Grit and Intrinsic Motivation for Language Learning: Instrument validation using the Rasch model

Michael J. Giordano
mikegio123@gmail.com
Kwansei Gakuin University

Abstract

University student grit and intrinsic motivation were measured by survey to determine the viability of grit as a language learning construct in this pilot study. The instruments were analyzed using the Rasch model in Winsteps (version 4.0) for construct validity, person and item fit, and unidimensionality. Logit scores were used in a correlation analysis to determine the relationship between grit, intrinsic motivation, and autonomous language learning dependent variables. The sample-dependent results suggest that the grit and intrinsic motivation instruments are unidimensional. Additionally, they are moderately correlated (r =.40) but not significantly related to autonomous language learning measurements. However, small sample size and poor person reliability limit the generalizability of the results. Person reliability and item targeting of the two constructs are discussed and disattenuated correlations are used to suggest avenues for future research.

Keywords: Rasch model, instrument validation, grit, motivation

Grit, as defined in the previous literature, is a psychological construct comprised of "a perseverance and passion for long-term goals" (Duckworth, Peterson, Matthews, & Kelly, 2007, p. 1087). Duckworth et al. (2007) originally hypothesized that the individuals who are grittier than others would view achievement as a marathon and would be able to power-through boredom and overcome hardship more than less gritty individuals. This construct, if applicable, would fit in well as an additional individual difference variable alongside motivation, willingness to communicate, self-efficacy, or anxiety in second language (L2) learning. That being said, little research has been conducted on the relationship between grit, motivation, and language learning.

Grit as a construct has been evaluated many times in educational, psychological, sociological research, though it has rarely been examined in foreign language research. In their original study, Duckworth et al. (2007) evaluated the construct validity using confirmatory factory analysis to create the original 12-item grit scale. They discovered a two-factor construct; consistency of interest and persistence of effort. These items were further refined to make a short grit scale (Grit-S) containing only eight items yet effectively and more efficiently measuring the same construct and facets (Duckworth & Quinn, 2009). Additionally, they discovered that grit related significantly to each of the Big Five personality traits; grit is highly positively related to conscientiousness, negatively related to neuroticism, and weakly correlated with agreeableness, extraversion, and openness to experiences. In summary of these two studies, grittier adults progress farther in their education and make fewer career changes; grittier adolescents earn higher GPAs; grittier West Point cadets are less likely to drop out after the first training session; and grittier spelling bee finalists are more likely to advance to later rounds (Duckworth & Quinn, 2009, p. 172-173). These original studies sparked a plethora of research into grit as predictor of success and achievement.

Duckworth and Eskreis-Winkler (2013) argue that one mechanism of grit is deliberate practice. Those people who are able to suffer through the challenges, focus intently on a long-term goal, and repeatedly put in the effort are the most successful. Deliberate practice has been argued by DeKeyser (2007) to be a fundamental aspect of automatization and overall second language acquisition. It seems natural then that grit should also be considered an L2 individual difference (ID) variable alongside other variables like anxiety, willingness to communicate, and, perhaps most similarly, motivation.

Learning a language is a lifelong pursuit with many ups and downs. Students and language learners of all ages are often bombarded with a variety of achievement goals and standards and may suffer fluctuations in motivation. Deci and Ryan (1985), in their discussion of Self-Determination Theory, suggest that motivation for human behavior and the regulation of that motivation is divided into categories and factors. Some people are motivated to act through external pressure or fear of being watched (extrinsic motivational factors). Whereas others are motivated by self-authored or internal factors (intrinsic motivational factors). Those people who are more intrinsically motivated toward a goal or an action tend to show "more interest, excitement and confidence, which in turn is manifest both as enhanced performance, persistence, and creativity and as heightened vitality, self-esteem and general well-being" than those who are extrinsically motivated (Ryan & Deci, 2000, p. 69). The connections between grit—perseverance and passion for long-term goals—and intrinsic motivation are plentiful. Self-driven people who accomplish long-term goals are likely to be both gritty and intrinsically motivated. The extent to which grit and intrinsic motivation are related is the intended outcome of this validation pilot study.

Method

Participants

Thirty-nine first-year students (F = 34, M = 5) from a private university in western Japan participated in this study by responding to a questionnaire via Google Forms. Of these 39 students, two needed to be removed from the study due to careless answering. This reduced to total convenience sample to 32 women and five men ($N = 37$) in three intact classes, including students who were no longer enrolled in the course. Their enrollment in the course did not affect their ability to respond or the presence of the two constructs. Their enrollment status only affected my ability to compare their grit and intrinsic measurements with their autonomous language learning scores on English Central. English Central is an online language learning tool designed to provide learners with videos, vocabulary learning management and quizzes, and speech and pronunciation practice ("English Central", n.d.). Participants in this study were required to watch five videos, study 50 new vocabulary items, and speak 50 lines every week using their personal computer and smartphone. The Google Form containing the questionnaire was available on the Learning Management System (LMS). Participants also received an email notification with the same message and reminder in class to complete the questionnaire in their free time. Ten minutes at the end of one class was allocated to help students access and complete the questionnaire if they wished.

