

Developing and Evaluating a Questionnaire to Measure EFL Learners' Vocabulary Learning Motivation

Mitsuko Tanaka

Introduction

Self-determination theory (SDT; Deci and Ryan 1985, 2000, 2002) is one of the most influential motivational theories in the field of educational psychology. Although numerous studies have been conducted to investigate the motivation to learn a second language (L2) using the SDT framework (e.g., Hiromori 2006; Noels et al. 2000; Pae 2008; Tanaka 2013; Vandergrift 2005), there is no research that examines the motivation for vocabulary learning when studying English as a foreign language (EFL) from this perspective. As there is no SDT questionnaire focusing on EFL vocabulary learning motivation, this study aims to develop and evaluate an SDT questionnaire for EFL vocabulary learning using Rasch analysis.

Self-determination Theory

SDT (Deci and Ryan 2002) categorizes motivation into three broad categories: intrinsic motivation, extrinsic motivation, and amotivation. Intrinsic motivation refers to motivation to engage in an activity for the sake of one's own enjoyment. Extrinsic motivation refers to motivation driven by external rewards. Amotivation is a state of lack of motivation. Extrinsic motivation is further classified into four types of regulation, three of which (identified, introjected, and external regulation) have been employed in empirical studies in L2 motivational literature (e.g., Noels et al. 2000; Tanaka 2013). Identified regulation is a state regulated by the importance and

M. Tanaka (✉)
Ritsumeikan University, Shiga, Japan
e-mail: mtanaka@fc.ritsumei.ac.jp

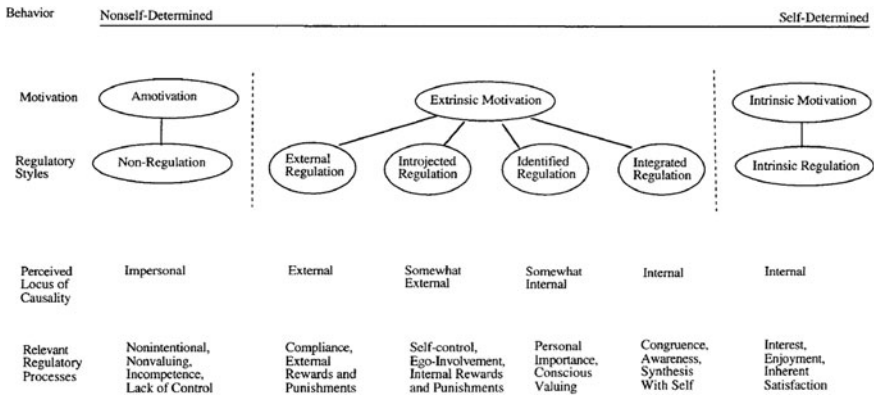


Fig. 1 The self-determination continuum (Ryan and Deci 2000, p. 72)

values of learning. For example, learners with identified regulation study English vocabulary because they believe that English vocabulary is useful or important to accomplish their life goals. Introjected regulation concerns the maintenance of a person’s self-worth. For instance, students study English vocabulary because they do not want their classmates to think that they are poor at English or slow in acquiring English vocabulary. External regulation is a state regulated by rewards or punishments. For example, learners with external regulation study English vocabulary because they want to get course credits, grades, or high test scores. According to SDT, these five types of motivation and regulation are ordered from intrinsic motivation to amotivation on a continuum (Fig. 1) and have a simplex-like structure. Theoretically, adjunct regulations on the continuum should be correlated more highly than regulations situated further apart.

Research Purposes

As discussed above, there is no questionnaire for EFL vocabulary learning motivation using the SDT framework. This study aims to develop and evaluate an SDT questionnaire for EFL vocabulary learning using Rasch analysis.

1. Does each item function properly?
2. Is each of the five constructs reliable?
3. Is each of the five constructs unidimensional?
4. Is the 6-point rating scale psychometrically optimal?
5. Do the five constructs form the simplex-like structure that SDT postulates?

Method

Creation of a Questionnaire for SDT Vocabulary Learning Motivation

An SDT questionnaire for English vocabulary learning motivation was developed drawing primarily from Tanaka (2013, 2014). The developed questionnaire consists of five constructs (intrinsic motivation, identified regulation, introjected regulation, external regulation, and amotivation for learning English vocabulary), with five items in each construct. The questionnaire is a 6-point Likert scale ranging from 1 (Strongly disagree) to 6 (Strongly agree). See Appendix for the English translation of the questionnaire.

Data and Sample

The data for this study comes from first-year science and engineering students ($N = 179$; mostly male students aged between 15 and 16) at a public technical college in Japan. They took five English classes per week (45 min per session) and vocabulary lists were assigned as weekly homework over the year. At the time of data collection, the students had completed approximately four years of compulsory English learning at secondary schools (i.e., at junior high school and college). The questionnaire was administered in Japanese in classes around the end of the 2012 academic year.

Data Analysis Procedures

Rasch analyses were performed using Winsteps 3.80.0 as follows. First, a Rasch fit analysis was conducted to examine items and the reliability of each construct measured by the questionnaire. Second, a Rasch principal components analysis (PCA) of item residuals was conducted to examine dimensionality of each construct. Third, the rating scale categories of each construct were assessed and optimized based on Linacre's (2002) six criteria. In addition to these Rasch analyses, Pearson product-moment correlation coefficients were calculated among the five constructs using SPSS 19.0 to examine the simplex-like structure that SDT postulates.

