東海大学文理融合学部紀要1（2022年）1頁～10頁
Bull. of The School of Humanities and Science of Tokai Univ ., 1 (2022) pp.1-10

Validity and Reliability of the Interaction and Audience Anxiousness Scale in the Japanese SLA Context

by
Larry J. Xethakis* and Michael J. Rupp*

(Received: September 28, 2022, Accepted: December 18, 2022)

Abstract

This measurement study reports on the validity and reliability of Okabayashi and Seiwa's (1991) version of the Interaction and Audience Anxiousness Scale (Leary, 1983) in the second language acquisition context. The target population of this study was Japanese university students. Responses from university students at two universities in western Japan (N = 307) comprised the dataset for this study. Normality of test items was examined, and reliability estimates (Cronbach's alpha) for the two subscales that make up the instrument were calculated. The fit of the two-factor model hypothesized by Okabayashi and Seiwa (1991) was tested using confirmatory factor analysis (CFA), and the results showed good fit for the proposed model. The practical and theoretical implications of these results for future research are discussed.

Key Words : social anxiety, English education, group-work, confirmatory factor analysis

1. Introduction

The social dynamics inherent in the use of group work in the language classroom—the novelty of the activities, the uncertainty of working with unfamiliar partners and the need to express themselves in a foreign language—have the potential to engender feelings of anxiety in learners, much more so than in a more traditional, teacher-centered classroom. Up to now, one of the primary concerns with regards to anxiety in the English as a Foreign Language (EFL) context has been that of foreign language anxiety (FLA; see Horwitz, 2010, for a timeline of research in this area). However, as interaction-centered approaches such as Communicative Language Teaching and Task-Based Language Teaching have taken a larger role in language learning classrooms (Leeming, 2011), increased attention needs to be given to the role that social anxiety plays in student attitudes towards language learning. Both King and Smith (2017) and Zhou (2016) have noted a lack of research into the impact of social anxiety on language learning. However, for research in this area to progress in Japan, there is a need for reliable and valid instruments to measure social anxiety in the language learning context (Xethakis, 2020). This study represents an attempt to address this need by examining the reliability and validity of the Japanese version of the Interaction and Audience Anxiousness Scale (IAS-AAS; Okabayashi & Seiwa, 1991).

2. Literature Review

For the researcher or practitioner interested in social anxiety, there exist a large number of measures, focused on a wide range of aspects that can be subsumed under this overarching construct (see Leary, 1991 or Orsillo, 2002 for summaries and reviews of a number of these measures). Among the more prominent of these is the instrument chosen for inclusion in this study, the Interaction and Audience Anxiousness Scale (IAS-AAS; Leary, 1983). As Nichols and Webster (2015) note, the IAS-AAS "addressed and overcame limitations of the earlier scales, and consequently became a widely adopted measure of social anxiety" (p. 110).

One reason this instrument is particularly useful for the study of social anxiety in the classroom is the fact that it consists of items which are concerned with both contingent and noncontingent situations. Leary and Kowalski (1995) make a distinction between these two classes of social situations that can elicit feelings of unease or distress. The first of these are social situations where an individual's responses depend on, or are contingent upon, the actions of others. Examples of contingent social situations include conversations, interviews, mixing at parties, and interacting in groups—situations similar to the group-work situation in the language classroom. The second category of situations are those where an individual's actions are primarily planned out beforehand, and so their actions are for the most part not contingent upon others' actions. As a result, they vary little in response to the actions of others. Noncontingent situations are exemplified by events such as giving speeches or presentations and performing in front of an audience.

*Associate Professor, Department of Community and Social Studies. Tokai University

Both contingent and noncontingent situations occur with regularity in group work contexts in the language classroom, and for this reason, it would be advantageous for any instrument adopted for use in research or practice in this area to contain items or subscales which purport to examine both the contingent and noncontingent situations which might engender feelings of distress in learners who are socially anxious.

As the name implies, the Interaction and Audience Anxiousness Scale, comprises two subscales, one concerned with feelings of unease in contingent social encounters, the Interaction Anxiousness Scale (IAS), and the other, the Audience Anxiousness Scale (AAS), concerned with circumstances that could be considered noncontingent in nature. The IAS-AAS, in its original long-form (Leary, 1983), was initially developed from a pool of 87 items selected on the basis of two criteria: 1) that the situations described in the items were concerned with forms of social interaction, rather than evaluation or the performance of a task; and 2) that the item denoted feelings of anxiety, such as nervousness, worry or discomfort, or its opposite, such as feelings of comfort or relaxation, but at the same time did not portray a behavioral element. This pool of items was piloted with 112 undergraduates, who responded on a 5-point Likert scale indicating the degree to which each of the items was true of them. Items were correlated with their hypothesized category (interaction or audience anxiety), and those with a correlation less than .40 were removed. This resulted in a total of 37 items. These items were given to a second group of university students (n = 123), and based on the correlations between items, 10 items were removed, leaving a total of 27 items divided between the two categories of interaction anxiousness, 15 items, and audience anxiousness, 12 items. For this sample, the reliability estimate (Cronbach's alpha) for each scale was .88. This 27-item version was then tested on a third sample of university students (n = 363). The items correlated with their respective scales at greater than .50, while the alpha values for each of the scales were .89 (interaction anxiousness) and .91 (audience anxiousness), respectively. In this sample, the degree of correlation between the two scales was found to be .44.

