2.2 *Autonomy support and structure*

Self-determination theory (SDT; Ryan & Deci, 2017) is based on the principle that human beings flourish and function best when they experience intrinsic motivation. Intrinsic motivation likewise grows in nurturing social environments, where people experience satisfaction of their autonomy (a need for personal

psychological freedom and sense of volition), relatedness (a need to feel socially connected to other people), and competence (a need to feel capable of influencing the surrounding world). These theoretical concepts are particularly well suited to the elementary foreign language activities context in Japan (Oga-Baldwin & Nakata, 2014b). In formal language education, learners thrive in situations where their teachers provide a need-supportive learning environment (McEown & Oga-Baldwin, 2019). The principles by which teachers create this environment are organized around constructs titled *autonomy support* and *structure*.

According to SDT, autonomy supportive teaching is central to nurturing students' motivational needs. Reeve (2012) defines autonomy support as what teachers do and say to facilitate students' sense of personal volition and investment in learning. Likewise, structure is the way in which teachers clearly demonstrate their expectations toward achieving educational outcomes (Skinner & Belmont, 1993). In broad terms, autonomy support may be recognized as the *quality* of the teachers' interactions with students, while structure represents the *form* of the instruction. More concretely, autonomy support may potentially be a way in which teachers remove barriers to action by understanding students' perspective and supporting interest; structure is how teachers make activities achievable through scaffolding (e.g., organizing materials, simplifying task steps, modelling expectations). Previous work shows that autonomy support and structure are strongly, and perhaps indelibly, linked (Oga-Baldwin & Nakata, 2015; Jang, Reeve, & Deci, 2010; Sierens et al., 2009).

While the above definition of autonomy supportive instruction is broad, studies have attempted to provide clear examples of what types of behavior promote students' autonomy in the classroom. Reeve and Jang (2006) found positive effects for teachers who listen to students, allow students to work in their own way, increase

student talk time, offer hints and encouragement, and acknowledge students' perspectives. Likewise, teachers who monopolized time and materials, gave controlling commands, and did not allow students the opportunity to discover answers on their own were perceived more negatively. Studies focusing on the quality of interactions between teachers and foreign language students have also demonstrated that teachers' autonomy support leads to greater intrinsic motivation to learn a foreign language (Oga-Baldwin et al., 2017; Carreira, Ozaki, & Maeda, 2013; Noels, 2001).

The other side of teaching in order to motivate students is the instructional form. How teachers structure their guidance in class to lead students toward learning goals is a key feature of engaging instruction (Skinner & Belmont, 1993). By helping students to recognize the next step, teachers reduce the difficulty of the learning task (Kirschner, Sweller, & Clark, 2006). Through appropriate guidance, explanations, activities, and materials, teachers can help to motivate students and improve their sense of volition and autonomy in class (Brophy, 2010; Cornelius-White & Harbaugh, 2009). This aspect of teaching has strong overlap with classroom management (Good & Brophy, 2008).

Past research has indicated that in many cases, the form and quality of instruction are linked. Jang and colleagues (2010) demonstrated that autonomy support and structure together lead to the best student learning. Using both classroom observation and student self-report, this study showed that the most engaged students had teachers who provided both support for students' inner motivational resources *and* were clear and authoritative in their guidance. Sierens and colleagues (2009) found similar positive effects on self-regulated learning.

Recognizing that teachers express their educational goals in specific ways for specific subjects (Brophy, 2010, p. 29–30), what helps to create the sense of purpose

and forward motion in class may differ across domains. Foreign language classes, especially foreign language activities classes (Kang & Crandall, 2014), require interaction and language use to be properly communicative, while other subjects might not require the same level of interpersonal communication between teachers and students. Thus, how foreign language teachers express effective structure and autonomy support may necessarily differ from other subjects. How these principles are expressed in effective language education is important for contextualizing young language learners' motivation (Butler, 2015, p. 319).

2.3 Engaging instruction in foreign language education

The research on autonomy support and structure overlaps with that of motivational strategies (i.e., practices) in foreign language classes. Language education has worked from the framework of these motivational practices (Dörnyei & Csizér, 1998; Dörnyei, 2001), many of which concur with autonomy supportive teaching, such as supporting learner autonomy through choice, developing positive relationships with learners, and making classes interesting. Other practices resemble structure, including directives to present tasks properly, increase learners' self-confidence, and create a personal example through teachers' behavior.

While these strategies may theoretically have numerous positive effects, empirical research has had difficulty consistently demonstrating the effects of any single teaching strategy (McEown & Takeuchi, 2012). Though many of these individual practices have shown weak or non-significant effects, other research has demonstrated that employing these practices together may engage students in learning activities (Guilloteaux & Dörnyei, 2008). As noted, the use of multiple strategies resembles supportive and structured teaching (Reeve & Jang, 2006), which are linked

to positive motivation (Oga-Baldwin & Nakata, 2014b; Wu, 2003; Butler, 2015; Carreira et al., 2013).

Recognizing that any single practice or strategy by itself may not be sufficient to motivate students, we sought to describe how highly engaging elementary foreign language teachers provide structure and autonomy support to their lessons. Prior studies have worked from a pre-existing framework of expected practices (Guilloteaux & Dörnyei, 2008; McEown & Takeuchi, 2012), which may exclude practices not previously theorized or recognized as relevant to the classroom setting. These theoretical language learning strategies have not yet been tested in settings with YLLs, thus justifying the use of an exploratory approach. Using bottom-up methods to investigate classroom principles, we hope to offer a more dynamic perspective on teaching practices influencing engagement and motivation (Dörnyei, MacIntyre, & Henry, 2015), further following the call for more non-linear approaches to motivation in classrooms with YLLs (Butler, 2015). In order to address this lacuna, we documented the practices of teachers in highly engaging, highly supportive classes.

3. Methods

We used a mixed-methods approach to triangulate findings from a large-scale quantitative study investigating young language learners' longitudinal motivational development (Oga-Baldwin et al., 2017; Oga-Baldwin & Fryer, 2018). The previous survey study found that, unlike in similar settings (Carreira, 2011), the current cohort of pupils' autonomous motivation did not significantly decrease over time. Based on these conclusions, further inquiry is necessary to understand the reasons behind this trend. Using a separate set of data from the larger study, the current research is an extension of the previous work using a *quan-QUAL-quan* design (Creswell, 2009),

where *quan* represents the quantitative grounding justification for the study, QUAL represents the core qualitative component of analysis, and the final *quan* represents quantitative follow-up to confirm the validity of qualitative results.

3.1 Participants and setting

The student subjects for this study came from seven suburban schools in southwestern Japan. A total of 515 students in 16 classes agreed to participate in the study. Cooperation for this research was provided by the principals and teachers at each school, with the support of the local board of education. All studies were granted approval by the [University] Ethics Review Board. Local boards of education provided permission for the research, coordinating with school principals and teachers. All participating students' parents, teachers, and principals were informed of the scope and aims of the study before agreeing to sign permission forms. Principals communicated with parents and obtained permission to gather student data.