Results

Items

First, a Rasch fit analysis was conducted to examine items designed to measure each construct in the questionnaire. The criteria for acceptable items were the infit and outfit mean square (MNSQ) statistics of 0.50–1.50 (Linacre 2012, p. 553). Table 1 shows the summary of the Rasch item fit statistics. All infit and outfit MNSQ statistics were between 0.50 and 1.50 (Max. = 1.38, Min. = 0.68 for infit MNSQ statistics; Max. = 1.48, Min. = 0.69 for outfit MNSQ statistics). The point-measure correlations of the items were adequately high ($M = 0.82$, Max. = 0.91, Min. = 0.75). Taken together, each item functioned properly and adequately contributed to measuring the intended construct.

Reliability and Separation of Measures

The reliability of each construct was examined based on Rasch person reliability and separation estimates and Rasch item reliability and separation estimates. The criteria for person estimates are above 0.80 for reliability and above 2.0 for separation, as person reliability of 0.80 indicates the presence of 2 or 3 statistically distinct levels in the sample (Linacre 2012, p. 574). As shown in Table 2, while three constructs (intrinsic motivation, introjected regulation, and amotivation) satisfied the criteria, two constructs (identified and external regulation) showed person reliability and separation estimates slightly lower than the criteria (reliability: 0.7, separation: 1.81 and 1.82). Consequently, eight misfitting people (5 % of the total of 179 participants) were temporarily eliminated based on the analysis of the most unexpected responses, and person reliability (separation) were recalculated for identified and external regulation. As a result, person reliability (separation) improved into 0.82 (2.12) for identified regulation and 0.81 (2.05) for external regulation, satisfying the criteria. Given that the elimination of a very small number of misfitting people improved the reliability (separation) estimates, person reliability of each construct was adequately high.

With respect to item estimates, a reliability estimate above 0.90 and a separation above 3.0 are considered ideal values, as this confirms the item difficulty hierarchy (low, medium, and high difficulties) of the instrument (Linacre 2012, p. 575). Most of the item reliability (separation) estimates were very high, being above or very close to the ideal values of 0.90 (3.00). However, one construct (external regulation) showed low item reliability (0.67) and separation (1.47) estimates. When the eight misfitting people were temporally eliminated, item reliability (separation) improved to 0.77 (1.82). Although these values are still lower than the ideal values, the reliability of 0.77 is not considered very problematic as it is close enough to the value of 0.80 where items are stratified between 2 and 3 levels in terms of difficulty.

Table 1 Rasch item fit statistics for the five SDT motivation items

Item	Measure	SE	Infit	Infit	Outfit	Outfit	PMC
			MNSQ	ZSTD	MNSQ	ZSTD	
<i>Intrinsic motivation for learning English Vocabulary (IM)</i>							
IM1	0.49	0.11	0.98	-0.17	0.94	-0.45	0.82
IM2	0.73	0.12	0.75	-2.35	0.69	-2.56	0.84
IM3	-0.90	0.11	1.38	3.02	1.48	3.51	0.79
IM4	-0.91	0.11	1.01	0.09	1.00	0.01	0.83
IM5	0.59	0.12	0.82	-1.59	0.81	-1.57	0.83
<i>Identified regulation for learning English Vocabulary (ID)</i>							
ID1	-0.49	0.12	1.02	0.21	1.04	0.43	0.76
ID2	0.62	0.12	1.25	2.11	1.24	1.87	0.75
ID3	-0.49	0.12	0.84	-1.51	0.80	-1.86	0.80
ID4	0.57	0.12	0.86	-1.27	0.80	-1.68	0.81
ID5	-0.21	0.11	1.03	0.28	1.09	0.83	0.75
<i>Introjected regulation for learning English Vocabulary (IJ)</i>							
IJ1	-0.71	0.12	1.32	2.51	1.29	2.28	0.88
IJ2	0.04	0.12	1.07	0.65	1.09	0.79	0.88
IJ3	0.21	0.12	0.68	-3.02	0.70	-2.75	0.91
IJ4	0.12	0.12	1.01	0.14	0.98	-0.10	0.88
IJ5	0.34	0.12	0.86	-1.24	0.90	-0.82	0.89
<i>External regulation for learning English Vocabulary (EX)</i>							
EX1	0.07	0.12	1.02	0.24	0.99	-0.05	0.79
EX2	0.22	0.12	0.78	-2.10	0.81	-1.73	0.83
EX3	0.19	0.12	1.03	0.34	0.99	-0.05	0.79
EX4	-0.29	0.12	1.09	0.88	1.09	0.84	0.77
EX5	-0.19	0.12	1.06	0.61	1.04	0.35	0.78
<i>Amotivation for learning English Vocabulary (AM)</i>							
AM1	0.18	0.13	1.11	1.00	1.12	1.04	0.80
AM2	-0.80	0.13	1.00	0.03	0.99	0.00	0.85
AM3	0.61	0.13	1.03	0.31	1.02	0.21	0.79
AM4	-0.29	0.13	0.82	-1.65	0.84	-1.38	0.85
AM5	0.30	0.13	1.05	0.48	1.05	0.43	0.80
M	0.00	0.12	1.00	-0.08	0.99	-0.10	0.82
Max.	0.73	0.13	1.38	3.02	1.48	3.51	0.91
Min.	-0.91	0.11	0.68	-3.02	0.69	-2.75	0.75

Note: PMC Point-measure correlation

However, some improvement in reliability and separation is recommended for a revised version. As low item reliability indicates “a narrow range of item measures, or a small sample” (p. 575), a revised version should have more items with a wider difficulty range or should be tested with a larger sample with wider ability variance.