Instrument

Grit items. The grit items on the instrument were developed by borrowing heavily from Duckworth and Quinn's (2009) Short grit (Grit-S) survey and—taking Kramer, McLean, and Martin's (2017) suggestions to heart—were altered to specifically represent language learning grit. All items on the Grit-S survey appear in a similarly worded version in the same order in the current questionnaire. Negatively keyed items which appear in the original survey also appear in the altered version for the current study, altered to better represent the construct of language learning grit rather than general grit. For instance, the first item on the original survey, *New ideas and projects sometimes distract me from previous ones.* was changed to *New assignments or projects in my language class distract me from older assignments or projects.* Two additional items designed to measure participant grittiness were added to the end of the questionnaire to help separate students into groups of high, medium, and low grit. These two additional items were: *When I am studying a language, I cannot be distracted from my task* and *I get disappointed and give up when I am unsuccessful.* In total, there were 10 language learning grit items including five reverse or negatively keyed items.

Intrinsic items. The inspiration for the intrinsic motivation items came from a study regarding intrinsic, extrinsic and integrative motivation by Noels, Clément, and Pelletier (2001). Many of the intrinsic motivation items in the current study were adapted and altered to make "I" statements of varying degrees of difficulty to represent varying degrees of intrinsic motivation. Some examples of the intrinsic motivation items include, *I enjoy learning new things in another language* and *I want to be someone who can speak more than one language.* None of the intrinsic motivation items were reversely keyed following the advice of previous research (Schmitt & Stults, 1985; Swain, Weathers, & Neidrich, 2008; Bond & Fox, 2015; Credé, 2018). Both constructs were measured using a five-point Likert-like scale, ranging from *Exactly like me* to *Not like me at all.* An odd numbered scale was chosen to match the original grit instrument. Construct maps and Item specifications can be found in Appendix A and B.

Dependent variables. In addition to the grit scale and intrinsic motivation measurement scores in the questionnaire, three additional measurements were included for each participant where available: English Central video count, English Central vocabulary count, and English Central spoken lines count. As part of the participants' weekly homework assignments, they are required to watch five videos, speak 50 lines (pronunciation practice), and study 50 vocabulary items found in the videos (a type of dictation activity). At the time of data collection, students were required to have completed 40 videos, 400 spoken lines, and 400 vocabulary items in total. Many of the students had completed much more than was required for the course. One of the hypotheses of the current study is to see if students who complete more than is required from their language learning course are actually grittier and intrinsically motivated than those who are just meeting the requirements.

Research Questions

1. Does the 10-item grit questionnaire reliably separate the participants into varying degrees of grit (construct validity)?
2. Are the grit items unidimensional?
3. Does the 10-item intrinsic motivation questionnaire reliably separate the participants into varying degrees of intrinsic motivation (construct validity)?
4. Are the intrinsic items unidimensional?
5. To what degree, if any, does grittiness correlate to intrinsic motivation measures?
6. Do gritty and intrinsically motivated students study English on their own more than their less gritty and less intrinsically motivated classmates (autonomous language learning)?

Results

The instrument and data were analyzed with Winsteps 4.0 software (Linacre, 2018) using the Rasch Rating Scale model for categorical data. The Rasch analysis consisted of person and item fit analysis, item-person maps, and Rasch principal component analysis (PCA) of item residuals (Apple, 2013; Kramer, McLean, & Martin, 2017).

Person and Item Fit (Grit)

The person data were analyzed using Winsteps to ensure the participants' responses were conforming to the model. The Rasch reliability of the person responses was approximately .59 with a separation of 1.21. With participants 1 and 18 removed from the grit analysis for careless responses, the final N-size for the grit questionnaire became (N =37). This means that the model could only stratify the person responses into one level across the construct. Table 1 shows the descriptive statistics for Grit items. Aside from the

two students removed for erratic responses, three additional students had high outfit MNSQ (>2.0). Their suspect responses were investigated and one misfitting response from each person—all different items—were removed and treated as missing. After removing these responses, the person infit and outfit fell within acceptable ranges (.5 - 1.5). The small sample size in this pilot study could be the cause of person fit issues, as even a one misfitting response can cause model fit issues.

Table 1.
Descriptive statistics for Grit Questionnaire items (N = 37)

Item	Item Description	M	SD
01	(R) New assignments or projects in my language class distract me from older assignments or projects.	2.87	0.86
02	Making mistakes in another language encourages me to study harder.	4.05	0.83
03	(R) When I am studying a language, I lose interest quickly.	3.74	0.94
04	I am a hard worker.	3.28	1.02
05	(R) When I have a language learning goal, I often choose to follow a different goal later.	2.64	0.99
06	(R) I have difficulty focusing on long term projects or goals.	2.44	0.94
07	I always accomplish my language learning goals.	3.00	0.92
08	I often do more than is required when I am studying a language.	3.46	0.85
09	When I am studying a language, I cannot be distracted from my task.	2.92	0.90
10	(R) I get disappointed and give up when I am unsuccessful.	2.85	1.23

Note. A Likert-like scale from (1) Not like me at all to (5) Exactly like me

As for item fit, the data look much more positive. The Rasch item fit analysis found that the reliability of the instrument was .93 with a separation of 3.71 (Table 2). A separation of 3.7 shows that the participants were able to distinguish three separate levels of the construct being measured by the items. All item mean-square scores and z-scores are within the appropriate criteria (0.5-1.5 MNSQ and ±3.0 for ZSTD), indicating that no items need to be removed from the questionnaire (Bond & Fox, 2015; Linacre, 2013).

