Examining the Kambara Locus of Control Scale

Michael James RUPP
Kumamoto University

Abstract

This measurement study reports on the results of statistical analyses of scores generated by the Kambara (1982, 1987) Locus of Control Scale (K-LoCS). The scale has two versions: the short form, K-LoC18 (18 items, 1982) and the long form, K-LoC43 (43 items, 1987). A data set (N=1125) was collected from Japanese high school students, and four models were tested a priori; two were the originally hypothesized models for the short and long forms, and two were based on EFA derived models present in the literature. The purpose of this study was to evaluate the plausibility of these models for an instrument which has been widely used in the Japanese research context; thus far in the absence of such a priori testing. The background rationale for this is also that measuring locus of control might serve as a viable proxy for measuring learner autonomy, given that the two are notionally related. The results of this study indicated that all models represented unsatisfactory fit to the dimensionality of the data. This has implications for past and future research, which relies upon structurally valid measurement, and also has implications for the practitioner.

Keywords: Locus of Control, Learner Autonomy, EFL, Second Language Acquisition

1.0 Introduction

This psychometric study is part of a larger effort, using both quantitative and qualitative means, towards providing an alternative approach to the measurement of learner autonomy. Previous attempts (Horai, 2013b) at direct measurement of learner autonomy have not yielded satisfactory results, thus prompting the author to pursue an alternative approach involving the use of constructs which are notionally related to learner autonomy such as locus of control, self-efficacy and personal responsibility; with this study focusing on locus of control. A similar approach has been taken by Tournat (2014; p. 217) who had some success examining the psychometrics of an adapted version of the Causal Dimension II Scale which uses constructs from attribution theory.

2. 0 Literature Review

Learner autonomy has attracted increasing interest in the field of applied linguistics over the last few decades (e.g. Holec, 1981; Little, 1991; Little & Dam, 1998; Oxford, 2003). According to Palfreyman (2003, p. 1) an early use of the word occurs in 1981, by Holec, who described it as “the ability to take charge of one’s own learning.” In the context of language education, it is also sometimes referred to as learner independence, as the learner’s ability to be responsible for his or her learning, or self-direction. Oxford (2003, p. 75) notes that the theoretical framework for learner autonomy is “far from coherent” and is “beset by conflicting ideologies, roiling inconsistencies, and fragmentary theories.”

A model of learner autonomy recently advanced by Tassinari (2012) characterized learner autonomy as comprising three dynamic dimensions: (a) the action dimension, (b) the cognitive and metacognitive dimension, and (c) the affective and motivational dimension. The action dimension refers to activities such as planning, material and method choice, completion of tasks, monitoring, evaluating, cooperating and management of learning. The cognitive and metacognitive dimension refers to the structuring of knowledge. Finally, the affective and motivational dimension refers to feelings and self-motivation. All three of these dimensions include a social dimension integrated into them. Tassinari points out that the distinctions between the dimensions are merely theoretical and that the various aspects are closely interrelated and thus can have an influence on each other.

Oxford (2003), somewhat earlier, attempted to develop a more systematic and comprehensive theoretical model in which autonomy can be viewed through technical, psychological, sociological, and political-critical prisms. Each of these prisms are further divided into the four themes of context, agency, motivation, and learning strategies. In the technical view of autonomy the focus is on the physical situation, such as the varieties of independent learning that can be afforded by self-access centers. In the psychological view the focus is on a combination of attitudes, abilities, strategies and styles. From the sociocultural view, the emphasis is shifted to a Vygotskian developmental attainment of autonomy which is mediated through social interaction, or through participation in communities of practice. Finally, in the political-critical view, the focus is on ideologies, access and power structures (Oxford, 2003).

In terms of the four themes that these viewpoints are further subdivided into, context of autonomy refers to looking at the entire situation, background, and environment of second language learning; agency refers to the capacity to act intentionally; motivation is that which provides the impetus or goals to the agent; and strategies are the plans which lead the agent towards the goals. In the study reported here, the author is focused on the psychological perspective of autonomy, in the context of EFL in Japanese secondary education. This context has a history of promoting high extrinsic motivation due to English education in Japan being primarily focused on examination results, creating a large population of students who could be said to be gaining little more than knowledge about English (e.g. grammar rules) rather than acquiring practical skills in the use of the language (e.g. possessing fluency or listening ability). The second theme, agency, as seen from the psychological view, can be understood as a characteristic of the individual, as well as the starting point of autonomy (Horai, 2013b). The third theme of motivation, while being viewed as a relatively stable characteristic, does include one changeable aspect in that self-efficacy can be modified through strategy instruction.

According to Oxford (2003), psychological research indicates autonomous learners as having high a high degree of self-efficacy (representing agency), positive attitudes, a need for achievement, a desire to seek meaning and lastly motivation; both extrinsic and intrinsic. Ryan and Deci (2000) note that although, traditionally, intrinsic motivation has been seen as resulting in higher quality learning, with extrinsic motivation being viewed as a more impoverished form of motivation, their Self-Determination Theory proposes that there are some forms of extrinsic motivation which represent active and agentic states. Oxford also points out that motivation to direct one's own learning can be related to learning styles, which in turn can be associated with a particular cultural context, for example with some cultures having learners who prefer to have an authority figure guiding them.

Smith (2003) echoes this sentiment, noting the critiques of inappropriate transfer of Western approaches, while at the same time lamenting the lack of discussion of appropriate teaching methodologies due to avoidance of making a priori generalizations. He argues for a different view wherein autonomy is seen as being possessed by learners in varying degrees ranging from weak to strong, thus requiring different approaches and goals. While this appears reasonable, it could also be criticized for being obvious due to the inherent nature of most psychological constructs representing dispositions of degree. It becomes apparent from this wide variety of perspectives that learner autonomy is a widely contested theoretical concept with no small amount of ambiguity in its various definitions. Smith (2003) conducted long-term action research involving case study of an implementation of the strong version of autonomy in the Japanese university context, and the success of this implementation led him to suggest that looking at regional or national characteristics is not a fruitful endeavor; noting that other factors such as institutional constraints, previous learning experiences and the degree of control offered to students were more fertile areas in which to search for explanations about appropriate pedagogies for autonomy.