Table 2 Reliability and separation of the five constructs

Components checked	Values				
	IM	ID	IJ	EX	AM
Item separation	6.23	4.00	2.68	1.47 (1.82)	3.60
Item reliability	0.97	0.94	0.88	0.67 (0.77)	0.93
Person separation	2.07	1.82 (2.12)	2.69	1.81 (2.05)	2.00
Person reliability	0.81	0.77 (0.82)	0.88	0.77 (0.81)	0.80

Note Values within the brackets represent estimates when the eight misfitting people are removed
IM Intrinsic Motivation; *ID* Identified Regulation; *IJ* Introjected Regulation; *EX* External Regulation; *AM* Amotivation

Dimensionality

Dimensionality of each construct was examined using the Rasch PCA of item residuals. Construct unidimensionality is assessed in terms of variance explained and variance unexplained by the Rasch measures. Table 3 shows the results of the analysis. The four constructs (intrinsic motivation, identified regulation, introjected regulation, and amotivation) had an adequate amount of variance explained by the Rasch measures as they were more than the half the total variance. Concerning unexplained variance, the ideal eigenvalue of the first contrast in the residuals was less than 2.0 (Linacre 2012, p. 353). In practice, however, the eigenvalue should be less than 3.0, as the strength of at least 3 items (i.e., eigenvalue of 3.0) is necessary to form a secondary dimension (p. 496). As shown in Table 3, all the five constructs had eigenvalues less than 3.0 and thus satisfied the practical criterion. However, the unexplained variance in percentages appeared to be somewhat large. In particular, external regulation had a large amount of residuals for the first contrast (30 %), which was greater than the variance explained by the item difficulties (20.1 %). The total unexplained variance (58.1 %) was also larger than the total variance

Table 3 Rasch PCA of item residuals of the five constructs

Variance component	Values				
	IM	ID	IJ	EX	AM
Raw variance explained by measures	65.60	51.50	66.30	41.90	60.50
Raw variance explained by persons	40.90	28.4	50.9	21.8	41.6
(eigenvalue)	6.0	2.9	7.6	1.9	5.3
Raw variance explained by items	24.7	23.1	15.4	20.1	18.9
(eigenvalue)	3.6	2.4	2.3	1.7	2.4
Raw unexplained variance	34.40	48.5	33.7	58.1	39.5
Unexplained variance in 1st contrast	16.80	19.60	10.90	30.00	11.90
(eigenvalue)	2.50	2.00	1.60	2.60	1.50

Note: *IM* Intrinsic Motivation; *ID* Identified Regulation; *IJ* Introjected Regulation; *EX* External Regulation; *AM* Amotivation

Table 4 Positively and negatively loading items in the Rasch PCA of item residuals for external regulation

Items	Loading	Measure	Infit MNSQ	Outfit MNSQ	Item Description
EX4	0.85	-0.29	1.09	1.09	Because I want to get English course credits
EX5	0.85	-0.19	1.06	1.04	Because I don't want to fail the English course
EX2	-0.68	0.22	0.78	0.81	Because I want to get good grades
EX1	-0.62	0.07	1.02	0.99	Because I want to get high scores on tests
EX3	-0.54	0.19	1.03	0.99	Because I don't want to get bad grades

explained by measures (41.90 %). Given that the total explained variance should ideally be four times larger than the total unexplained variance (p. 496), residuals in the construct of external regulation is very large. As such, item loadings were examined to explore a possible secondary dimension.

As shown in Table 4, the five items were separated into two clusters. Whereas two items with high positive loadings (EX4 and EX5) concern course credits, three items with high negative loadings (EX1, EX2, and EX3) are grade-related items. The disattenuated person measures from these two clusters of items showed a mere medium correlation ($r = 0.53, p < 0.001$).

Although some degree of multidimensionality is suggested for this construct, “[m]ultidimensionality always exists to a lesser or greater extent” (Linacre 2012, p. 497). Examination of the content of items is also important to determine unidimensionality. Linacre (p. 489) suggested the following guidelines for determining unidimensionality:

[L]ook at the content (wording) of the items. If those items are different enough to be considered different dimensions (similar to “height” and “weight”), then split the items into separate analyses. If the items are part of the same dimension (similar to “addition” and “subtraction” on an arithmetic test), then no action is necessary. You are seeing the expected co-variance of items in the same content area of a dimension.

The theoretical content of external regulation represents a “broad” motivational state regulated by external factors such as rewards and punishment, which include grades, scores, and credits. Although the removal of either grade- or credit-related items improves the Rasch unidimensionality of this construct, both clusters are part of the same external regulation. As such, it is not necessary to separate the items into two constructs.