Table 2
Item fit statistics for Grit Items (N = 37)

Item	Measure	S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
6	1.08	0.21	0.99	0.04	1.01	0.13
5	0.82	0.21	1.03	0.23	1.05	0.31
10	0.61	0.25	1.40	1.72	1.37	1.61
1	0.43	0.21	0.79	-0.98	0.79	-1.01
9	0.39	0.21	1.08	0.44	1.09	0.47
7	0.26	0.21	0.70	-1.48	0.69	-1.57
4	-0.09	0.21	0.88	-0.48	0.86	-0.58
8	-0.55	0.22	0.92	-0.27	0.95	-0.16
3	-1.12	0.24	0.92	-0.28	0.92	-0.25
2	-1.83	0.28	1.23	0.97	1.22	0.96

Note. Measure in Rasch logits, S.E. = standard error, MNSQ = mean squared, ZSTD = standard z-scores.

Additionally, Figure 1 shows the person-item map in which person ability is indicated by *X* marks on the left and item difficulty is measured from low to high on the right. The higher the person on the map, the grittier they are; and the higher the item on the right, the more difficult it is to endorse. As can be seen on the map in Figure 1 and in the item fit statistics in Table 2, the perseverance sub-set of the grit construct were the for most difficult to endorse items. These items were also the items with reverse valances.

A Rasch Principle Components Analysis (PCA) of the standard residuals was conducted to check the unidimensionality of the grit construct. Results indicated that 41.8% of the variance (eigenvalue = 7.19)

was explained by the grit construct and the principal contrast explained 11.1% of the variance (eigenvalue = 1.92). According to dimensionality guidelines Linacre (2013), if the unexplained variance explained by the 1st contrast has an eigenvalue less than 2.0, the possible second dimension has fewer than 2 items and is not likely an issue. These results suggest that that grit construct here is probably unidimensional, but some items might be problematic. Table 3 shows the results of the Rasch PCA. Analysis of the items with negative loadings show that most of these items are "perseverance" sub-scale items with reversed valances. Items 8 and 4 address diligence or passion to complete tasks or do more than is necessary, whereas Items 1 and 9 address inability to focus or distractibility. These two facets seem to be complimentary and might not represent separate constructs. These results, while tentative and sample-dependent, do not support Credé's (2018) suggestion that the grit variable is actually comprised of two independent constructs, passion and perseverance. A larger, more varied sample and a confirmatory factor analysis would be required to definitively contradict the claims made by Credé (2018). That being said, the results of the PCA suggest that the grit instrument is acting unidimensionally and person ability measurements from the Rasch analysis can be used in further analyses.

```
MEASURE PERSON - MAP - ITEM
        <more>|<rare>
    2          +
               |
               |T
           X T|
               |
         XXXX  |
    1      X  +  (R)I have difficulty focusing on long term projects or study goals
           XX S|S (R)When I have a language learning goal, I often choose to follow a different goal later
          XXX  |  (R)I get disappointed and give up when I am unsuccessful
           XX  |  (R)New assignments in my language classes sometimes distract me from older assignments
          XXX  |  I always accomplish my language learning goals
               |  When I am studying a language, I cannot be distracted from my task
      XXXXXXX M|
    0     XXX  +M
           XX  |  I am a hard worker
          XXX  |
            X  |  I often do more than is required when I am studying a language
              S|
            X  |S
   -1       X  +
           XX  |  (R)When I am studying a language, I lose interest quickly
            X T|
               |
               |T
               |  Making mistakes in another langauge encourages me to study harder
   -2          +
         <less>|<freq>
```

Figure 1. Grit item-person map

Table 3
Principle Component Analysis of Standard Residuals for Grit Construct

Item	Loading	Measure	Infit MNSQ	Outfit MNSQ
8	**.61**	-0.55	0.92	0.95
4	**.56**	-0.09	0.88	0.86
3	**.51**	-1.12	0.92	0.92
7	.27	0.26	0.70	0.69
2	.20	-1.83	1.23	1.22
1	**-.58**	0.43	0.79	0.79
9	**-.55**	0.39	1.08	1.09
5	-.34	0.82	1.03	1.05
10	-.25	0.61	1.40	1.37
6	-.18	1.08	0.99	1.01

Note. Measure is in Rasch logits. Positive loadings indicate the items likely measure the intended construct whereas negative loadings indicate a possible subdimension.

Person and Item Fit (Intrinsic Motivation)

A Rasch analysis was conducted with Winsteps 4.0 to assess the reliability and construct validity of the intrinsic motivation items. The participants belong to an international studies course and must study multiple foreign languages. Additionally they are required to study abroad in a country of their choosing in the 2nd or 3rd year in university. It is not surprising that the mean responses for each of these intrinsic motivation items would be so high (Table 4). With participants 1 and 18 removed, person reliability for the intrinsic motivation items was .77 with a person separation of 1.81. These results suggest a targeting problem in the items. Furthermore, the Rasch analysis showed that the item reliability of the intrinsic items was .90 with an item separation of 2.93. These results suggest that the people were able to identify at least two different levels in the intrinsic motivation construct. Similar to the grit instrument, some of the participants gave unexpected responses. One participants' response to one item was removed from the study and treated as missing. Removing this response did not significantly change the reliability measurements or the item fit statistics found in Table 5.