Leary and Kowalski (1993) examined the construct and criterion-related validity of the IAS, using a large pool of data (1,864 respondents) compiled from several separate studies conducted by the authors and concluded that the IAS possessed a high degree of reliability, with Cronbach alpha values between .87 and .89 in the various studies considered. The concurrent validity of the IAS was confirmed in this study using responses on a large number of instruments, 23 in total, examining a wide range of subjective and behaviorally related aspects, such as self-consciousness, shyness, embarrassability, blushing and body consciousness. Responses on the IAS were found to correlate highly with other scales assessing generalized social anxiety, while they correlated less strongly with measures focused on more specific aspects of the construct. Discriminant validity was also ascertained, and it was determined that the IAS measures a construct other than general anxiety or neuroticism.

While no similar review for the AAS has, to the best knowledge of the authors, been published, the AAS has shown a very high degree of reliability in studies conducted by other authors. Sabini, Siepmann, Stein and Meyerowitz (2000) calculated a reliability estimate (Cronbach's alpha) of .93 for this scale, while in Hazel, McMahon and Schmidt's (2011) study this value was only slightly lower at .91. Moreover, in the study conducted by Sabini and his colleagues, the AAS was found to correlate highly with a scale the authors used to measure respondents' feelings of unease at being the center of attention in a social situation, which suggests a degree of construct validity in the AAS, as well.

Furthermore, while the use of the AAS has been employed primarily in contexts where public performance is a significant factor (e.g., Hazel, et al., 2011; Sabini et al., 2000), the IAS has been used to examine the impact of interactional anxiety on a number of other factors related to working in groups, such as social skills (Miller, 1995), leadership and collaboration (Steed, Slater, Sadagic, Bullock & Tromp, 1999), interpersonal relations (Heerey & Kring, 2007), brainstorming in groups (Camacho & Paulus, 1995), and online social interactions (Tian, 2013).

The IAS-AAS has also been widely used in the Japanese population. Okabayashi and Seiwa (1991) adapted the instrument for use in the Japanese population, and their version of the IAS-AAS (or as it is often referred to in the Japanese literature, the I-AA scale) has been employed in studying the impact of social anxiety on a number of factors related to pair- and group-work in the language classroom, such as, communication in face-to-face and online environments (Nishimura, 2005), performance anxiety (Yoshie & Shigematsu, 2007), and speech anxiety (Matsumoto, 2014).

Okabayashi and Seiwa's version of the IAS-ASS was developed on the basis of an exploratory factor analysis (EFA) conducted on a dataset of responses from university students (n =140), using the 27 items from Leary's (1983) original scale.

The results from the initial varimax rotation indicated a three-factor solution, with items from the AAS subscale loading primarily on the first factor, and those from the IAS subscale loading primarily on the third factor. The second factor comprised six items: four from the IAS and two from the AAS. Of these items, five were considered to express positive feelings towards a situation (these are, interestingly, five of the six reversed scored items on the IAS-AAS), while one item was considered to express a broad general tendency, rather than a more specific feeling of anxiety. For these reasons, the six items were removed from the analysis, and a second EFA was conducted. This resulted in a two-factor structure, similar to that of Leary's (1983) original instrument. In order to shorten and simplify the scale, seven items from each factor were chosen on the basis of factor loading and item content and a third EFA was conducted. This resulted in a 14-item, two-factor, simple structure. The brevity of Okabayashi and Seiwa's version of the IAS-ASS can be seen as a distinct advantage for the use of this instrument in classroom settings.

With the greater use of pair- and group-work in the English language classroom, the greater degree of interpersonal interaction that comes with it, and the strong impact that feelings of social anxiety can have on such interactions, there is a need for evidence-based measurement of learner's social anxiety. The IAS-AAS was chosen as a suitable instrument for use in the Japanese EFL context due to its focus on both contingent as well as noncontingent social situations, evidence of its reliability and validity, as well as the brief nature of the Japanese version and its previous use in the Japanese population. This study, which has undertaken an investigation of the IAS-AAS in the Japanese EFL context, is an incremental step in the establishment of secure empirical foundations for further research in this area.