Although Japan's economy is highly dependent on international trade, the Japanese people are often perceived as being poor at English communication; and indeed, this is often a self-perception as well. Recently, and against the background of this perception and self-perception, the importance of learner autonomy in the Japanese context has been emphasized in policies published by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT). This emphasis represents part of an effort to improve communicative competence, which includes the goal of "fostering of autonomy toward lifelong learning" (MEXT, 2003). Autonomy is seen as a necessary trait to help Japanese people become more adaptive in an international linguistic environment.

In spite of this importance, there has been little solid empirical research clearly showing the predictive power of learner autonomy on learning outcomes. In order to address this issue, Horai (2013a; 2013b) attempted to adapt three different autonomy instruments into the Japanese context in order to establish a psychometric foundation so that research can proceed with demonstrating the predictive power, or lack thereof, of learner autonomy on learning outcomes. Establishing such a psychometric foundation is a necessary first step to the pursuant questions concerning the predictive power of the independent variable (autonomy) on the dependent variable (learning outcome), due to the fact that lacking a secure measurement foundation any findings related to the pursuant questions are threatened by charges of invalid measurement. Horai found that with respect to each of the instruments she evaluated, the CFA results showed that the models hypothesized by the authors (as well as other models posited in other literature via post hoc EFAs) were not satisfactory and did not fit the dimensionality of the data produced by the instruments. Overall, this amplifies the deficit with regard to measuring the construct of autonomy – a construct that has generated volumes of discussion and debate but which conspicuously has not been properly measured to date.

Given this persistent failure with respect to a direct approach to measuring autonomy, the author proposes an alternative solution to the measurement problem, along similar lines to Tournat (2014), through the investigation of instruments designed to measure constructs which are notionally or theoretically related to learner autonomy. The constructs under consideration include: (a) locus of control (LoC), (b) self-efficacy, and (c) personal responsibility. Tournat also followed this approach of exploring proxy constructs for autonomy, but focused singularly on attribution theory (four instruments in a series of studies). Exploring this research rationale, chosen from among the three constructs mentioned above, this study focuses on the first construct, locus of control, and investigates an instrument which has been widely used in the Japanese secondary educational context, including recently in studies on English achievement. The author of the instrument, Kambara (1987) also confirmed (Personal Communication, May 2014) that the instrument would easily be adaptable to the tertiary context as well, although such adaptation is the subject of later research, and is not reported in this study.

Locus of control (LoC) is a psychological construct that originated from Rotter, having its roots in his Social Learning Theory (1954). He later developed the first LoC scale in 1966 (Rotter, 1966). Locus is the Latin term meaning place, and a person's locus of control can be described as being placed on a continuum ranging from external to internal. LoC refers to the extent to which someone believes they are in control of their lives, and was conceived of as having the sub-dimensions of internal locus of control (belief that one is in control of outcomes in one’s own life) and external locus of control (belief that outcomes in one’s life are externally-controlled). In the case of external locus of control, there are two further sub-dimensions: (a) random events or fate as the external source of control, and (b) powerful others as the external source of control. LoC, according to Neill (2006), is also conceptualized as referring to a unidimensional continuum, ranging from external (E) to internal (I).

Students in EFL curriculums scoring higher on the internal side of LoC have been shown to be better able to take control of their own learning which is a key component of learner autonomy (Ghonsooly, 2010). Though having a high degree of internal locus of control is generally seen as being desirable, there are cases where it can be too much of a good thing, such as when being overly internal can have the adverse consequences of a tendency to excessively blame oneself for failure or to become neurotic. Likewise, although external orientation is usually seen as too passive and fatalistic for high levels of success, highly external students can sometimes benefit from having a happy-go-lucky, easygoing attitude towards life.

Rotter (1975, 1990) emphasized that locus of control is not a binary typology, but rather represents points on a continuum; i.e. it is not an either/or proposition. Although locus of control is a generalized expectancy, and predicts behavior across situations, there may be domain specific situations in which people switch from behaving like externals to behaving like internals, such as domains in which the person has notable skills and experience, or a lack thereof. Some of the domains in which LoC scales are used include: (a) health, i.e. to predict patient outcomes through protocol compliance, (b) business, i.e. for employee evaluations, (c) education, i.e. for student and teacher evaluations, and (d) psychology, i.e. for the psychiatric evaluation of prisoners, parents, and children.

The LoC theory is one a number of theories, including attribution theory (e.g. Heider, 1958; Kelley, 1967; McLeod, 2010), self-efficacy theory (Bandura, 1977), and self-determination theory (Deci & Ryan, 1985), which are notionally related to the construct of learner autonomy (Duttweiler, 1984). Deci and Ryan (1987), in particular, have linked self-determined behaviors with other autonomy-supportive behaviors which are associated with increased intrinsic motivation, interest, and learning.

In the Japanese literature, the most prominent LoC scale that has emerged was a scale created by Kambara (1982, 1987). This scale was developed for investigating the LoC levels of students in the Japanese high school context. It was originally an 18-item questionnaire that was later expanded to 43 items in order to more specifically measure LoC in the context of secondary education. It is this expanded version, the Kambara 43-Item LoC Scale (K-LoC43) (see Appendix), which has been predominantly used for the last 30 years in Japan; with studies using it appearing in a wide variety of domains, ranging from developmental psychology (e.g. Kanda, 2006), educational studies (e.g. Hosaka, 2007; Kambara, 1987), to studies about employee psychological distress (e.g. Fushimi, 2011). The significant presence of this instrument in the literature made it a good candidate for investigation. None of the studies above have reported results for a CFA on the scores generated by the instrument, and only EFA results have been reported. This is important because EFA is a post hoc form of analysis and cannot confirm a hypothesized measurement model for the instrument a priori. Notably, the reported EFAs all produced three factors rather than the expected two which, consistent with theory, would be External (E-LoC) and Internal (I-LoC); a result which would seem to indicate a problem with the scale, though this has not been directly addressed in the literature.

