Psychometric Properties of Scores on a Japanese Translation of the Junior Version of the Metacognitive Awareness Inventory

Ryosuke UENO

Abstract

The purpose of the research reported in this paper was to examine the model fit of an adapted version (English-speaking population to Japanese-speaking population) of the Junior Metacognitive Awareness Inventory (JMAI; Sperling, Howard, Miller, & Murphy, 2002). The full 18-item version of the instrument was translated and then administered to N=464 high school students. Confirmatory factor analysis (CFA) was used to test for model fit, and the 18-item model was rejected. Following this modification indexes were used to test a series of changes in the model which arrived at an abbreviated and balanced model of two constructs measured by four items each. This model produced good fit, and is short and convenient to administer. However, the model will require further a priori testing in new samples from the same population in the future. Other implications of the research are discussed in the paper.

Introduction

In the Japanese educational situation, students are required to have various competencies through their learning in school. The Ministry of Education, Culture, Sports, Science and Technology (MEXT) has suggested a concept called "proactive, interactive, and deep learning [translated]" (MEXT, 2018a, p. 19) for schools including kindergarten, elementary school, junior high school, high school, and special needs education schools. As for the first term in the above translated quote, "proactive learning," this is regarded as important for students to recognize the manner of their learning and its transformation through planning and reviewing (MEXT, 2018b). In addition, the Central Council for Education (2021), which is related to the MEXT, proposes that teachers should develop their abilities to teach students how to regulate themselves and cultivate an attitude that also leads to independent and self-regulated learning. Such proactive ability in students toward their own learning, including regulating their emotions, behaviors and attitudes are, according to the MEXT, associated with the concept of metacognition (MEXT, 2021). Of course, the MEXT regards this requirement as applying to all subjects, and this would include foreign language education, especially, English education. This means that evaluating students' metacognitive skills in their learning of English is necessary. We cannot understand and improve the metacognition of students if we are unable to measure the ability itself.

According to Flavell (1976, p. 232), "[m]etacognition refers to one's knowledge concerning one's own cognitive processes or anything related to them, e.g., the learning-relevant properties of information or data." Louca (2003) has explained metacognition in similar terms to Flavell. She argues that metacognition is a second-order level of cognition, and has also referred to it as "thoughts about thoughts, knowledge about knowledge, or reflections about actions" (p. 10). Flavell also explained that goals and subgoals are important for metacognitive knowledge and that selecting the right cognitive process to achieve these goals and subgoals is an ability which falls under the concept of metacognition. Learners who have high levels of metacognitive strategies have good learning results. For example, Okamoto (1991) reported results that showed different metacognitive levels cause different outcomes using a set of math worksheets and a method of stimulus recall to check the metacognitive levels of 45 fifth-grade students. According to the results from this study, it was clear that learners who had developed metacognitive strategies for learning were able to more effectively solve problems. In English education, Çubukçu (2008) researched the effect of metacognitive strategies on English reading skills using a sample of 130 university students. Half of the students (the experimental group) had taken classes focused on reading expansion skills for a period of five weeks. This experimental group showed better scores for both vocabulary and reading when compared to the control group. In a more recent study on metacognition and reading comprehension, Okamoto (2012) theorized that metacognitive activity during task performance could be further subdivided into the constructs of online metacognition and offline metacognition. He goes further to argue for the possibility of interaction in which online metacognition is facilitated by training the learner's offline metacognition. He also argues that metacognitive experiences of what strategies work or do not work are involved in the acquisition of metacognitive knowledge, and the level and use of metacognitive knowledge in reading may influence reading comprehension.

To make rigorous observations of learners' metacognition and actually measure the ability, various inventories have been developed. Schraw and Dennison (1994) developed the original version of the Metacognitive Awareness Inventory (MAI) to measure adults' metacognitive awareness. The MAI, which has 52 items in its original format, was hypothesized by the authors to have two components: Knowledge of Cognition (17 items) and Regulation of Cognition (35 items). The Knowledge of Cognition component includes subcomponents of declarative knowledge (8 items), procedural knowledge (4 items), and conditional knowledge (5 items). The Regulation of Cognition component includes subcomponents of planning (7 items), comprehension monitoring (7 items), information management strategies (10 items), debugging strategies (5 items), and evaluation (6 items). Although Knowledge of Cognition and Regulation of Cognition look unbalanced in terms of number of items (17 versus 35 items), it should be noted that Regulation of Cognition has more subcomponents than Knowledge of Cognition. The MAI has been used with high school and college learners for research purposes and for validation studies of the instrument itself (e.g., Abe & Ida, 2010; Akin, Abaci, & Cetin, 2007; Miyamoto, 2010; Moxon, 2022; Xethakis, 2020; Yasuda, 2016).

Focusing on the MAI in Japan, Abe and Ida (2010) adapted the instrument for the Japanese population by translating it and then examining the psychometric properties of the scores of 283 private university students. Abe and Ida used exploratory factor analysis (EFA) as the method and obtained three factors: Monitoring, Control, and Metacognitive Knowledge. Notably, these constructs and the associated structure are different from the constructs and structure proposed by Schraw and Dennison (1994), and although some of the same items are used (or translations of them), the instrument is heavily revised and essentially different. Factors in Abe and Ida's model had moderate positive correlation with each other (Monitoring, Control and Knowledge of Cognition) and the model was confirmed using CFA. The results for the fit indexes were as follows: GFI, .865; AGFI, .842; and RMSEA, .048. Abe and Ida also obtained the results for Cronbach's α and the results were as follows: Metacognitive Knowledge, .749; Control, .788; and Monitoring, .878. On the basis of this evidence, Abe and Ida claimed validity for what is essentially a highly abbreviated (28 items) version of the MAI under a new factor structure.

There are two translations of the MAI in the Japanese literature. One is the translation conducted by Abe and Ida (2010) and referred to above, and the other was conducted by Miyamoto (2010). The translation of Abe and Ida has been cited and used in a variety of research (e.g., Moxon, 2022; Niwa & Yamaji, 2018; Xethakis, 2020; and Yasuda, 2016). Yasuda (2016), for example, used the translation of Abe and Ida to research the metacognitive abilities of Japanese learners of English using 435 Japanese university students. She employed three procedures: EFA; appropriateness of content; integration of content, and induced construct validity. She also arrived at an instrument with 28 items, but with six factors. Importantly, the model was not confirmed using CFA.

Outside of the Japanese population, the MAI also has been adapted for younger learners. Sperling, Howard, Miller and Murphy (2002) developed the Junior Metacognitive Awareness Inventory (JMAI) for such learners. In their study, two modified versions of the instrument were advanced: the JMAI (Version A) was for grades 3 through 5 and comprised 12 items; and the JMAI (Version B) was for grades 6 through 9 and comprised 18 items. Both versions were hypothesized to have two factors (Knowledge of Cognition and Regulation of Cognition) similar to the original version of the MAI proposed by Shraw and Dennison (1994). In subsequent research, Kim, et. al (2016) studied the factor structure of Version B in the English-speaking population using 6th- to 12th-grade students.

The JMAI, therefore, has a place in the literature and has been used for young learners in English-speaking countries. However, there are few studies reporting results for the instrument in other populations where it might be useful. The purpose of this study was to adapt the JMAI developed by Sperling, Howard, Miller and Murphy (2002) into the Japanese population and examine the psychometric properties of scores which it produces for validity.

The primary research question for the study was therefore whether a translated version of the instrument would produce a good-fitting model in the Japanese population. Were the model not to fit, attempts would then be made to arrive at a revised model which would fit.