The schools ranged in size from just over 100 students in grades one through six to close to 1000 students in all grades. The largest participating school had 4 classes of 35+ students, while the smallest had only 1 class of roughly 25 students. District per capita income was similar to the per capita GDP for all of Japan (Japan Statistics Bureau, 2016). The schools observed were similar to those described in previous ethnographies (*cf.* Cave, 2007); they can be considered roughly representative of non-urban schools throughout Japan. Profiles of participating schools are presented in Table 1.

All schools employed a native- or near-native English speaking assistant language teacher (ALT) who was primarily responsible for executing class plans.

Classes were planned in conjunction with homeroom teachers (HRTs) and Japanese-nationals trained as English teachers (Japanese Teachers of English; JTEs). While some of the ALTs had practical licenses for teaching English (i.e., CELTA, TESOL, etc.), none of the teachers indicated specialized knowledge of motivation or motivational strategies. Schools 3, 4, and 7 employed a JTE, while 1, 2, 5, and 6 did not have a trained specialist on staff. None of the participating homeroom teachers had any qualification for English proficiency or language teaching. Teachers' individual classroom roles varied by school and situation; some HRTs were active in managing and planning curriculum, while others allowed ALTs, JTEs, or other HRTs in the school to do the planning, with a similar variety of roles as those noted in previous research (Aline & Hosoda, 2006). Likewise, teachers' working relationships could be highly varied, ranging from full collaboration to near complete independence (Nakao, 2009; Nakao, Oga-Baldwin, & Fryer, 2019). According to Ministry of Education policy, specific role guidelines for JTEs, HRTs, and ALTs remain vague (MEXT, 2008).

Classes generally followed the Ministry endorsed communicative language teaching approach (MEXT, 2008). Teachers did not use a task-based curriculum, though some tasks (as real-world communication focused activities) were occasionally brought in as part of a unit (e.g., a fast food shop roleplay). Nearly all of the activities observed across all classes were repetition and identification drills, game-like or game-oriented activities (e.g., passing a ball, completing a BINGO sheet), or highly controlled interviews (e.g., ask 5 friends their favorite color). These activity choices align with the materials provided by the textbooks (MEXT, 2012). This approach has largely become standard throughout Japanese elementary schools, and there was very little differentiation between schools on the types of activities

chosen. In line with this curriculum, regular summative assessments were not provided to the children.

Table 1. Profiles of participating schools.

School	Number of Grade 5 Classes	Specialized English Classroom	Specialist Japanese Teacher of English
1	4	No	Not employed
2	3	Yes	Not employed
3	3	No	Employed
4	1	Yes	Employed
5	2	Yes	Not employed
6	1	No	Not employed
7	2	Yes	Employed

For this study, fifth-year classes were chosen as this is the first year targeted for foreign language study in Japanese elementary schools (MEXT, 2008). While previous findings have shown that motivation often decreases in upper elementary school (Carreira, 2011), these students generally showed no significant decrease in motivation over the course of a single school year (Oga-Baldwin et al., 2017), and in many cases motivation improved (Oga-Baldwin & Fryer, 2018). Given that motivation is both a predictor and outcome of engagement (Oga-Baldwin & Nakata, 2017; Oga-Baldwin et al., 2017; Oga-Baldwin, 2019), the practices which support engagement in class are crucial for understanding how motivation develops. Thus, the current study may offer explanations as to why these specific students maintained and improved their motivation to learn a foreign language. In order to explain this trend, we used nearly 50 hours of classroom video, three class periods from each participating class filmed three times during the 2013-2014 school year.

3.2 Instruments

Supportive and structured teaching. Students rated their classes for autonomy supportive and structured teaching using a 5-item measure developed for elementary

foreign language classes (see Oga-Baldwin & Nakata, 2015 for instrument creation procedures). Items measured the teachers' clarity, pacing, English usage, guidance, and elicited interest in the subject. Internal reliability for the scales was acceptable (Cronbach's $\alpha = .75$; Devellis, 2012). Surveys were taken during the Fall of 2013. Students' survey answers were then used to create tercile ranks for each class; classes were ordered from highest to lowest rating, with the top five ranked in the top tercile, the middle six assigned to the middle tercile, and the lowest five placed in the bottom tercile.

Intrinsic motivation. Students reported their motivation using the three-item intrinsic motivation scale from a Japanese translation of the Self-Regulation questionnaire–Academic (SRQ-A; Ryan & Connell, 1989; Oga-Baldwin & Nakata, 2017). Surveys were taken at the beginning and end of the 2013 school year. Internal reliability was good (Cronbach's $\alpha_{\text{TIME 1}} = .83$; Cronbach's $\alpha_{\text{TIME 2}} = .84$).

External ratings of engagement. Observers were asked to rate the 48 class videos for collective engagement on a 5-point scale. Raters assessed students' collective engagement during each minute of the classes. Each rater assessed engagement alone. The mean score for each class was then calculated. Inter-rater agreement was calculated using Pearson's correlation coefficient ($r = .93, p < .001$).

Examples of all instruments are presented in Appendix 1.

3.3 Observation procedures

Observers were two language teacher trainees, selected based on their training as English specialists. They applied to be research assistants for this project funded by a JSPS research grant, and were trained in qualitative observation. Observations were conducted using a paradigm of naturalistic inquiry (Cozby & Bates, 2012; Creswell,

2009), where the observers were allowed to make observations about a natural setting without a priori expectations. Observers made these observations after completing ratings for students' collective engagement, reviewing each of the 16 classes and nearly 50 hours of video multiple times to observe patterns and code the practices according to the principles of autonomy support and structure. Observers were paid for their time. The design for qualitative observation and coding is presented in Figure 1.