The study analyzes four models for Kambara's instrument. The first two models are informed by Kambara's a priori conception for the LoC in the case of both the original 18-item instrument (Short Form; Kambara, 1982), Model A, and the 43-item instrument (Long Form; Kambara, 1987), Model B. Both of these models comprise two factors labeled E-LoC and I-LoC, consistent with theory and conception. Model A has 9 items for E-LoC and 9 items for I-LoC. Model B has 22 items for E-LoC and 21 items for I-LoC. The next two models are informed by EFAs conducted by Kambara (Long Form; 1987), Model C, and Hosaka (Long Form; 2007), Model D. Model C has 43 items and comprises three factors derived by Kambara in an EFA: Factor 1, Factor 2 and Factor 3 (these factors were not interpretively labeled in the Kambara [1987] study). Factor 1 has 18 items, Factor 2 has 15 items and Factor 3 has 10 items. Model D has 13 items and also comprises three factors obtained from an EFA conducted by Hosaka. Hosaka labeled Factor 1 as Effort (5 items), Factor 2 as Contingency (4 items) and Factor 3 as Environment (4 items). Models C and D are notable in that they are no longer aligned with the original theory where only two factors are hypothesized, namely, E-LoC and I-LoC. Furthermore, Model D represents a significant abbreviation of the Long Form to only 13 items which is actually even shorter than the original Short Form.

3.0 Methodology

The methodology is reported in terms of the instrument under study, the participants and data collection procedure, and the analytical procedure.

3.1 Instrument

The Kambara Locus of Control Scale (Long Form or KLoC-43) comprises 43 statements (see Appendix) for which students answer on a 4-point Likert scale which ranges from strongly disagree to strongly agree. The scale was originally created in Japanese by Professor Kambara, a native-speaker of Japanese, and thus there was no need to alter the language of the original scale. As the scale was originally created for Japanese high school students it was decided to collect the sample from this population. The Short Form comprises the first 18 items from the Long Form.

3.2 Participants and Procedure

There were 1223 total participants in this study with 98 responses removed for having missing data. This data was missing at random and therefore removing these cases did not systematically affect the sample. This process left 1125 usable responses (N=1125). The participants were Japanese high school students of both sexes with 57% male and 43% female respondents. The percentage of responses according to high school grade was 56% for first year, 10% for second year and 34% for third year, with a mean age of 16.44 years. The data was collected from two private high schools and one public high school. Participation was voluntary and had no effect on the students' grades. The consent form was printed at the top of each questionnaire which clearly stated in Japanese that those who did not wish to consent could freely do so by merely not completing the questionnaire. Permission to administer the questionnaires was given by each of the high school's principals and the questionnaires were anonymous. The time required for administration was approximately 15 minutes.

3.3 Analytical Procedure

The data collected from the students was stored in a Microsoft Office Access 2007 database. Descriptive statistics and reliability estimates (Cronbach's alphas) were calculated using the IBM/Statistical package for the Social Sciences (SPSS) software Version 16.0. The CFA was performed using AMOS 18.00. The data was first considered from the point of view of descriptive statistics focusing on univariate normality (i.e. skew and kurtosis). Following this, reliability estimates were calculated before performing a CFA of the hypothesized model in each case (i.e. for all the models: Model A, Model B, Model C, and Model D).

The method for evaluating skew and kurtosis was to calculate the critical ratio by dividing the value for skew and value for kurtosis by the respective standard error in the case of each of the 43 items. In order to evaluate the skew and kurtosis, the author stipulated, in advance, a minimum evaluation criterion of 3.0, as well as a stricter criterion of 2.0 to identify meritorious results.

With respect to Cronbach's alpha, the confidence intervals (95%) were also calculated in accordance with the recommendations of Fan and Thompson (2001), with the cut off threshold set at .70 based on Nunnally and Bernstein's (1994) criterion. It is important to note however that these estimates are not as important as the results for the CFA which is more powerful than alpha as a diagnostic for valid measurement. CFA is particularly important in terms of demonstrating the unidimensionality (Gerbing & Anderson, 1988) of the sub-scales for the instrument, in the case of each model.

4.0 Results

Results are initially reported for the descriptive statistics for scores generated by the K-LoC43, and in terms of each item making up the scale. Following this, the results for Cronbach's alpha and the CFAs are reported by Model (i.e. for Model A, Model B, Model C and Model D).

4.1 Descriptive Statistics

Table 1 indicates the means and standard deviations for each item, as well as the value for skew and kurtosis in each case, with the associated standard errors. These were used to calculate the critical ratio for skew and kurtosis for each item, which can be inspected in Table 2.

Table 1
Item Means, Standard Deviations (SD), Skew and Kurtosis for Scores Derived on Items Comprising the K-LoC43 (N=1125)

Test Items	M	SD	Skew		Kurtosis	
			Statistic	SE	Statistic	SE
Item 01	2.23	0.896	0.16	0.073	-0.835	0.146
Item 02	3.09	0.909	-0.804	0.073	-0.145	0.146
Item 03	2.4	0.906	0.023	0.073	-0.81	0.146
Item 04	2.94	0.817	-0.417	0.073	-0.357	0.146
Item 05	2.3	0.929	0.24	0.073	-0.797	0.146
Item 06	2.24	0.947	0.261	0.073	-0.872	0.146
Item 07	2.7	0.876	-0.327	0.073	-0.54	0.146
Item 08	2.34	0.965	0.22	0.073	-0.909	0.146
Item 09	2.42	1.029	0.112	0.073	-1.127	0.146
Item 10	3.34	0.79	-1.176	0.073	1.024	0.146
Item 11	2.52	0.945	-0.095	0.073	-0.897	0.146
Item 12	2.59	0.833	0.024	0.073	-0.601	0.146
Item 13	3.05	0.872	-0.703	0.073	-0.15	0.146
Item 14	2.17	0.937	0.403	0.073	-0.72	0.146
Item 15	2.58	0.856	-0.088	0.073	-0.623	0.146
Item 16	2.28	0.804	0.434	0.073	-0.164	0.146