Methodology

The methodology for the study is reported in terms of the instrument itself, the participants, and the procedures (which covers both administration issues as well as analytical procedures).

Instrument

The JMAI (Sperling, et. al. 2002) is based on the MAI (Schraw & Dennison, 1994), and both of them are hypothesized to have the same two first-order factors: Knowledge of Cognition and Regulation of Cognition. Because the JMAI has been adapted for younger users and comprises simple English sentences it does not correspond with translations of the MAI itself such as those undertaken by Abe & Ida (2010) and Miyamoto (2010). Thus, the original version of the JMAI offered by Sperling et al. was translated for the purposes of this study by the current author with the assistance of two other people.

The procedure for this was as follows. First, the current author and another Japanese-speaking academician who specialized in English education conducted the translation. This translation was then independently checked by another Japanese speaker with experience in translation. Finally, to check for content consistency of the Japanese version and also for comprehensibility, ten Japanese high school students checked to see whether there were any vague items which threatened understanding. These initial student participants were not included in the final sample.

Participants

A sample of data was obtained from 464 students at a public high school in western Japan. Due to missing responses (which were checked and determined to be at random and not systematic), 28 cases were deleted leaving a final sample of N=436. In terms of gender distribution, there were 242 females, 191 males, and 3 non-responses. All of the students were enrolled in the general course at the high school, and according to past patterns almost all of the students on this course would continue their education at university after graduating from high school.

Procedures

Data was collected from all 464 students at the same time at the beginning of August 2022 in the students' respective homeroom classes. They responded to a paper version of the JMAI (Sperling et al., 2002) which had been translated into Japanese according to the approach covered above. A five-point Likert scale was used with the following semantic anchors: 1 (never), 2 (seldom), 3 (sometimes), 4 (often), and 5 (always). Administration of the inventory took less than 10 minutes. Before answering the 18-item instrument, they read the instructions, and they were asked to respond to items in terms of the following frame of reference: "I [the author] would like to ask you about your English classes and study efforts." After data collection, all data was entered into Microsoft Excel Version 2019, and this spreadsheet data was subsequently imported into IBM Statistical Package for the Social Sciences (SPSS) Version 29.0 where most of the descriptive statistics and results for reliability were computed. In addition, and for the purpose of CFA, scores for the 18 items comprising the JMAI were analyzed using IBM AMOS Version 28.0.

Results

The results are reported in terms of the descriptive statistics (which provide information on the properties of the sample and the score distributions) and then model tests (which provide information on the primary model tested as well as subsequent models which were analytically derived for better fit).

Descriptive Statistics

The mean, standard deviation, critical ratios for skew, and critical ratios for kurtosis for the 18 items comprising the JMAI were calculated (Table 1). The highest mean was Item 12 (4.43), while on the other hand, the lowest one was Item 9 (2.8). As for the standard deviation, this showed a range from 0.867 to 1.25.

With respect to the results for the critical ratios (and these are interpreted in terms of absolute values with the negative or positive sign disregarded), the criteria used for judging adequacy were as follows. A value of 1.96 is associated with significant nonnormality in a statistical test at the 0.05 probability level. Therefore, this was rounded off to 2.0 and any value for the critical ratio for skew below this 2.0 level (absolute value) was interpreted as good. Statistical tests for nonnormality are, however, often regarded as overly sensitive, and because of this a more relaxed threshold was set of less than 3.0, and therefore values greater than 2.0 but less than 3.0 were interpreted as adequate. Values greater than 3.0 were considered problematic depending on how much higher they were than 3.0.

In terms of skew, Items 6, 7, 9, 16, and 17 had a value less than 2.0, which shows a good degree of normality. Items 10, 13, and 14 had results between 2.0 and 3.0 indicating some level of problem in this regard, but not an extreme problem. There were ten items (Items 1, 2, 3, 4, 5, 8, 11, 12, 15, and 18) exceeding the threshold of 3.0 indicating they were problematic in terms of normality depending on how much higher they were than 3.0. As for the critical ratios for kurtosis, Items 2, 3, 4, 13, and 15 had results less than 2.0 indicating good distributional properties for kurtosis. Items 8, 9, 10, 14, 16, and 18 were between 2.0 and 3.0, again indicating some level of problem but not an extreme problem. There were seven items (Items 1, 5, 6, 7, 11, 12, and 17) exceeding the threshold of 3.0 indicating a problematic result.

Table 1
Means, Standard Deviations, Critical Ratios for Skew, and Critical ratios for Kurtosis for the 18 Items Comprising the JMAI

Item Number	Item Content (Japanese Version)	Mean	Std. Deviation	Critical Ratio for Skew	Critical Ratio for Kurtosis
Item 1 (K)	学習していることがわかった瞬間を感じることができる。	4.100	0.963	-11.079	6.919
Item 2 (K)	学習しなければならないとき、進んで取り組むことができる。	3.750	0.976	-5.926	-0.065
Item 3 (K)	うまくいった学習方法を使うようにしている。	3.910	0.978	-6.979	1.759
Item 4 (K)	先生が生徒に学んでほしい学習のポイントがわかる。	3.780	0.999	-7.170	1.643
Item 5 (K)	学んでいるテーマについてすでに知識があるとき、学習がはかどる。	4.190	1.022	-12.603	8.057
Item 6 (R)	学習に取り組んでいるとき、理解を助けるために絵や図を描くことがある。	3.020	1.250	-0.791	-4.385
Item 7 (R)	授業が終わったとき、自分が学びたいと思ったことが身についているか振り返っている。	2.950	1.061	0.030	-3.278
Item 8 (R)	課題や問題に対して、取り組み方や解き方をいくつか考え、一番良い方法を選ぶ。	3.400	1.081	-3.220	-2.726
Item 9 (R)	授業が始まるとき、何を学ぶことになるのか予想する。	2.800	1.069	0.433	-2.638
Item 10 (R)	新しいことを学んでいるとき、しっかり学べているのか自問している。	3.180	1.041	-2.020	-2.083
Item 11 (R)	学習するとき、重要な情報には十分注意している。	4.160	0.885	-10.524	7.728

Item 12 (K)	学習する内容に興味があるとき、学習がはかどる。	4.430	0.867	-16.338	17.719
Item 13 (K)	苦手なことに対しては、得意な学習方法でカバーしている。	3.380	1.000	-2.656	-1.101
Item 14 (K)	学ぶ内容によって学習方法を変えている。	3.360	1.174	-2.549	-2.871
Item 15 (R)	制限時間のある課題には、時間を確認しながら取り組む。	3.810	1.137	-7.042	-0.873
Item 16 (K)	どんな学び方がよいか考えないで、学習に取り組むことがある。	3.280	1.056	-1.995	-2.648
Item 17 (R)	課題を終えた後、もっと簡単に終わらせることができる方法がなかったか考える。	3.000	1.143	-0.791	-3.508
Item 18 (R)	課題に取り組む前に、最初に手順を考えてから始める。	3.530	1.125	-3.816	-2.573

Note. (K) = Knowledge of Cognition; (R) = Regulation of Cognition

Reliability

Results were obtained for Cronbach's alpha. The coefficient for Regulation of Cognition was .80 and for Knowledge of Cognition was .81. Coefficient alpha, however, does not show that items are unidimensional (Cortina, 1993; Green, Lissitz, & Mulaik, 1977), and the results for the CFA tests below are more important.

