MEASURING ACADEMICALLY ORIENTED HIGH SCHOOL STUDENTS' MOTIVATION TO LEARN ENGLISH USING ITEM RESPONSE THEORY ANALYSIS: THE JAPANESE EFL CONTEXT

CHIKA TAKAHASHI – *Ehime University &*

SEONGAH IM – *University of Hawai'i at Mānoa*

ABSTRACT

This study was conducted primarily to examine the scale intended to measure academically oriented high school students' motivation to learn English within self-determination theory by utilizing item response theory (IRT), and secondly to capture it with a sample of Japanese high school students. Although high school students have not been the focus in the majority of second language (L2) motivation studies in Japan, they constitute an area worthy of inquiry, given their short-term goal of passing entrance exams and the long-term goal of nurturing L2 communicative competence. In order to examine the scale measuring academically oriented high school students' L2 motivation and to capture its structure, a questionnaire was administered to 273 high school students. In analyzing the data, the graded response model, an IRT model for polytomous responses, was applied. Analyses indicated that identified regulation, a type of internalized extrinsic motivation, was relevant and seemed to play an important role for the participants. The psychometric quality of the scale items were presented using functions of the IRT model. Implications and suggestions were discussed.

INTRODUCTION

Motivation to learn a second/foreign language (hereafter called L2) is one of the so-called individual difference variables that have been widely researched. L2 motivation is considered to entail the *direction* and *magnitude* (Dörnyei & Ushioda, 2011, p. 4, original emphasis) of L2

learning, i.e., why learners learn an L2 and how intensely they do so. These factors have shown relationships in L2 learning to such important variables as L2 achievement (Gardner, 1985; Shaikholeslami & Khayyer, 2006) and L2 self-confidence (e.g., Pae, 2008; Yashima, 2002). Various models have been proposed to date, such as the socio-educational model (Gardner, 1985, 2001) and the L2 motivational self system (Dörnyei, 2005, 2009).

Researchers have examined various aspects of L2 motivation regarding Japanese learners of English. Japan is a typical English as a foreign language (EFL) context where communicative competence in English is strongly emphasized by the government despite the lack of daily opportunities to communicate in English. In examining L2 motivation, many researchers have utilized Likert-scale questionnaires, and the ways in which they validate and improve them have included calculating Cronbach alpha coefficients for reliability, calculating inter-correlations among different sub-constructs within L2 motivation to examine convergent and discriminant validity, and performing exploratory and confirmatory factor analyses to verify latent constructs.

One area that still needs more discussion in L2 motivation studies is L2 motivation among high school students (see, for example, Apple, Da Silva, & Fellner [2013] for a collection of recent L2 motivation studies in Japan; none of the chapters focus exclusively on high school students). High school students may constitute a population worthy of more investigation considering the complex situations surrounding their English studies: On the one hand, they may feel the necessity to nurture their communicative competence in English now that the role of English as the primary international language has been emphasized in the era of globalization; on the other hand, high school students, particularly academically oriented ones, need to go through a series of university entrance examinations and study English quite intensely for them. Under such circumstances, high school students might be driven by various types of orientations, each exerting distinct motivational powers.

This study is an attempt to examine a scale intended to measure academically oriented high school students' motivation to learn English in an EFL context and capture it with a sample

of Japanese high school students. In doing so, self-determination theory (SDT) (Deci & Ryan, 1985, 2002), a motivational theory originally developed in psychology that might be particularly insightful for analyzing high school students' L2 motivation, was applied. The data were analyzed within an item response theory (IRT) framework, which has been utilized widely in fields such as education and psychology, but not in L2 motivation research. As the IRT approach enables in-depth analyses of items and an entire scale composed of them, the study adopted an IRT model to evaluate psychometric characteristics of the items and examined L2 motivation among high-achieving Japanese high school students.

Self-Determination Theory

Among the many motivational models proposed to date, SDT (Deci & Ryan, 1985, 2002) is one of the most widely applied. In SDT, it is postulated that all human beings have an innate tendency toward growth and integration. This nature interacts with social contexts that either "nurture or impede the organism's active nature" (Deci & Ryan, 2002, p. 6). Human beings are considered to have three fundamental psychological needs: autonomy, competence, and relatedness. Social environments can be categorized ranging from supportive (insofar as they satisfy these three needs) to antagonistic (insofar as they conflict with these needs).

First, to further explain the three fundamental psychological needs, autonomy refers to the perceived origin or source of one's own behavior. Second, competence is concerned with "feeling effective in one's ongoing interactions with the social environment and experiencing opportunities to exercise and express one's capacities" (Deci & Ryan, 2002, p. 7). Third, relatedness refers to feeling connected to others.

Regarding motivation, there is a distinction between intrinsic and extrinsic motivation in addition to amotivation, which is "a sort of antithesis to motivation" (Noels, 2009, p. 297). Deci and Ryan (2002) suggest that humans become intrinsically motivated when the three fundamental psychological needs are supported. When learners are intrinsically motivated, they engage in an activity (e.g., L2 learning) purely for their own enjoyment. In contrast, learners

may have more extrinsic types of motivation, i.e., external regulation, introjected regulation, or identified regulation, and these are the types of motives that are external to the enjoyment of the activity itself. Within L2 learning, external regulation happens when someone studies an L2 to avoid punishment or to seek rewards. For example, a student may study an L2 only because of pressure from his/her parents. Introjected regulation is more internalized than external regulation. Students high on introjected regulation might study an L2 in order to avoid guilt or shame. In such a case, the pressure comes from within. However, learners are still reacting to pressure and therefore are not self-determined. In contrast, at the stage of identified regulation a student may study an L2 because of personal importance, such as his/her personal development. In addition, amotivation represents the state of lacking the intention to act. In this framework, these types of motivation lie along a continuum from the least internalized (amotivation) to the most internalized (intrinsic motivation). In sum, SDT provides a useful frame of reference as described above, particularly for learners in classroom situations.

Past studies have shown important relationships between types of L2 motivation within SDT and among L2 motivation and other variables in L2 learning. For example, Pae (2008) performed structural equation modeling with a sample of Korean university students learning English. The study demonstrated that of the different types of L2 motivation, only intrinsic motivation showed an acceptable fit to the data in the model. Intrinsic motivation led to the magnitude of motivation (motivational intensity, desire to learn English, and attitudes toward learning English) and L2 self-confidence. These two variables then led to L2 achievement. Second, Nishida (2013) examined the relationships between different types of motivation within SDT on the one hand and other L2-related variables (L2 ideal self, L2 ought-to self, international posture, can-do, and willingness to communicate) on the other hand with a sample of Japanese university students. Intrinsic motivation and identified regulation had positive and significant correlations, and intrinsic motivation had higher correlations than identified regulation with these L2-related variables, except L2 ought-to self. Third, Vallerand and Bissonnette (1992) investigated the differences between persistent and non-persistent learners

in a French course at a junior college in Canada. In this study, the researchers found that not only intrinsic motivation but also identified regulation was significantly different between the two groups. Noels (2009) elaborated upon the similarities and differences between intrinsic motivation and identified regulation. They are both internalized types of motivation, which consist of "high levels of involvement, positive affect, and internal sense of control" (Noels, 2009, p. 308). They are different, however, in that personal importance is the motivating force behind identified regulation, whereas for intrinsic motivation, attraction or interest in the activity is the motivating force (Noels, 2009, p. 308). Although there have been some competing results in terms of the importance of identified regulation, the importance of intrinsic motivation in terms of its close relationships to important variables in L2 studies can be pointed out from past studies.