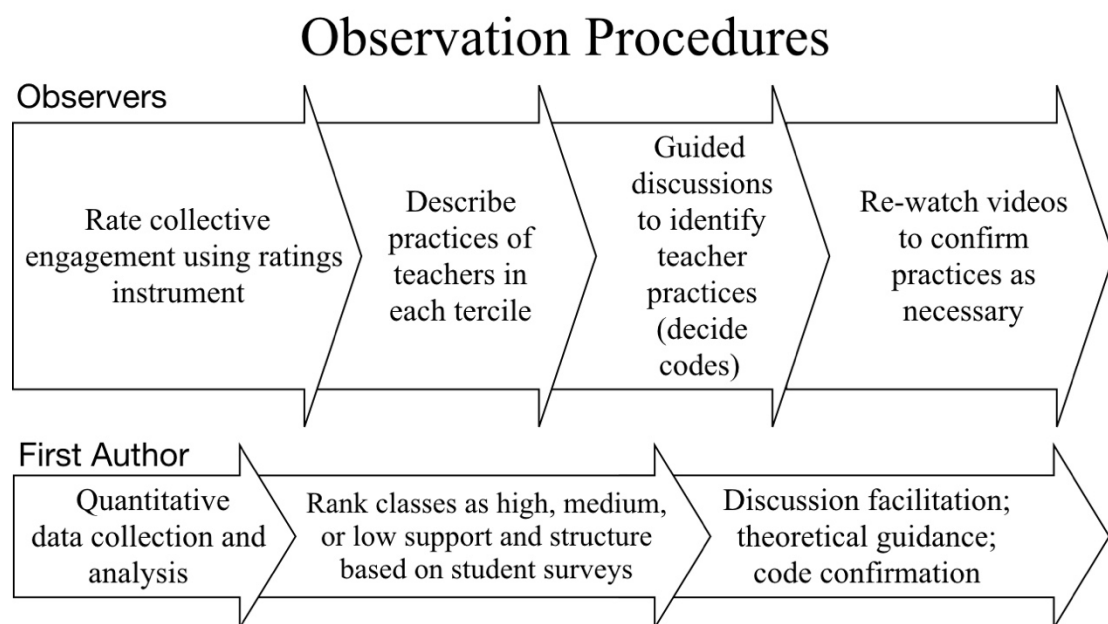


Figure 1. Observation and coding procedures

During the first round of rating and observation, raters were allowed to make naïve assessments in order to best simulate the type of observations made by teachers, parents, and supervisors without training in motivational theory. They were not given pre-existing observation instruments, as items listed on many observational schemes (e.g., Guilloteaux & Dörnyei, 2008) corresponded with indicators from the classroom surveys, and might have unintentionally biased observations. Instead, observers watched each class video at least three times, reached mutual agreement on practices that influenced students' observed engagement, then organized these into inferred

principles of autonomy support and structure.

During the documentation process, each observer had veto power with regard to individual inferred practices; either observer who disagreed with the validity of a specific code application discussed could have it removed. For example, one of the observers believed that specific feedback on activities engaged the children, while the other disagreed based on their observations. As the pattern could not be consistently demonstrated in the highest tercile classes, the practice of *feedback* was removed from consideration. When observers reached a consensus, conclusions were then compared to students' quantitative ratings. Practices were then logged and recorded, and finally coded into the principles of either autonomy support, structure, or a combination of the two.

As the principal investigator, the first author moderated all discussions but was not directly involved in the assignment of principles beyond arbitration. Due to the relationships developed between the principal investigator and the teachers and students through direct participation in the data gathering, a fresh perspective was necessary to avoid personal bias. Because these observers had developed an understanding of these classes without personal contact with the teachers or students, observers lead the documentation while the first author managed the data and connected the relevant codes.

3.4 Analyses

Qualitative and quantitative analyses were conducted double-blind in order to prevent unintentional bias. The authors were not present during the rating, and the observers did not have access to the quantitative survey data until after the completion of their ratings and qualitative coding.

Quantitative analyses. Inter-rater agreement for tercile rankings was assessed using

Cohen's Kappa statistics calculated in Stata 14 (StataCorp, 2014). Inter-rater agreement used both observers' external ratings and terciles according to students' self-reports. Both raters' combined agreement with the actual data (using students' ratings of classroom support) was calculated using Cohen's kappa, .62, $p < .000$, indicating substantial agreement (Landis & Koch, 1977).

Student rating terciles were compared with raters' engagement scores and self-reported motivation using multivariate analysis of variance (MANOVA), followed by ANOVA tests to look for finer differences. Higher rated terciles were hypothesized to show higher engagement and motivation.

Qualitative analyses. The current research coded different principles according to autonomy support, structure, and a combination of the two. While the initial codes were documented and described by the rater-observers without knowledge of theoretical background, the first author later condensed the codes through discussion with observers and sorted them according to theoretical principles. The second author then offered peer-debrief to confirm the codes and provide confirmability and credibility to the coding.

4. Results

In order to reflect the mixed-methods used to analyze this data, we have divided the results into three parts: the quantitative results that provide a basis for the naturalistic observation; the corresponding qualitative factors which influenced students' engagement; and two specific incidents illustrating the teachers' use of these instructional principles.

4.1 Quantitative results

Raters' combined agreement indicated the tercile rankings to be an effective

description of the data. Table 2 displays the raters' rankings of classroom structure compared to rankings created by self-report results. In discussing the data, high, middle, and low terciles will be used as a shorthand. Students' rankings based on their mean scores were adopted as representative as they reflect the consensus of the class regarding teachers' support.

Table 2. Rater rankings compared with student reported score rankings, class-level self-report mean for support, engagement, and student motivation at the beginning and end of the school years. 1 = Lowest Tercile, 2 = Middle Tercile, 3 = Top tercile. External engagement is the mean of both raters' scores.

School	Class	Student Rating of Supportive Teaching (Mean Score)	Student Rating Score Tercile Rank	Externally Rated Engagement Mean	Student Motivation Time 1	Student Motivation Time 2
1	A	4.24	3	4.18	3.24	3.39
	B	3.73	1	3.11	3.50	3.15
	C	3.68	1	3.69	3.37	3.22
	D	3.65	1	3.27	3.53	3.58
2	E	4.44	3	3.76	3.67	3.86
	F	4.04	2	3.64	3.83	3.93
	G	4.10	2	3.53	3.31	3.52
3	H	3.81	1	3.24	3.35	3.11
	I	3.94	2	3.22	3.69	3.53
	J	3.81	1	3.06	3.31	3.16
4	K	4.42	3	3.78	3.69	4.06
5	L	4.04	2	3.95	3.84	3.95
	M	4.31	3	3.92	3.88	4.29
6	N	4.29	3	3.61	3.76	3.80
7	O	4.13	2	3.59	3.67	3.79
	P	4.11	2	3.86	3.78	3.86

Students' self-reported motivation and raters' assessment of students'

engagement also reflected students' rankings. MANOVA tests found significant differences between the groups, Pillai's trace = 0.81, $F(6, 24) = 2.70, p = .038$. Follow-up ANOVA tests confirmed the differences across groupings as significant for observed engagement $F(2, 15) = 7.19, p = .008, \eta^2 = .53$; motivation at the beginning of the year, $F(2, 15) = 5.36, p = .02, \eta^2 = .45$, and motivation at the end of the year, $F(2, 15) = 6.72, p = .01, \eta^2 = .51$. A within-subjects ANOVA test further revealed significant differences between motivation at the beginning and end of the year, $F(2, 15) = 10.19, p = .002, \eta^2 = .61$. The effect sizes for each of the differences between and within groups were large (Kline, 2009), with the tercile differences explaining between 45% and 61% of the model variance. Figure 2 shows an increasing trend in engagement and motivation in the highest tercile, and a decreasing trend in the lowest, while Table 3 provides the correlations between student-rated teacher support, externally rated engagement, and students' self-reported motivation at two time points.