Table 4
Descriptive statistics for Intrinsic Motivation Questionnaire items (N = 37)

Item	Item Description	*M*	*SD*
1	I study languages to improve myself.	4.33	0.66
2	I enjoy learning new things in another language.	4.26	0.75
3	I get pleasure from using another language.	4.15	0.81
4	I enjoy the challenge of trying to learn another language.	4.15	0.74
5	I feel excited when I can use something I have learned recently.	4.46	0.76
6	I want to be someone who can speak more than one language.	4.69	0.69
7	I study other languages in order to understand the culture better.	3.77	1.01
8	I often feel that studying a language will help me in the future.	4.82	0.45
9	I enjoy interacting with people in other languages.	4.36	0.78
10	I feel satisfied when I complete challenging activities in a foreign language.	4.33	0.77

Note. A Likert-like scale from (1) Not like me at all to (5) Exactly like me

Looking at the item fit statistics (Table 5) and the person-item map (Figure 2), it is fairly clear that the items did not target these particular participants' intrinsic motivation well. Most of the participants were able to able to easily endorse even the most difficult item, *I study other languages in order to understand*

the culture better. This is a problem of instrument design and targeting with this specific group of students. It is not recommended to make general conclusions of item fit for such a small sample size.

Table 5
Item Fit Statistics for Intrinsic Items (N = 37)

Item	Measure	*S.E.*	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
7	1.58	0.31	1.29	1.18	1.27	1.12
3	0.83	0.29	0.62	-1.63	0.59	-1.82
4	0.74	0.29	0.78	-0.83	0.82	-0.65
2	0.49	0.30	0.75	-0.95	0.68	-1.29
1	0.22	0.30	0.82	-0.67	1.00	0.09
10	0.22	0.36	1.44	1.58	1.38	1.33
9	0.13	0.31	0.96	-0.06	0.97	-0.03
5	-0.27	0.35	1.21	0.87	1.09	0.39
6	-1.78	0.48	1.34	1.15	1.02	0.22
8	-2.15	0.48	1.10	0.39	0.90	0.08

Note. Measure in Rasch logits. S.E. = standard error, MNSQ = mean squared, ZSTD = standard z-scores.

A principle components analysis conducted on the intrinsic items showed that the measures explained 49.7% of the variance with an eigenvalue of 9.86, while the first contrast explained 10.5% of the unexplained variance (eigenvalue = 2.08). Table 6 shows the loadings for intrinsic items. These results indicate that the intrinsic items explain a significant amount of the variance and is likely unidimensional. Item 7, *I study other languages in order to understand the culture better,* and Item 4, *I enjoy the challenge of trying to learn another language*, might represent different aspects of motivation, for instance, instrumental or extrinsic rather than intrinsic motivation since these two items are dealing with purposes for learning rather than feelings about learning. However, when comparing these items to Item 10 and Item 8, there does not seem to be a significant deviation in theme. They seem to describe satisfaction gained by studying and perceived utility of studying a foreign language.

Table 6
Principle Component Analysis of Standard Residuals for Intrinsic Motivation Construct

Item	Loading	Measure	Infit MNSQ	Outfit MNSQ
10	**.83**	0.22	1.44	1.38
8	**.63**	-2.15	1.10	0.90
5	.14	-0.27	1.21	1.09
9	.13	0.13	0.96	0.97
3	.09	0.83	0.62	0.59
1	.04	0.22	0.82	1.00
7	**-.74**	1.58	1.29	1.27
4	**-.57**	0.74	0.78	0.82
2	-.25	0.49	0.75	0.68
6	-.04	-1.78	1.34	1.02

Note. Measure is in Rasch logits. Positive loadings indicate the items likely measure the intended construct whereas negative loadings indicate the likely presence of a secondary dimension.