Rating Scale Categories

The effectiveness of the original 6-point rating scale categories (1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Slightly disagree*, 4 = *Slightly agree*, 5 = *Agree*, and 6 = *Strongly agree*) was examined and optimized based on Linacre's (2002) six guidelines:

1. Each category should have more than 10 observations;
2. Each category should have a peak in the probability curve;
3. The average category measures should progress with the rating scale categories;
4. Outfit mean squares should be smaller than 2.0;
5. Threshold calibration should progress with the rating scale category; and
6. The category threshold should be between 1.4 and 5.0 logits apart.

Concerning the sixth criterion, the minimum threshold separation was assessed based on Wolfe and Smith's (2007) criteria: 0.59, 0.81, 1.1, and 1.4 for a 6-, 5-, 4-, and 3-point scale, respectively. When the above six criteria were not satisfied, the rating scale categories were optimized by combining categories.

Table 5 shows the summary of the category structure for intrinsic motivation. In the 6-point rating scale, the separation between the first and second thresholds ($\tau_1 = -2.01$, $\tau_2 = -1.71$) was 0.30, which was well below the required 0.59 logits

Table 5 Summary of the category structure for intrinsic motivation

Category label	Count	(%)	Average measure	Infit MNSQ	Outfit MNSQ	Structure measure
The 6-point rating scale						
1 Strongly disagree	184	(21)	-2.60	1.09	1.15	NONE
2 Disagree	120	(14)	-1.69	0.73	0.69	-2.01
3 Slightly disagree	201	(23)	-0.81	0.78	0.70	-1.71
4 Slightly agree	212	(24)	0.37	0.76	0.77	-0.35
5 Agree	97	(11)	1.40	1.07	1.23	1.53
6 Strongly agree	72	(8)	2.40	1.77	1.60	2.55
The 5-point rating scale						
1 Disagree	304	(34)	-2.70	1.03	1.03	NONE
2 Slightly disagree	201	(23)	-1.62	0.82	0.81	-2.12
3 Slightly agree	212	(24)	-0.21	0.78	0.70	-1.03
4 Agree	97	(11)	0.92	1.09	1.24	1.01
5 Strongly agree	72	(8)	2.08	1.49	1.46	2.13

Note Boldface indicates values that did not meet the criteria

for a 6-point rating scale. In the 5-point rating scale, when categories 1 and 2 were combined, the separation between the thresholds became 1.09 ($\tau_1 = -2.12$, $\tau_2 = -1.03$), which was greater than the required 0.81 for a 5-point rating scale. The other criteria were also satisfied as all the categories had more than 10 observations; the outfit mean square statistics were below 2.0; the average category measures were ordered, progressing from -2.70 for category 1 to 2.08 for category 5; the shape of the probability curves was peaked for each category (Fig. 2). Thus, the 5-point rating scale was considered optimal for the construct of intrinsic motivation.

Table 6 shows the summary of the category structure for identified regulation. In the 6-point rating scale, threshold measures were disordered between categories 1 and 2. In the 5-point rating scale when these categories are combined, the threshold measures were ordered but the separation between the thresholds (0.35 , $\tau_1 = -1.40$, $\tau_2 = -1.05$) was well below the required 0.81 for a 5-point rating scale. Consequently, categories 1 and 2 were combined again. The separation between the first and second thresholds ($\tau_1 = -1.49$, $\tau_2 = -0.12$) became 1.37, which was larger than the required 1.1 for 4-point rating scale. The other criteria were also satisfied (see Table 6 and Fig. 3). Thus, the 4-point rating scale was considered optimal for the construct of identified regulation.