Table 3. Correlation table for support, engagement, and motivation

	1	2	3	4
1. Student rated teacher support	-	.67**	.44**	.72**
2. Externally rated engagement		-	.31**	.61*
3. Intrinsic motivation time 1			-	.85***
4. Intrinsic motivation time 2				-
Mean	4.05	3.59	3.59	3.64
95% Confidence Interval	[3.91, 4.18]	[3.41, 3.76]	[3.47, 3.70]	[3.45, 3.82]
SD	.26	.32	.22	.36
Cronbach's Alpha	.75	-	.83	.84

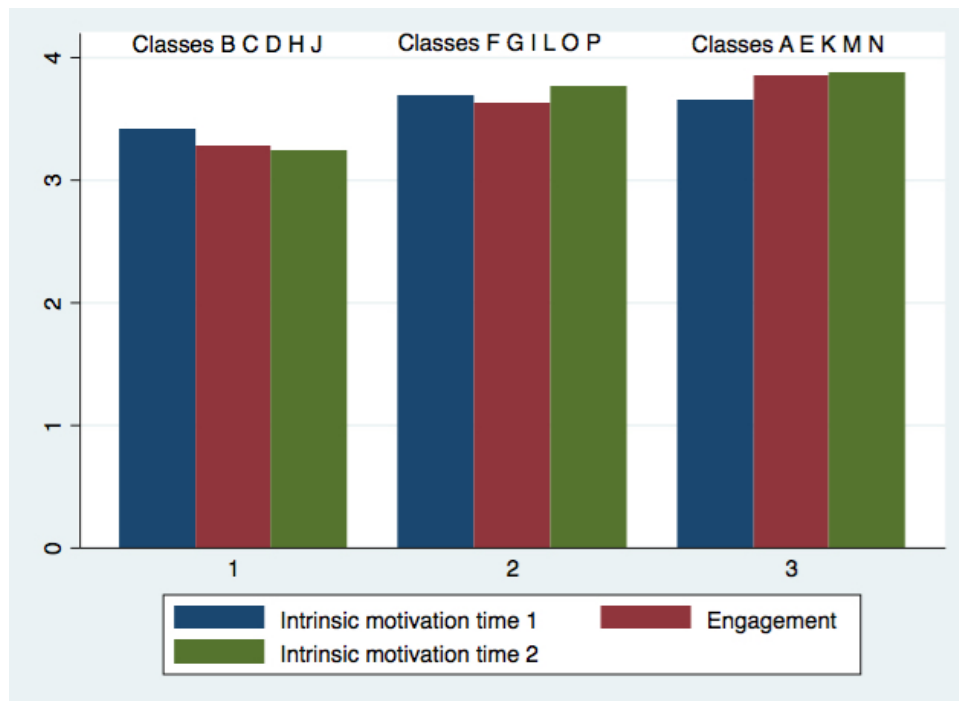


Figure 2. External engagement rating over tercile ranks according to student classroom ratings. X-axis: 1 = lowest tercile, 2 = middle tercile, 3 = highest tercile.

Based on these results, we concluded that students' ratings of teacher support sufficiently represented the realities of classroom life, in addition to functioning in theoretically consistent ways (Reeve, 2012). Students experience instruction on a daily basis and are keenly aware of how instruction may meet or thwart their needs. Likewise, the raters saw all 16 classes multiple times, representing a macro-outsider perspective on all classes, while the children's perspective was that of an insider participant (Nakata, 2014). At the intersection of these insider and outsider perspectives, the qualitative factors emerged from the observations. Table 4 presents the final categorizations of these practices according to autonomy support, structure, and the overlap between the two. Each of these factors is detailed in the following qualitative results, and confirmed with quantitative analysis where possible. The practices noted are those which were used regularly and consistently in the top tercile, and could be used to clearly differentiate those classes from the other two. We have chosen to focus primarily on those practices found in the top tercile as we feel they offer the best hope for instructors hoping to emulate an engaging teaching style.

Table 4. Categorizations for each observable teaching practice

Principle Categorization	Observed classroom practice
Autonomy Support	Homeroom teacher involvement Warm / strict teaching
Autonomy support and structure	Balance of activities Progressive games Keep it short Predictable interactive routines
Structure	Signals for meaning Repetition, demonstration, simplification

4.2 Qualitative results: Autonomy support

The practices described in this section refer to ways in which teachers recognize students' perspective, build interest, create positive student emotions, and help students to feel comfortable in the classroom.

4.2.1 Homeroom Teacher Involvement

In the top tercile classes, homeroom teachers played a large role in students' and observers' perceptions of structured teaching. These observations confirm the importance of the HRT, as the top classes were those where HRTs were consistently in the room and shared teaching responsibility with ALTs. This feature was consistent; only low tercile classes had passive homeroom teachers.

Homeroom teachers, through daily contact, are more likely to have an existing positive relationship with their students. While ALTs may develop positive relations with students, their status as an outsider assistant and the infrequency of classes (once per week) make it difficult to build the same level of relationship with homeroom teachers.

To confirm these findings, a follow-up comparison was conducted to test the number of homeroom teacher utterances (in any language) in each observed class. Table 5 presents the differences in homeroom teachers' utterances. In the lowest tercile, HRTs spoke rarely, with an average of 22 utterances per class. The middle

tercile spoke more, with an average of 29 utterances per class. The top tercile showed the most active homeroom teachers, with an average of 72 utterances. A follow-up ANOVA test indicated differences between the terciles were significant, $F(2, 15) = 7.37, p = .007, \eta^2 = .53$. Looking at the proportion of English (versus Japanese) usage, no consistent pattern was found across terciles in the percentage of homeroom teacher English utterances (see Table 5), though more involved teachers in the top tercile made more overall English utterances. These utterances may further show an indication of the HRTs' attitudes toward learning English via their classroom involvement.

Table 5. Class with tercile rank and homeroom utterance counts

Class	Tercile Rank	Homeroom teacher utterances (count)	English utterances (count)	English Usage (% of Utterances)
A	3	105	49	47%
B	1	19	3	16%
C	1	6	3	50%
D	1	17	4	24%
E	3	50	27	54%
F	2	50	29	58%
G	2	22	9	41%
H	1	23	7	30%
I	2	41	30	73%
J	1	45	29	64%
K	3	46	28	61%
L	2	40	17	43%
M	3	55	40	73%
N	3	105	41	39%
O	2	9	3	33%
P	2	3	3	100%

The quantitative ratings offer further evidence of this point. As shown in Table 2, the same ALTs in Classes A and I were rated significantly lower in classes B, C,

and D or H and J, where the primary difference was the degree of involvement on the part of the homeroom teacher, indicated by their number of utterances. This agrees with previous research on Japanese classrooms (Authors, 2013), where HRTs serve as models for their students. Japanese teachers in low tercile classes often occupied the more passive roles of translator or interpreter (*cf.* Aline & Hosoda, 2006), potentially leading to lower student engagement.

4.2.2 *Warm/strict teaching*

Discussions among the observers further revealed that strict but warm teachers were thought to be the best, while teachers with a controlling style were seen in the most negative light. Teachers' warmth is well documented as an important factor in building positive relationships and managing classrooms (Cornelius-White & Harbaugh, 2009). As students had no access to lessons taught by other teachers, this feature was based on the observations of the research assistants, who saw all of the classes several times. While this may relate to other unobserved contextual factors, rating was consistent; teachers in the highest tercile were seen as warm and strict, and the bottom showed a controlling attitude. Top tercile teachers did not let students get off task, and had a clear idea of how to manage behavior and misbehavior. Both raters ranked classes run in this style in the top tercile. While these teachers were personable, they also allowed little room for off-task behavior.

The middle tercile classes were generally warm and focused, but occasionally lacked one aspect of the *warm/strict* practice. Some were warm, but somewhat permissive of student off task behavior. Others maintained order, but did so without a sense of geniality. These classes were not unpleasant or disorganized, but were missing one element of this practice.