Model Testing

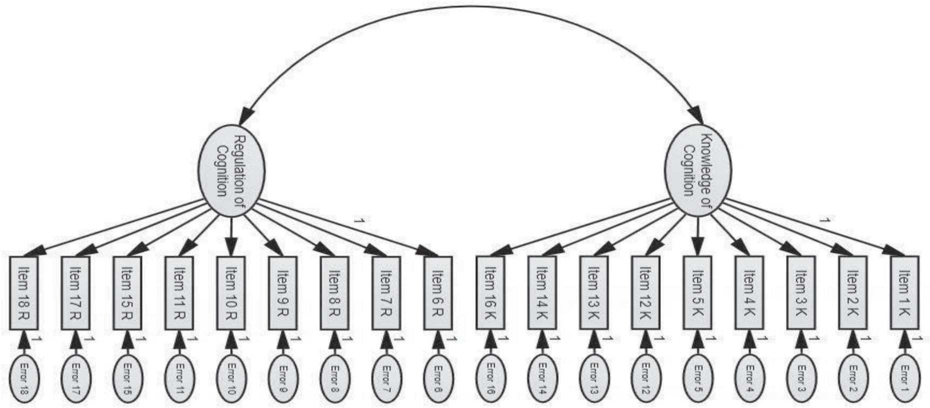
Model testing was conducted in two phases. The first model tested in the first phase addressed the primary research question which was whether a Japanese translation of the JMAI would produce scores with adequate fit when tested in the Japanese high school population. As the results below indicate, the fit was not adequate, and this led to a second phase of analysis (sequential CFAs) where a better fitting model was sought (with this ultimately turning out to be an eight-item abbreviated model) and then tested. The results of these sequential CFAs are therefore also reported, in turn, below.

Primary Eighteen-Item JMAI (Sperling et al., 2002)

The full eighteen-item model was originally advanced by the primary authors in their seminal article (Sperling, Howard, Miller & Murphy, 2002). The model comprises two factors, namely, Knowledge of Cognition and Regulation of Cognition with each of these factors indicated by nine items. The Knowledge of Cognition factor is indicated by Items 1, 2, 3, 4, 5, 12, 13,14, and 16. The Regulation of Cognition factor is indicated by Items 6, 7, 8, 9, 10, 11, 15, 17, and 18. The unidimensional configuration was tested which is normal practice. In the unidimensional model, each item is allowed to indicate only one factor, and error is not permitted to correlate, and the graphic representation of this model (produced in AMOS) is presented in Figure 1. According to the configuration represented in Figure 1, the model had 171 sample moments, 37 free parameters, and 134 degrees of freedom, thus meeting the necessary criterion of overidentification. All models tested in this paper (here and below) met this criterion.

Figure 1

AMOS 21 Graphical Representation of Eighteen-Item Unidimensional Model Tested Using CFA as the Method.



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

The Chi-square (χ^2) test was examined and the result was significant ($\chi^2 = 684.024$, df = 134, $p < 0.01$). In CFA, a significant result indicates that the underlying dimensionality of the scores in the data matrix is not aligned with the model specified in the CFA model (through AMOS). In other words, the expected model does not actually fit the dimensionality of the data. However, in CFA, it is well-recognized that the Chi-square test is overly sensitive (Byrne, 2001), and tends to almost always reject models even if the difference between the model itself and the underlying data is trivial. This is especially true if the sample is relatively large.

Because the Chi-square test usually rejects the model owing to its sensitivity, model fit indexes were also examined (as is normal practice) and the results for these are shown in Table 2. These are indexes and not statistical tests, and the results are interpreted using cutoff values recommended in the literature, usually Hu and Bentler (1999) but also other authors such as Byrne (2001). In terms of this literature the following cutoffs were adopted. For the TLI and CFI a value of >.90 was regarded as adequate, and following Hu and Bentler a value of >.95 was regarded as good. For the RMSEA a value of < .08 was regarded as adequate and, again following Hu and Bentler, a value of < .06 was regarded as good. For the SRMSR a value of < .08 was regarded as acceptable and good scores were considered as <.06. These criteria or cutoffs were decided in advance.

Table 2

Model Fit Indexes and Criteria of Japanese-Translated JMAI and its Results

TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(*p*)
.782	.809	.097	.0727	684.024	.000

Note. TLI: Tucker-Lewis index; CFI: comparative fit index; RMSEA: root mean squared error of approximation; SRMSR: standardized root mean square residuals; and Chi-square test statistic.

All of these indexes showed a lack of good fit in terms of the associated cutoff value, and in fact three of the results were less than adequate. Only the SRSMR was adequate with a result of .0727.

As Hu and Bentler (1999) point out, the results from the indexes need to be triangulated. In other words, an adequate or good result is required on all of them and not just one or two of them. From this point of view, the results obviously indicate that the model should be rejected. The two-factor, eighteen-item model did not fit the dimensionality of scores in the data matrix.

New Model for the Japanese-translated JMAI

Because the original 18-model did not fit the data, alternative models were explored and tested. This was done sequentially to reach a better-fitting model. The original model has nine items in Knowledge of Cognition, and nine items in Regulation of Cognition. Initially, these constructs were analyzed separately (as completely independent constructs) to look for improvements in their specification as independent constructs. These were then put together again in a two-construct (correlated model) which was further tested and adjusted. The separate examination of the independent constructs is reported immediately below, and then reporting of adjustments after they were put back together follows.

Knowledge of Cognition as an Independent Construct.

To begin with, the modification indexes were examined (as produced by AMOS) for the Knowledge of Cognition construct. The results for these indexes are displayed in Table 3, and they indicate improvement in model fit that would be obtained by correlating the various error estimates. In principle, error cannot be allowed to correlate because this sacrifices the unidimensional model, and because error is presumed to be random and therefore not able to be correlated (only systematically varying measurement points can potentially correlate). If the modification index value for the correlation of two error estimates is high, then this in fact shows that the error is not random, and that there is systematic variance hiding in the error estimate. This systematic variance is likely caused by some other construct (not specified in the model), and this unspecified construct is threatening the unidimensional properties of the items for which the error has been estimated. Items with error estimates which were highly correlated were therefore targeted for deletion.

Table 3

Modification Indexes for Knowledge of Cognition

Error for Item Number	Allowed to Correlate	Error for Item Number	M.I.	Change
Error14	<-->	Error16	26.084	-0.277
Error13	<-->	Error16	7.908	-0.121
Error13*	<-->	Error14	60.429	0.342
Error12	<-->	Error16	13.048	0.115
Error12	<-->	Error13	11.376	-0.086
Error5	<-->	Error14	5.342	-0.088
Error5	<-->	Error13	11.665	-0.103
Error5*	<-->	Error12	47.973	0.152
Error4	<-->	Error14	6.963	-0.104
Error4	<-->	Error12	8.187	-0.066
Error3	<-->	Error16	12.868	-0.143

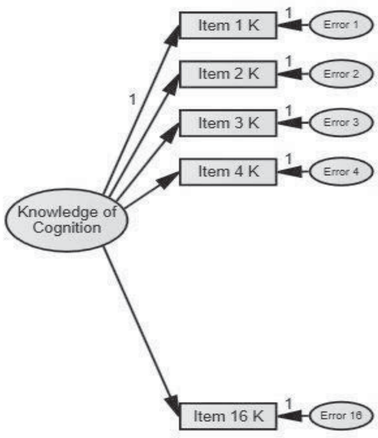
Error3	<-->	Error14	26.750	0.210
Error3	<-->	Error13	16.971	0.133
Error2	<-->	Error14	6.077	-0.099
Error2	<-->	Error13	6.684	0.082
Error2	<-->	Error12	10.293	-0.075
Error2	<-->	Error5	8.372	-0.079
Error2	<-->	Error4	10.488	0.092
Error1	<-->	Error2	7.748	0.070

Note. * The symbol “<-->” in Column 2 indicates that the error for the pair of items adjacent to each instance of the symbol is permitted to correlate and the associated modification index in the fourth column indicates the improvement in the overall model when this correlation is permitted. Only pairs where the MI is greater than 4.0 are shown in this table, and MI values less than 4.0 were suppressed because they were presumed to be trivial. The Error 13 and Error 14 pair as well as the Error 5 and Error 12 pair showed a high value for the modification index (MI) when permitted to correlate. In other words, when these pairs of error terms are allowed to correlate (which sacrifices the unidimensional model) the overall model improves a lot. Rather than sacrifice the unidimensional model by allowing these pairs of error terms to correlate, their respective items were instead targeted for deletion.