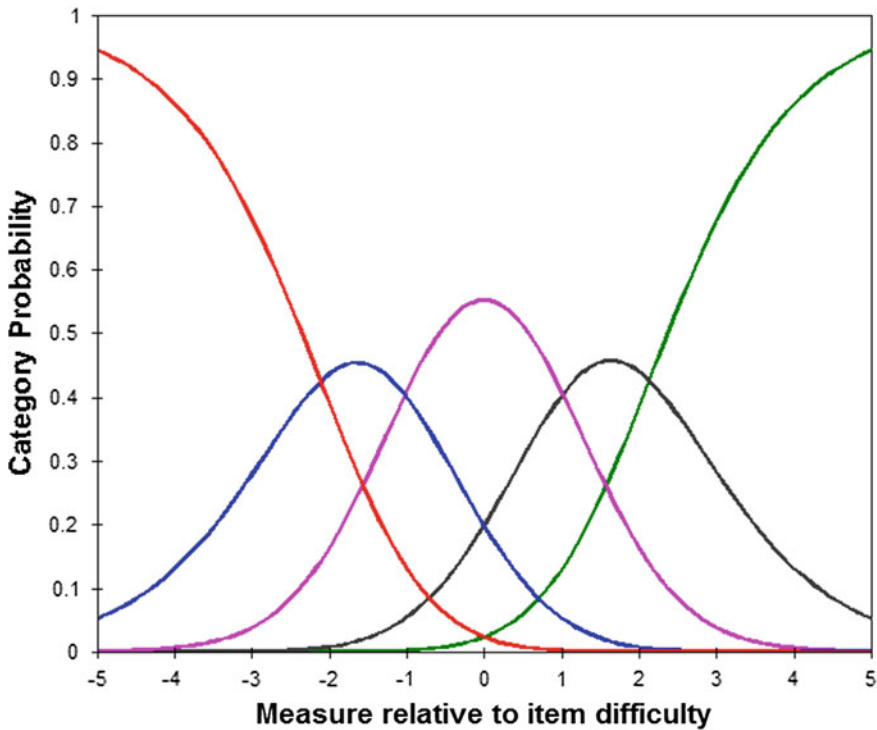


Fig. 2 The 5-point rating scale performance for intrinsic motivation

Table 6 Summary of the category structure for identified regulation

Category label	Count	(%)	Average measure	Infit MNSQ	Outfit MNSQ	Structure measure
The 6-point rating scale						
1 Strongly disagree	74	(8)	-0.98	1.30	1.42	NONE
2 Disagree	45	(5)	-0.83	0.75	0.68	-0.82
3 Slightly disagree	126	(14)	-0.42	0.68	0.63	-1.48
4 Slightly agree	247	(28)	0.31	0.74	0.84	-0.66
5 Agree	228	(26)	1.17	0.70	0.73	0.70
6 Strongly agree	168	(19)	1.76	1.71	1.32	2.26
The 5-point rating scale						
1 Disagree	119	(13)	-1.32	1.22	1.22	NONE
2 Slightly disagree	126	(14)	-0.92	0.68	0.67	-1.40
3 Slightly agree	247	(28)	-0.04	0.82	0.87	-1.05
4 Agree	228	(26)	0.91	0.72	0.76	0.42
5 Strongly agree	168	(19)	1.61	1.45	1.30	2.03
The 4-point rating scale						
1 Disagree	245	(28)	-1.74	1.09	1.07	NONE
2 Slightly disagree	247	(28)	-0.78	0.87	0.91	-1.49
3 Slightly agree	228	(26)	0.42	0.78	0.82	-0.12
4 Agree	168	(19)	1.30	1.20	1.19	1.62

Note Boldface indicates values that did not meet the criteria

Table 7 shows the summary of the category structure for introjected regulation. All the categories had more than 10 observations; the outfit mean square statistics were below 2.0; the average category measures were ordered, progressing from -3.47 for category 1-3.19 for category 6, and the smallest separation between the thresholds was 1.71 ($\tau_5 = 2.17$, $\tau_6 = 3.88$), which was greater than the required 0.59 logits for a 6-point rating scale. Moreover, the shape of the probability curves peaked for each category (Fig. 4). Thus, the 6-point rating scale was optimal for this construct.

Table 8 shows the summary of the category structure for external regulation. Categories 1 and 2 were combined twice, as the separation between the first and the

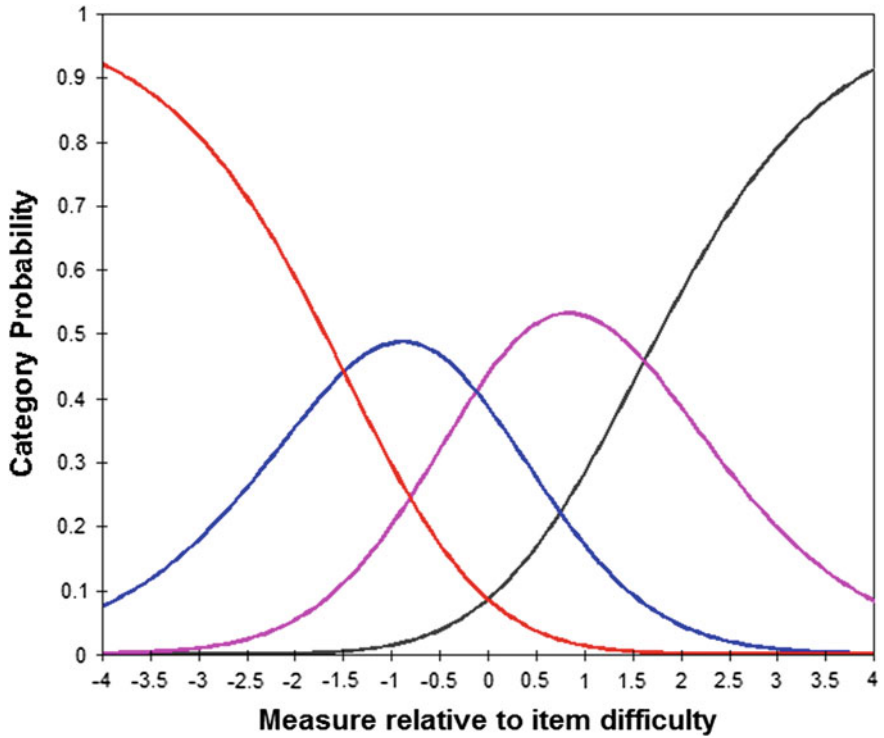


Fig. 3 The 4-point rating scale performance for identified regulation

Table 7 Summary of the category structure for introjected regulation

Category label	Count	(%)	Average measure	Infit MNSQ	Outfit MNSQ	Structure measure
The 6-point rating scale						
1 Strongly disagree	222	(25)	-3.47	1.56	1.42	NONE
2 Disagree	195	(22)	-2.51	0.62	0.64	-4.28
3 Slightly disagree	211	(24)	-0.87	0.74	0.72	-1.75
4 Slightly agree	167	(19)	0.57	0.91	0.97	-0.02
5 Agree	62	(7)	1.66	1.48	1.53	2.17
6 Strongly agree	31	(3)	3.19	1.17	1.15	3.88

