Development of the Associative Memory Intelligence Test for High School Students Based on Cattel-Horn-Carrol Theory

Hidayati[a] and Dewi Maulina[b*]

[a]Faculty of Psychology, Universitas Indonesia, Depok, Indonesia; [b]Psychology Research Method Department, Faculty of Psychology, Universitas Indonesia, Depok, Indonesia

*Corresponding author:
Dewi Maulina
Psychology Research Method Department
Faculty of Psychology, Universitas Indonesia
Jl. Lkr. Kampus Raya, Depok, Jawa Barat
Indonesia, 16424
Tel.: +62 217270004
Email address: dewi_maulina@yahoo.com

Development of the Associative Memory Intelligence Test for High School Students Based on Cattel- Horn-Carrol Theory

The objective of the study was to develop and validate the Associative Memory Test (*AM* Test), as a subtest of the intelligence test (*Tes Inteligensi Siswa SMA, TISS),* based on Cattel-Horn-Caroll (CHC) theory of intelligence. The *AM* Test was used as a psychological assessment tool for determining high school students' specializations. A total of 96 first-year students from Jakarta specializing in natural and social sciences participated in the study. In the *AM* Test, the participants were asked to memorize pairs of pictures and words and then recall the words based on the picture. The reliability testing showed that the *AM* Test had a good internal consistency. The construct validity testing showed that the *AM* Test can effectively measure associative memory because it correlated significantly with the *IST* test, which can also associative measure memory. Meanwhile, the item analysis indicated that most items had a low difficulty level. Most of the items also had good discriminative index based on the discrimination items. The findings of the study demonstrate that the *AM* Test is a promising test that can be used in assessing high school students in a classroom setting. However, item revisions, better retesting, and improved sample number are still needed to increase the quality of the *AM* Test and further develop its norms.

Keyword: intelligence test, associative memory, CHC theory

Introduction

Currently, Indonesia's Ministry of Education has changed the curriculum for its educational systems. Hence, in the past years, a significant transformation from the old to the new curricula has occurred in the high school level. The changes require students to choose their specialization when entering high school. Referring to Government Regulation Number 64, Year 2014, a "specialization" refers to a specific program to accommodate a student's ability and interest in a particular discipline. The specializations are split into three groups: natural science, social science, and languages (Kementrian Pendidikan dan Kebudayaan, 2014). Students must choose one from these three specializations according to their own interests and abilities.

Ability and interest significantly influence a student's performance. Consequently, the failure to explore a student's interest can lead to the absence of motivation (Legault, Pelletier & Green-Demers, 2006). Psychological assessment is an alternative approach in determining students' abilities and interests. In a psychological assessment involving several methods, such as interview, observation, psychological testing, and others, we integrate psychological data to evaluate one's psychological condition (Cohen, Swerdlik & Sturman, 2013). A psychological test, specifically an intelligence test, is a powerful tool providing a comprehensive description, which can be used to give recommendations about school specializations for students. Previous studies found a significant correlation between a student's intelligence and academic achievement at school. According to Watkins, Lei, and dan Canivez (2007), there exists a significant correlation between intelligence and academic achievement. In addition, students with

a high level of intelligence have a 90% chance to garner good academic achievements at school (Deary, Strand, Smith & Fernandez, 2007).

School psychologists in Indonesia use different intelligence tests, including the *Tes Kemampuan Differensial 2* (*TKD 2*), *Intelligenz Structure Test* (*IST*), and *Tes Kolektif Inteligensi Indonesia* (*TIKI*). However, most of the test items have been leaked and can already be accessed from various sites online, such as in psikotesyuk.com. Consequently, this has influenced the reliability and validity of such tests. In addition, the items of previous intelligence tests are already obsolete and need to be updated. Some instruments also do not measure all the necessary abilities needed in a high school context. For example, *TKD* 2 does not measure memory ability, which is very important for learning in high school (Wilhite, 1990). Therefore, developing a new intelligence test that is particularly aimed for high school referral is an important research undertaking.

In the present study, we used Cattel-Horn-Caroll (CHC) theory as the basis upon which a new intelligence test was constructed. CHC is a contemporary theory that describes the structure of human intelligence in detail (Flanagan & Dixon, 2013) and has gained the approval of experts in the field (McGrew, 2009). CHC structures consist of three strata: the first stratum is called *narrow ability,* which contains specific abilities and is very useful in daily problem solving, the second stratum is broad ability, and is more specific than the third stratum, which is general ability (Carroll, 1993).