According to this procedure, Items 13 and 14 were deleted (their respective error estimates had high correlations), and also Items 5 and 12 were deleted (their respective error estimates also had high correlations). This resulted in the following abbreviated model for Knowledge of Cognition (Figure 2).

Figure 2

AMOS 21 Graphical Representation of Five-Item Unidimensional Model of Knowledge of Cognition Tested Using CFA as the Method



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

As the results displayed in Table 4 indicate, the fit of the model is exemplary for this five-item version of the Knowledge of Cognition Construct. The CFI and TLI were well above the threshold of .95, the RMSEA was below the threshold of .06 (indicating good fit), and the SRMSR was also very low. The Chi-square was non-significant. It is rare to get a non-significant result on the Chi-square because it is so sensitive and it almost always indicates significant differences between the model and the data. Therefore, the fact that such a rare result (non-significant result) was actually obtained for this model also supported good fit of the model to the data. It should be noted, however, that the result was non-significant at the $p < .01$ level rather than the $p < .05$ level (where it remained significant because the p value was .035 which is less than .05).

Table 4
Model Fit Indexes and Criteria of Five-Item Knowledge of Cognition Japanese-Translated JMAI and its Results

TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(p)
.972	.986	.057	.0325	12.017	.035

Regulation of Cognition as an Independent Construct.

Having examined the Knowledge of Cognition construct independently the analysis turned to a similar examination of the Regulation of Cognition construct. The analytical procedure for the Regulation of Cognition construct was exactly the same as for the Knowledge of Cognition construct (reported immediately above). The construct was treated separately, and modification indexes (Table 5) were examined for cases of high correlation between error estimates.

Table 5
Modification Indexes for Regulation of Cognition

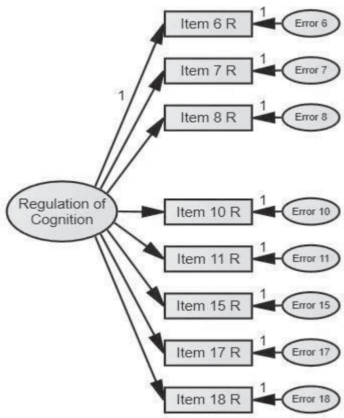
Error for Item Number	Allowed to Correlate	Error for Item Number	M.I.	Par Change
Error15	<-->	Error18	5.947	0.114
Error11	<-->	Error18	21.667	0.163
Error9*	<-->	Error18	5.287	-0.099
Error9*	<-->	Error17	12.065	0.163
Error9*	<-->	Error15	5.295	-0.101
Error9*	<-->	Error11	9.254	-0.099
Error9*	<-->	Error10	6.274	0.096
Error7	<-->	Error18	5.395	-0.093
Error7	<-->	Error17	6.960	-0.114
Error7	<-->	Error9*	5.902	0.090
Error6	<-->	Error18	6.941	-0.142
Error6	<-->	Error7	12.932	0.167

Note. * The symbol "<-->" in Column 2 indicates that the error for the pair of items adjacent to each instance of the symbol is permitted to correlate and the associated modification index in the fourth column indicates the improvement in the overall model when this correlation is permitted. Only pairs where the MI is greater than 4.0 are shown in this table, and MI values less than 4.0 were suppressed because they were presumed to be trivial. All pairs where one member of the pair is Error 9 are presented with an asterisk* here (a total of six pairs), and this indicates that Item 9 is problematic. Item 9 was therefore deleted.

According to the above indexes (Table 5), the first analytical point was that the error estimate for Item 9 (Error 9) was correlated to a reasonable degree with error estimates for six other items (Items 10, 11, 15, 17 and 18). The first step was therefore to remove this item. The model without this item was again tested (see Table 6 for results, and Figure 3 for model) and there was an improvement, but it was still not entirely satisfactory.

Figure 3

AMOS 21 Graphical Representation of Eight-Item Unidimensional Model of Regulation of Cognition Tested Using CFA as the Method



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

Table 6

Model Fit Indexes and Criteria of Eight-Item Regulation of Cognition Japanese Translated-JMAI and its Results (Item 9 Deleted)

TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(*p*)
.922	.944	.068	.043	60.619	.000

Subsequent modification indexes (see Table 7) indicated that the Item 6 and Item 7 pair were the most problematic. Furthermore, Item 7 also had some correlation with Item 18, which in turn also had a significant correlation with Item 11, and this was supported by results for modification indexes for regression coefficients. As a result, Item 7 and Item 18 were deleted. This resulted in the model seen in Figure 4.

Table 7

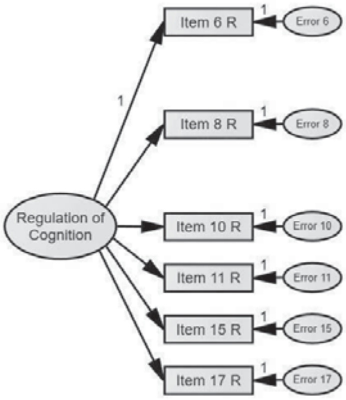
Modification Indexes for Knowledge of Cognition with Item 9 Removed

Error for Item Number	Allowed to Correlate	Error for Item Number	M.I.	Par Change
Error11	<-->	Error18*	14.839	0.131
Error8	<-->	Error17	4.808	0.098
Error7*	<-->	Error18	6.020	-0.099
Error6	<-->	Error18	8.228	-0.154
Error6	<-->	Error7*	16.314	0.193

Note. * The symbol “<-->” in Column 2 indicates that the error for the pair of items adjacent to each instance of the symbol is permitted to correlate and the associated modification index in the fourth column indicates the improvement in the overall model when this correlation is permitted. Only pairs where the MI is greater than 4.0 are shown in this table, and MI values less than 4.0 were suppressed because they were presumed to be trivial. The Error 6 and Error 7 pair had the highest MI when permitted to correlate, and furthermore Error 7 was involved in a further correlated relationship with Item 18 (and with Item 18 in turn having the second highest MI in its relationship with Item 11). Therefore, both Item 7 and Item 18 were deleted.

Figure 4

AMOS 21 Graphical Representation of Six-Item Unidimensional Model of Regulation of Cognition Tested Using CFA as the Method



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

The above model represented in Figure 4 and with Item 7 and Item 18 removed (and Item 9 removed at an earlier stage) was tested, and the results (as with the Knowledge of Cognition construct reported above) were exemplary (See Table 8). Both the TLI and CFI produced results well above the >.95 threshold and the RMSEA was also very good with a result well below the <.06 threshold (for good fit). The SRMSR also produced a good (very low) result. Importantly the p value was non-significant (even at the $p < .05$ threshold which is harder to satisfy than the $p < 0.01$ threshold) which also supported the case for exemplary fit.