3. Methodology

3.1 Participants

A total of 345 responses were collected from students enrolled in one public and one private university in western Japan. There were 178 males and 156 females (eleven respondents did not provide a gender) among the respondents, and their ages ranged from 18 to 29, with a median age of 19 years old. Two responses were removed for missing data.. The analysis described below is based on the data from the remaining 343 responses.

Prior to commencing the survey, participants were informed by the administrator of the survey that they were not required to take part in the survey, and that participation was completely voluntary. Informed consent was obtained by the inclusion of a form at the top of the survey paper asking for participants' informed consent and clearly stating in Japanese that those not wishing to participate could do so merely by leaving the form blank. There was no specified time limit within which participants were expected to complete the survey, however most completed the form within 10 minutes.

3.2 Instrument

As described above, Okabayashi and Seiwa's (1991) version of the IAS-AAS comprises 14 items in total, with the items divided between two subscales. The first of these, the IAS, comprises Items 1-7, while the second subscale, the AAS, comprises Items 8-14. Responses to each of the items are on a 5-point Likert scale, with 1 being semantically anchored to *not at all characteristic of m*e, and 5 to *extremely characteristic of me*. As each subscale concerns a specific class of social situation that might trigger anxiety in respondents, i.e., contingent situations in the case of the IAS and noncontingent situations for the AAS, scores for each of the subscales are computed separately, and there is no composite score for the entire instrument.

3.3 Analytical Procedures

Data collected from participants (scores on the IAS-AAS, age, and gender) was entered into a Microsoft Access 2016 database. For the purpose of calculating descriptive statistics and reliability estimates (Cronbach's alpha), the data was imported into IBM/Statistical Package for the Social Sciences (SPSS) software (Version 21). First, means for each item, their standard deviations, and degree of skew and kurtosis were calculated. The univariate normality of the scores was determined following the recommendation of Tabachnick and Fidell (2013), i.e., values > -2 or < 2 for skew and kurtosis. Next, an estimate of reliability, Cronbach's alpha, was calculated for each subscale. Following the recommendations of Fan and Thompson (2001), reliability estimates with confidence intervals (95%) were computed. A value of .70 or greater for the reliability of the scale was adopted (Nunnally and Bernstein, 1994). Finally, confirmatory factor analysis (CFA) was conducted on Okabayashi and Seiwa's (1991) proposed 14-item, two-factor structure for the IAS-AAS. using AMOS (Version 21). Four fit indices—the Tucker-Lewis index (TLI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root

mean squared residual (SRMSR)—were utilized in conjunction with the chi-square to determine the degree of model fit. Combining results from the four fit indices together with the chi-square is one means to overcome the latter's tendency to over-reject models. Hu and Bentler (1999) have recommended cut-off values for each the four fit indices (TLI and CFI >.95; RMSEA <.06; SRMSR <.08), which are used in conjunction to evaluate the fit of the model, and these values were adopted to evaluate the fit of the hypothesized model for the IAS-AAS.

4. Results

4.1 Descriptive Statistics, Skew and Kurtosis

Descriptive statistics were derived for the 14 items on the IAS-AAS. The descriptive statistics for each item on the IAS scale are presented in Table 1 below. As can be seen in the table, the highest mean among the items was 3.58 for Item 9, while the lowest mean was that for Item 3 with, with a value of 2.25. The standard deviations ranged from 1.097, for Item 3, to 1.410, for Item 14. Values for skew and kurtosis for all items were within the acceptable range of > -2 or < 2, and thus the scores were determined to possess a sufficient degree of univariate normality for use in factor analysis.

Table 1: Item Means, Standard Deviation, Skew and Kurtosis for the Items on the IAS-AAS

Item	M	SD	Skew	Kurtosis
1	2.98	1.182	.151	-.882
2	2.32	1.140	.549	-.530
3	2.25	1.097	.698	-.185
4	2.65	1.185	.397	-.727
5	2.45	1.156	.479	-.646
6	2.90	1.329	.020	-1.225
7	3.06	1.187	.050	-.961
8	2.85	1.239	.341	-.976
9	3.58	1.172	-.324	-.994
10	2.76	1.298	.309	-1.005
11	2.93	1.218	.169	-.971
12	2.63	1.247	.399	-.866
13	2.96	1.214	.133	-.994
14	3.24	1.410	-.186	-1.281

4.2 Reliability Estimates

The reliability estimates (Cronbach's alpha) and respective confidence intervals for scores on the two subscales of the IAS-AAS are presented in Table 2. The reliability estimates for the subscales exceed the .70 cut-off value by a fair margin, with even the lower bounds of the 95% confidence intervals above .80. These values are similar to the alpha values for the IAS and AAS reported in other studies. Values for the IAS have been between .81 (Renshaw, 2004) and .89 (Sasaki & Tanno, 2006), with the AAS exhibiting values between .81 (Yoshie & Shigematsu, 2007) and .91 (Sasaki & Tanno, 2006). The results in this study, together with those of past studies, provide a degree of evidence that the subscales of the IAS-AAS exhibit a high degree of reliability.