The lowest classes had a negative atmosphere at times. In these classes, while

teachers were organized, they were angered by misbehavior and often issued commands like “sit down” and “be quiet,” or even “shut up.” Within the same school, students’ ratings of teachers differed largely based on the warmth of the teacher. While foreign ALTs taught in all the classes, in some classes these same teachers took a more controlling attitude and were thus ranked in the lowest tercile. This more negative attitude may have been related indirectly to the involvement of the HRTs and their working relationship between the ALTs. Homeroom teachers less willing to involve themselves in the classes may not have been active in planning classes with the ALTs, and may thus have modeled negative communication behaviors toward English, thus leading to diminished student enthusiasm, and the need to adopt a more controlling stance to engage students and keep them on task.

4.3 Qualitative results: Structure

The factors presented in this section refer to ways in which teachers facilitate students’ activity through clear instruction and providing students with the confidence to act.

4.3.1 Signals for Meaning vs. English “paint job”

The teachers in top tercile classes all included methods for signaling the meaning of instructions. By accompanying classroom instructions with gestures and demonstrations, teachers gave students support for comprehending the English used in class. Context, gestures, and other visual aids buttressed students’ understanding and allowed them to spend more time using English. Visual support for instruction assisted task presentation. Crucially, teachers used Japanese in a controlled fashion. While the quantitative ratio of Japanese to English was not a consistent indicator of quality instruction (see Table 5), in top tercile classes, English was a tool to share meaning and impart understanding.

In some middle and all low tercile classes, visual signals for meaning were weak, irregular, or absent. These classes needed more Japanese usage on the part of the HRT or JTE after the English instruction, providing the majority of teacher utterances from the *Homeroom Teacher Involvement*. While this allowed greater involvement on the part of the Japanese teachers, it also contributed to the “othering” of English; by virtue of the fact that only ALTs used English in these classes while HRTs and JTEs primarily used Japanese, teachers showed that English should be quickly translated. In these classes, English was not a structural feature of the class. Students were not required to comprehend teachers’ utterances in English, being given translation soon after. In these classes, students would listen to English utterances from their teachers, but quickly turn their heads to their Japanese teachers searching for a translation. The English used was not for communicative purposes, but rather something “painted on” the surface. When teachers failed to use demonstration and signals to scaffold meaning, they contributed to this phenomenon.

4.3.2 *Repetition, Demonstration, and Simplification*

The classes rated highest included a high degree of repetition, demonstration, and simplification to support students’ comprehension. As noted in the previous section, the most effective teachers did not rely on words alone to communicate, but contextualized their speech. These teachers used multiple modalities and forms of expression to help scaffold students’ understanding of the new language. More than simply using high frequency language, these teachers facilitated understanding through the classroom context.

Conversely, in classes in the bottom tercile, English was not often repeated as a part of interaction or modeling. Teachers would often make statements in English, but give no time or support for their English. Much as in the *English “paint job”*

feature, teachers would often translate without repeating or attempting to clarify English medium utterances.

4.4 Qualitative results: Combined Autonomy Support and Structure

The final category of teaching principles involve ways that are not clearly theoretically separable into autonomy support and structure, but involve elements of both.

4.4.1 Predictable Interactive Routines

The top classes' activities followed a predictable routine. Students had repeated activities previously, knew how to respond, and could call upon existing linguistic resources to complete tasks. Previous research has shown that students may feel secure in classes where teachers do similar routines each time (Oga-Baldwin & Nakata, 2014a).

A crucial element not previously documented is that the most successful routines were not simply rote production. Most classes had a series of predictable routines to one extent or another; however, the highest rated classes prompted real, rather than automatic, interaction and responses. In one top tercile class, physical response activities were executed in a fashion that required students to pay attention, either by adding new words or purposefully mismatching teacher gestures and instructions. By creating a situation where students needed to listen carefully, teachers made the contents of this routine interesting and challenging.

Routines were similarly featured in low tercile classes, though these routines lacked interaction and challenge. Choral repetition routines were predominant, and were sometimes even forcibly passive in their implementation; students were told to sit still and watch or listen. In the bottom tercile, the ALT in classes B, C, and D used a two-stage vocabulary presentation where students first listened silently and watched

the teacher present every word, and only after repeated the target language. Students noticeably increased off-task behaviors, such as fidgeting or looking out windows. Though clearly a routine intended to support recognition, the extra step of watching silently diminished students' interest. Conversely, this same ALT did not use this two-stage presentation in class A (prior to the other three), and was rated noticeably higher by the students. Thus, routines are not in and of themselves positive, but rather *meaningful* routines build both a sense of competence and draw student interest.

4.4.3 Progressive Games vs. Game upon Game

Much has been discussed regarding the importance of engaging students through an enjoyable classroom atmosphere (Cornelius-White & Harbaugh, 2009). To this end, games were a major part of all lessons. By framing activities as game-like, teachers were often able to satisfy students' needs and draw interest (Oga-Baldwin et al., 2017). However, the key to these activities was that the games and activities led were learning-centered (Cameron, 2001), rather than simply games for their own sake. In top classes, students were given the goal of the lesson, and each game was chosen to teach a new point, gradually reducing teacher support. Classes started with simple review and vocabulary identification activities, but pushed students to respond on their own as lessons progressed. In these classes, review was never passive, but rather elicited through use and recall.

Not every high tercile class included a so-called "joy factor" (Lemov, 2015). Classes in the bottom tercile were classes primarily comprised of games, while the top tercile sometimes involved activities that were not game like. Many commentators on English activities (e.g., Naoyama, 2011) stress the idea of fun as necessary and sufficient for elementary foreign language classes. However, in the top classes the games were not themselves used to drive engagement. Teachers instead drew out

engagement through scaffolding—in this case, providing clear examples of expectations, ordering activities in logical steps, and focusing children’s attention on key language points (Cameron, 2001; Pinter, 2011). This may relate to the combination of activities which are enjoyable with those which are meaningful, challenging, or personally valuable (Authors, 2009; Brophy, 2010).

In classes in the lowest tercile, games were simply piled on one another, either without a clear goal in mind or with the idea that simply doing a game would actively engage students in learning. While the teachers in all terciles used games, not all the games were clearly designed to educate. Low tercile games were redundant, simply repeating words and phrases specified by the teacher. This pattern was seen in all classes in the lowest tercile, as well as some in the middle tercile. The teachers planned activities as games with the intent of providing enjoyment, though this did not always come to fruition. Recognizing that not every class activity is necessarily fun (Brophy, 2010, p. 208), overemphasis on this aspect detracted from both learning and enjoyment.

4.4.4 Keep it Short vs. Overextension

Top tercile teachers kept students on task using short activities. Using activities that took less than 5 minutes to complete, teachers maintained interest in the learning tasks. Students were given opportunities to complete free interaction tasks, but were not given excess time to complete these activities. Not all activities were carried out until their final completion. This continued engagement may be interpreted as a manifestation of the Zeigarnik effect, where students show high task engagement after an incomplete task (Reeve, Cole, & Olson, 1986).