Table 8

Model Fit Indexes and Criteria of Six-Item Regulation of Cognition Japanese Translated-JMAI and its Results

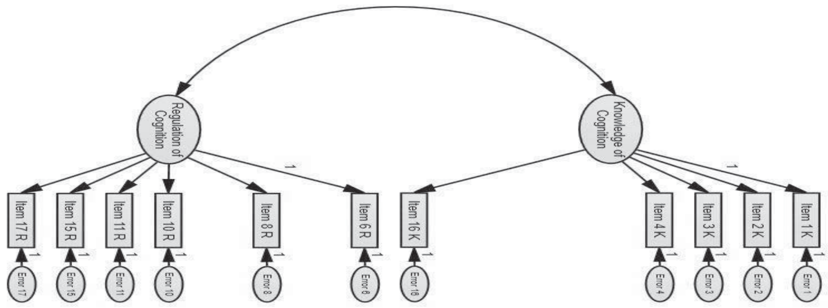
TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(p)
.980	.988	.035	.0271	13.690	.134

Revised Two-Construct Model.

The above two revised versions of the Knowledge of Cognition and Regulation of Cognition constructs (as independent constructs) were then again placed together into a single, two-factor, correlated model (see Figure 5).

Figure 5

AMOS 21 Graphical Representation of Eleven-Item Unidimensional Model of the JMAI (Using Modified Versions of the Regulation of Cognition and Knowledge of Cognition Constructs)



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

This model which comprised modified versions of the Knowledge of Cognition and Regulation of Cognition constructs (see results reported immediately above) was then tested using CFA as the method. The results are reported in Table 9 (see below).

Table 9

Model Fit Indexes and Criteria of Eleven-Item Unidimensional Model of the JMAI

TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(*p*)
.919	.936	.067	.0473	126.190	.000

The above results (see Table 9) are clearly an improvement on the results from the original 18-item model hypothesized by Sperling et al. (2002) and tested in this study. However, and while an improvement, these results are not exemplary (as was the case when the two constructs were treated independently and modified). The TLI, CFI and RMSEA all produce values which could be interpreted as adequate in terms of the criteria set in advance for this study, but none of them could be considered good in terms of those same criteria. This would indicate that the problem is not within each of the two component constructs (which performed well when examined independently) but rather between them. Therefore, the analytical procedure from here was to look for items which were not unidimensional; meaning, items which were caused by not only the construct they were hypothesized to be caused by, but also by the other construct which they were hypothesized to be independent of. Put another way, items were examined for being caused by both of the constructs rather than one of them. For this purpose, the modification indexes (see Table 10) were again examined.

Table 10

Modification Indexes for 11-Item Unidimensional Model of the JMAI (Using Modified Versions of the Regulation of Cognition and Knowledge of Cognition Constructs)

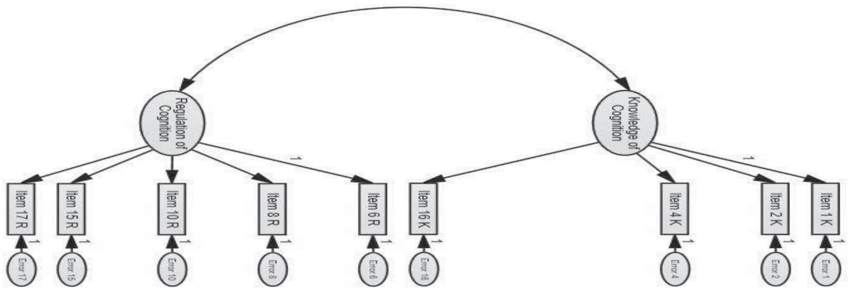
Error for Item Number	Allowed to Correlate	Error for Item Number	M.I.	Par Change
Error3	<-->	Knowledge of Cognition	8.811	-0.056
Error3*	<-->	Regulation of Cognition	10.190	0.040
Error3	<-->	Error16	8.327	-0.115
Error2	<-->	Error3	6.120	-0.068
Error17	<-->	Error1	4.708	-0.080
Error15	<-->	Error3	4.793	0.081
Error11*	<-->	Knowledge of Cognition	9.650	0.048
Error11	<-->	Regulation of Cognition	9.063	-0.029
Error11	<-->	Error17	5.007	-0.077
Error8	<-->	Error16	11.555	-0.149
Error8	<-->	Error3	9.563	0.101
Error8	<-->	Error1	4.704	-0.064
Error8	<-->	Error17	9.792	0.141
Error8	<-->	Error11	11.408	-0.093
Error6	<-->	Error16	6.018	-0.146
Error6	<-->	Error4	12.082	-0.149
Error6	<-->	Error8	5.231	0.112

Note. * The symbol "<-->" in Column 2 indicates that the adjacent non-observables (which could be error or constructs in this case) are permitted to correlate, and the value in the fourth column is the modification index (MI) which indicates improvement in the model when this correlation is permitted. Only pairs where the MI is greater than 4.0 are shown in this table, and MI values less than 4.0 were suppressed because they were presumed to be trivial. Both Error 3 and Error 11 are marked with an asterisk* for association with the construct for which they were not hypothesized to measure, and they were deleted for this association.

On the basis of the above results, it was noted that the error for Item 3 had a relationship with the Regulation of Cognition construct, and Item 3 was not hypothesized to be caused by the Regulation of Cognition construct but rather (and exclusively) by the Knowledge of Cognition construct. Therefore Item 3 was deleted. Furthermore, and in a similar fashion, the error for Item 11 displayed similar crossover relations with the construct it was not hypothesized to be caused by, and it was also deleted. The resulting modified model is presented below (Figure 6).

Figure 6

AMOS 21 Graphical Representation of Nine-Item Unidimensional Model of the JMAI (Using Modified Versions of the Regulation of Cognition and Knowledge of Cognition Constructs) and With Non-Unidimensional Items Removed



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

The results for the above model are presented below in Table 11, and are exemplary overall. The TLI and CFI produce results well above the >.95 threshold, and the RMSEA is well below the <.06 threshold. The *p* value has not reached the point of being non-significant, but a positive value is recorded at the third decimal place indicating improvement on the *p* value for the original 18-item model which recorded no value at the third decimal place and which was probably much smaller. The SRMSR was also good.

Table 11

Model Fit Indexes and Criteria of Nine-Item Unidimensional Model of the JMAI

TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(*p*)
.960	.971	.046	.0383	50.331	.003

While the above results were exemplary, the modification indexes were examined for remaining problems. These indexes are presented below (Table 12). It can be seen that there is significant correlation between the error estimates for Item 8 and Item 16 (with both of these items indicating separate constructs). To accommodate for this, Item 8 was deleted producing a balanced eight-item final model with each construct measured by four items.