Note Boldface indicates values that did not meet the criteria

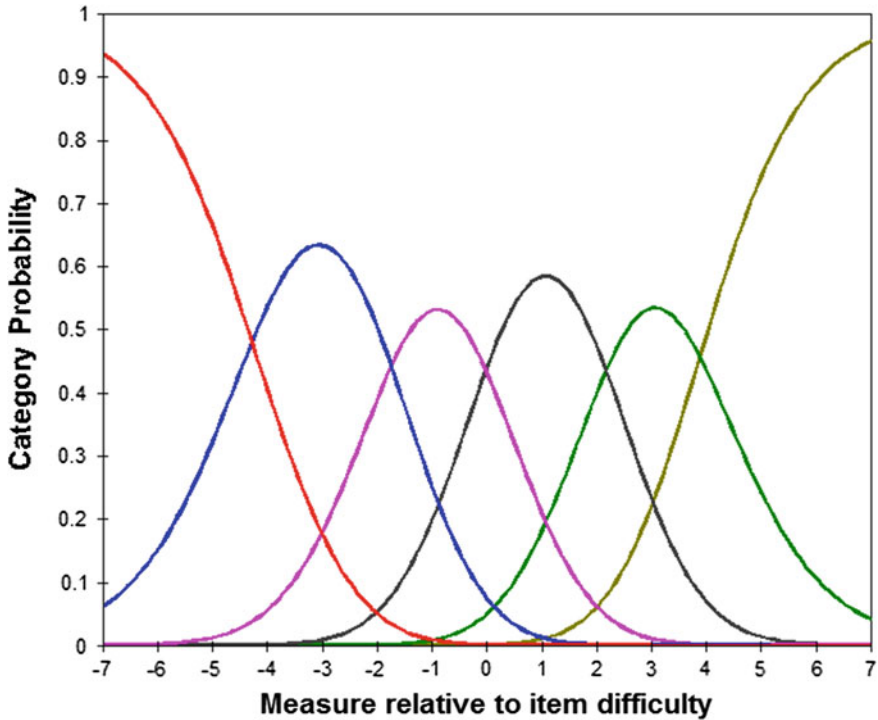


Fig. 4 The 6-point rating scale performance for introjected regulation

second thresholds was well below the required value of 0.59 logits for a 6-point rating scale and 0.81 logits for a 5-point rating scale. In the 4-point rating scale, the smallest separation between the thresholds (1.42, $\tau_3 = 0.08$, $\tau_4 = 1.46$) was greater than the required 1.1 for a 4-point rating scale. The other criteria were also satisfied (see Table 8 and Fig. 5). Thus, the 4-point rating scale was considered optimal for the construct of external regulation.

Table 9 shows a summary of the category structure for amotivation. Categories 5 and 6 in the 6-point and categories 4 and 5 in the 5-point rating scales were combined as threshold measures were reversed. In the 4-point rating scale, the threshold measures were ordered. The other criteria were also satisfied (see Table 9 and Fig. 6). Thus, the 4-point rating scale was considered optimal for the construct of identified regulation.

Table 10 shows the results of rating scale optimization. The 6-point rating scale was retained only for introjected regulation. Scales were reduced into 5-point rating scales for intrinsic motivation, and 4-point rating scales for identified regulation, external regulation, and amotivation.

Table 8 Summary of the category structure for external regulation

Category label	Count	(%)	Average measure	Infit MNSQ	Outfit MNSQ	Structure measure
The 6-point rating scale						
1 Strongly disagree	55	(6)	-1.09	1.41	1.74	NONE
2 Disagree	50	(6)	-0.69	0.98	1.01	-1.31
3 Slightly disagree	103	(12)	-0.18	0.84	0.87	-1.06
4 Slightly agree	259	(29)	0.40	0.77	0.78	-0.74
5 Agree	213	(24)	1.14	0.72	0.77	0.93
6 Strongly agree	212	(24)	1.61	1.23	1.11	2.19
The 5-point rating scale						
1 Disagree	105	(12)	-0.91	1.30	1.38	NONE
2 Slightly disagree	103	(12)	-0.66	0.88	0.92	-1.27
3 Slightly agree	259	(29)	0.00	0.80	0.81	-1.16
4 Agree	213	(24)	0.80	0.74	0.76	0.57
5 Strongly agree	212	(24)	1.32	1.13	1.09	1.86
The 4-point rating scale						
1 Disagree	208	(23)	-1.27	1.27	1.25	NONE
2 Slightly disagree	259	(29)	-0.66	0.81	0.83	-1.54
3 Slightly agree	213	(24)	0.35	0.78	0.74	0.08
4 Agree	212	(24)	0.99	1.01	1.01	1.46

Note Boldface indicates values that did not meet the criteria

The Theoretical Tenets of the Simplex-Like Structure of the SDT Scale

As discussed earlier, SDT (Deci and Ryan 2002) postulates a simplex-like pattern on the continuum of the five subscales, where adjunct regulations have a stronger and positive correlation with each other. The results of the correlation analysis showed that the five constructs have the simplex-like structure that SDT postulates (Table 11). As such, the measurement of the five constructs adequately represents SDT.