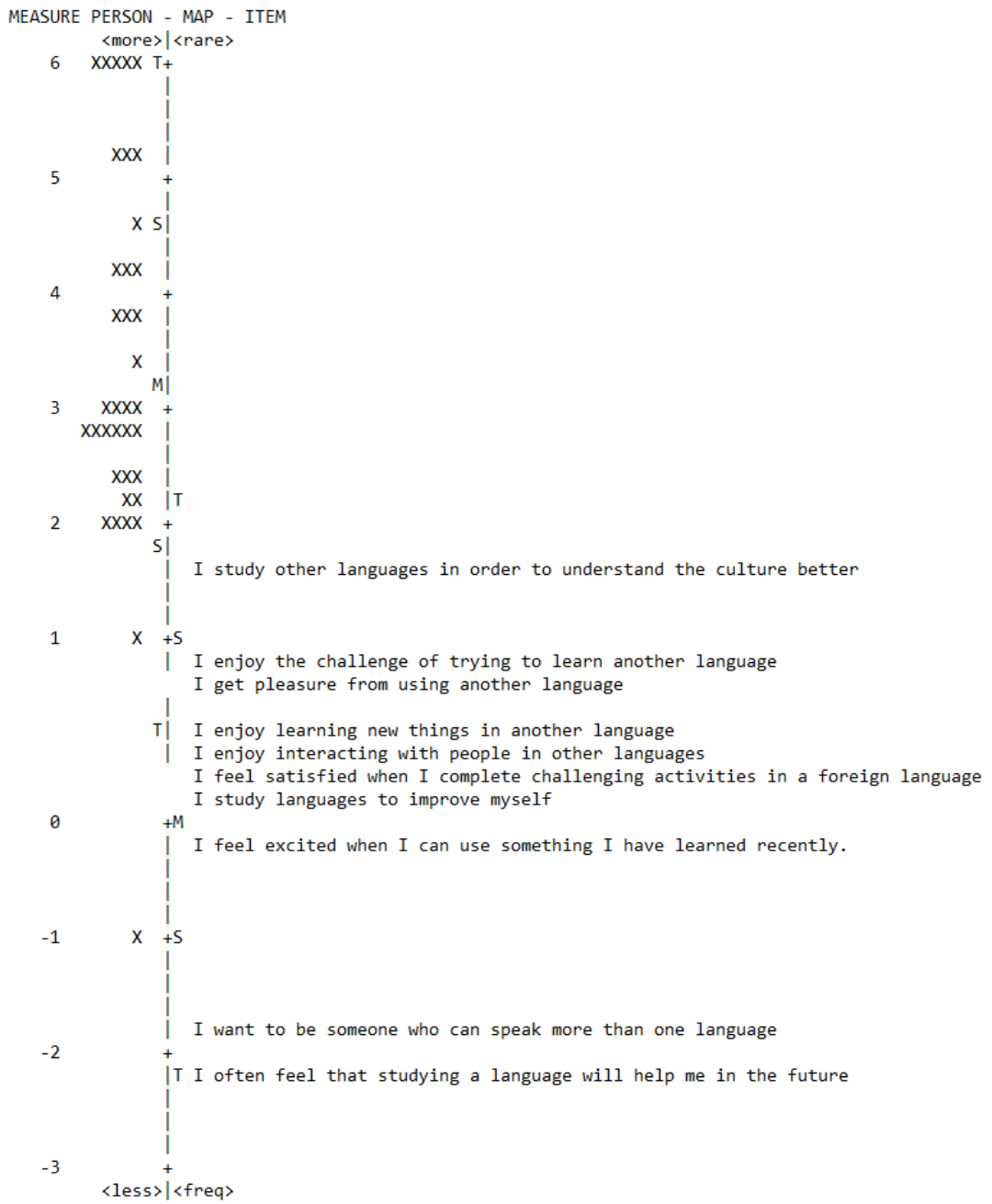

Figure 2. Intrinsic motivation item-person map

Correlation Analysis

In order to see if grit and intrinsic motivation are connected in any way, the person logit measurements for both constructs were exported into SPSS and a correlation analysis was conducted. Pearson correlation coefficient suggested that the two constructs are moderately positively related (r = .40, p = .014). Cohen (1988) offers general psychological effect size interpretation guidelines which suggest that a significant Pearson correlation between r =.30-.50 be classified as having a medium effect. A more recent set of

guidelines for interpreting correlation effect sizes has been suggested by Plonsky and Oswald (2014). In their meta-analysis, they suggest that *r* correlation statistics in the second language research should be amended to represent their actual distribution in the research. They conclude that in L2 research, "*r*s close to .25 [should] be considered small, .40 medium, and .60 large" (p. 889). The Pearson correlation coefficient of .40 suggests that 16% of the variance in the measurement is explained by the relationship between grit and intrinsic motivation, and by either standard, this relationship is one of medium effect.

However, this correlation might not tell the whole story due to the low person reliability of the grit instrument. The use of disattenuated correlations can show if the correlation between two measurements is a result of measurement error in the instrument or if the two scores are not actually correlated. Following a formula to calculate disattenuated correlations from a Pearson correlation (Statistica Help, n.d.), the disattenuated correlation between grit and intrinsic motivation showed that the two constructs are highly correlated when taking error into account ($r = .59$). These results should be taken with a grain of salt, and are not to be used to assess the correlation between the two measures. Disattenuation is presented here as a suggestion to future researchers interested in grit and motivation. The low person reliability of these instruments in this pilot study might be causing an artificially smaller relationship.

The next step in determining the value of these two measurements is to see if they correlate to student effort. Data regarding participants' English Central usage was plotted in a correlation analysis to see if the participants who are grittier or more highly intrinsically motivated actually study more on the app. Participants in the current study were required to watch 40 videos, study 400 vocabulary words, and speak 400 lines at the time of data collection. Some students completed noticeably more than the required amount while others did not use the program at all. Table 7 shows the descriptive statistics for English Central video usage, vocabulary words studied, and lines spoken.