Prevalence of English and Learning English in Japan

Japan is a typical EFL context, without many opportunities to communicate in English on a daily basis. As in other EFL contexts, however, with the globalization of the world today, English is considered the primary international language. One of the main areas where English is needed is in business, as businesspeople communicate with both native and non-native speakers in English: Some of the major Japanese companies have even decided that their official language is no longer Japanese, but English. The government has also emphasized the importance of improving the communicative competence of Japanese learners of English.

Japanese learners of English, particularly high-achieving high school students, need to study hard, not only for university admissions, but also for communication purposes in the future. That is, in order to get into a prestigious university, students need to go through a series of intense entrance examinations and therefore study English as well as other academic subjects; thus, passing these examinations is a strong reason for studying English. However, some learners may also have a long-term goal of obtaining the ability to communicate in English with people around the world. Some learners, especially after passing the university entrance

examinations, equate English learning with improving practical communication skills in English, which results in "extraordinary interest in voluntary English language education at the adult level" (Berwick & Ross, 1989, p. 207). This may mean that they are driven from dual orientations, one related to a short-term goal of passing university entrance examinations and the other related to a long-term goal of gaining "practical" communicative competence (Yashima, Zenuk-Nishide, & Shimizu, 2004).

From a motivational perspective this might mean that Japanese high school students learning English are motivated for various orientations within the SDT framework. Some may intrinsically enjoy learning English while others study it only because of external regulation, e.g., just to pass university entrance examinations. Still others may understand the personal importance of learning English for their future careers and thus have high identified regulation. These variations might be keener for students that are academically oriented and seek admission to prestigious universities. Thus, the SDT framework may offer valuable insights into the motivational structure among high-achieving high school students in Japan. These students might represent a population worthy of investigation for the wider readership outside Japan as well for they are a case of L2 learners with "dual orientations" in a typical EFL context.

Item Response Theory and the Graded Response Model

IRT contains a large family of models where the probability of endorsing an item is expressed as a mathematical function of person and item parameters. The simplest model is the Rasch (1960) model, also known as the one-parameter logistic (1PL) model, where the item characteristic is expressed in terms of item difficulty or item location. In the two-parameter logistic (2PL) model for dichotomous responses, another item characteristic is considered. The following equation of the 2PL model represents a probability function with item difficulty and item discrimination parameters.

$$p_{ij}(\theta_j \mid x_{ij} = 1) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

The θ_j indicates the person *j*'s trait level. The trait is set on a standard scale with a mean of 0 and a standard deviation of 1. The p_{ij} is the person *j*'s probability of endorsing an item *i* correctly/positively given the person *j*'s trait level. The a_i and b_i are the discrimination parameter and the threshold (or difficulty) parameter of the item *i,* respectively.

The graded response model (or GRM, Samejima, 1969) is an extension of the 2PL model that was developed for ordered polytomous responses such as those from Likert-scale items. The GRM explores the relationship between the trait scale and the series of probabilities of being assigned to each category *k* at a given trait level (e.g., L2 motivation level in this study). For each Likert-scale item, category response functions (CRFs) can be calculated as follows, for example: The probability of selecting category 1 is the probability of selecting 2 or higher (2+) categories subtracted from the probability of selecting 1 or higher categories, ~~$P_1 = P^*(X \geq 1) - P^*(X \geq 2)$~~. Similarly, the probability of selecting category 2 is the probability of selecting 3 or higher (3+) categories subtracted from the probability of selecting 2 or higher categories, ~~$P_2 = P^*(X \geq 2) - P^*(X \geq 3)$~~.

As the GRM offers functions corresponding to each category, it is especially useful for (a) exploring the relationship between the trait scale and a series of probabilities of being assigned to category options across trait θ level and (b) examining how many categories are optimal to encompass the entire θ range with appropriate orders and distances among the category options. When constructing a Likert scale, researchers often have a rather arbitrary number of answer options. This may seem like a trivial matter, but whether or not these categories indeed appropriately distinguish among respondents' traits is worthy of further investigation. If a certain category shares a substantially similar range of the latent trait with an adjacent category or is even nested within the range of the adjacent category, a researcher might have to consider modification of the number of category options to most appropriately distinguish among each category in the questionnaire. Careful consideration should be given to the number of response categories of the whole questionnaire, should such results be gained.

The Present Study Using the GRM to Improve an SDT Scale

Although high school students have not been the focus of many studies examining L2 motivation in Japan, it is important to understand what drives high school EFL students to study English in the era of globalization, even without daily opportunities to communicate in English outside the classroom. Furthermore, Likert-scale questionnaires measuring L2 motivation within SDT have been validated mainly by performing factor analysis (e.g., Hiromori, 2003; Noels, Pelletier, Clément, & Vallerand, 2000), but not from the perspective of psychometric evaluation of individual items. To validate and improve the SDT scale with high school students in Japan, the present study utilizes GRM analysis for micro-level psychometric examination of individual items and categories as well as both exploratory and confirmatory factor analytic approaches for macro-level examination of factor structure of the scale. Specifically, we seek to answer the following research questions:

1. Do the five constructs within SDT (intrinsic motivation, identified regulation, introjected regulation, external regulation, and amotivation) emerge as distinct factors as expected from the theory?

2. Which questionnaire items do the GRM analyses suggest are psychometrically sound?

3. Which types of motivation within SDT are relevant for the sample?

METHOD

Participants

The participants for this study were 273 Japanese learners of English at an academically oriented coeducational high school in Tokyo. After graduating, many students go to top universities, such as the University of Tokyo. This means that they study hard for the university entrance examinations, as explained above. In this sense, the participants were generally likely to be motivated to learn English, in particular in terms of passing university entrance examinations (i.e., scoring high on external regulation) and understanding the personal

importance of learning English for their future careers (i.e., scoring high on identified regulation). The data come from a study (Takahashi, 2013) which investigated the structure of L2 motivation and its relationship to persistence in L2 learning. The present study focuses on the pilot study, which aimed at improving a questionnaire to be used in the main study.