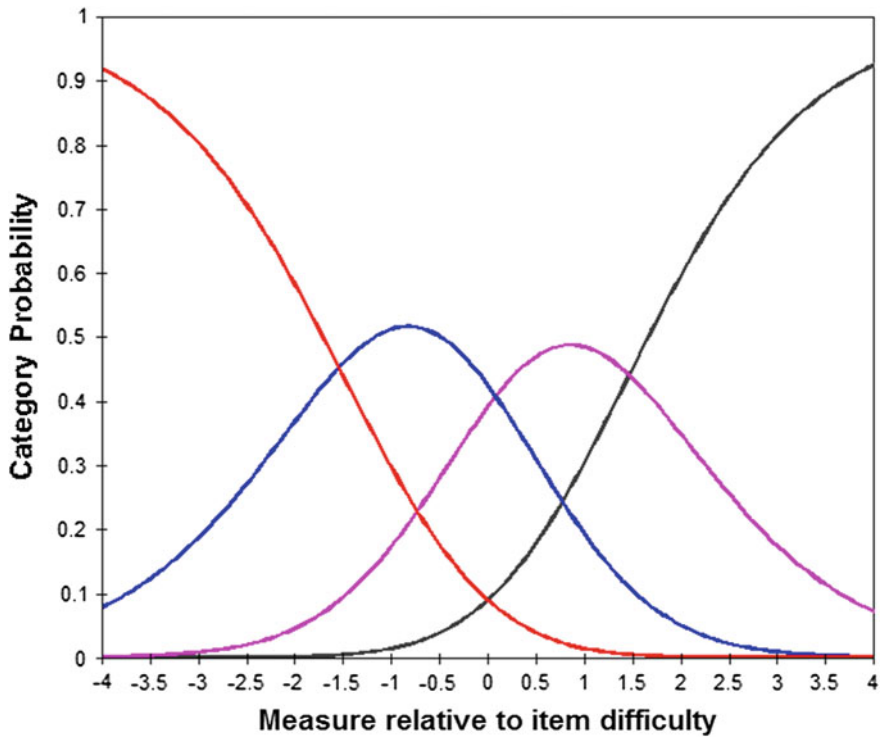


Fig. 5 The 4-point rating scale performance for external regulation

Table 9 Summary of the category structure for amotivation

Category label	Count	(%)	Average measure	Infit MNSQ	Outfit MNSQ	Structure measure
The 6-point rating scale						
1 Strongly disagree	269	(30)	-2.73	1.42	1.11	NONE
2 Disagree	264	(30)	-1.62	0.66	0.74	-2.69
3 Slightly disagree	213	(24)	-0.32	0.69	0.71	-0.72
4 Slightly agree	77	(9)	0.45	0.76	0.73	0.97
5 Agree	34	(4)	0.84	0.87	0.87	1.37
6 Strongly agree	32	(4)	0.89	1.54	2.62	1.06

(continued)

Table 9 (continued)

Category label	Count	(%)	Average measure	Infit MNSQ	Outfit MNSQ	Structure measure
The 5-point rating scale						
1 Strongly disagree	269	(30)	-2.54	1.27	1.11	NONE
2 Disagree	264	(30)	-1.32	0.70	0.76	-2.46
3 Slightly disagree	213	(24)	0.09	0.73	0.72	-0.39
4 Slightly agree	77	(9)	1.02	0.82	0.83	1.47
5 Agree	66	(7)	1.35	1.39	1.80	1.38
The 4-point rating scale						
1 Strongly disagree	269	(30)	-2.23	1.13	1.12	NONE
2 Disagree	264	(30)	-0.84	0.75	0.75	-2.08
3 Slightly agree	213	(24)	0.84	0.78	0.79	0.18
4 Agree	143	(16)	2.05	1.34	1.44	1.91