Table 2: Reliability Estimates, Confidence Intervals for Alpha (95%), Scale Means, and Scale Standard Deviations for Scores on the IAS-AAS

Subscale	Cronbach's alpha	95% Confidence Intervals		Scale Mean	SD for Scale
		Lower Bound	Upper Bound		
IAS	.834	.805	.859	18.62	5.87
AAS	.885	.865	.903	20.94	6.78

4.3 Confirmatory Factor Analysis

A CFA was conducted in order to directly test the two-factor, 14-item structure postulated by Okabayashi and Seiwa (1991). For the purpose of determining the degree of fit between the proposed models and the scores in the data, a selection of goodness-of-fit indices, the RMSEA, the SRMSR, the TLI, and the CFI, were employed in addition to the $\chi 2$ test statistic. In order to ascertain the degree of multivariate non-normality in the data, Mardia's co-efficient was employed.

The model tested possessed 105 distinct sample moments, 29 distinct parameters to be estimated, with 76 degrees of freedom, and thus was overidentified. For this model, the results of the fit indices were as follows (Hu and Bentler's [1999] cut-off values in parentheses): TLI .938 (>.95), CFI .949 (>.95); RMSEA .063 (<.06); SRMSR .0469 (<.08). The value for the $\chi 2$ was 180.508 with a probability level of .000. A degree of multivariate non-normality in the data was indicated by the value for Mardia's coefficient for this model, 17.650, which is close to the threshold value of 5.0. Taken in combination, which is the procedure recommended by Hu and Bentler (1999), the values for the goodness-of-fit indices for the model, which border closely on the recommended values, strongly suggest at least an adequate degree of fit between the

underlying structure of the data and the structure specified by the model.

Inspection of the goodness-of-fit indices is the first step in determining the overall, or global, fit of a model. When these indices indicate a sufficient degree of fit between the model and the scores themselves, Brown (2015), Byrne (2016), and Kline (2011) all suggest that the researcher examine the modification indices and standardized residuals in order to determine if specific relationships in the model exhibit signs of misspecification.

The modification indices for the error covariances of this model can be seen in Table 3 below. This table shows only those indices with a value of 10 or greater, as relationships with values less than this are often considered to have little to no effect on the fit of the model (Byrne, 2016). Only three of the covariance modification indices for this model exceed this threshold, and even the largest of these by a mere 1.695. Among the regression weight modification indices, only one value, that between Item 18 and Item 8, produced a value greater than 10, at 11.828. The small magnitude of these indices and their relative dearth suggest that there are few, if any areas of misspecification in the measurement model.

Table 3: Covariance Modification Indices for the IAS-AAS Measurement Model

Covariance			Modification Index
e8	<-->	e18	11.695
e7	<-->	e16	10.803
e4	<-->	e8	10.378

After the inspection of the modification indices was completed, the standardized residuals for this model were examined. In determining the significance of the values for the residuals, an absolute value greater than 1.96 is often adopted as the cut-off for residuals, as this corresponds to a statistically significant z score (p =.05), meaning that the unaccounted-for covariance is more than likely not due simply to chance (Brown, 2015). Out of a total of 91 residuals for this model, only four (4.3%) had a value greater than 1.96, with the highest absolute value of these being 2.325 (The complete table of standardized residuals for the model are available from the authors upon request).

The small number of residuals surpassing the cut-off can be interpreted as a further indication that there is not a significant degree of misspecification found in this model. This result, taken together with that for the modification indices, and the values of the goodness-of-fit indices, strongly suggest that the structure proposed by Okabayashi and Seiwa (1991) for the IAS-AAS fits the actual underlying structure of the scores in the dataset to a sufficient degree.

The implications of the results outlined above in regard to the postulated structures of the IAS-AAS and the use of this instrument in the Japanese EFL context are considered in the Discussion section below.

5. Discussion

Since its development by Leary (1983), the IAS-AAS has been a widely-used instrument for examining subjective affective and cognitive aspects of social anxiety in both Japanese and non-Japanese populations. Despite its popularity in both populations, the hypothesized factor structure of the IAS-ASS had yet to be confirmed. Evidence either for or against the structural validity of the 14-item adapted version developed by Okabayashi and Seiwa (1991) would aid in determining the viability of the IAS-AAS for use in the Japanese EFL context. The availability of an evidence-based measure of social anxiety would enable research in this area to progress with greater confidence in its derived results. The primary purpose of this paper has been to provide such evidence through the use of CFA to directly test the structure postulated by Okabayashi and Seiwa.