Table 7
Descriptive Statistics for English Central Usage

	N	Min	Max	*M*	*SD*	Skew	SES	Kurt	*KES*
Videos	37	0.00	104	41.55	22.85	0.75	0.38	1.36	0.75
Vocabulary	37	0.00	810	357.05	191.75	0.08	0.38	0.25	0.75
Lines	37	0.00	927	345.63	202.55	0.39	0.38	1.07	0.75

Note. Required amount was 40 videos, 400 vocabulary, and 400 spoken lines.

Spearman's rank correlation coefficients (r_s) were calculated for grit, intrinsic motivation, and the three English Central usage variables because it is more robust to violations of normal distribution and combinations of variable type, in this case continuous and ordinal count data (Larson-Hall, 2016). The results in Table 8 show that grit and intrinsic motivation are moderately correlated at $r_s = .46$, $p = .004$. The three English Central usage counts were all highly correlated with each other (all greater than $r_s > .70$, $p < .001$), as was to be expected. None of the EC measures correlated significantly with the grit measures. However, intrinsic motivation was significantly correlated to one EC variable, spoken lines ($r_s = .39$, $p = .016$). This suggests that participants with higher intrinsic motivation are more inclined to practice speaking using the English Central application.

In order to get a clearer view of the relationships between the five variables, they were all investigated using a scatterplot matrix (Figure 3). Using this figure, the relationship between grit and intrinsic motivation seems to be following a central, upward tendency. Additionally, all three counts of English Central data show a clear, positive relationship as noted in Table 8. Finally, the relationship between intrinsic motivation and the number of lines spoken is less apparent.

Table 8
Spearman's Rho Correlation Coefficients for Grit, Intrinsic Motivation, and English Central Usage

	Intrinsic	Videos	Vocab	Lines
Grit	.46*	.15	.23	.26
Intrinsic		.30	.12	.39*
Videos			.71**	.74**
Vocabulary				.74**

Note.*. $p < 0.05$. **$p < 0.01$.

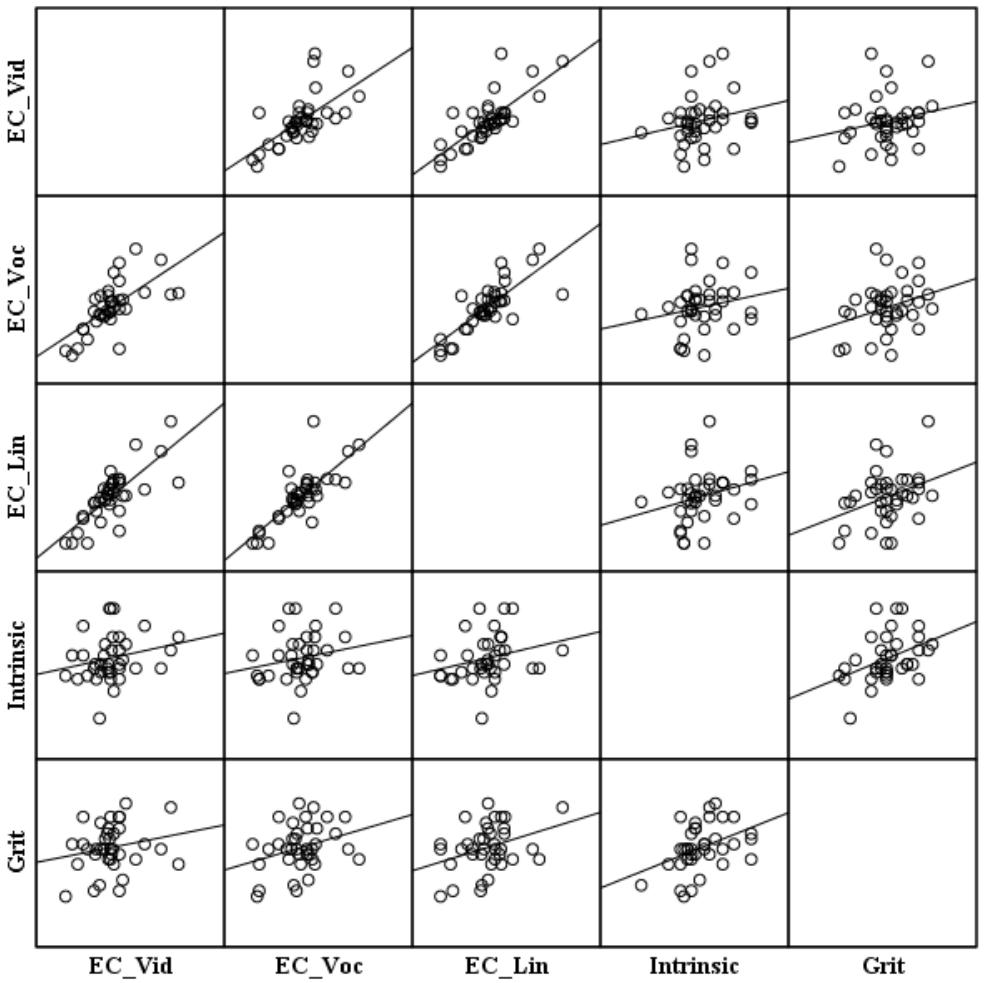

Figure 3. Scatterplot of grit, intrinsic motivation, and EC variables

Discussion

It is necessary to discuss each of the research questions one-by one. The first research question was designed to test one aspect of construct validity of the grit instrument. Do the 10 items adapted from Duckworth and associates reliably separate the participants into varying degrees of language learning grit? The answer seems to be yes and no. The person reliability of the responses was approximately .59 with a

separation of 1.21, yet the item reliability was .93 with a separation of 3.71. Seemingly the participants were able to distinguish between at least three different levels of the construct, but the instrument is not sensitive enough to separate the people as well. This may be a problem of item wording. Clearly some items were more difficult to endorse than others, as Figure 1 shows. However, the low person reliability of instrument is an issue with instrument design. Increasing the number of unidimensional and well-targeted items would likely increase the person reliability on future versions of this instrument as would increasing sample size by surveying individuals with various degrees of grit.