Table 12

Modification Indexes for Nine-Item Unidimensional Model of the JMAI (Using Modified Versions of the Regulation of Cognition and Knowledge of Cognition Constructs)

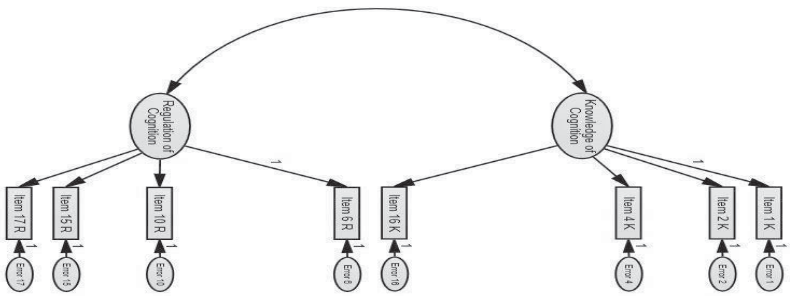
Error for Item Number	Allowed to Correlate	Error for Item Number	M.I.	Par Change
Error16	<-->	Knowledge of Cognition	4.076	0.065
Error16	<-->	Regulation of Cognition	6.601	-0.062
Error17	<-->	Error4	4.467	0.084
Error8*	<-->	Error16	11.967	-0.151
Error6	<-->	Error16	5.350	-0.136
Error6	<-->	Error4	8.401	-0.126

Note. * The symbol “<-->” in Column 2 indicates that the adjacent non-observables (which could be error or constructs in this case) are permitted to correlate, and the value in the fourth column is the modification index (MI) which indicates improvement in the model when this correlation is permitted. Only pairs where the MI is greater than 4.0 are shown in this table, and MI values less than 4.0 were suppressed because they were presumed to be trivial. Error 8 and Error 16 have the highest MI (with both of these indicating separate constructs), and so the item for Error 8 (Item 8) was deleted, and this also produced a balanced instrument with items measuring each construct.

The final model with Item 8 removed is represented in Figure 7. It is a balanced model with four items on each of the two constructs.

Figure 7

AMOS 21 Graphical Representation of Final Eight-Item Unidimensional Model of the JMAI (Using Modified Versions of the Regulation of Cognition and Knowledge of Cognition Constructs)



Note. The large ellipses indicate constructs (non-observables), the squares indicate items on the instrument (observables) and the small ellipses indicate error (non-observables). Each item is caused by only one construct and error is not permitted to correlate, and this is consistent with the specification of a unidimensional model.

The results for the final eight-item model are presented below in Table 13. The TLI and CFI produced results well above the >.95 threshold and the RMSEA is well below the <.06 threshold. The *p* value for χ^2 has not reached the point of being non-significant at the $p < .05$ level (95%), but has reached the point of being non-significant at the $p < .01$.

Table 13
Model Fit Indexes and Criteria of Final Eight-Item Unidimensional Model of the JMAI

TLI (>.95)	CFI (>.95)	RMSEA (<.06)	SRMSR (<.08)	X^2	(p)
.969	.979	.040	.0340	32.553	.027

The final model, the eight-item Japanese-translated JMAI, showed a slight but better improvement over the nine-item model. Both of the two factors, Knowledge of Cognition and Regulation of Cognition, have four items, and this shows a balance of operationalization.

Discussion

The results for the descriptive analysis didn't show good score distributions for many of the items in the full 18-item, Japanese-translated JMAI (Table 1). There were problems with ten items (at the 3.0 threshold) with regard to the critical ratio for skew and with seven items with regard to the critical ratio for kurtosis. In the research conducted by Kim, et al., (2016), the values for skew itself, rather than the critical ratio for skew are interpreted, and reported to be acceptable. However, if these values for skew reported by Kim et al. are divided by the standard error for skew (also reported by Kim et al. in an adjacent column of the reporting table), the critical ratios can be computed, and these computations indicate that similar problems existed with the distributional properties of items on their version of the instrument for many items.

In terms of the confirmatory aspect of this study, the 18-item Japanese-translated MAIJ was tested for model fit with AMOS using CFA as the method. The results for this 18-item model were not satisfactory, and this was inferred on the basis of the Chi-square test (which was significant) and the four indexes recommended by Hu and Bentler (1991), namely, the TLI, CFI, RMSEA and SRMSR. However, other than for the result for the SRMSR, the results for the model were not satisfactory (see Table 2 for these results). Consistent with the model hypothesized for the MAI by Schraw & Dennison (1994), the model hypothesized for the JMAI by Sperling, et. al. (2002) was specified to comprise two factors, namely Knowledge of Cognition and Regulation of cognition, with this theoretical distinction being based on the work of Brown (1978). Therefore, in this research, the instrument was tested using this hypothesized model, and the model was rejected.

Given that the model (Figure 1) suggested by the literature did not fit (see Table 2), the next steps were to examine items which were problematic with the aim of arriving at a good-fitting model (see Table 3). Initially, each subscale, Knowledge of Cognition and Regulation of Cognition, was examined independently for items not contributing to proper measurement, and modification indexes were used for this.

With respect to the Knowledge of Cognition subscale, Items 13 and 14 as well as Items 5 and 12 were eliminated because of high error correlations. High correlation between the error for particular items indicates that some of the error is systematic, and therefore not error (error should be random and produce low correlations). This usually suggests that latent constructs exist which are not represented in the model, and these are causing some of the "error" in the items for which the error correlates highly. The remaining five-item model (after the above items were deleted) was examined again (as a single unidimensional construct), and it showed a good fit (see Table 4 and Figure 2). The results for both the TLI (.972) and CFI (.986) were comfortably above the .95 threshold adopted from Hu and Bentler (1999). In addition, the value for the RMSEA (.057) just met the threshold for good fit being below .06. This index gives better results for simpler models, and the model here is very simple (one construct with five items indicating it). Finally, the SRMSR which is a measure of residuals after the model is fitted was very low (.0325) and this was a good result. The value for the Chi-square was 12.017 and the p value was .035 (<.05 but > .01). This means that at the .05 level, the model was significantly different from the data matrix, and at the .01 level it was not significantly different from the data matrix. This shows that even with the Chi-square, which is very sensitive and almost always finds all models significantly different from the data matrix (and thus the use of model indexes instead), the model is still approaching the point where the null hypothesis is accepted at the .05 level, and in fact the null hypothesis was accepted at the .01 level. In CFA, and different to typical statistical research, accepting the null hypothesis is preferable because this shows that the model is very much the same as the dimensionality of data in the data matrix. Unfortunately, accepting the null hypothesis is very rare because the Chi-square is so sensitive that it almost always finds a difference between the matrix and the model even if that difference is small. Nonetheless, in this case the null hypothesis is accepted at the .01 level and is close to being accepted at the .05 level. Overall, these results taken in triangulation show that this 5-item Knowledge of Cognition scale had good fit when analyzed separately from the Regulation of Cognition scale.

With respect to Regulation of Cognition, the error for Item 9 correlated considerably with the error for six other items (Table 5), especially with Item 17, and it was removed. Following this, the 8-item model (with Item 9 deleted) was tested again as a single unidimensional construct (Figure 3) and it improved in terms of model fit (see Table 6). The results for the TLI (.922) and CFI (.944) were between .90 and .95, with the level of >.90 being set by this study as the threshold for an acceptable fit and the level of >.95 as the threshold for good fit. This indicates that for these two indexes, the fit improved with the deletion of Item 9 but was still not good. The value for the RMSEA (.068) was between the values of .06 and .08 with the value of <.06 being the threshold for good fit and the value of <.08 being the threshold for acceptable fit. This value, therefore, is acceptable but still in need of improvement. Finally, the value for the SRMSR (.043) was good. The value for the Chi-square was 60.619 and the p value associated with this was .000 (<.01). This means that in terms of the Chi-square, the model was rejected and does not fit the data matrix, but of course, the usual criticism of the Chi-square applies which is that it is too sensitive and always rejects models. Thus, the values for the indexes are more important which here suggests that the eight-item model has improved but is still not as good as it should be.