Several psychological tests, such as Woodcock Johnson IV (WJ-IV) and Weschler Memory Scale (WMS), also used CHC theory as the basis of their design. However, these tests are not appropriate in assessing a student's ability for the purpose of choosing a specialization. According to McGrew, Shrank, and Laforte (2014), WJ-IV must be administered individually and is not effective and efficient in testing students in a large group setting. Meanwhile, WMS is best for clinical usage, such as for brain impairment diagnosis (Drozdick, Holnack & Hilsabeck, 2011). Therefore, it is necessary to develop a new intelligence test that can provide accurate specialization-related recommendations for high school students and can be administered in groups. The new test is called the *Tes Inteligensi Siswa (TISS) SMA.*

The new intelligence test (*TISS*) should consist of several broad abilities that are needed in a specialization referral. One important broad ability is long-term storage and retrieval (*Glr*). According to Solso (2005), long-term memory is defined as a storage to keep the information for long periods of time. It has an unlimited space, so the information stored here is kept permanently and coded based on visual, sound, and semantic stimuli (Eysenck & Keane, 2005).

Harel et al. (2014) stated that memory capacity and knowledge increase according to age. Therefore, older students have more information stored in their long-term memory compared with younger students. The former can effectively use several retrieval strategies, such as rehearsal, meaningful learning, organization, elaboration, and visual imagery. Such strategies enable older students to encode and retrieve information easily (Ormrord, 2003).

From the CHC point of view, *Glr* is defined as "the ability to learn, store, consolidate, and retrieve information over periods of time measured in minutes, hours, days, and years" (McGrew2005; McGrew 2009; McGrew 2014). Associative memory is a kind of narrow ability

of *Glr*, which is very important in learning activities. This ability is placed within the realm of learning efficiency. Senior school students are supposed to memorize and master great amounts of lesson materials; therefore, it is important for them to learn effectively. According to McGrew (2014, p.1), associative memory refers to "the ability to remember previously unrelated information as having been paired." There are two phases in measuring associative memory (Caroll, 1993). In the first phase, participants are given some materials for them to study. The materials consist of several paired items, which reflect stimulus and response, respectively. In the second phase, participants are asked to remember the response to a stimulus. Some participants have the ability to encode information effectively by relying on imagery memory and mnemonics. The second phase is very important in observing participants' ability to remember and study unrelated materials through association.

According to O'donnell, Reeve, and Smith (2012), information would be easier to remember if frequently associated with other information. Therefore, high school students must have associative memory because teaching materials used in high school are more complicated than those used in the middle school. Given that one material has an interrelationship with other materials in the school lesson, association ability is thus highly needed. Referring to the 2013 curriculum implemented in Indonesia, the Ministry of Education integrated the thematic learning method in a certain field of science with other disciplines (Kementerian Pendidikan dan Kebudayaan, 2014). Students are supposed to be able to make associations between the concepts of knowledge taught in various subjects in school. Therefore, the ability of associative memory is important in helping the learning effectiveness either in natural science, social science, or languages.

The current study is a preliminary study that aims to develop an associative memory test (*AM* Test), as a subtest of the new intelligence test (*TISS*) to measure memory ability among high school students. As a new test, the *AM* Test could be a reliable test with appropriate items as well as good internal consistency and validity in measuring associative memory ability.

Methods

Participant
A total of 96 first-year high school students from three different schools in Jakarta participated in the study. The participants' ages ranged from 14 to 17 years old (M = 16.38, SD = 0.26). There were almost equal proportions of male (53.1%) and female (46.9%) students. The participants consisted of those having natural science (66.7%) and social science (33.3%) specializations. The participants were recruited by using convenience sampling technique.

Measures
The *AM* Test is a subtest of the *TISS*, which uses CHC theory as its main basis. The *AM* Test can be categorized as a power test that is administered in groups. The Test has two indicators that indicate associative memory ability: remembering unrelated information and retrieving the previously paired information. Each item of the *AM* Test is presented as a pair of images and words. Stacy, Ames, aand Grenard (2006) stated that when a concept is presented, it is more likely associated with a relevant image rather than a word. A previous study also found that the

interaction between a word and a picture could facilitate the retrieval of information in the learning process (Lippman and Shanahan, 1973).