In order to control any confounding factors and gain a clear picture of the data, several steps were necessary. First, when asked if they used English for communication with their family, four answered yes and were eliminated, as this was not considered to be typically Japanese. Second, one participant turned out to not be Japanese and was also eliminated. Among the 268 respondents left, some did not fully complete the questionnaire. Thus, the 26 respondents who did not answer more than eight of the 31 questions (approximately 30% of the items) were eliminated.

Instrument

The questionnaire items on L2 motivation were adapted from studies within the SDT framework conducted on both high school and college students (Hiromori, 2003, 2005; Noels, Pelletier, Clément, & Vallerand, 2000; Sakai & Koike, 2008; Tanaka & Hiromori, 2007). Each item stated a possible reason for learning English, such as for exams, to avoid feeling ashamed, for the enjoyment of learning English, etc., each targeting one of the types of L2 motivation within SDT. Participants were to rate the extent to which they agreed with the statement. Each of the aforementioned types of motivation in SDT was measured with five to seven items in a six-point Likert scale, with 1 being "completely disagree" and 6 being "completely agree."

Procedures

The questionnaire was given to English teachers at the high school, who signed a consent to participate and agreed to cooperate and distribute it to students. Because the whole questionnaire was relatively long, taking about 40–50 minutes to complete, students were told to take it and fill it out at home. They were informed of their anonymity as well as the voluntary

nature of the study. The questionnaire was conducted in Japanese, the respondents' first language. The completed questionnaires were collected and subsequently analyzed.

Data Analysis

The data analysis for psychometric examination of the items and scale were composed of two phases, exploratory and confirmatory. The first exploratory phase included reliability analysis, exploratory factor analysis (EFA), and the IRT analysis using GRM. The confirmatory phase consisted of confirmatory factor analysis (CFA). The reliability and EFA at the exploratory stage were performed using SPSS version 23. Then, a principal component analysis (PCA) was conducted using SPSS version 23 in order to identify the number of construct dimensions. After confirming the constructs, an EFA was performed in order to examine the specific relationships between the items and factors and residuals. Items that were confirmed were then analyzed within the GRM framework. The IRT analysis using GRM was performed with MULTILOG version 7.0.3 (Thissen, Chen, & Bock, 2003). A set of effective good items was selected after deleting items that showed poor psychometric properties at the exploratory stage. Thereafter, the structure of sub-constructs of motivation was confirmed using a confirmatory factor analysis (CFA) with AMOS version 23.

RESULTS

The univariate and multivariate outliers were detected by checking *z*-scores (Tabachnick & Fidell, 2007). After deleting 29 cases, either as univariate or multivariate outliers, 213 cases remained in the new data set. Table 1 presents the descriptive statistics of each subscale. As seen, identified regulation was somewhat negatively skewed and amotivation was slightly positively skewed. This may be because high-achieving high school students who are preparing for university entrance exams are generally motivated to learn English.

Table 1

Descriptive Statistics of Each Subscale (n = 213)

Scale	k	Min.	Max.	*M*	*SD*	SKW	*SES*	KUR	*SEK*	*Alpha*
IM	7	1.00	6.00	3.53	.97	-.02	.17	-.10	.33	.91
IDR	6	1.00	6.00	4.26	.95	-.46	.17	.24	.33	.87
ITJ	6	1.00	6.00	3.39	.78	-.03	.17	.35	.33	.77
EXT	7	1.14	5.29	3.30	.73	-.11	.17	.31	.33	.70
AMT	5	1.00	4.40	2.26	.78	.34	.17	-.52	.33	.83

Note: IM = Intrinsic Motivation; IDR = Identified Regulation; ITJ = Introjected Regulation; EXT = External Regulation; AMT = Amotivation.

Exploratory Factor Analysis

PCA was performed to examine the nature and number of factors, and how items measured intended sub-constructs. The data were factorable, as evidenced by the Kaiser-Meyer-Olkin (KMO) measure of sample adequacy being .89. This means that the items showed acceptable correlations with one another and that the data were suitable for being factor analyzed. Having established strong support for KMO, the number of factors was determined. Judging from the eigenvalues and the scree plot, a four-factor solution was selected. Then, a four-factor EFA was performed in order to closely examine the items. Table 2 presents the factor matrix with varimax rotation, which includes the loadings on each factor, communalities, and proportion of variance explained in each factor.

Table 2

Rotated Factor Matrix with Varimax Rotation of SDT Items

	IM	EXT	AMT	IDR	h^2
Q5 IM	**.74**	-.02	-.29	.10	.64
Q18 IM	**.77**	.11	-.08	.13	.64
Q20 IM	**.79**	.12	-.07	.17	.68
Q21 IM	**.81**	.13	-.16	.20	.74
Q22 IM	**.66**	.15	-.12	.31	.57
Q48 IM	**.71**	.23	-.12	.20	.61
Q51 IM	**.64**	.12	-.16	.27	.53
Q7 EXT	-.19	**.43**	.11	-.07	.24
Q15 EXT	.22	**.51**	.20	.14	.37
Q23 EXT	.26	**.69**	-.04	.24	.60
Q29 EXT	.16	**.48**	.21	.02	.30
Q33 EXT	.07	**.80**	.09	.23	.71
Q47 EXT	**-.54**	.01	.36	.09	.42
Q49 EXT	.06	**.64**	.01	.22	.46
Q6 ITJ	**.52**	**.40**	.05	.10	.44
Q8 ITJ	-.07	**.61**	-.28	.29	.54
Q12 ITJ	.14	.22	.10	**.44**	.27
Q32 ITJ	.22	**.67**	.20	.04	.54
Q36 ITJ	**.42**	**.55**	.07	.25	.55
Q50 ITJ	**.41**	.38	-.09	.26	.39
Q3 AMT	-.09	.01	**.68**	-.31	.57
Q9 AMT	-.11	.11	**.75**	-.24	.65
Q13 AMT	-.12	.17	**.77**	-.22	.68
Q28 AMT	-.29	.15	**.53**	.08	.39
Q46 AMT	-.24	.15	**.72**	-.07	.60
Q17 IDR	.37	**.45**	-.03	.31	.44
Q38 IDR	**.45**	.12	-.22	**.57**	.58
Q40 IDR	.35	.22	-.32	**.59**	.62
Q42 IDR	.25	.16	**-.42**	**.61**	.64
Q52 IDR	.16	.23	-.25	**.73**	.67

Q53 IDR	.17	.24	-.20	**.71**	.63
Proportion of Variance	.18	.14	.11	.11	

Note: Factor loadings > .40 are in boldface. IM = Intrinsic Motivation; EXT = External Regulation; AMT = Amotivation; IDR = Identified Regulation; ITJ: Introjected Regulation.