Item 17	2.5	1.037	-0.018	0.073	-1.161	0.146
Item 18	1.94	0.838	0.695	0.073	-0.009	0.146
Item 19	2.97	0.826	-0.566	0.073	-0.112	0.146
Item 20	2.53	0.927	-0.085	0.073	-0.843	0.146
Item 21	2.84	0.963	-0.445	0.073	-0.752	0.146
Item 22	2.92	0.881	-0.533	0.073	-0.374	0.146
Item 23	2.73	0.929	-0.208	0.073	-0.843	0.146
Item 24	2.96	0.853	-0.593	0.073	-0.168	0.146
Item 25	2.8	0.909	-0.239	0.073	-0.815	0.146
Item 26	1.85	0.834	0.85	0.073	0.257	0.146
Item 27	2.26	0.898	0.252	0.073	-0.704	0.146
Item 28	2.46	0.832	-0.035	0.073	-0.571	0.146
Item 29	2.59	0.938	-0.108	0.073	-0.872	0.146
Item 30	2.7	0.782	-0.06	0.073	-0.477	0.146
Item 31	2.32	0.931	0.13	0.073	-0.88	0.146
Item 32	2.99	0.859	-0.567	0.073	-0.305	0.146
Item 33	2.19	0.953	0.377	0.073	-0.791	0.146
Item 34	2.86	0.893	-0.444	0.073	-0.526	0.146
Item 35	2.79	0.858	-0.31	0.073	-0.534	0.146
Item 36	2.01	0.861	0.618	0.073	-0.194	0.146
Item 37	2.57	0.871	-0.12	0.073	-0.657	0.146
Item 38	3.42	0.79	-1.383	0.073	1.444	0.146
Item 39	3.3	0.814	-1.112	0.073	0.769	0.146
Item 40	2.33	0.77	0.36	0.073	-0.148	0.146
Item 41	2.75	0.963	-0.263	0.073	-0.913	0.146
Item 42	2.98	0.895	-0.615	0.073	-0.345	0.146
Item 43	2.81	0.834	-0.346	0.073	-0.402	0.146

As noted above, Table 2 indicates the results for skew and kurtosis for each item. The asterisks indicate whether the absolute value of the critical ratio fails to meet the more relaxed (3.0, two asterisks) or the stricter (2.0, one asterisk) criterion; and thus the sign for the critical ratio in the case of each item is not important.

Table 2
Calculated Values for Critical Ratio for Skew and Kurtosis

Test Items	Calculated Values		Test Items	Calculated Values	
	Skewness	Kurtosis		Skewness	Kurtosis
Item 01	*2.19	**-5.73	**Item 23**	*-2.843	**-5.785
Item 02	**-10.995	-1.025	**Item 24**	**-8.113	-1.184
Item 03	0.309	**-5.561	**Item 25**	**-3.264	**-5.591
Item 04	**-5.706	*-2.47	**Item 26**	**11.624	1.712
Item 05	**3.283	**-5.471	**Item 27**	**3.449	**-4.835
Item 06	**3.569	**-5.982	**Item 28**	-0.477	**-3.926
Item 07	**-4.465	**-3.718	**Item 29**	-1.474	**-5.978
Item 08	**3.009	**-6.233	**Item 30**	-0.825	**-3.286
Item 09	1.527	**-7.717	**Item 31**	1.78	**-6.036
Item 10	**-16.077	**6.944	**Item 32**	**-7.757	*-2.113
Item 11	-1.299	**-6.147	**Item 33**	**5.157	**-5.427
Item 12	0.329	**-4.132	**Item 34**	**-6.069	**-3.62
Item 13	**-9.612	-1.06	**Item 35**	**-4.235	**-3.679
Item 14	**5.516	**-4.942	**Item 36**	**8.456	-1.362
Item 15	-1.202	**-4.282	**Item 37**	-1.643	**-4.516
Item 16	**5.935	-1.156	**Item 38**	**-18.918	**9.807
Item 17	-0.245	**-7.948	**Item 39**	**-15.212	**5.202
Item 18	**9.506	-0.099	**Item 40**	**4.917	-1.042
Item 19	**-7.746	-0.803	**Item 41**	-3.6	**-6.263
Item 20	-1.167	**-5.782	**Item 42**	**-8.408	*-2.39
Item 21	**-6.08	**-5.164	**Item 43**	**-4.727	*-2.774
Item 22	**-7.293	*-2.587			

Note. *Test item is skewed at a threshold of 2.0. ** Test item is skewed at a threshold of 3.0.

As can be seen in Table 2 above, with respect to skew, 13 items (30.2%) fell below the 2.0 threshold, with 2 items (4.6%) meeting the 3.0 threshold and the remaining 28 (65.1%) items failing to meet the 3.0 threshold. The calculated values for kurtosis are 9 items (20.9%) meeting the 2.0 threshold, 5 items (11.6%) meeting the 3.0 threshold and the remaining 29 items (67.4%) failing to meet the 3.0 threshold. It should be noted, that this scale is very coarse for this kind of analysis, having only four points of discrimination, and the results should be critically understood in those terms.

4.2 Results by Model

The models reported below were all tested using the same analytical procedure. A combination of the chi-square and a range of model fit indexes were used; with these indexes, and their associated criteria or cutoffs, being offered by Hu and Bentler (1999). The chi-square has the problem of over-rejecting models and that is why the other indexes are also used.

4.2.1 Model A

Model A is derived from the original conception for the short form of the instrument (K-LoC18). The short form, (Kambara, 1982) includes the first 18 items from the long form. Items 19 through 43 were added to the instrument later to create the K-LoC43 (Kambara, 1987). The original conception for the instrument was that there were two constructs, namely, I-LoC and E-LoC, and different items were hypothesized to measure each of these two constructs as indicated in the 1982 article. The items measuring I-LoC included Items 2, 3, 4 10, 11, 12, 13, 14, and 17, and the items measuring E-LoC included Items 1, 5, 6, 7, 8, 9, 15, 16, and 18. Multivariate non-normality was assessed using Mardia's coefficient (41.19), and was found to be high. In terms of Cronbach's alpha, the results were as follows (95% confidence intervals in brackets): I-LoC, .74 (.72-.76) and E-LoC, .69 (.66-.72).