The degree of fit displayed by Okabayashi and Seiwa's (1991) version of the IAS-AAS was interpreted to be good on the basis of three sources of evidence: 1) the goodness-of-fit indices; 2) the modification indices for the model; and 3) the standardized residuals for the model.

The goodness-of-fit indices for this version of the IAS-AAS provide global indications of a sufficient degree of fit between the underlying structure of the data and the structure specified by the model. The values for the first of these indices, the TLI and CFI were very close indeed to Hu and Bentler's (1999) recommended values, with the CFI (.949) within a mere .001 of the recommended value (.95), and the TLI (.938) only .012 less than this value. This degree of proximity is of note for two reasons. The first being that both of these indices are approximate fit indices, and thus their values should be interpreted on a continuum (Hu & Bentler, 1999, p. 2). Second, and moreover, as Brown (2015) makes clear, Hu and Bentler (1999) intentionally include the words "close to" in their recommendations for acceptable values for goodness-of-fit indices (e.g., Hu & Bentler, 1999, p. 27). They state that, "it is difficult to designate a specific cutoff value for each fit index

because it does not work equally well with various conditions," (p.27), and thus values "close to" those of their recommendations, when used in combination with other indices, allow researchers to "have more confidence about the goodness of fit of the model" (p. 28). One example of this interpretation along a continuum can be found in Brown (2015). While clearly stating that values less than .90 for the TLI and the CFI can be interpreted as evidence for the rejection of a model, Brown cites Bentler (1990), who proposes that values of .90 or greater can be interpreted as evidence of good fit. Hair et al. (2014) also suggest that the values of goodness-of-fit indices should not be taken as absolutes, but rather need to be interpreted on the basis of sample size as well as the characteristics of the model in question. In the case of this model, where there are more than 250 respondents and more than 12 observed variables, values for the TLI and CFI above .92 are indicative of good fit following their guidelines (Hair et al., 2014). Finally, as Byrne (2016) points out, when there is a degree of non-normality present in the data, there is a tendency for the values of CFI and TLI to be underestimated. On the basis of the above rationales, the values of the TLI and CFI for the model specifying Okabayashi and Seiwa's (1991) version of the IAS-AAS, were interpreted as providing evidence of good fit between the model and the empirical structure underlying the scores in this dataset.

The result for the SRMR, which indicates the degree of difference between the correlations in the dataset and those predicted by the model (Kline, 2011), was significantly less than the recommended value, suggesting that there was not a large degree of difference between these two sets of correlations, and this again suggests a good degree of fit between the model and the scores.

The value for the RMSEA could be seen as more problematic, as it surpassed the recommended value of .06, albeit by a mere .003. However, as with the TLI and CFI, this result is very "close to" the value recommended by Hu and Bentler (1999), which suggests that the value for the RMSEA indicates, at least, an acceptable degree of fit. In addition to Hu and Bentler's recommendations, a value of less than .08 has been postulated to suggest adequate fit (See Browne & Cudeck, 1993), with values between .08 and 1.0 suggesting mediocre fit (See MacCallum et al., 1996). These more relaxed criteria come from more dated sources, but there is currently a pattern in the general literature that below .06 is good and below .08 is adequate. Further evidence for the degree of fit expressed by the RMSEA can be found in the values for the 90% confidence intervals calculated for this index. In the case of the model tested in this study, the values for lower and upper bounds of the confidence interval were .052 and .075, respectively. A value below .08 for the upper bound of the confidence interval can be interpreted as additional evidence for the degree of fit suggested by the value for the RMSEA (Brown, 2015). Hair et al. (2014) also recommend that for models with larger sample sizes ($n > 250$) and more than 12 indicators, RMSEA values of less than .07, together with a CFI value of .92 or higher suggest a good degree of fit. Thus, the RMSEA value for the model can also be interpreted as suggesting good fit.

It should be noted that the result of the $\chi 2$ test statistic was significant, and this in turn suggests that degree of fit between the scores in the dataset and the postulated model may be "not entirely adequate," (Byrne, 2011, p.76). There are, however, a number of factors other than the degree of consistency between the covariance structure found in the data and that predicted by the model that can influence the behavior of this statistic, such as the degree of multivariate non-normality (Kline, 2011), the sample size (Hair et al., 2014) and the complexity of the model (Brown, 2015). According to Hair et al. (2014), for a model with between 12 and 30 indicators and a sample size greater than 250, as is the case for the model in question here, significant p-values for $\chi 2$ should be expected, and thus the fit of the model should be interpreted in light of the results of the other indices employed. For this reason, notwithstanding the result of the $\chi 2$, the values for the goodness-of-fit indices, when taken together, were interpreted as evidence for the conclusion that the model possessed a sufficient degree of fit.