Factor 1 received heavy loadings from items measuring intrinsic motivation. Factor 2 received heavy loadings from most external regulation items and some introjected regulation items. Factor 3 received heavy loadings from items measuring amotivation, and Factor 4 received heavy loadings from identified regulation items. Overall, the items measuring intrinsic motivation, identified regulation, external regulation, and amotivation performed in accordance with the theory. With regard to introjected regulation we decided not to analyze the items within the GRM framework due to the poor results of the EFA. Implications of this result on introjected regulation are explored in the discussion section. Furthermore, Q47 on external regulation, which did not load on the intended factor (Factor 2), was deleted before further analyses.

Graded Response Model Analysis

Items measuring the subtypes of motivation within SDT, except introjected regulation, were analyzed in four separate GRM analyses. Four PCAs for each subscale were performed to see the percent of variance explained by the first dimension and how dominant the first dimension is, which ensures the unidimensionality assumption (Embretson & Reise, 2000). The first eigenvalues of each scale were substantially large and explained the total variance of each scale: 61.70% for amotivation; 62.42% for identified regulation; and 66.09% for intrinsic motivation. As for external regulation, the total variance explained by the first dimension was only 41.80%, followed by 24.26% by the second dimension. Thus, after deleting Q7, which loaded heavily on the second component, another PCA was performed. The first eigenvalue was substantially large and explained 56.11% of total variance. This means that for external

regulation at this stage, five of the seven original items without Q7 and Q47 remained for further analyses. In each scale, the remaining eigenvalues were substantially smaller than the first value (Reckase, 1979). Another assumption of IRT, local independence among items, was also examined using a local dependence (LD) χ2 test (Chen & Thissen, 1997) for each scale. The range of LD χ2 values were -0.8 to 3.6 for amotivation; -0.6 to 2.8 for identified regulation; -0.7 to 3.5 for intrinsic motivation; and -0.7 to 4.5 for external regulation. All of LD χ2 values were substantially smaller than 10, indicating all items in each scale were locally dependent (Cai, Thissen, & du Toit, 2011).

Next, the missing values were imputed with the multiple stochastic regression method for multiple imputations because there were some missing values due to incomplete answers (1.12% of all the answers). The logic behind multiple imputations is that missing values are imputed by drawing at random from their conditional distribution given the observed data and assuming that the parameter values are equal to their maximum likelihood estimates (Arbuckle, 2005).

Two models with fixed and free *a*-parameters were compared to examine which fit better with the given data. For amotivation, the difference between negative twice the log-likelihood values of free and fixed *a*-parameter models was Δ-2LL = 142.6, *df* = 5, *p* < .001, suggesting the model with freely estimated *a*-parameters fit the data significantly better than the more restricted model with the fixed *a*-parameter. We also obtained the same conclusion with the other three scales: For external regulation, Δ-2LL = 130.7, *df* = 5, *p* < .001; for identified regulation, Δ-2LL = 218.1, *df* = 6, *p* < .001; and for intrinsic motivation, Δ-2LL = 320.1, *df* = 7, *p* < .001. Hence the GRM with free *a*-parameters better fits the data for each subscale. Table 3 presents the parameter estimates of all the items using the GRM. It should be noted that for the first three items measuring amotivation and the third item measuring external regulation no participant answered with category 6, or "completely agree." Thus, only four location (*b*) parameters are estimated.

Table 3

Item Parameter Estimates of Subscales Using the GRM

Item Number	a	b_1	b_2	b_3	b_4	b_5
Amotivation						
3	2.45	-.53	.79	1.94	3.18	---
9	2.91	-.20	.92	2.31	2.98	---
13	4.34	-.38	.69	1.69	2.30	---
28	1.15	-2.28	-.51	.53	2.20	3.55
46	1.94	-.96	.00	1.37	2.79	2.95
External Regulation						
15	1.42	-1.35	-.04	1.00	2.59	3.68
23	3.68	-1.77	-1.09	-.27	.76	1.66
29	.93	-1.40	.44	2.04	4.61	---
33	3.90	-1.34	-.54	.15	1.17	2.34
49	1.26	-2.84	-1.56	-.38	1.29	2.72
Identified Regulation						
17	1.19	-2.77	-1.37	-.45	1.49	2.92
38	2.53	-2.55	-1.70	-1.07	.02	1.32
40	2.40	-2.20	-1.81	-1.09	-.11	.91
42	2.96	-3.18	-1.94	-1.37	-.06	.77
52	3.23	-2.36	-1.63	-.99	.10	1.02
53	2.91	-2.12	-1.53	-.74	.19	.93
Intrinsic Motivation						
5	2.08	-2.20	-1.16	.04	.97	1.69
18	2.45	-1.88	-.93	.03	1.07	2.09
20	4.86	-1.79	-.99	-.12	.86	1.52
21	5.76	-2.09	-.94	-.33	.70	1.57
22	1.96	-2.15	-.76	.42	1.37	2.36
48	1.89	-2.05	-.77	.32	1.40	2.16
51	2.23	-2.21	-1.44	-.53	.73	1.79

A subsequent analysis for all items was run using the GRM with free discrimination parameters. Overall, the items had good psychometric properties with high discrimination parameters. However, for some items the CRFs did not have equal distances between one and the next, and some categories were nested within the range of the adjacent categories. For example, as shown in Figure 1 for Q28: AMT, category 3 was not likely to be chosen by participants at any theta levels, just as category 3 in Q17: IDR and category 2 in Q40: IDR. As for Q29: EXT, no participants at any theta levels were likely to answer with categories 5 or 6. Subsequently the amount of information offered by these items is low, as evidenced by their item information curves (see Figure 2). These are in contrast to another example item presented in Figure 1 (Q21: IM), for which the CRFs have about equal distances between one CRF and the next, and well-ordered parameter distances associated with the entire latent continuum of intrinsic motivation with probabilities of endorsing each category option of the Likert scale. This clarifies that participants with different levels of intrinsic motivation were likely to respond to this item by choosing different answers in a Likert scale, which is how researchers hope every questionnaire item performs.