Low tercile classes used activities which focused on completion, decreasing students’ energy. Even enjoyable activities appeared tiresome when carried on too

long. Similar to the idea of *game upon game*, the game-like features of activities lose their value and effectiveness when the fun is lost (Brophy, 2010), and the effect appears to persist beyond that single activity. Students showed diminished engagement both during and after *overextended* activities.

4.4.5 Balance of Activities

Multiple modalities for scaffolding language were helpful in promoting student understanding. In order to provide these modalities, the most successful teachers also provided a balance of different types of activities. Learners who received the language through a listen and repeat, a chant, a song, a dance, and a game showed greater enjoyment. Students may have benefitted in the classes where teachers mixed and matched ways of presenting new vocabulary and expressions through the use of both interactive physical games (Tomlinson & Masuhara, 2009) and songs (Schön et al., 2008).

Low- and mid-tercile classes used listen and repeat activities to present the new language phrases and vocabulary before playing games. This scaffolding pattern used primarily passive recognition, and did not require students to use the material on their own. In more successful classes, teachers presented activities in multiple ways, allowing students to experience information in their preferred modality, be that physical movement, visual processing, or auditory stimulation. While different modalities do not change the basic information that needs to be learned or reduce cognitive burdens for processing the information (Willingham, 2009), they provide added exposure without rote repetition.

The *balance-of-activities* feature of classes had a secondary organizational effect in that it pushed teachers to shorten activities and use tighter transitions (Lemov, 2015; Oga-Baldwin & Nakata, 2014a), as noted in the *Keep it short* feature.

Students engaged with the language more deeply both by repeating it and encountering it in multiple ways (Willingham, 2009). Teachers in top classes scaffolded language learning and maintained students' energy levels by regularly changing from a chant to a game to a song to a physical activity.

4.5 Exemplary Incidents

To illustrate the principles involved in autonomy support and structure described here, we selected two extracts from top tercile classes to show how many of the ideas work in concert as an organic whole. Transcription conventions are presented in Appendix 2. These class excerpts were deemed to show the largest number of the practices informed by the principles detailed above within a single incident or activity—given that none of the practices function independent of the others, we hope to emphasize their interconnected nature. Though not available in this set of extracts, in both situations the teachers had finished reviewing or presenting the target vocabulary, offering *predictable interactive routines* and *progressive games* to build students' understanding of the language; transcriptions of listen and repeat or elicited response activities did not effectively demonstrate anything that could not be seen in other classes, and so we have not included them. As is culturally standard in Japanese classrooms, activities were primarily teacher directed. In the first classroom extract, teachers worked as a team to scaffold students' first exposure to language, confirmed understanding in Japanese without using Japanese, and allowed students to build comprehension through repetition and demonstration.

Extract 1, Class E, Time 07:33–08:27:

ALT:	Let's put it into Japanese. So what's "food" in Japanese?
HRT:	"Food" in Japanese.
ALT:	"Food" in Japanese please.
HRT:	[Student 1 name].

Student 1: *Donna furuutsu ga suki desu ka?* (What fruit do you like?)
 Student 2: *Chigau.* (Wrong.)
 ALT: Food. ::points to numerous food pictures::
 HRT: In Japanese, food.
 ALT: Food, food. Food. What's food in Japanese?
 Student 1: Food?
 ALT: Yes.
 Student 1: *Tabemono. Donna tabemono ga suki desu ka?* (Food. What food do you like?)
 ALT: OK, that's right. Thank you. Next one, fruits in Japanese. Fruits, fruits. Fruits in Japanese please.
 HRT: [Student 3 name], Stand up.
 ALT: OK. Fruit in Japanese.
 HRT: Fruit.
 Student 3: *Kudamono ne.* (Fruit, right?)

This excerpt shows both the JTE and the HRT teaching together, scaffolding English through the support of Japanese. Both teachers were actively involved; the ALT modeled the language, and the HRT assists in comprehension primarily using English. The additional explanation from the teacher was minimal; the ALT stated and repeated the question, followed by the HRT (*repetition, demonstration, simplification*). Teachers directly addressed the students, and helped them to achieve the correct answer through repetition and signals (*repetition, demonstration, simplification*). When the student made a mistake, teachers also did not correct it, but provided support with gestures and other signals until the student was able to comprehend (*warm/strict, signals for meaning*). Finally, the students' comprehension was checked by the HRT through limited use of their own language (*homeroom teacher involvement*).

In Extract 2, the foreign teacher in Class K demonstrated each action and tied his instruction to a physical representation of the language. The teachers coordinated as a team, and judiciously used students' own language to facilitate the activity.

Extract 2, Class K, Time 18:09–23:27:

- ALT: OK so, take your “Hi Friends [textbook]” and pencil case to the back.
::Demonstrates taking books and pencils to back of the room. Students imitate. JTE writes numbers 1-5 on board::
- ALT: OK. So, next, we’re going to play the? alphabet game.
HRT: Alphabet game.
ALT: OK? So, we are going to form one, two, three, four, five groups.
::Counts off five on fingers, gestures with five spread fingers to students::
Five groups.
::Turns to JTE::
Explain the game in Japanese, right?
- JTE: You explain first?
- ALT: All right. OK so, where is group one ? Group one ?
::Counts students. JTE and HRT help students make groups::
One, two, three, four. Make a circle. Sit down.
::Gestures in a circle::
Group 2.
Make a circle and sit down. Group three, four, and five .
::Gestures to each group::
And sit down, please.
These cards go from A to Z.
::Holds up cards, pauses to show to students::
First, shuffle.
::Shuffles cards::
Shuffle, shuffle.
::Puts down cards and spreads them::
Then arrange them on the floor. Then I want to count.
Arrange from A to Z. A to Z, in the group. A, B, C, D...OK? A to Z.
::Taps alphabet cards on the board. HRT holds up cards::
The first group to finish, stand up. OK? I will the check time.
::Shows stopwatch::
- HRT: *Minna san wakatta?* (Did everyone understand?)
Students: *Wakatta. OK.* (We understand. OK.)
ALT: Three, two, one, start.

::ALT starts stopwatch. Students arrange cards. ALT, JTE, and HRT walk around. After a minute, one group stands up::

Student group: Finish!

JTE: What group is this?

Students: Four.

JTE: OK. Thank you, sit down.

ALT: ::Taps stopwatch:: 58 seconds!

HRT: Fifty-eight!

::Writes “58” next to Group Four on the blackboard::

This section displayed the teamwork that the teachers use, illustrating how the HRT’s involvement in class helped to move activities along (*homeroom teacher involvement*). Though the HRT did not speak much beyond repeating the ALT’s phrases, his active presence was a model for the students. The ALT did his part by demonstrating the expectations of the task using gestures and repetition (*signals for meaning; simplification, demonstration, repetition*). While the JTE had the opportunity to translate, she did not do so before the English explanation, and then did not translate once it was clear that students had sufficiently understood the English explanation (*signals for meaning*). The JTE and HRT were also instrumental in organizing the groups, preparing the blackboard, and watching for when students had finished the timed activities. The activities were timed and had a specific end point (*keep it short*). By dispersing roles, maximizing the teacher resources, and helping students complete tasks efficiently and quickly, these teachers engaged students on emotional, cognitive, and behavioral levels.