The second research question referred to dimensionality. Does this questionnaire address one construct, grit? The PCA seems to suggest that the grit instrument is unidimensional but some of the items suggest a small subdimension accounting for approximately 25% of the variance. This is not significant enough to suggest that grit is actually two separate constructs. Passion and perseverance for long term goals could actually be two separate but connected constructs, however the limitations of the current pilot study prevent further investigation. Factor analysis would be a better methodology to address Credé (2018) suggestion that grit is two constructs. Within the current study, it is possible that the negative valance of the items distracted participants, resulting in a noticeable subdimension in the instrument. A social desirability threat to validity could also have an effect on the responses.

The third research question, do the ten intrinsic motivation items reliably separate the participants into varying degrees of intrinsic motivation, did not yield conclusive results. Even though the intrinsic items were adapted from previous intrinsic motivation research, there was not enough spread to successfully separate the participants into varying motivation levels. Most of the items were far too easy to endorse by these participants. This is likely due to the participants surveyed in this study. As previously mentioned, they are all international studies students. The majority of them are women and they are all currently preparing to study abroad next year. They are all highly motivated students. More difficult items or a wider variety of participants would provide better spread of person ability and item targeting.

Research question four dealt with intrinsic items' unidimensionality. The results of the PCA showed that the intrinsic items were unidimensional with the exception of one or two items which may not be tapping the intrinsic motivation construct. Deleting Item 7 and Item 4 from the intrinsic motivation instrument reduced item reliability to .88 and did not change the explained variance in the PCA. Leaving those items in the questionnaire does not seem to have a detrimental effect on unidimensionality or reliability.

Addressing the fifth research question, the test of relationship between grit and intrinsic motivation resulted in a significant positive correlation of medium effect (r = .40). Both grit and motivation have an aspects of perseverance, passion, and stick-to-itiveness. It seemed natural at the outset of this investigation that these two constructs would be significantly correlated. As the results showed, grit and intrinsic motivation are highly correlated with a disattenuated Pearson correlation of r = .59. There are generalizability concerns with this study and as such the results should be interpreted carefully. The participants in this study might not be representative of the population. Further research is required.

Lastly, do these construct measurements correlate to a real-world, autonomous language learning measurement? When comparing the person grit and intrinsic logit measurements to English Central usage statistics, there did not seem to be any significant relationship. Increasing person reliability, as discussed above, would rectify this discrepancy. It would allow for a more direct comparison without the need of disattenuation. Alternatively, it is also possible that English Central homework counts are not a valid proxy measurement of grit or motivation. This is a required homework assignment. I had hoped that the wide range in achievement (see min, max, and mean in Table 7) for the English Central data would provide enough variance in the data to overcome the fact that these tasks were required by the curriculum and not completed out of desire to study only.

Limitations

There are many limitations to the current pilot study. First, the participant sample and size was a matter of convenience. If the instrument were to be given to larger sample with varying degrees of motivation and grit, the results might vary. The lack of representation of males in the study is also an issue to be addressed. Additionally, these students are all international studies majors. They all intend to study abroad next year and many have already had study abroad experiences. These experiences may color the participants' motivation responses.

Finally, the grit questionnaire contained items with negative valance. Research and experience tends to agree that people do not want to answer negatively about themselves. Kramer, McLean, and Martin (2017) found that even when translated into Japanese, these items were problematic. The current study used only English items. Future grit research cannot continue to use these psychometrically awkward items.

Conclusion

In conclusion, the current study explored the relationship of language learning grit and intrinsic motivation to learn a language while at the same time validating two questionnaires. The lack of research on grit in the second language learning makes this exploration worthwhile. However, the similarity of the grit construct to other more well-established personality constructs like conscientiousness seem to provide doubt regarding the value of the construct at all (Credé, Tynan, & Harms, 2017; Credé, 2018; Rimfeld, Kovas, Dale, & Plomin, 2016). The conclusion of this pilot study—grit and intrinsic motivation are moderately correlated—suggest that with improvements in methodology, item design, and targeting, the grit construct could be a valuable addition to individual differences toolbox for L2 researchers.

References

Apple, M. T. (2013). Using Rasch analysis to create and evaluate a measurement instrument for foreign language classroom speaking anxiety. *JALT Journal*, *35*(1), 5-28.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York, NY: Routledge.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences* (2nd ed.). New York, NY: Lawrence Erlbaum Associates.

Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology, 113*(3), 492-511. DOI:10.1037/pspp0000102

Credé, M. (2018). What shall we do about grit? A critical review of what we know and what we don't know. *Educational Researcher, 47*(9), 606-611. DOI:10.3102/0013189X18801322

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

DeKeyser, R. (2007). Introduction: Situating the concept of practice. In: DeKeyser, R. (ed.) *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge, UK: Cambridge University Press.

Duckworth, A. L., & Eskrels-Winkler, L. (2013). True Grit. *Observer, 26*(4). Retrieved from https://www.psychologicalscience.org/observer/true-grit

Duckworth, A. L., Peterson, C., Matthews, M., & Kelly, D. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087-1101. DOI: 10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment, 91*(*2*), 166-174. DOI:10.1080/00223890802634290

English Central. (n.d.). About English Central. Retrieved from https://www.englishcentral.com/static/corporate/section/

Kramer, B., McLean, S., & Martin, E. S. (2017). Student grittiness: A pilot study investigating scholarly persistence in EFL classrooms. *大阪女学院短期大学紀要第, 47,* 25-41.

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R* (2nd ed.). New York, NY: Routledge.

Linacre, J. M. (2018). WINSTEPS (Version 4.0). Retrieved from http://www.winsteps.com/

Linacre, J. M. (2013). *A user's guide to WINSTEPS*. Chicago, IL: MESA. Retrieved from http://www.winsteps.com/

Noels, Clément, and Pelletier (2001). Intrinsic, extrinsic, and integrative orientations of French Canadian learners of English. *The Canadian Modern Language Review, 53*(3), 424-442. DOI: 10.3138/cmlr.57.3.424

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912. DOI:10.1111/lang.12079

Rimfeld, K., Kovas, Y., Dale, P. S., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of personality and social psychology*, *111*(5), 780-789. DOI: 10.1037/pspp0000089

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. DOI:10.1037/0003-066X.55.1.68

Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, *9*, 367–373. DOI:10.1177/014662168500900405

Statistica Help (n.d). *Reliability and item analysis introductory overview – Correction for attenuation.* Retrieved from http://documentation.statsoft.com/STATISTICAHelp.aspx?path=Reliability/ReliabilityandItemAnalysis/Overview/ReliabilityandItemAnalysisIntroductoryOverviewCorrectionforAttenuation

Swain, S. D., Weathers, D., & Niedrich, R. W. (2008). Assessing three sources of misresponse to reversed Likert items. *Journal of Marketing Research*, *45*, 116–131. DOI:10.1509/jmkr.45.1.116

Appendix A: Construct Maps

Construct Map:

Grit for Autonomous Language Learning

High Grit	+1	• (R) When I have a language learning goal, I often choose to follow a different goal later. • (R) I get disappointed and give up when I am unsuccessful. • When I am studying a language, I cannot be distracted from my task. • (R) New assignments or projects in my language class distract me. from older assignments or projects. • I always accomplish my language learning goals • I am a hard worker • I often do more than is required when I am studying a language • (R)When I am studying a language, I lose interest quickly.
Moderate Grit	0	• Making mistakes in another language encourages me to study harder
Low Grit	-1	

Note. (R) means the valance of the items is reversed. These codes were reversed at data entry.

Construct Map:

Intrinsic Motivation for Autonomous Language Learning

High Intrinsic	+1	• I study other languages in order to understand the culture better. • I enjoy the challenge of trying to learn another language. • I get pleasure from using another language. • I enjoy learning new things in another language. • I enjoy interacting with people in other languages.
Moderate Intrinsic	0	• I feel satisfied when I complete challenging activities in a foreign language. • I study languages to improve myself. • I feel excited when I can use something I have learned recently. • I want to be someone who can speak more than one language.
Low Intrinsic	-1	

Appendix B: Test Specifications

Test Specification Table	
Time allowed	As much time as needed; approximately 10 minutes
Test Delivery	Google Forms; Delivered via link on participants' learning management system
Skill Focus	Self-reflection; ability to read statements and discern the degree with which the statements describe the participants
Task Description	Read a short statement or situation and decide how much the situation describes their own opinion
Language level	Statement language should contain frequent words that do not require a dictionary; language should fall within the first 2k vocabulary band of English on the BNC
Expected response	Exactly like me (5); Mostly like me (4); Somewhat like me (3); Not much like me (2); Not like me at all (1)
Discourse purpose	To discover the participants' level of grit and intrinsic motivation for language learning.
Scoring Parameters	Likert-like; degree to which the statement describes their opinion; Some items require score reversal. Scores will be reversed upon data entry.
Instructions to candidates	Here are a number of statements that may or may not apply to you concerning your language learning experiences. Think about how you compare to most language learners in the world, not just your friends and classmates. There are no wrong answers, just answer honestly.
Guiding language	• "I" Statements are acceptable. • Avoid verbs with negative connotations when possible • Reverse coded questions should be careful not to imply negatively or bias the students against choosing the item.. No one wants to say that "I give up easily" is very much like themselves • Avoid passive and complicated grammatical structures, L2 learners may not be able to understand them. • Do not include the scores for the items next to the choices, this may bias the participants' choices.
Additional details	Participants can provide additional comments and questions regarding the questionnaire in English or Japanese. Students may also sign up to participate in a follow-up interview regarding their responses