The second and third sources of evidence for this conclusion were the modification indices and standardized residuals, which were utilized to look for indications of localized misspecification in the measurement model. Such misspecification can exist even in models which exhibit good fit according to the values of their goodness-of-fit indices, and thus these values should be examined before making a determination of a model's fit (Kline, 2011). In the case of this model, the modification indices (exceeding the cut-off value of 10) were few in number (three) and surpassed the cut-off by only small amounts. Byrne (2016) points out that modification indices with small values, such as those for the model in this study, are "of little concern," (p. 104), and thus these results suggest that the relationships between the scores in the dataset are being reflected in a sufficiently accurate manner by the model. Regarding the standardized residuals, only four exceeded the value of 1.96 that Brown (2015) suggests may

indicate a source of misspecification other than chance, and none surpassed the value to be considered large (2.58). These results can be interpreted as a further indication that the model hypothesized by Okabayashi and Seiwa (1991) is adequately accounting for the covariance between the vast majority of items.

The values of the goodness-of-fit indices, which suggest good overall fit between the scores in the dataset and the structure of the model, together with the lack of misspecification indicated by the modification indices and the values of the standardized residuals can be interpreted as providing strong evidence for the viability of Okabayashi and Seiwa's (1991) version of the IAS-AAS in measuring respondent's subjective feelings of social anxiety in the Japanese EFL context.

This result has both practical and theoretical implications. In terms of practice, the degree of fit, and thus structural validity, exhibited by Okabayashi and Seiwa's (1991) version of the IAS-AAS allows both researchers and classroom teachers to have confidence in the interpretation of scores according to the constructs the respective items are claimed to measure. This is because evidence of good fit in a measurement model provides evidence of the unidimensional nature of the scales that comprise the instrument. Interpretation of scores on an instrument is based upon the assumption of unidimensionality, and thus, evidence of structural validity allows for confidence in the interpretation of scores.

In the theoretical realm, the degree of fit exhibited by this version of the IAS-AAS also provides a degree of evidence for the validity of Leary's distinction between contingent and noncontingent situations as an aspect of social anxiety. A psychological instrument is a systematic representation of the relationships between the underlying construct and the items on the instrument. If these hypothesized relationships are shown to be accurate, within a reasonable measure, this argues for not only the viability of the instrument as a measure of the underlying construct, but also for the plausibility of the underlying theory itself. The results for Okabayashi and Seiwa's (1991) version of the IAS-AAS in this study by no means confirm Leary's hypothesis concerning the division of social situations into contingent and noncontingent classes, as the research design here is not experimental but rather psychometric, however they do suggest that this theoretical, and common-sense, conception may be well-grounded and supported by evidence.

As a matter of further qualification, these results also do not suggest that the classes of contingent and noncontingent situations are themselves unitary factors that cannot be further subdivided. Conceptually, it is possible to envision that each class of situations could be partitioned into finer distinctions of distinct contingent and noncontingent situations with their own, unique anxiety provoking aspects. Among the subcategories of contingent events, it may be plausible to assume that speaking with members of the opposite sex, may involve different a dynamic and thus likely a different source of anxious thoughts and feelings as opposed to for example, speaking with persons of authority. Similar distinctions could be drawn among the noncontingent situations as well. For example, performing in a play and giving a speech both involve speaking in front of an audience. In the former situation, one is speaking from a script and playing a part, in a way taking on a new identity and even hiding one's true self, whereas the latter situation tends to be more extemporaneous and involves a greater degree of self-presentation, and thus risk to one's self-image. These different aspects might trigger different forms of anxious thoughts and feelings, which while still being subsumed under the larger class of noncontingent situations, might be independent of each other nonetheless.

The exploration of these theoretical and conceptual issues may be aided by the evidence for the structural validity of the IAS-AAS presented in this paper, and thus its theoretical underpinnings as well. The positive results for this instrument represent the substantial contribution of this study to the literature, allowing future research employing this instrument to go forward with greater confidence.

6. Conclusion

This study is, to the best of the authors' knowledge, the first study to utilize confirmatory methods to examine the structural validity of the IAS-AAS as a composite instrument in either the Japanese or the non-Japanese population. For this reason, the good degree of fit displayed by Okabayashi and Seiwa's (1991) 14-items version of the IAS-AAS represents not only a positive finding of this study, but also an incremental step forward to grounding the use of this instrument in both practice and research on an evidence-based foundation. In addition to these findings providing evidence for this version of the IAS-AAS as a viable model for the measurement of social anxiety in the Japanese EFL context, they also provide evidence for the plausibility of Leary's (1983) conception of contingent and noncontingent classes of social situations. It must be remembered, however, that the evidence for an instrument's

validity comes about through a cumulative process of evidence gathering, and thus, this study provides only a single piece of such evidence.