5. Discussion

In answer to the research question *What instructional practices in elementary foreign language classrooms support and engage students?*, teachers used a coinciding series of practices to engage and support students. The most effective

classes had interactive routines and involved warm and strict HRTs, JTEs, and ALTs. Teachers worked as a team to facilitate a vigorous pace through shared responsibilities. They used English as a structurally integral part of the lesson, and employed multiple strategies to make that English comprehensible. Teachers in these classes balanced activities using multiple modalities, organized games around a final goal, and provided clear signals for the meaning of the English used in class. They used short activities and provided comprehensible instruction through simplified English. The same teachers in different classes used these different principles in different situations.

Low tercile classes used more mechanical and rote instruction, had controlling or angry teachers, and homeroom teachers were sometimes not even in the same room. Their activities involved little sense of clear progress; they were often collections of teacher-led games. A single activity sometimes took up a large portion of class. Teachers' English was often beyond students' comprehension without considerable translation, and teachers offered few opportunities for students to independently comprehend the language.

In this study, we used the idea that certain universal principles, in this case autonomy support and structure, undergo some surface-level changes for effective localization. Thus, structure and autonomy support organize the results, while the final practices proposed here offer more concrete ideas for what may constitute effective teaching in this and similar settings. Recognizing that the form and quality are simultaneously crucial for engaging teaching (Jang, Reeve, & Deci, 2010), we propose that teachers and students benefit from the combination of the principles detailed in the results.

The results provide a qualitative understanding of these students' continued autonomous motivation (Oga-Baldwin et al., 2017), in spite of previously noted trends toward decreases in motivation with similar samples (Carreira, 2011). These results provide further evidence for the idea that motivation improves as engagement increases (Oga-Baldwin, 2019). More than being positive, more than using games, and more than organization, an interconnected sense of meaningfulness across all of the classroom activities was a powerful force in promoting engagement and learning in these classes. The positive influence of active homeroom teachers further shows the importance of close significant others in foreign language classes. These results confirm the ideas presented by Cameron (2001) and Pinter (2011), which indicate that effective foreign language pedagogy is learning-centered and focuses on improving students' abilities and motivations in unison.

5.1 Limitations

Several caveats should be made about the findings. While the corollary quantitative data (Oga-Baldwin et al., 2017) indicate the reliability of these findings, care must be taken in interpreting these findings beyond similar cultural settings. The use of a mixed-method approach with a large representative sample of Japanese elementary-school children hints at transferability; as with all qualitative analyses, further quantitative exploration of these practices is needed to fully generalize.

Further, none of the practices here should be taken as singularly sufficient for engaging students. It may help to think of the classroom as both organic and mechanical; some of these micro-level features should be considered as parts of a natural ecosystem, integrated with other features, others as mechanical parts which may be swapped in and out as needed. Through careful judgment and integration, teachers may effectively engage students in learning activities. For teachers looking to

emulate these practices, careful consideration should be made of how each feature functions in conjunction with the others.

6. Conclusions

In this paper, we aimed to answer the question of what effective teachers do and say to support engagement during learning activities in a natural classroom setting.

Elementary teachers often struggle with issues of engaging and managing students (Copland, Garton, & Burns, 2014), especially when not thoroughly trained in foreign language pedagogy (Nishino & Watanabe, 2008). However, through effective application the principles presented, teachers in the top tercile classes were best able to promote student engagement, and thereby maintain students' sense of internally regulated motivation (Oga-Baldwin et al., 2017; Oga-Baldwin & Fryer, 2018). In this study, we presented several important theoretically organized practices for promoting engagement and motivation in elementary school foreign language classes. The principles of lesson form and lesson quality offer teachers of YLLs practical options for improving their instructional quality and enhancing students' positive in-class behavior, affect, and cognition. With the expansion of EFL instruction in elementary schools across the world, further localized explorations of effective instructional principles may yet still expand teachers' repertoire of skills for teaching in these contexts.

7. References

Aline, D., & Hosoda, Y. (2006). Team teaching participation patterns of homeroom teachers in English activities classes in Japanese public elementary schools.

JALT Journal, 28(1), 5–22.

Brophy, J. E. (2010). *Motivating students to learn* (3rd ed.) Routledge: New York.

- Butler, Y. G. (2005). Comparative perspectives towards communicative activities among elementary school teachers in South Korea, Japan and Taiwan. *Language Teaching Research*, 9(4), 423–446. doi:10.1191/1362168805lr176oa
- Butler, Y. G. (2015). English language education among young learners in East Asia: A review of current research (2004–2014). *Language Teaching*, 48(03), 303–342. doi:10.1017/S0261444815000105
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge, UK: Cambridge.
- Carreira, J. M. (2011). Relationship between motivation for learning EFL and intrinsic motivation for learning in general among Japanese elementary school students. *System*, 39(1), 90–102. doi:10.1016/j.system.2011.01.009
- Carreira, J. M., Ozaki, K., & Maeda, T. (2013). Motivational model of English learning among elementary school students in Japan. *System*, 41(3), 706–719. doi:10.1016/j.system.2013.07.017
- Cave, P. (2007). *Primary school in Japan: Self, individuality and learning in elementary education*. New York: Routledge.
- Copland, F., Garton, S., & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly*, 48(4), 738–762. doi:10.1002/tesq.148
- Copland, F., & Garton, S. (2014). Key themes and future directions in teaching English to young learners: Introduction to the Special Issue. *ELT Journal*, 68(3), 223–230. doi:10.1093/elt/ccu030

- Cornelius-White, J. D. & Harbaugh, A. (2009). *Learner centered instruction*. Thousand Oaks: Sage.
- Cozby, P., C. & Bates, S., C. (2012). *Methods in behavioral research*. New York: McGraw-Hill Companies.
- Creswell, J., W. (2009). *Research Design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: Sage Publications, Inc.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. NY: Plenum.
- Devellis, R. F. (2012). *Scale Development: Theory and Application*. (3rd ed.) Thousand Oaks: Sage.
- Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge, UK: Cambridge University Press.
- Dörnyei, Z., & Csizer, K. (1998). Ten commandments for motivating language learners: results of an empirical study. *Language Teaching Research*, 2(3), 203–229. doi:10.1177/136216889800200303
- Dörnyei, Z., MacIntyre, P., & Henry, A. (2015). *Motivational dynamics in language learning*. Bristol: Multilingual Matters.
- Fennelly, M., & Luxton, R. (2011). Are they ready? On the verge of compulsory English, elementary school teachers lack confidence. *The Language Teacher*, 35(2), 19–24.
- Fredricks, J. A., Blumenfeld, P., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <http://doi.org/10.3102/00346543074001059>
- Garton, S., Copland, F., & Burns, A. (2011). *Investigating global practices in*

- teaching English to young learners* (British Council ELT Research Paper 11-01). London: British Council.
- Good, T., & Brophy, J. (2008). *Looking in classrooms* (10th Ed.). New York: Longman.
- Guilloteaux, M.-J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 42(1), 55–77.
- Hiver, P., & Al-Hoorie, A. H. (2019). Reexamining the role of vision in second language motivation: A preregistered conceptual replication of You, Dörnyei, and Csizér (2016). *Language Learning*, 1–56.
<http://doi.org/10.1111/lang.12371>
- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy-support or structure but autonomy-support and structure. *Journal of Educational Psychology*, 102(3), 588–600. doi:10.1037/a0019682
- Japan Statistics Bureau (2016). *Japan statistical yearbook 2016*. Available from <http://www.stat.go.jp/english/data/nenkan/index.htm>
- Kang, J. S., & Crandall, J. (2013). *Teaching young learners English*. New York: Heinle Cengage.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1) 159–174.
- Lemov, D. (2015). *Teach like a champion 2.0*. San Francisco, CA: Jossey-Bass.