Figure 1. Categorical Response Functions. AMT = Amotivation; EXT = External Regulation; IDR = Identified Regulation; IM = Intrinsic Motivation; c1 = category 1; c2 = category 2; c3 = category 3; c4 = category 4; c5 = category 5; c6 = category 6.
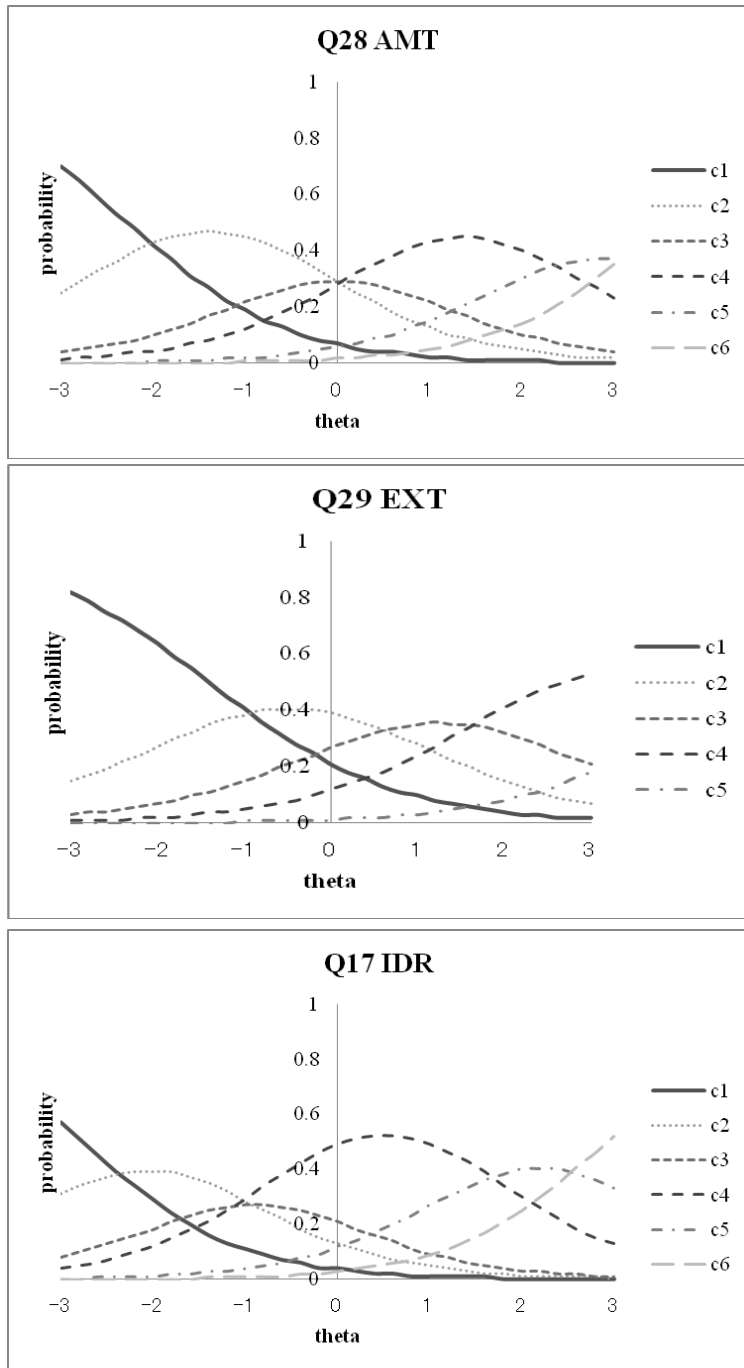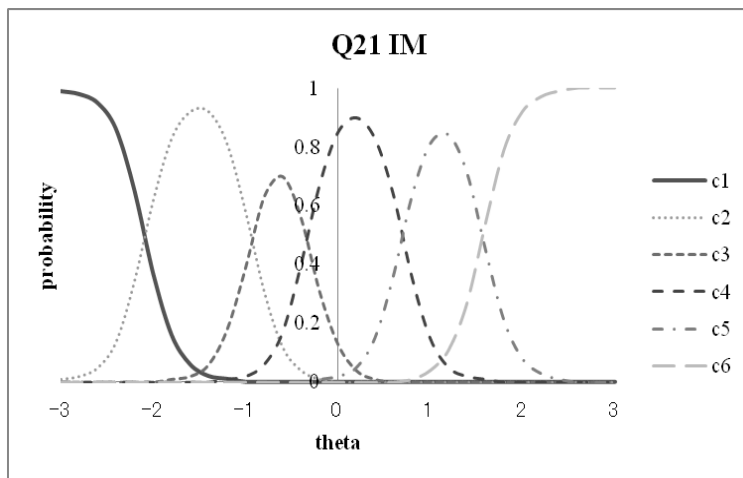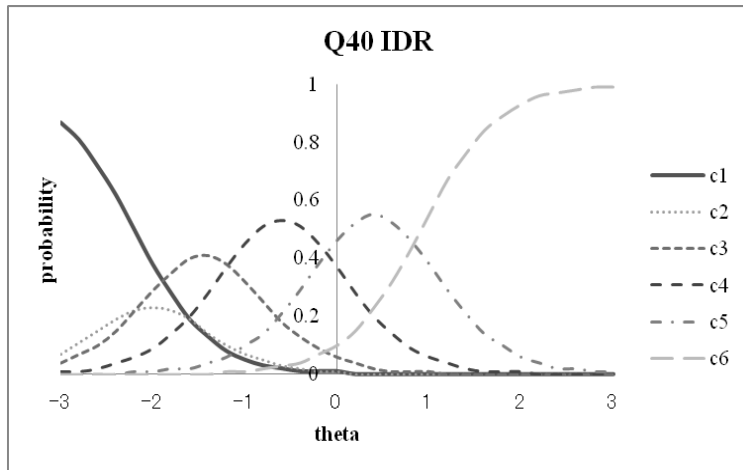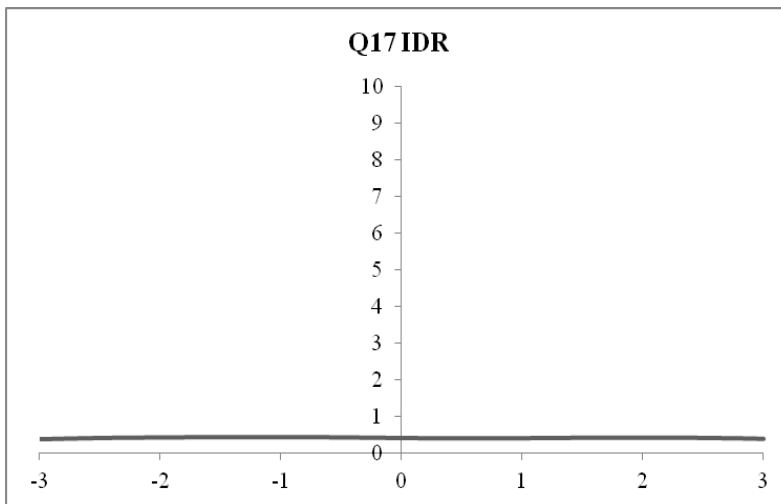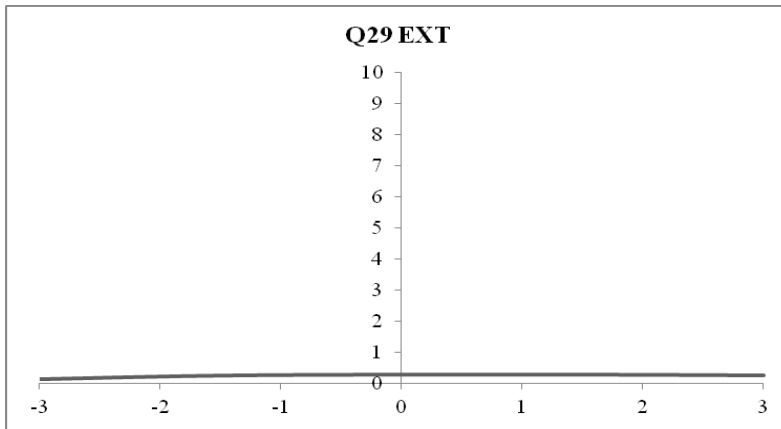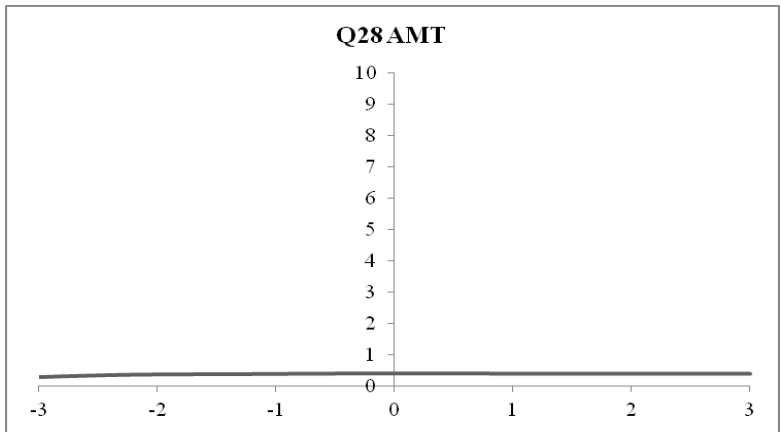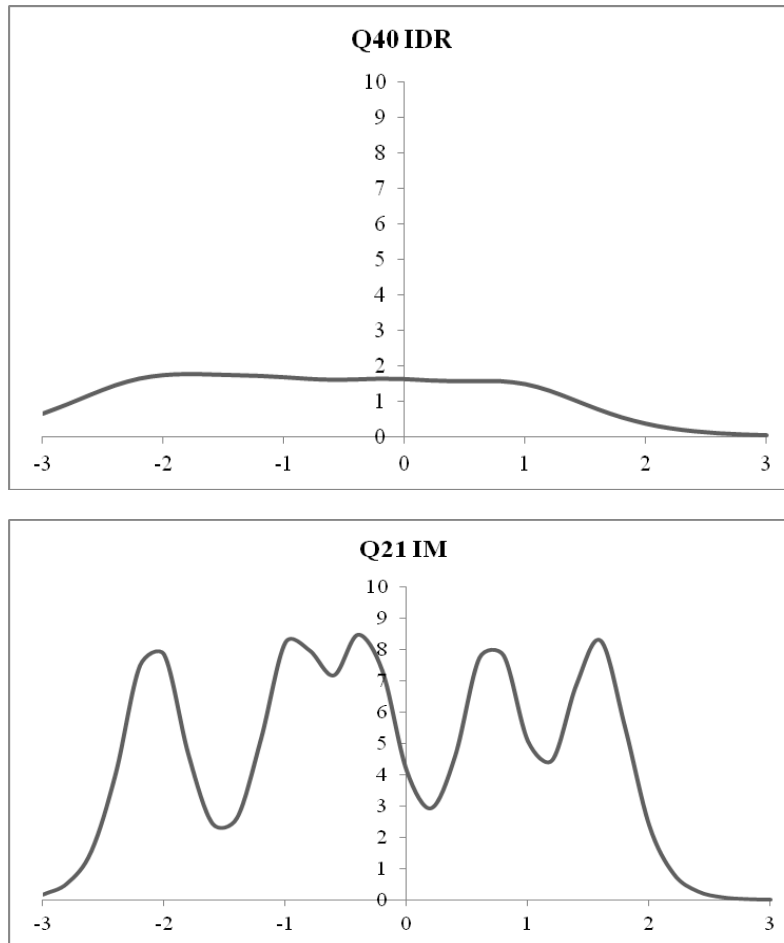
Figure 2. Item Information Curves. AMT = Amotivation; EXT = External Regulation; IDR = Identified Regulation; IM = Intrinsic Motivation.