Finally, with regards to the limitations of this study, there are two factors that must be noted. First, the degree of non-normality found in the dataset employed in this study may represent a limit on the generalizability of its results. However, due to a lack of reporting on the degree of normality found in the datasets employed in other studies on the structure of the IAS-AAS, it is difficult to know if the degree of non-normality found in this study is unique to the scores in this dataset, is a characteristic of the Japanese adaptation of the instrument, unique to the Japanese population, or indeed if it is an invariant characteristic of the instrument itself.

The second limitation has to do with the sample investigated in this study, which was not a truly random sample of Japanese university EFL students, but rather a sample of convenience. It must be noted here, however, that this limitation is common to the majority of studies in the literature of this field, as well as research in the social sciences more generally. If researchers were to wait for perfect samples, far less research would be done. One obvious means of overcoming this limitation is repeated sampling of the target population. While it is most likely the case that these further studies will also rely on samples of convenience, the limitations inherent in each of these possible future studies could be overcome through the utilization of meta-analysis as a means to average out the sample specific properties which hamper generalization to the population in the case of each study.

7. References

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* New York, NY.: Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen & J.S. Long (Eds.), *Testing structural equation models*, (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* New York, NY: Routledge.

Camacho, L. M., & Paulus, P. B. (1995). The role of social anxiousness in group brainstorming. *Journal of Personality and Social Psychology, 68*(6), 1071-1080.

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.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis: Pearson new international edition.* Essex, UK: Pearson Education Limited.

Hazel, M., McMahon, C., & Schmidt, N. (2011). Immediate feedback: A means of reducing distracting filler words during public speeches. *Basic Communication Course Annual, 23*(1), 6-28.

Heerey, E. A., & Kring, A. M. (2007). Interpersonal consequences of social anxiety. *Journal of Abnormal Psychology, 11*6(1), 125-134.

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching, 43*, 154-167. https://doi.org/10.1017/S026144480999036X

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.

King, J., & Smith, L. (2017). Social anxiety and silence in Japan's tertiary foreign language classrooms. In C., Gkonou, M., Daubney & J-M., Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*, (pp. 91-109). Bristol, UK: Multilingual Matters.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Leary, M. R. (1983). Social anxiousness: The construct and its measurement. *Journal of Personality Assessment, 47*(1), 66-75.

Leary, M. R., & Kowalski, R. M. (1995). The self-presentation model of social phobia. In R. G. Heimberg (Ed.), *Social phobia: Diagnosis, assessment, and treatment*, (pp. 94-112). New York, NY: Guilford Press.

Leary, M. R. (1991). Social anxiety, shyness, and related constructs. In J.P. Robinson, P.R. Shaver, & L. W. Wrightsman (Eds.), *Measures of personality and social psychological attitudes: Measures of social psychological attitude*s, (Vol. 1, pp. 161-195). San Diego, CA: Academic Press.

Leeming, P. (2011). Japanese high school students' use of L1 during pair-work. *International Journal of Applied Linguistics, 21*(3), 360-382.

MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods,*

1(2), 130-149.

松本明生 [Matsumoto Akio]. (2014). アクセプタンス自己教示がスピーチ不安に及ぼす影響 [Effects of Acceptance-Based Self-Instructions on the Speech Anxiety of Undergraduate Students]. 教育心理学研究 [Japanese Journal Of Educational Psychology], 62(1), 38-49.

Miller, R. S. (1995). On the nature of embarrassability: Shyness, social evaluation, and social skill. *Journal of Personality, 63*(2), 315-339.

Nichols, A. L., & Webster, G. D. (2015). Designing a brief measure of social anxiety: Psychometric support for a three-item version of the Interaction Anxiousness Scale (IAS-3). *Personality and individual Differences, 81*, 207-212. https://doi.org/10.1016/j.paid.2015.03.013

西村洋一 [Nishimura Youichi]. (2005). コミュニケーション時の状態不安および不安生起に関連する要因の検討 [An examination of state communication anxiety, communication medium, and trait social anxiety]. パーソナリティ研究 [The Japanese Journal of Personality], 13(2), 183-196.

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

岡林尚子 [Okabayashi Naoko], & 生和秀敏 [Seiwa Hidetoshi]. (1992). 対人不安感尺度の信頼性と妥当性に関する一研究 [A study on the reliability and validity of Leary's social anxiety scale]. 広島大学総合科学部紀要. III, 情報行動科学研究 [Memoirs of the Faculty of Integrated Arts and Sciences, Hiroshima University, III, Studies in Information and the Behavioral Sciences], 15, 1-9.