Note Boldface indicates values that did not meet the criteria

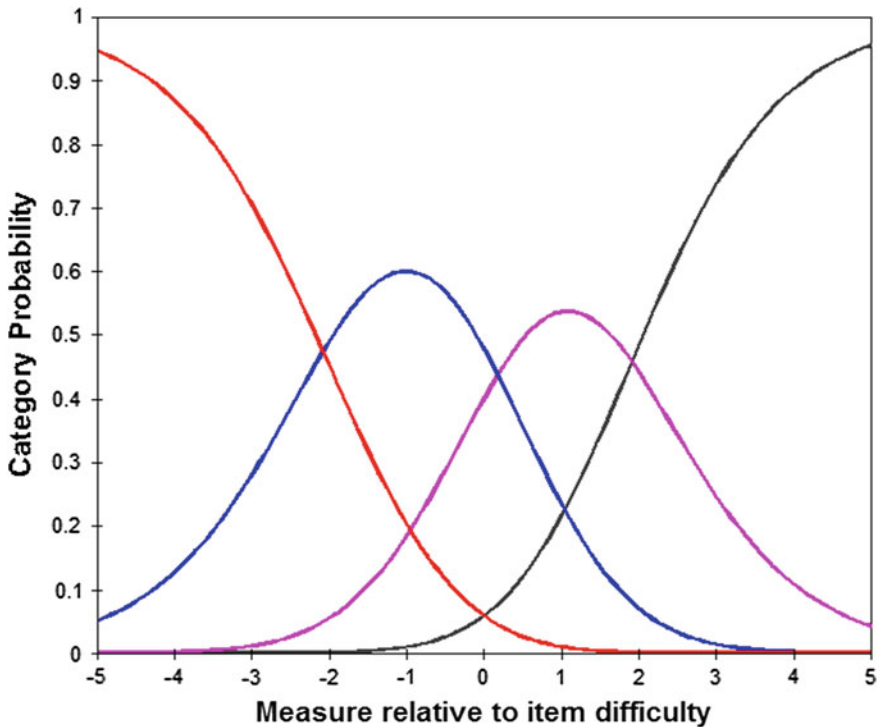


Fig. 6 The 4-point rating scale performance for amotivation

Table 10 Summary of the rating scale optimization

Constructs	The resulting rating scale
Intrinsic motivation	5-point scale (112345)
Identified regulation	4-point scale (111234)
Introjected regulation	6-point scale (123456)
External regulation	4-point scale (111234)
Amotivation	4-point scale (123444)

Table 11 Correlation matrix of the five constructs

Factor	Intrinsic motivation	Identified regulation	Introjected regulation	External regulation
Identified regulation	0.57	–		
Introjected regulation	0.22	0.13	–	
External regulation	0.02	0.15	0.14	–
Amotivation	–0.33	0.52	0.26	–0.12

Conclusion

The present study aimed to develop and evaluate an SDT questionnaire for EFL vocabulary learning motivation using Rasch analysis. First, the results of a Rasch fit analysis showed that each item functions properly and adequately contributes to measuring the intended construct. Second, the results of Rasch reliability and separation analyses revealed that both person and item reliability (separation) were adequately high, although some improvement was recommended for external regulation. Third, the results of the Rasch PCA of item residuals showed that constructs were adequately unidimensional. Fourth, the results of rating scale analysis showed that the original 6-point rating scale was retained only for introjected regulation. The rating scale categories of the remaining four constructs were properly optimized by reducing categories. Fifth, the results of the correlation analysis showed that the measurement of the five constructs adequately represented the self-determination theory. Taken together, the developed SDT questionnaire instrument for EFL vocabulary learning motivation was adequately valid and reliable for the participants of the present study.

Appendix: Questionnaire Items (English Translation)

Why do you study English vocabulary?

<i>Factor 1: Intrinsic Motivation for Learning English Vocabulary (IM)</i>	
IM1	Because learning English vocabulary is enjoyable
IM2	Because learning English vocabulary is interesting
IM3	Because I feel pleasure when I discover new things through learning English vocabulary
IM4	Because I feel pleasure about increasing my English vocabulary
IM5	Because I like learning English vocabulary
<i>Factor 2: Identified Regulation for Learning English Vocabulary (ID)</i>	
ID1	Because English vocabulary is useful
ID2	Because English vocabulary is important to make my dreams come true
ID3	Because English vocabulary will be necessary in the future
ID4	Because English vocabulary is necessary to attain my life goals
ID5	Because it is important to acquire English vocabulary
<i>Factor 3: Introjected Regulation for Learning English Vocabulary (IJ)</i>	
IJ1	Because I'd feel ashamed if I have smaller amount of English vocabulary than my classmates
IJ2	Because I'd feel ashamed if my classmates think that I am an incapable student
IJ3	Because I don't want my classmates to think that I am poor at English
IJ4	Because I don't want my classmates to think that I don't have an adequate amount of English vocabulary
IJ5	Because I don't want my classmates to think that I am slow in acquiring English vocabulary compared to others
<i>Factor 4: External Regulation for Learning English Vocabulary (EX)</i>	
EX1	Because I want to get high scores on tests
EX2	Because I want to get good grades
EX3	Because I don't want to get bad grades
EX4	Because I want to get English course credits
EX5	Because I don't want to fail the English course
<i>Factor 5: Amotivation for Learning English Vocabulary (AM)</i>	
AM1	I won't get anything out of learning English vocabulary
AM2	I don't know what I am getting out of learning English vocabulary
AM3	Learning English vocabulary is useless
AM4	I cannot see why I have to study English vocabulary
AM5	Learning English vocabulary is meaningless

Note All the questionnaire items are randomly ordered 6-point Likert scale items

References

- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.
- Deci, E. L., & Ryan, R. M. (2000). The “What” and “Why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.
- Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- Hiromori, T. (2006). *Theories and practices that increase the motivation of foreign language learners*. Tokyo: Taiga Shuppan.
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, *3*, 85–106.
- Linacre, J. M. (2012). *A user’s guide to WINSTEPS: Rasch-model computer program*. Chicago, IL: MESA.
- Noels, K. A., Pelletier, L., Clément, R., & Vallerand, R. (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.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78.
- Tanaka, M. (2013). Examining kanji learning motivation using self-determination theory. *System*, *41*, 804–816.
- Tanaka, M. (2014). The effects of affective variables and kanji growth on L1 Chinese JSL learners’ kanji learning (Doctoral dissertation, 2014). Available from ProQuest Dissertations and Theses database. (UMI No. 3611180).
- Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics*, *26*, 70–89.
- Wolfe, E. W., & Smith, E. V., Jr. (2007). Instrument development tools and activities for measure validation using Rasch models: Part II—validation activities. *Journal of applied measurement*, *8*, 204–234.