These items were tested using CFA in a two-factor correlated model. The model had 171 distinct sample moments, 37 distinct parameters to be estimated, and therefore 134 degrees of freedom. This met the criterion of overidentification. The chi-square value was 967.20 with an associated probability level of < .01. Thus, according to the chi-square test statistic the model should be rejected. The results for the model indexes were as follows (with the cutoffs provided by Hu and Bentler [1999] next to each result in parentheses): TLI, .71 (>.95); CFI .75 (>.95); RMSEA .07 (<.06); and SRMSR .07 (<.08). These results, overall, indicate that there is insufficient fit of the data to the model, and therefore the model has to be rejected.

4.2.2 Model B

Model B is the long form of the instrument (K-LoC43), Kambara (1987), includes 43 items and is the version predominantly used in the Japanese context. The first 18 items are identical to the short form analyzed in Model A. The expanded form, or long form with its 43 items, has a total of 21 items to measure I-LoC and 22 items to measure E-LoC. The items measuring I-LoC included Items 2, 3, 4 10, 11, 12, 13, 14, 17, 19, 21, 23, 24, 29, 30, 34, 35, 38, 39, 40, and 42, and the items measuring E-LoC included Items 1, 5, 6, 7, 8, 9, 15, 16, 18, 20, 22, 25, 26, 27, 28, 31, 32, 33, 36, 37, 41 and 43. Multivariate non-normality was assessed using Mardia's coefficient (118.52), and was found to be high. In terms of Cronbach's alpha, the results were as follows (95% confidence intervals in brackets): I-LoC, .83 (.82-.85) and E-LoC, .77 (.75-.79).

These items were tested using CFA in a two-factor correlated model. The model had 946 distinct sample moments, 87 distinct parameters to be estimated, and therefore 859 degrees of freedom. This met the criterion of overidentification. The chi-square value was 4344.26 with an associated probability level of < .01. Thus, according to the chi-square test statistic the model should be rejected. The results for the model indexes were as follows (with the cutoffs provided by Hu and Bentler [1999] next to each result in parentheses): TLI, .60 (>.95); CFI .62 (>.95); RMSEA .06 (<.06); and SRMSR .07 (<.08). These results, overall, indicate that there is insufficient fit of the data to the model, and therefore the model has to be rejected.

4.2.3 Model C

Model C is derived from Kambara's (1987) EFA, using data collected with the K-LoC43, which produced 3 factors: Factor 1, Factor 2 and Factor 3. Factor 1 has 18 items, Factor 2 has 15 items and Factor 3 has 10 items. The items measuring Factor 1 included Items 2, 3, 8, 10, 11, 13, 17, 18, 19, 21, 23, 24, 29, 33, 34, 38, 39, and 42. The items measuring Factor 2 included Items 1, 5, 6, 7, 9, 15, 16, 20, 22, 25, 26, 27, 31, 36 and 41. The items measuring Factor 3 included Items 4, 12, 14, 28, 30, 32, 35, 37, 40 and 43. Multivariate non-normality was assessed using Mardia's coefficient (118.52), and was found to be high. In terms of Cronbach's alpha, the results were as follows (95% confidence intervals in brackets): Factor 1, .73 (.70-.75); Factor 2, .73 (.70-.75); and Factor 3, .35 (.29-.41).

These items were tested using CFA in a three-factor correlated model. The model had 946 distinct sample moments, 89 distinct parameters to be estimated, and therefore 857degrees of freedom. This met the criterion of overidentification. The chi-square value was 4425.00 with an associated probability level of < .01. Thus, according to the chi-square test statistic the model should be rejected. The results for the model indexes were as follows (with the cutoffs provided by Hu and Bentler [1999] next to each result in parentheses): TLI, .59 (>.95); CFI .61 (>.95); RMSEA .06 (<.06); and SRMSR .08 (<.08). These results, overall, indicate that there is insufficient fit of the data to the model, and therefore the model has to be rejected.

4.2.4 Model D

Model D is derived from Hosaka's (2007) EFA, using data collected with the K-LoC43, and is comprised of 3 factors and 13 items. Factor 1 (LoC 1, Effort), Factor 2 (LoC 2, Contingency) and Factor 3 (LoC 3, Environment) were the labels given to the three factors. Effort has 5 items, Contingency has 4 items, and Environment has 4 items. The items measuring Effort included Items 2, 3, 11, 17 and 21. The items measuring Contingency included Items 1, 9, 26 and 38. The items measuring Environment included Items 7, 20, 27 and 41. Multivariate non-normality was assessed using Mardia's coefficient (26.54), and was found to be high. In terms of Cronbach's alpha, the results were as follows (95% confidence intervals in brackets): Effort, .67 (.63-.70); Contingency, .22 (.14-.29); and Environment, .47 (.41-.51).

These items were tested using CFA in a three-factor correlated model. The model had 91 distinct sample moments, 29 distinct parameters to be estimated, and therefore 62 degrees of freedom. This met the criterion of overidentification. The chi-square value was 468 with an associated probability level of < .01. Thus, according to the chi-square test statistic the model should be rejected. The results for the model indexes were as follows (with the cutoffs provided by Hu and Bentler [1999] next to each result in parentheses): TLI, .72 (>.95); CFI .78 (>.95); RMSEA .08 (<.06); and SRMSR .07 (<.08). These results, overall, indicate that there is insufficient fit of the data to the model, and therefore the model has to be rejected.

4.3 Summary of Reliability Estimates

In order to provide a clearer view of the reliability estimates, they have been summarized below in Table 3. It is important to keep in mind that the first two models (A and B) are conceptions informed by a priori theory for the structure of the LoC construct, whereas the final two models (C and D) are from post hoc EFAs which deviate from the normal two-construct model. Thus, even though Models C and D are being tested for reliability, were they to have been shown to have higher reliability, it would still be a question as to what these models are trying to achieve, as they are already no longer in conformance with the instruments original conception.