Q28 AMT

Q29 EXT

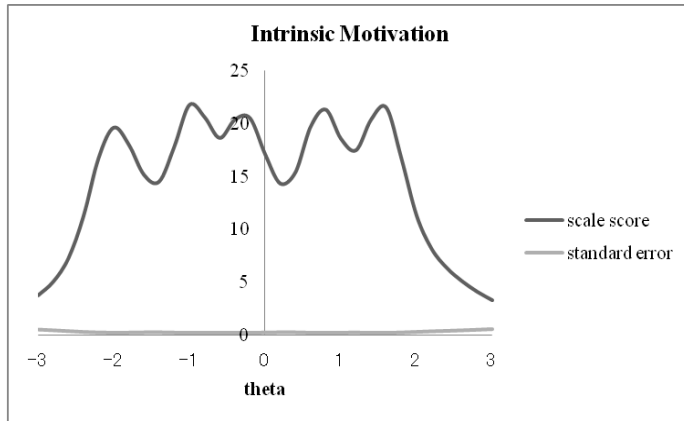**Q17 IDR**

Q40 IDR

Q21 IM

Second, with respect to the amount of information the items offered in each subscale the trend seemed to be different depending on subscales. As presented in Figure 3, whereas for external regulation and intrinsic motivation the amount of information was approximately equal across all theta levels, this was not the case for amotivation or identified regulation. For amotivation, the items offered little information with regard to participants with low theta levels, and for identified regulation the items offered little information with regard to participants with high theta levels. It should be noted that because the number of items varies depending on subscales, the test information varies as well.