Orsillo, S. M. (2002). Measures for social phobia. In M. M. Antony, S. Orsillo, & L. Roemer (Eds.), *Practitioner's guide to empirically based measures of anxiety*, (pp. 165-187). Boston, MA: Springer.

Renshaw, S. L. (2004). Subjective measures of communication anxiety: An assessment of the reliability and structural integrity of three scales. 神田外語大学紀要 [The Journal of Kanda University of International Studies], 16, 35-59.

Sabini, J., Siepmann, M., Stein, J., & Meyerowitz, M. (2000). Who is embarrassed by what? *Cognition & Emotion, 14*(2), 213-240.

Sasaki, J., & Tanno, Y. (2006). Two cognitions observed in Taijin-kyofusho and social anxiety symptoms. *Psychological reports, 98*(2), 395-406.

Steed, A., Slater, M., Sadagic, A., Bullock, A., & Tromp, J. (1999, March). Leadership and collaboration in shared virtual environments. In L. Rosenblum, P. Astheimer, and D. Teichmann (Eds.), *Proceedings IEEE Virtual Reality*, (pp. 112-115). Los Alamitos, CA; IEEE.

吉江路子 [Yoshie Michiko], & 繁桝算男 [Shigematsu Kazuo]. (2007). 対人不安傾向と完全主義認知が演奏状態不安に及ぼす影響 [Effects of Trait Social Anxiety and Cognitions of Perfectionism on State Music Performance Anxiety]. パーソナリティ研究 [The Japanese Journal of Personality], 15(3), 335-346.

Tabachnick, B. G., and Fidell, L. S (2013). *Using multivariate statistics* (6th ed.). Pearson.

Tian, Q. (2013). Social anxiety, motivation, self-disclosure, and computer-mediated friendship: A path analysis of the social interaction in the blogosphere. *Communication Research, 40*(2), 237-260.

Xethakis, L. (2020) *The Psychometrics of Affective, Cognitive and Social Factors Impacting Group-Based Learning Activities* [Unpublished doctoral dissertation]. Kumamoto University

吉江路子 [Yoshie Michiko], & 繁桝算男 [Shigematsu Kazuo]. (2007). 対人不安傾向と完全主義認知が演奏状態不安に及ぼす影響 [Effects of Trait Social Anxiety and Cognitions of Perfectionism on State Music Performance Anxiety]. パーソナリティ研究 [The Japanese Journal of Personality], 15(3), 335-346.

Zhou, M. (2016). The roles of social anxiety, autonomy, and learning orientation in second language learning: A structural equation modeling analysis. *System, 63*, 89-100.

Appendix: Items on Okabayashi and Seiwa's (1991) version of the IAS-AAS

	Item
1	私は知らない人の集まりの中にいると、いつも居心地が悪い。I usually feel uncomfortable when I am in a group of people I don't know..
2	私は先生や上司と話をしなければならないと、そのことが負担になる。I get nervous when I must talk to a teacher or boss.
3	私はパーティなどで、しばしば不安になったり不快な気持ちになったりする。Parties often make me feel anxious and uncomfortable.
4	私は同性の人でも、あまり親しくない人と話すと時々緊張する。I sometimes feel tense when talking to people of my own sex if I don't know them very well.
5	私がもし仕事で人と会わなければならないとしたら、そのことがかなり気がかりとなる。I would be nervous if I was being interviewed for a job.

6 私はあまり親しくない人に電話をかける時、そのことが苦になる。I often feel nervous when calling someone I don't know very well on the telephone.

7 私は偉い人に話しかける時、いつも緊張する。I get nervous when I speak to someone in a position of authority.

8 私は人前で話をしている間中、ずっと緊張している。I usually get nervous when I speak in front of a group.

9 私は人前に出て行かなければならない時、緊張する方だ。I tend to experience 'stage fright' when I must appear before a group.

10 私がもし、たくさんの聴衆の前に出て行かなければならないとしたら、考えただけでも恐い。I would be terrified if I had to appear before a large audience.

11 私は人前で話したり、何かをしなければならない時、そわそわして落ち着かなくなる。I get 'butterflies' in my stomach when I must speak or perform before others.

12 カメラで写されることが分かると、緊張してぎこちなくなる。I would feel awkward and tense if I knew someone was filming me with a movie camera.

13 私は人前で話をする時、自分の考えがまとまらなくなってしまう。My thoughts become jumbled when I speak before an audience.

14 私は人前で話すことがこんなに苦にならなければいいのにと思う。I wish I did not get so nervous when I speak in front of a group.