Table 3
Reliability Estimates according to Model and Construct

Model	No. of Items	Construct	Cronbach's α	95% Confidence Lower Limit	95% Confidence Upper Limit
A	9	I-LoC	0.74	0.72	0.76
A	9	E-LoC	0.69	0.66	0.72
B	21	I-LoC	0.83	0.82	0.85
B	22	E-LoC	0.77	0.75	0.79
C	13	Factor 1	0.73	0.7	0.75
C	15	Factor 2	0.73	0.7	0.75
C	10	Factor 3	0.35	0.29	0.41
D	5	Effort	0.67	0.63	0.7
D	4	Contingency	0.22	0.14	0.29
D	4	Environment	0.47	0.41	0.51

By inspection of this table, it becomes apparent that models with higher numbers of items tend to have higher alphas as well. Also, with the EFA derived models (C and D) the factors extracted earlier in the process tend to have higher alphas than factors extracted later.

5.0 Discussion

The purpose of this study was to investigate, the K-LoC43, a widely used instrument for assessing LoC in the Japanese secondary education context, in addition to addressing the abbreviated form of the instrument, the K-LoC18. A first crucial step towards the aim of verifying whether the instrument is actually measuring what it purports to measure was taken, and a CFA conducted a priori, via a new set of data, can determine whether the dimensionality of the data produced by the instrument fits with models hypothesized either by the instrument's author under its initial conception or with subsequent models which have emerged through a posteriori EFA analyses. If Kambara's instrument is able to yield structurally valid results then it would be a valuable tool for indirectly measuring learner autonomy, and it would also provide confirmation of the measurement assumptions underpinning significant previous research in the literature. Based on the results of this study, however, all four of the models tested were shown not to produce structurally valid scores.

Descriptive statistics revealed that the data was non-normal for a number of the items. As was seen in Table 2 above, the majority (65.1% for skewness and 67.4% for kurtosis) of items failed to meet even the less stringent threshold of 3.0, which is problematic. It is recognized, however, that the four-point scale is very coarse for this kind of analysis. It is also recognized that this is problematic for the normal theory analyses conducted in this study via Cronbach’s alpha and CFA; although in practice there is significant precedent for relaxation of these assumptions, and of course the EFAs conducted to date have occurred under similar departures from these assumptions. In this regard, changes in the Likert scale from four response points to five or six response points in potential revisions of the instrument would allow for the scale to more closely approximate a continuous scale. This would also allow for non-normal items to be removed or modified with greater precision. Changing the Likert scale to a five-point scale was also an outcome in a focus group study (Rupp, in press) of university students analyzing the K-Loc43 instrument. This qualitative research was being conducted in an effort to adapt the instrument to the tertiary level and also to improve it.

As can be seen in Table 3 above, Cronbach's alpha is above the interpretive threshold of .70 for: Model A, I-LoC; Model B I-Loc and E-LoC; Model C, Factors 1 and Factor 2. It is noteworthy that the two EFA-based, three-factor models, i.e. Model C and Model D, produced results that fell far below the threshold for some of the constructs, which arguably should not be surprising given that the original conception of the instrument only has two constructs, which should express themselves as two unidimensional factors. However, it has been demonstrated that alpha can have a favorable bias when there are a large number of items on a scale (Cortina, 1993; Green, Lissitz, & Muliak, 1977). This can be seen in the upward trend of alpha values when moving from Model A to Model B, which involves a twofold increase in the number of items per construct. Conversely, we also see that in Model D, which has a drastically reduced number of items, that alpha is also seen to be generally lower. Thus, given the shortcomings of this measure of reliability, we are better served by looking at the CFA results to analyze the dimensionality of the data.

When interpreting the CFA results for all the models, it is important to note that the chi-square result was unsatisfactory in all cases. In the case of CFA, if you get a significant result, it means that the model in question is significantly different from the dimensionality of the data, and thus it does not fit, which may seem counterintuitive to non-CFA statistical interpretation. That being the case, there is another problem with the chi-square (Hu & Bentler, 1999) which is that it has a tendency to over-reject models. So there could be cases where the difference between the model and the dimensionality of the data is quite trivial, but the model is rejected due to the oversensitivity of the chi-square statistic. Therefore, although we cannot disregard the chi-square, we need to be aware of this limitation. In the case of this study, the chi-square is negative evidence for the fit of the model, but this result alone is not enough for interpretation of model fit. We need to pay additional attention to the CFA indexes as these are typically used to triangulate the results. When judging the results of the indexes, and in order to be deemed a satisfactory fit, the model must satisfy all of the indexes, rather than just one or two (Hu & Bentler, 1999). Also, it is important to keep in mind that the SRMSR is unique among the indexes in that it is an index of the residuals or what is left over when the model is fitted. Finally, when thinking about the four models, there is an important distinction between the first two (Model A and Model B) and the last two (Model C and Model D), which is that the first two are based on a priori conceptions of the instrument and represent the original two constructs of LoC (I-LoC and E-LoC) whereas the final two models are based on post hoc conceptions derived from EFAs and which have already ventured out of the realm of measuring LoC as it was originally conceived. Nevertheless, these two post hoc models exist in the Japanese literature and it is worthwhile investigating their plausibility.

Looking at Model A, the values for the two incremental indexes, the TLI of .71 (>.95) and the CFI of .75 (>.95), provide strong evidence against the model. In the case of the absolute fit indexes, the RMSEA of .07 (<.06) is just above the threshold, and the SRMSR of .07(<.08) is just below the threshold. However, it is necessary to have a satisfactory result on all of the indexes in order to accept the model, and therefore, overall, and in triangulated interpretation, the model must be rejected.