Figure 3. Test Information Curves.
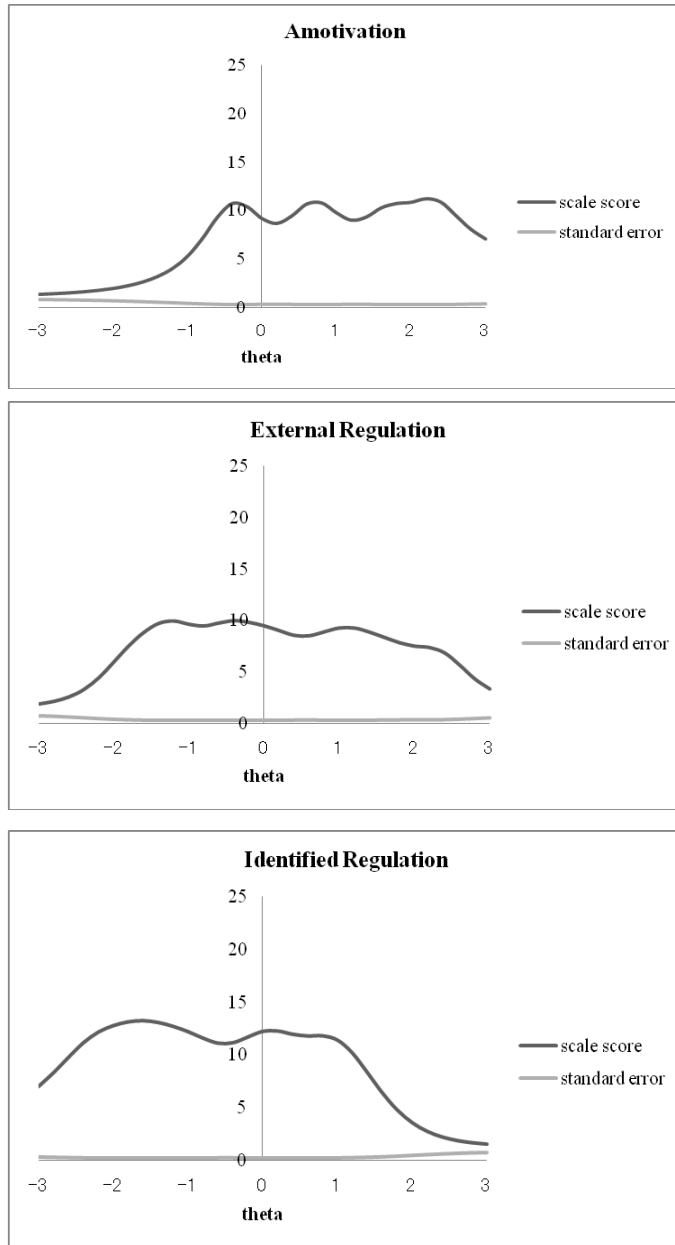
Confirmatory Factor Analysis and Inter-Correlations

In order to further confirm the structure of motivation, a CFA was performed. The fit of the model with the four subscales after deleting items with poor psychometric properties (Q17: IDR; Q40: IDR; Q29: EXT; and Q28: AMT) was acceptable with CFI = .91 and RMSEA = .08 (see Appendix for the list of questionnaire items included in the CFA). The inter-correlations among the four subscales are presented in Table 4.

Table 4

Inter-Correlations Among the Four Subscales

	Intrinsic Motivation	Identified Regulation	External Regulation	Amotivation
Intrinsic Motivation	1.00	.62	.43	-.41
Identified Regulation		1.00	.57	-.60
External Regulation			1.00	-.02
Amotivation				1.00

Overall the correlations were in accordance with the theoretical framework of SDT: The correlations between adjacent subscales were higher than those farther apart, and the correlations between self-determined and non-self-determined types of motivation were negative. Two things should be noted: the correlations between external regulation and intrinsic

motivation, and between external regulation and identified regulation, were higher than expected; and the magnitude of the correlation between identified regulation and amotivation was higher than between intrinsic motivation and amotivation. Implications of these results will be further discussed in the next section.

DISCUSSION AND CONCLUSION

The purpose of this study was to examine the scale intended to measure academically oriented high school students' motivation to learn English in an EFL context and capture it with a sample of Japanese high school students. The theoretical framework of the questionnaire adopted in this study was based on SDT, which has been used and examined in diverse fields such as education, psychology, and health care.

As a preliminary step before the GRM analysis, we performed a PCA and checked whether the five subscales within SDT would emerge as distinct factors as expected from SDT. Of the five theorized constructs (intrinsic motivation, identified regulation, introjected regulation, external regulation, and amotivation), the results showed that the four subscales for intrinsic motivation, identified regulation, external regulation, and amotivation were distinctive. The items intended to measure introjected regulation showed cross loadings and were difficult to interpret. It should be noted that in previous studies with samples of Japanese high school students learning English and Canadian university students learning French (Hiromori, 2003; Noels, Pelletier, Clément, & Vallerand, 2000), the subscale of introjected regulation in particular had low Cronbach alpha coefficients as well (e.g., Cronbach's alpha = .47 in Hiromori, 2003 and .67 in Noels, Pelletier, Clément, & Vallerand, 2000). Therefore, the conceptual framework for the subscale needs to be further examined so that each item represents the type of regulation that comes from within but is not internalized. Then it should be examined with other L2 samples learning various languages, in particular those learning an L2 and planning to take high-stakes exams.

The comparison of -2LL values showed that the models with free a-parameters fit the data better for the four constructs. The CRFs of each item informed us whether the threshold parameters of categories of each item overlay evenly and entirely across the θ levels, which suggested the GRM analyses were valuable in helping to envision what category options should be modified or eliminated. Overall, the items analyzed within the GRM framework fit the model well and seemed to do a good job of discriminating participants with distinct theta levels. There are exceptions, however, including Q28: AMT, which did not show good psychometric properties. This item is phrased as, "I feel that even if I study English, I will not accomplish much." Compared to other items on amotivation, which targeted the lack of understanding the purpose of studying English, this item might have meant something different to the participants. That is, Japanese high school students study English diligently, particularly in order to pass entrance examinations, but in a typical EFL context it might be difficult to feel that they have achieved high competence, particularly communicative competence, in English. In that sense, even when they still understood the purpose and the importance of studying English (i.e., low endorsement to other items on amotivation), the participants might have endorsed Q28: AMT because the item represents their honest feelings.

With respect to Q17: IDR and Q40: IDR, which also showed poor psychometric properties, these items might have been unclear as to the degree of internalization. Q17: IDR is phrased as, "I study English because I want to be able to speak at least one foreign language," and Q40: IDR is phrased as, "I study English because I want to be someone who can use English." These might not have represented the regulation as internalized as, "I study English because I think it is good for my personal development," for example, because being able to speak English may represent distinct purposes for different L2 learners, such as for one's personal growth, to feel superior, etc.

In further analyzing the items intended to measure identified regulation, they had the highest mean among the five subscales (see Table 1). Japanese high school students who go to top universities may think that they will need English communication ability in the future and

thus it is personally important to them. Identified regulation, in this sense, might have been the type of motivation within SDT that was particularly relevant for the participants. Thus, with this type of student, it may very well be that "regulatory guides are learned in the case of identified" regulation (Noels, 2009, p. 308) by learning the role of English given the situations in Japan. This might also explain the higher magnitude of correlation between identified regulation and amotivation than between intrinsic motivation and amotivation; identified regulation is situated on the opposite side of amotivation, which represents the lack of intention to act. This trend coincides with the fact that the questionnaire items were not good at discriminating participants with low levels of amotivation and high levels of identified regulation because the overall level of amotivation was low and that of identified regulation was high among the participants (see Figure 3). Although the role of identified regulation was not clear in past studies (e.g., Pae, 2008), if identified regulation is the type of L2 motivation particularly relevant to this type of high school student, the relationship between identified regulation and L2-related variables such as L2 achievement should be further investigated after the improvement of the questionnaire. Furthermore, if we were to examine students with high identified regulation in more detail, it would be necessary to carefully design the tone of the statement and create items that are more difficult to endorse for this type of population and better at discriminating participants with high levels of identified regulation.