As a power test, the item difficulty level of the *AM* Test is determined by the types of words (abstract or concrete) and types of images (uninteractive or interactive) employed. According to Paivo (1965), concrete nouns are easier to memorize compared with abstract nouns, because the former can elicit images from a word stimulus. On the other hand, an interactive image can integrate the two items in some mutual or reciprocal action (Lutz & Lutz, 1977).

In the present study, the administration of the *AM* test was divided into two phases. In the first phase, the participants were presented several pairs of items and were requested to remember these items for a certain amount of time. After that, they received an answer sheet containing the stimuli (pictures) of the pairs of items presented in the previous phase. Then, the participants were requested to recall the pairs of items (word) listed on the answer sheet.

Table 1
AM **Test Specification**

Item Category	Criterion Level of Difficulty	Final Item	Original Item
Easy	Interactive image (image or concrete noun)	4	6
Intermediate	Concrete noun with an uninteractive image	7	12
Difficult	Image and abstract noun	4	7
Total		15	25

Procedure

The procedure of test construction followed several steps according to Cohen, Swerdlik, and Sturman (2013). The first step started from the conceptualization of the instrument. Then, we pooled the test items so that they were 1.5 times more than the original number of targeted items. Hence, there were 25 original items. Next, we relied on expert judgment and received suggestions regarding the content of the test, level of difficulty, and test instruction. The readability of items was tested by recruiting 10 high school students from Jakarta and Bogor. The purpose was to determine the time needed to finish the test and whether the participants understood the test instruction and test items. After this process, we conducted a trial with 84 high school students from Depok City. The results of the trial are presented in the table below.

Table 2
Descriptive Analysis of the *AM* Test Trial Results

Descriptive Statistics	
Mean	13.68
Standard Deviation	5.94
Max	25
Min	1
Time average (minutes)	5
N finished in 5th min	87%
Coefficient alpha	0.89

The average score obtained from the *AM* Test trial was 13.68, indicating that the *AM* Test can be categorized as having a medium level of difficulty. There were 87% of participants who completed the task within 5 minutes. From this result, we considered 5 minutes as the optimal time limit for the *AM* Test. The reliability testing showed that the *AM* Test had good internal consistency ($\alpha=0.89$). The item difficulty level analysis showed that most items were categorized

as easy and moderate items. Meanwhile, item discrimination analysis using the item-total correlation method (cr_{it}) showed that all items had good discrimination index ($cr_{it} \geq 0.3$).

Data Analysis

The reliability testing used Cronbach's alpha method to examine the internal consistency of the test. Validity testing used construct validity, specifically the correlation with other tests, i.e., correlating the *AM* Test with the *IST* subtest *ME* as a criterion that measures verbal memory. Item difficulty analysis (p) was conducted to analyze the item difficulty level. Item discrimination analysis (cr_{it}) was performed to determine the ability of the test items to distinguish participants with a high or low ability to memorize information. According to Nunally and Beirnsten (1994), a good item discrimination index is ≥ 0.3.

Results

Table 3
Descriptive Statistics

Descriptive	
Mean	19.35
SD	5.15
Max	25
Min	4
Average time (minutes)	5
N finished within 5 minutes (optimum time)	100%

N item=25; N= 96

The results in Table 1 showed that the *AM* test was quite easy for the participants (Mean > Median). The time limit of the *AM* test also proved to be optimal, because 100% of the participants accomplished the task within 5 minutes.

The reliability test showed that the coefficient alpha was 0.88. Therefore, the *AM* test is a reliable test with good internal consistency in measuring the same construct. The test also showed that 88% of the observed score comprised the true score variance, whereas 12% comprised the error variance derived from content sampling and content heterogeneity error. Calculating the Standard Error of Measurement (SEM) is also important in determining the individual true score estimation. The SEM value obtained from the *AM* Test was 1.78, which meant that the true score range was not too far from observed score. Thus, it can be concluded that the *AM* Test has a low error in measuring a certain construct.

The validity testing of the *AM* Test by correlating it with the *IST* subtest *ME* showed a significant correlation $r = 0.42$, $n = 96$, $p < 0.01$. Thus, the *AM* Test proved to be valid in measuring associative memory because it correlated significantly with *IST*, which also showed validity in measuring memory. The shared variance between the *AM* Test and the *IST* Test was 18%. The analysis of the item level of difficulty showed that almost all of the *AM* test items were categorized as easy, indicating that the proportion of the degree of difficulty level of the test items is not suitable with the targeted item because the test should contain various degrees of item difficulty.