With respect to Model B, the values for the two incremental indexes provide even stronger evidence against the model with the TLI of .60 (>.95) and the CFI of .62 (>.95) being quite low. However the RMSEA of

.06 (<.06) is just on the threshold and the SRMSR of .07 (<.08) is just below the threshold. As stated above, the principle of triangulation requires that the model be rejected because not all of the indexes are satisfactory. The RMSEA is a little better than in Model A, but this could be because the RMSEA is an index which rewards for model parsimony; and Model B is more parsimonious because more data points are being reduced onto the same two constructs that Model A hypothesizes but with few data points. Put another way, Model B has 43 items loading onto two constructs, whereas Model A only has 18 items loading on to the same two constructs. The length of Model B came under criticism from the university focus groups (Rupp, in press) analyzing this instrument, and it was suggested that there was a great deal of fatigue with repeated or redundant item content.

Moving on to Model C, which as noted above is based on an EFA that was notable for producing an a posteriori model inconsistent with the original conception for the instrument, satisfactory results were also not produced. The value for the TLI was .59 (>.95) and for the CFI .61 (>.95), both of which are poor outcomes, even given the fact that both the RMSEA value of .06 (<.06) and the SRMSR value of .08 (<.08) are both just on the threshold of acceptability. As with the first two models, the results require an interpretation rejecting the model.

Finally an examination of Model D, also an a posteriori, EFA-based, three-factor model (i.e. deviating from the original conception of LoC), shows that the value for the TLI of .75 (>.95) and for the CFI of .78 (>.95) is slightly better than in Model C, while the RMSEA of .08 (<.06) is significantly worse. This could be because there are very few items in this model and there is less data reduction occurring and therefore less model parsimony. The SRMSR of .07 (<.08) is just below the threshold of acceptability. Another aspect of this model which stands out from the others is that it represents an extremely abbreviated version of the original instrument, having only 13 items. This version has discarded 30 out of the original 43 items, which raises questions about why this would have to be the case.

6.0 Conclusion

While the enterprise to look at LoC as a proxy to measure learner autonomy is still worth pursuing, it is clear that this endeavor cannot be pursued with this instrument in its current form and that this instrument requires further revision. One of the starting points would be to change the

Likert scale to a five- or six-point scale in order to provide more refined Likert scales that are amenable to a more sensitive evaluation for normal distribution. This would allow for the better exclusion of skewed and kurtotic items; besides giving respondents a better range over which to express their disposition. Following such changes, the data could then be analyzed for which operationalizations are the most effective for measuring the two constructs for LoC. This would result in the loss of items which, given the results of Rupp (in press) where participants indicated that the instrument was too long, should also have a beneficial effect on the instrument.

References

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98.

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

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*(6), 1024.

Duttweiler, P. C. (1984). The internal control index: A newly developed measure of locus of control. *Educational and Psychological Measurement, 44*(2), 209-221.

Fan, X., & Thompson, B. (2001). Confidence intervals for effect sizes confidence intervals about score reliability coefficients, please: An EPM guidelines editorial. *Educational and Psychological Measurement, 61*(4), 517-531. http://dx.doi.org/10.1177/0013164401614001

Fushimi, M. (2011). The relationship between individual personality traits (internality-externality) and psychological distress in employees in Japan. *Depression Research and Treatment, 2011*. doi:10.1155/2011/731307.

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 186-192.

Ghonsooly, B., & Elahi, M. (2010). Validating locus of control questionnaire and examining its relation to general English (ge) achievement. *The Journal of Teaching Language Skills (JTLS), 2*(1), 117-143.

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*(4), 827-838.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York, NY: Wiley.

Holec, H. (1981). *Autonomy and foreign language learning*. Oxford, England: Pergamon.

Hosaka, Y. (2007). An empirical study on streaming in English class: With a focus on the relationship between LoC and achievement. *Ritsumeikan Studies in Language* *and Culture, 18*(3): 193-200.

Horai, K., (2013a). Measuring learner autonomy: A confirmatory factor analysis of scores on an EFL learner autonomy questionnaire. *Soujou Daigaku Kiyou, 38*, 59-72.

Horai, K., (2013b). Confirmatory factor analyses of measurement models for a Japanese adaptation of a learner autonomy instrument. *International Journal of Social and Cultural Studies, 6*, 57-76.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Kambara, M., Higuchi, K., Kiyomizu, N. (1982) *Locus of control no shakudo no sakusei to shinraisei, datousei no kentou* [Creating and Researching the Reliability and Validity for a Locus of Control Scale]. *Kyouiku Shinrigaku Kenkyuu, 30*(4), 302-307.

Kambara, M. (1987). *Koukousei no locus of control ni kan suru kenkyuu: kitai oyobi gakushuu douki to no kanren* [Locus of Control belief in high school students: On the relation to expectancy and motives for learning]. *Toukyou Daigaku Kyouiku Gakubu Kiyou. 26*, 107-117.

Kanda, N. (2006). Development change of perceived control in children. *Life Science Studies. Bunkyou University*, 28, 35-42, 2006-03.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (ed.), *Nebraska Symposium on Motivation* (Volume 15, pp. 192-238). Lincoln, NE: University of Nebraska Press.

Little, D. (1991). *Learner autonomy 1: Definitions, issues and problems*. Dublin, Ireland: Authentik Language Learning Resources Limited.

Ministry of Education, Culture, Sports, Science and Technology. (2003). *Regarding the Establishment of an Action Plan to Cultivate "Japanese with English Abilities."* Retrieved from http://warp.ndl.go.jp/info:ndljp/pid/286794/www.mext.go.jp/b_menu/houdou/15/03/03033101/001.pdf

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory, 3*, 248-292.

Oxford, R. L. (2003). Toward a more systematic model of L2 learner autonomy. In D. Palfreyman & R. C. Smith (Eds.), *Learner Autonomy Across Cultures: Language Education Perspectives* (pp. 75-91). Basingstoke, England: Palgrave Macmillan.

Palfreyman, D., & Smith, R. C. (Eds.). (2003). *Learner autonomy across cultures: Language education perspectives*. Basingstoke, England: Palgrave Macmillan.

Rotter, J. B. (1954). *Social learning and clinical psychology*. New York: Prentice-Hall.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), 1.

Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of Consulting and Clinical Psychology, 43*(1), 56.