The modification indexes for the eight-item model of Regulation of Cognition were therefore examined again (Table 7) and there appeared to be high correlations within Error 6 and 7, and Error 11 and 18. Therefore, Item 7 and Item 18 were deleted leaving only one of each pair (see Figure 4 for the resulting model). The remaining six-item model was tested again (as a single unidimensional construct), and it showed a good fit for the model (see Table 8). The results for the TLI (.980) and CFI (.988) were well above the threshold of .95 for good model fit. The value for the RMSEA (.035) was well below .06 which indicates good fit also. Finally, the value for the SRMSR (.0271) was far below the threshold of <.08 indicating good fit. The value for the Chi-square was 13.690 and the associated *p* value was .134 which is greater than both .05 and .01 indicating acceptance of the null hypothesis at both levels. This means that in spite of the over-sensitivity of the Chi-square to differences between the model and the data matrix, the null hypothesis was still able to be accepted which indicates very good fit for the eight-item model of the Regulation of Cognition subscale.

These two revised constructs (the five-item Knowledge of Cognition and six-item Regulation of Cognition) were put back together (see Figure 5) in a combined model which was tested using CFA. The results showed improvement (see Table 9) over the original 18-item model tested initially. The results for the TLI (.919) and the CFI (.936) were acceptable but not good. This is because they were greater than .90 (the threshold for acceptable fit) but less than .95 (the threshold for good fit). The value for the RMSEA (.067) was almost below .06 (the threshold for good fit) but not quite. Therefore, this value indicated room for improvement. The value for the SRMSR (.0473) was quite good. In terms of the Chi-square, the value returned was 126.190 and the associated *p* value was .000 (<.01) which was significant. As stated with similar cases above, the Chi-square overly rejects models, and so it is sufficient in this case to only consider the values for the indexes. Overall then, the 11-item model was much better than the original 18-item model, and refinements made when looking at the two constructs independently (discussed above) led to this improvement, however, further improvement was necessary. When two constructs which have been improved independently are put back together, there is the possibility of problems because some items may be indicated by constructs which they are not supposed to measure (in this case the other of the two constructs we are interested in, either Regulation of Cognition or Knowledge of Cognition).

The above eleven-item model was re-evaluated for potential modifications. The error for Item 3 and Item 11, showed correlations with both Knowledge of Cognition and Regulation of Cognition (see Table 10). Error should not be caused by other factors and should be random. Therefore, Item 3 and 11 were deleted. This resulted in a nine-item model (Figure 6) which was tested using CFA. It showed further improvement over the eleven-item model (see Table 11). The results for the TLI (.960) and CFI (.971) were securely above the threshold of >.95 which was interpreted, therefore, as good. The value for the RMSEA (.46) was a large improvement on the same value for the eleven-item model. The value for the index was well below the threshold of <.06 for a good fit. Finally, the value returned for the SRMSR (.0383) was also well below the threshold of <0.8 for good fit. In terms of the Chi-square, the value returned was 50.331 and the associated *p* value was .003. This value was less than both .05 and .01 indicating a difference between the model and the matrix. The same analysis is offered (as above) that this is the result of the over-sensitivity of the Chi-square. Notably, a non-zero value was recorded at the third digit (AMOS presents only up to the third decimal point) which was not the case with the eleven-item model which returned a value of .000 indicating a value less than .003. This is because the value for the Chi-square has improved moving from 126.190 for the eleven-item model to 50.331 for the nine-item model. Overall, the model fit here was good, but as explained below, it was noticed that there was the possibility of one further change.

While the improvement was significant, it was examined again in terms of modification indexes, and the error for Item 16 showed correlations with both Knowledge of Cognition and Regulation of Cognition, and there was also a correlation between the error for Items 8 and 16 (see Table 12). Therefore, Item 16 was deleted. This eight-item (four items on each of two constructs) model (see Figure 7) was again tested for the final time and the results were very good (see Table 13). The results for the TLI (.969) and CFI (.979) were both well above the threshold of >.95, and also showed a slight improvement on the nine-item model. The value for the RMSEA (0.40) was also good and showed a slight improvement on the nine-item model. Finally, the value returned for the SRMSR (.0340) was also good. The value for the Chi-square was 32.553 and the associated *p* value was .027. This value is significant at $p<.05$ and non-significant at $p<.01$. In other words, in terms of the .05 level, the null hypothesis is rejected, and in terms of the .01 level, the null hypothesis is accepted. While the null hypothesis was still not accepted at the .05 level, this result is still a big improvement on the nine-item model where the null hypothesis was rejected at both the .05 and .01 level. Furthermore, given the oversensitivity of the Chi-square and the high likelihood that it will lead to rejection of the null hypothesis in most cases, this is actually a good result, because the *p* value has moved into a range where acceptance of the null hypothesis occurs at the .01 level and is close to occurring at the .05 level.

Through the examination conducted in this study, preliminary evidence is provided for a balanced eight-item version of the Japanese-translated JMAI (four items on each of two constructs). The original JMAI (Sperling, Howard, Miller, & Murphy, 2002) is an eighteen-item version (nine items on each of two constructs), which was based on the MAI (Schraw and Dennison, 1994). The advantage of the abbreviated model is that students in junior high school or high school are able to answer the inventory in a short time, which enables them to avoid getting tired or bored doing it. On the other hand, operational bandwidth is sacrificed, that is, the less the number of items measuring a construct, the narrower the actual measurement of it. Most importantly in terms of future research, the CFA tests conducted in this study (other than the first test of the original model) to produce a better-fitting model were conducted in the same data set in which the modifications were diagnosed. This means these are not a priori tests. To address this, the final and best-fitting model arrived at in this study needs to be tested in new samples of data from the same population in future studies.

References

阿部真美子［Abe, M.］, & 井田政則［Ida, M.］.（2010）. 成人用メタ認知尺度の作成の試み：Metacognitive Awareness Inventory を用いて［An attempt to construct an adult metacognition scale based on the Metacognitive Awareness Inventory］, *立正大学心理学年報* ［*The Journal of Psychology, Rissho University*］, *1*, 23-34.

Akin, A., Abaci, R., & Cetin, B.（2007）. The validity and reliability of the Turkish version of the Metacognitive Awareness Inventory. *Educational Sciences: Theory and Practice, 7*(2), 671–678.

Brown, A. L.（1978）. Knowing when, where, and how to remember: A problem of metacognition. *Advances in Instructional Psychology, 1*, 77–165.

Byrne, B. M.（2001）. *Structural equation modeling with AMOS*. London: Lawrence Erlbaum Associates.

中央教育審議会［Central Council for Education］.（2021）.『「令和の日本型学校教育」の構築を目指して～全ての子供たちの可能性を引き出す、個別最適な学びと、協働的な学びの実現～（答申）［Toward the construction of Japanese-style school education in the Reiwa Era: Optimal individual and collaborative learning that brings out the potential of all children]』Retrieved from https://www.mext.go.jp/content/20210126-mxt_syoto02-000012321_2-4.pdf（accessed 2022.9.25）

Çubukçu, F.（2008）. How to enhance reading comprehension through metacognitive strategies. *Uluslararası Sosyal Ara_tırmalar Dergisi*［*The Journal of International Social Research*］, *1*（2）, 83-93.

Cortina, J. M.（1993）. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*（1）, 98-104.

Flavell, J. H.（1976）. Metacognitive aspects of problem solving. In L.B. Resnick（Ed）, *The Nature of Intelligence*（pp. 231-235）. Lawrence Erlbaum Associates.

Flavell, J. H.（1979）Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

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*, 827-838.

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

Kim, B., Zyromski, B., Mariani, M., Lee, S. M., & Carey, J. C.（2017）. Establishing the factor structure of the 18-item version of the junior metacognitive awareness inventory. *Measurement and Evaluation in Counseling and Development, 50*（1-2）, 48-57.

Louca, E.（2003）. The concept and instruction of metacognition. *Teacher Development, 7*（1）, 9-30.