Lastly, another item with poor psychometric properties was Q29: EXT. This item is phrased as, "I study English so that I can please my parents and teachers." Perhaps for these high-achieving students, regulation by their parents and teachers is distinct from regulation by examinations and future career, which other items on external regulation targeted. This might explain why the correlations were higher than expected between external regulation on the one hand and intrinsic motivation and identified regulation on the other hand after deleting Q29: EXT. For academically oriented students, exams and career might be something more internalized than mere rewards or punishment, and the regulation might not have been as external as the theory postulates. In order to tap into what exams and career mean to the

participants, it will be necessary in future research to employ other methods, such as a think-aloud protocol or interviews.

    This study is limited in a number of areas. The sample size in this study was small. Thus, the study should be replicated with other samples of high school students. In doing so, more items that will help to improve the subscales should be added and further analyzed. When we examine all five sub-constructs within SDT (i.e., amotivation, external regulation, introjected regulation, identified regulation, and intrinsic motivation), it will be possible to investigate which type of L2 motivation is most relevant to this population. In future research on L2 motivation, we can not only further analyze questionnaire data within the GRM but can also examine how the improvement of a questionnaire after these analyses changes the relationship between L2 motivation and important variables related to L2 learning. There has been much discussion of what English and English learning mean for Japanese students (Ushioda, 2013), but not enough has been empirically investigated regarding high school students. More research is needed to explain the intensity of studying English in contexts where communication opportunities are scarce and where the target language is still a *foreign language*.

REFERENCES

Arbuckle, J. L. (2005). *Amos 6.0 user's guide*. Chicago, IL: SPSS.

Apple, M. T., Da Silva, D., & Fellner, T. (Eds.). (2013). *Language learning motivation in Japan*. Bristol, UK: Multilingual Matters.

Berwick, R., & Ross, S. (1989). Motivation after matriculation: Are Japanese learners of English still alive after exam hell? *JALT Journal, 11*, 193-210.

Cai, L., Thissen, D., & du Toit, S. H. C. (2011). *IRTPRO guide.* Lincolnwood, IL: Scientific Software International.

Chen, W. H., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22*, 265-289.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research.* Rochester, NY: University of Rochester.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). Clevedon, England: Multilingual Matters.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Harlow, England: Longman.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, NJ; London: Lawrence Erlbaum Associates.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London: Edward Arnold.

Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei &

R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 1-19). Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Hiromori, T. (2003). What enhances language learners' motivation?: High school English learners' motivation from the perspective of self-determination theory. *JALT Journal, 25*, 173-186.

Hiromori, T. (2005). Three factors that motivate L2 learners: From the perspectives of general tendency and individual differences. *JACET Bulletin, 41*, 37-50.

Nishida, R. (2013). The L2 self, motivation, international posture, willingness to communicate and can-do among Japanese university learners of English. *Language Education & Technology, 50*, 43-67.

Noels, K. A. (2009). The internalization of language learning into the self and social identity. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 295-313). Clevedon, UK: Multilingual Matters.

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*, 57-85.

Pae, T. (2008). Second language orientation and self-determination theory: A structural analysis of the factors affecting second language achievement. *Journal of Language and Social Psychology, 27*, 5-27.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danish Institute for Educational Research.

Reckase, M. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*, 207-230.

Sakai, H., & Koike, H. (2008). Changes in Japanese university students' motivation to learn English: Effects of volunteering in an international event. *JALT Journal, 30*, 51-67.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph No. 17, 34* (4, Pt. 2).

Shaikholeslami, R., & Khayyer, M. (2006). Intrinsic motivation, extrinsic motivation, and learning English as a foreign language. *Psychological Reports, 99*, 813-818.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th Ed.). Needham Heights, MA: Allyn and Bacon.

Takahashi, C. (2013). *Motivation and persistence of learning among L2 learners in self-instruction.* Unpublished doctoral dissertation, University of Hawai'i at Mānoa, Honolulu, HI.

Tanaka, H., & Hiromori, T. (2007). The effects of educational intervention that enhances intrinsic motivation of L2 students. *JALT Journal, 29*, 59-80.

Thissen, D., Chen, W. H., & Bock, D. (2003). *MULTILOG 7* [Computer software]. SSI Scientific Software International: Lincolnwood, U.S.A.

Ushioda, E. (2013). Foreign language motivation research in Japan: An 'Insider' perspective from outside Japan. In M. T. Apple, D. Da Silva, & T. Fellner (Eds.), *Language learning motivation in Japan* (pp. 1-14). Bristol, UK: Multilingual Matters.

Vallerand, R., & Bissonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality, 60*, 599-620.

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal, 86*, 54-66.

Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influences of attitudes and affect on willingness to communicate and second language communication. *Language Learning, 54*, 119-152.

APPENDIX:

QUESTIONNAIRE ITEMS

(Answer options: 1 = "completely disagree," 2 = "disagree," 3 = "if I had to choose, then I disagree," 4 = "if I had to choose, then I agree," 5 = "agree", and 6 = "completely agree")

Amotivation

Q3:AMT. I can't understand why I study English.

Q9:AMT. Honestly, I think I am wasting my time in studying English.

Q13:AMT. I can't understand what I am doing studying English.

Q46:AMT. I do not understand what I am gaining from studying English.

External Regulation

Q15:EXT. I study English in order to pass English certificate exams such as the Eiken and TOEIC.

Q23:EXT. I study English because I want to have a good career in the future.

Q33:EXT. I study English because I want to secure well-paid employment in future career.

Q49:EXT. I study English because I want to get a good grade.

Identified Regulation

Q38:IDR. I study English because I think it is good for my personal development.

Q42:IDR. I study English because I think acquiring English conversation and writing skills is necessary for me.

Q52:IDR. I study English because I would like to have English skills that I can use in the future.

Q53:IDR. I study English because English is important for my future.

Intrinsic Motivation

Q5:IM. I study English because it is interesting.

Q18:IM. I study English because I enjoy finding out new things when studying English.

Q20:IM. I study English because I enjoy gaining a sense of accomplishment and satisfaction.

Q21:IM. I study English because I enjoy having more knowledge about English.

Q22:IM. I study English for the "high" I feel when hearing English spoken.

Q48:IM. I study English for the "high" feeling that I experience while speaking in English.

Q51:IM. I study English because I enjoy making progress with my English studies.