Rotter, J. B. (1990). Internal versus external control of reinforcement: A case history of a variable. *American Psychologist, 45*(4), 489.

Rupp, M. J. (in press). Using focus groups for qualitative feedback on psychometric instruments. In G. Brooks, M. Grogan, & M. Porter (Eds), *The 2015 PanSIG Conference Proceedings*. Kobe, Japan: JALT.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67.

Smith, R. C. (2003). Pedagogy for autonomy as (becoming-) appropriate methodology. *Learner autonomy across cultures: Language Education Perspectives*, 129-146.

Tassinari, M. G. (2012). Evaluating learner autonomy: A dynamic model with descriptors. *Studies in Self-Access Learning Journal, 3*(1), 24-40.

Tournat, T. (2014). Attribution theory: Measurement foundations for an emerging research area within applied linguistics. *JALT Journal: The Research Journal of the Japan Association for Language Teaching, 36*(2), 201-222.

Appendix

Author's Translation of Kambara Locus of Control Scale (K-LoC43) (Kambara, 1987)
Item (E=External; I=Internal)

1 E 何でも成り行きまかせが 1 番だ。[It is best to just go with the flow.]

2 I 努力すれは？立派な人間になれる。[I can be a great success if I work hard.]

3 I 一生懸命に話せは？誰にでも自分を分かってもらえる。[Anyone will be able to understand me if I try my best to communicate with them.]

4 I 自分の人生を自分自身で決定している。[I decide my own life.]

5 E 自分の人生は運命で決められている。[My life is decided by fate.]

6 E 自分が幸福なるか不幸になるかは偶然によって決められる。[My happiness and sadness are determined by chance.]

7E 自分の身に起こることは自分の置かれている環境によって決定されている。[What happens depends on the situation.]

8 E どんなに努力しても友人の本当の気持ちを理解することはできない。[My friends can't understand me no matter how hard I try.]

9 E 人生はキ？ャンフ？ルのようだ。[Life is a gamble.]

10 I 将来自分が何になるかを考えることは , 役に立つ (意味がある)。[It is useful (meaningful) to think about what I want to be in the future.

11 I 努力すればどんなことでも自分の力でできる。「If I try hard, I can do anything on my own.]

12 I たいていの場合 , 自分自身で決断した方が良い結果を生む。[Usually, things turn out better if I make my own decisions.]

13 E 幸福になるか不幸になるかは , 自分の努力次第だ。[My happiness or sadness is determined by my own efforts.]

14 I 自分の一生を思い通りに生きることができる。[I will be able to life my entire life as I plan to.]

15 I 自分の将来は運やチャンスによって決まる。[My future is determined by fate or chance.]

16 E 自分の身に起こることは自分の力ではどうすることもできない。[What happens does not depend on my efforts.]

17 I 努力すれは？誰とでも友人になることができる。[I can be friends with anyone if I try.]

18 E あなたの努力と成功とはあまり関係がない。[Your efforts and success are not really related to each other.]

19 I 自分の行動に注意していれは？いす？れは人から信頼される。[If I am careful about my actions, people will trust me.]

20 E 親友ができるかどうかは , クラスやクラフ？の雰囲気による。[My ability to make good friends depends on the class or club's atmosphere.]

21 I 努力すれは？希望の職につくことができる。[If I try hard, I will be able to get the job I want.]

22 E 理想的な相手と結婚できるかどうかは巡り合わせだ。[Marrying an ideal partner depends on fate or luck.]

23 I 予習復習をしておけばテストで良い成績を取るのは簡単だ。[It is easy to get a good score on tests if I prepare for lessons and review afterwards.]

24 I 自分の努力次第で異性の友人を作ることができる。[I can make friends with the opposite sex if I try.]

25 E 自分でも気付かずに衝動的に行動することがよくある。 [I often do impulsive things without being aware of it.]

26 E 希望する大学に進学できるかどうかは能力よりも偶然に左右される。[Getting into my first choice university depends more on luck than ability.]

27 E 友人とのつきあいが長く続くかどうかは周りの状況による。[Being able to maintain long friendships depends on the external situation.]

28 E あなたが何か行動する時 , 自分の希望と言うよりも人が言うからそうすることがよくある。[When you take actions, it is more often the case that others have suggested them rather than you acting upon your own desires.]

29 I 学校の授業が面白くないとすれば自分がその教科の勉強をあまりしないからだ。 [If a class in school is boring, it is because you are not interested in that subject.]

30 I 自分のすることはいつも自分で決める。[I always decide what I'm going to do.]

31 E テストの結果はあなたの場合 , 体調や偶然の出来事でしばしば左右される。
[In your case, when it comes to test results, they are often influenced by your physical condition or other random events.]

32 E 自分で決めたように行動することは難しい。[It is hard for me to do things as I have planned.]

33 E 頭の良し悪しは変えることはできない。[We can't change how smart or stupid we are.]

34 I 友情が続くかどうかはあなたの努力次第である。[Maintaining friendships depends on your effort.]

35 I 必要があればいつでも自分の欲求を抑えることができる。[If necessary, I can suppress my desires at any time.]

36 E 異性の友人ができるかは運によるので自分の行動をどうすべきか考えても仕方ない。 [There is no use in thinking about how to make friends with members of the opposite sex as such things are determined by fate/depend upon luck.]

37 E 自分の行動はまわりの状況によく流される。[My actions tend to end up going along with the flow of circumstances.]

38 I 前もって計画的に試験勉強をすれば結果はずっと良くなる。[The results are far better when I prepare for exams in advance.]

39 I 友人と仲良くやるために自分の行動を考えることは重要である。[It is important to think about my actions in order to have good relationships with my friends.]

40 I 友人と意見が違っても , 自分の行動を優先することが多い。 [Even if my friends have different ideas, I place a priority on my own actions.]

41 E 成績はつける先生によって変わる。[My grades depend on the teacher.]

42 I 友人に親切にしていればいつかは友人に助けてもらえる。[If I am kind to my friends, someday they will help me.]

43 E やりたくないと思っていても行動していることがよくある。[I often find myself doing things that I don't like to do.]