- McEown, M. S., & Oga-Baldwin, W. L. Q. (2019). Self-determination for all language learners. *System*, 102124. <http://doi.org/10.1016/j.system.2019.102124>
- McEown, M. S., & Takeuchi, O. (2012). Motivational strategies in EFL classrooms: How do teachers impact students' motivation? *Innovation in Language Learning and Teaching*, 8(1), 20–38. doi:10.1080/17501229.2012.741133
- MEXT. (2008). *Shogakkou gakushuu shidou youryou kaisetu: Gaikokugo katsudouhen [Explanatory commentary for the elementary school curriculum guidelines: Foreign language activities]*. Tokyo: Kyouiku Shuppan.
- MEXT. (2012). *Hi, friends! 1*. Tokyo: Kyouiku Shuppan.
- Nakao, K. (2009). Shogakko English learning environments: The ALT role. In A. M. Stoke (Ed.). Tokyo: JALT.
- Nakao, K., Oga-Baldwin, W. L. Q., & Fryer, L. K. (2019). Expanding Japanese elementary school English education: Native and nonnative speaking team-teachers' perspectives on team-teaching quality. *The Bulletin of the Graduate School of Education of Waseda University*, 29, 17–31.
- Nakata, Y. (2009). Intrinsic motivation in the EFL school context: A retrospective study of English learning experiences in Japanese elementary schools. *The Journal of Asia TEFL*, 6(4), 293–291.
- Nakata, Y. (2014). Insider–outsider perspective: revisiting the conceptual framework of research methodology in language teacher education. *International Journal of Research & Method in Education*, 1–18.
doi:10.1080/1743727X.2014.923835
- Naoyama, Y. (2011). *Eigo nooto 1 wo katsuyoushita eigo katsudou no jugyou [English activities lessons using English Note 1]*. Tokyo: Shogakukan.
- Nikolov, M. (1999). “Why do you learn English?” ‘Because the teacher is short.’ A

- study of Hungarian children's foreign language learning motivation. *Language Teaching Research*, 3(1), 33–56.
- Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, 50(1), 57–85.
- Noels, K. A. (2001). Learning Spanish as a second language: Learners' orientations and perceptions of their teachers' communication style. *Language Learning*, 51(1), 107–144.
- Oga-Baldwin, W. L. Q. (2019). Acting, thinking, feeling, making, collaborating: The engagement process in foreign language learning. *System*, 102128.
<http://doi.org/10.1016/j.system.2019.102128>
- Oga-Baldwin, W. L. Q., & Fryer, L. K. (2018). Schools can improve motivational quality: Profile transitions across early foreign language learning experiences. *Motivation and Emotion*, 42(4), 527–545. <http://doi.org/10.1007/s11031-018-9681-7>
- Oga-Baldwin, W. L. Q., & Nakata, Y. (2014a). Optimizing new language use by employing young learners' own language. *ELT Journal*, 68(4), 410–421.
<http://doi.org/10.1093/elt/ccu010>
- Oga-Baldwin, W. L. Q., & Nakata, Y. (2014b). Supplementing the elementary foreign language course of study with a self-determination framework. *International Journal of Curriculum Development and Practice*, 16(1), 13–26.
- Oga-Baldwin, W. L. Q., & Nakata, Y. (2015). Structure also supports autonomy: Measuring and defining autonomy-supportive teaching in Japanese elementary foreign language classes. *Japanese Psychological Research*, 57(3), 167–179.
<http://doi.org/10.1111/jpr.12077>

- Oga-Baldwin, W. L. Q., & Nakata, Y. (2017). Engagement, gender, and motivation: A predictive model for Japanese young language learners. *System*, *65*, 151–163. <http://doi.org/10.1016/j.system.2017.01.011>
- Oga-Baldwin, W. L. Q., Nakata, Y., Parker, P. D., & Ryan, R. M. (2017). Motivating young language learners: A longitudinal model of self-determined motivation in elementary school foreign language classes. *Contemporary Educational Psychology*, *49*, 140–150. <http://doi.org/10.1016/j.cedpsych.2017.01.010>
- Philp, J. & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics*, *36*, 50-72.
- Pinter, A. (2011). *Children learning second languages*. London: Palgrave Macmillan.
- Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149–172). Boston, MA: Springer US.
- Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218. doi:10.1037/0022-0663.98.1.209
- Reeve, J., Cole, S. G., & Olson, B. C. (1986). The Zeigarnik effect and intrinsic motivation: Are they the same? *Motivation and Emotion*, *10*(3), 233–245.
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, *57*, 749–761.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Press.

- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*(2), 975–983.
doi:10.1016/j.cognition.2007.03.005
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, *85*(4), 571–581. doi:10.1037/0022-0663.85.4.571
- StataCorp. (2013). *Stata statistical software: Release 13*. College Station, TX: StataCorp LP.
- Svalberg, A. M.-L. (2009). Engagement with language: Interrogating a construct. *Language Awareness*, *18*(3-4), 242–258. doi:10.1080/09658410903197264
- Tomlinson, B., & Masuhara, H. (2009). Playing to learn: A review of physical games in second language acquisition. *Simulation & Gaming*, *40*(5), 645–668.
doi:10.1177/1046878109339969
- Willingham, D. (2009). *Why don't students like school?* San Francisco, CA: Jossey-Bass.
- Wu, X. (2003). Intrinsic motivation and young language learners: the impact of the classroom environment. *System*, *31*(4), 501–517.
doi:10.1016/j.system.2003.04.001

Appendix 1.

The Supportive Structure Scales (Oga-Baldwin & Nakata, 2015)

Supportive Structure	My teacher gives clear explanations
	The pace of class is appropriate
	My teacher directs me as to what to do
	My teacher uses the foreign language
	My teacher appeals to my interests

The English Language Self-Regulation Questionnaire (SRQ-EL; Oga-Baldwin & Nakata, 2017; McEown & Oga-Baldwin, 2019)

Intrinsic Regulation	English is fun
	I'm interested in English
	English has value

Observed Engagement Instrument

Engagement	⑤	⑤	⑤	⑤	⑤	⑤	⑤
	④	④	④	④	④	④	④
	③	③	③	③	③	③	③
	②	②	②	②	②	②	②
	①	①	①	①	①	①	①
min	—	—	—	—	—	—	—

Sample section of the marksheet scanning optimized instrument for rating observed engagement. 1 = students off-task, listless, and/or distracted; 5 = students focused, interested, and engrossed in tasks.

Appendix 2: Transcription conventions

Italicized text

(Parenthesis)

:: ::

Japanese utterance

English translation

Actions