文部科学省［Ministry of Education, Culture, Sports, Science and Technology］.（2018a）『高等学校学習指導要領［Course of Study for High School］』Retrieved from http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2018/07/11/1384661_6_1_2.pdf（accessed 2022.9.25）

文部科学省［Ministry of Education, Culture, Sports, Science and Technology］（2018b）『高等学校学習指導要領解説総則編［Explanatory Comments for Course of Study for High School - in General Rules Part］』Retrieved from https://www.mext.go.jp/content/20211102-mxt_kyoiku02-100002620_1.pdf（accessed 2022.9.25）

文部科学省［Ministry of Education, Culture, Sports, Science and Technology］（2021）「学習指導要領の趣旨の実現 に向けた個別最適な学びと協働的な学びの 一体的な充実に関する参考資料［Reference Materials toward the Realization of the Objectives of the Course of Study Concerning the Integrated Enhancement of Individual Optimal Learning and Collaborative Learning］」Retrieved from https://www.mext.go.jp/content /20210428-mxt_kyoiku01-00014639_13.pdf（accessed 2022.9.23）

Miyamoto, T.（2010）. Metacognition over learning attitudes in EFL learning. In A. M. Stoke（Ed.）, JALT2009 Conference Proceedings. Tokyo: JALT.

Moxon, J.（2022）. Psychometric evaluation of abridged versions of the Metacognitive Awareness Inventory in the Japanese population. *Current Psychology, 42*, 13228–13230

丹羽量久［Niwa, K］, 山地弘起［Yamaji, H.］. & Bernick, P. J.（2018）. 成人用メタ認知尺度

Metacognitive Awareness Inventory の邦訳と活用 - 大学初年次学生のメタ認知と情報基礎科目における学習活動との関係［Metacognitive Awareness Inventory for Adults - metacognition of first-year university students and its relation to learning activities in basic information courses］. *情報コミュニケーション学会研究報告*［*Research Report of Japan Association for Communication, Information and Society*］, *15*(3), 39-46.

Okamoto, M. (1991). Influence of metacognitive knowledge and aptitude in solving word problems in arithmetic. *The Japanese Journal of Developmental Psychology, 2*(2), 78-87

岡本真彦［Okamoto, M.］. (2012). 教科学習におけるメタ認知－教科学習のメタ認知知識と理解モニタリング―［Metacognitive activities in school learning: a review of metacognitive knowledge and comprehension monitoring］. *教育心理学年報*［*Annual Report of Education Psychology in Japan*］, *51*, 131-142.

Schraw, G. & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460–475.

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*, 51-79.

Xethakis, L. (2020). The psychometric properties of the Metacognitive Awareness Inventory in the Japanese EFL context. *熊本大学社会文化研究*［*Kumamoto University Studies in Social and Cultural Sciences*］, *18*, 121-146.

安田利典［Yasuda, T.］. (2016). 日本人英語学習者を対象としたメタ認知尺度の基盤作成［Preliminary design of a metacognitive scale for Japanese learners of English］. *KATE Journal, 30*, 57-70.

Appendix A

English version of the Junior Metacognitive Awareness Inventory (JMAI; Sperling, Howard, Miller & Murphy, 2002).

We are interested in what learners do when they study. Please read the following sentences and circle the answer that relates to you and the way you are when you are doing school work or home work. Please answer as honestly as possible.

		1 Never	2 Seldom	3 Sometimes	4 Often	5 Always
1	I know when I understand something.					
2	I can make myself learn when I need to.					
3	I try to use ways of studying that have worked for me before.					
4	I know what the teacher expects me to learn.					
5	I learn best when I already know something about the topic.					
6	I draw pictures or diagrams to help me understand while learning.					
7	When I am done with my schoolwork, I ask myself if I learned what I wanted to learn.					
8	I think of several ways to solve a problem and then choose the best one.					
9	I think about what I need to learn before I start working.					
10	I ask myself how well I am doing while I am learning something new.					
11	I really pay attention to important information.					
12	I learn more when I am interested in the topic.					
13	I use my learning strengths to make up for my weaknesses.					
14	I use different learning strategies depending on the task.					
15	I occasionally check to make sure I'll get my work done on time.					
16	I sometimes use learning strategies without thinking.					
17	I ask myself if there was an easier way to do things after I finish a task.					
18	I decide what I need to get done before I start a task.					

Appendix B

Japanese version of the Junior Metacognitive Awareness Inventory (JMAI; Sperling, Howard, Miller & Murphy, 2002).
Note. The domain was restricted to English study specifically for this administration, and this is reflected in the instruction to participants which precedes the items themselves.

あなたの英語の授業や勉強への取り組みについてうかがいます。下記のそれぞれのアンケート項目に対して、1～5のいずれか1つを○で囲んでください。

		1 あてはまらない	2 あまりあてはまらない	3 どちらでもない	4 ややあてはまる	5 あてはまる
①	学習していることがわかった瞬間を感じることができる。	1	2	3	4	5
②	学習しなければならないとき、進んで取り組むことができる。	1	2	3	4	5
③	うまくいった学習方法を使うようにしている。	1	2	3	4	5
④	先生が生徒に学んでほしい学習のポイントがわかる。	1	2	3	4	5
⑤	学んでいるテーマについてすでに知識があるとき、学習がはかどる。	1	2	3	4	5
⑥	学習に取り組んでいるとき、理解を助けるために絵や図を描くことがある。	1	2	3	4	5
⑦	授業が終わったとき、自分が学びたいと思ったことが身についているか振り返っている。	1	2	3	4	5
⑧	課題や問題に対して、取り組み方や解き方をいくつか考え、一番良い方法を選ぶ。	1	2	3	4	5
⑨	授業が始まるとき、何を学ぶことになるのか予想する。	1	2	3	4	5
⑩	新しいことを学んでいるとき、しっかり学べているのか自問している。	1	2	3	4	5
⑪	学習するとき、重要な情報には十分注意している。	1	2	3	4	5
⑫	学習する内容に興味があるとき、学習がはかどる。	1	2	3	4	5
⑬	苦手なことに対しては、得意な学習方法でカバーしている。	1	2	3	4	5

⑭	学ぶ内容によって学習方法を変えている。	1	2	3	4	5
⑮	制限時間のある課題には、時間を確認しながら取り組む。	1	2	3	4	5
⑯	どんな学び方がよいか考えないで、学習に取り組むことがある。	1	2	3	4	5
⑰	課題を終えた後、もっと簡単に終わらせることができる方法がなかったか考える。	1	2	3	4	5
⑱	課題に取り組む前に、最初に手順を考えてから始める。	1	2	3	4	5

「「ジュニア層を対象とするメタ認知意識度インベントリー日本語版」から得られたスコアに見られる心理測定学的特性」

上野　令資

本研究の目的は、「ジュニア層を対象とするメタ認知意識度インベントリー日本語版」（Junior Metacognitive Awareness Inventory（JMAI): Sperling, Howard, Miller, & Murphy, 2002）のモデル適合性を検証することである。JMAI は全 18 個のアンケート項目から構成される。調査は 464 人の日本人高校生に実施した。モデル適合性の検証には確認的因子分析を用いた。解析で得られた結果から、JMAI のモデル適合度は不十分であることが判明した。修正指数に基づきモデルに一連の修正を加えた結果、いずれも 4 つの項目を持つ 2 つの構成要素から成る、バランスのとれたモデルを得ることができた。本研究において最終的に得られたこのモデルは適合度が高く、オリジナル版よりも項目数も少なく、修正版を用いての調査も短時間で可能となる。今後は新たな被験者からさらにデータを収集し検証を重ねる必要がある。