Item discrimination analysis showed that there were 20 items that can distinguish participants' ability ($c_{rit} > 0.3$). However, 5 items were unable to discriminate participants' ability because of their low c_{rit} index ($c_{rit} < .03$). The low item discrimination index was also influenced by the degree of item difficulty. The low-discriminating items were easy items; hence, almost all of the participants were able to answer the items correctly.

Next, integrative analysis was conducted to select the best items referring to the *AM* Test specification. Item selection was based on qualitative (item form and content) and quantitative (c_{rit}, alpha if item deleted) considerations. Out of the 15 items selected, 4 items already had good quality and 11 items still needed revision. Items test revision is thus required to improve the varying degrees of item difficulty.

According to the best selected items, we re-conducted the reliability and validity testing. The value of the coefficient alpha was 0.86 and the correlation with the *IST* Test showed $r = 0.45$, $n = 96$, $p < 0.01$. Thus, it could be stated that the *AM* test had a good internal consistency and validity in measuring associative memory.

Table 4
Item Difficulty Analysis

P	Number Item	Category	N Item
$p > 0.8$	8, 12, 14, 15, 17, 18, 19, 21, 22, 23, 25	Very easy	11
$0.6 < p \leq 0.8$	3, 4, 5, 6, 7, 9, 10, 11, 13, 16, 24	Easy	11
$0.4 \leq p \leq 0.6$	1, 2, 20	Medium	3
$0.2 \leq p < 0.4$	-	Difficult	-
$p < 0.2$	-	Very difficult	-
Total			25

Table 5
Item Discrimination

Crit	Category	Number of item	N item
> 0.3	Good discrimination	1,2, 3, 4, 5, 6,7, 8, 9, 10, 11, 12, 13, 14, 16, 18, 19, 20, 23, 25	20
< 0.3	Unable to discriminate	15, 17, 21, 22, 24	5

Discussion

This study aimed to develop and validate the proposed *AM* Test as a subtest of the intelligence test for high school students (*TISS*). As a new psychological test, especially for the purpose of determining high school specialization, the *AM* Test hasgood psychometric properties. According to psychometric testing, the test meets the criteria for a good psychological instrument to measure associative memory ability.

First, the *AM* Test has high internal consistency. Some errors may have been due to low inter-item inconsistency (Urbina, 2004). Moreover, validity testing showed that the *AM* Test has a significant correlation with the *IST* subtest *ME*. The results proved that the *AM* Test has good valididuty in measuring associative memory. The correlation with other tests that measure the same construct showed that the test is free from irrelevant factors (Anastasi & Urbina, 1997).

However, the number of participants limited the validity of the *AM* Test. Therefore, further studies should employ larger samples and include participants with diverse characteristics.

Based on item analysis, the results showed that the *AM* Test is relatively easy as most of the participants obtained high scores. This is most likely due to several factors. First, the students came from top-ranked high schools. Therefore, it could be assumed that the students already have good academic abilities. Second, the sample size is too small and can be considereed homogeneous; thus, the findings may not accurately represent the diverse high school population in Indonesia. However, almost all of the test items have good discriminating ability. It is highly likely that a low degree of item difficulty can be related to the types of words used in the *AM* Test. Half of the words used in the test contained concrete words. According to Paivo (1965), the type of word (abstract and concrete) can influence the level of item difficulty at the time of retrieval. The more abstract the words, the more difficult it is to use imagery strategies. Conversely, if the word is a concrete noun, it can bring the image to cognition, allowing the individual to visualize easily. For further revision, it is necessary to increase item difficulty by increasing the use of more abstract words.

For further test development, there is also a need to select participants from different quality schools (top-ranked vs. average-performing schools), so that the sample would be more diverse. To improve the validity of the *AM* Test, it is also necessary to conduct criterion validity testing. Academic achievement could be used as a suitable criterion given that there might be a link between academic ability and memory skills (Wilhite, 1990). Criterion validity testing can be donw by using school exam score before remedial or by correlating the *AM* Test score with the achievement test. In addition, it is also necessary to consider a construct validity testing by using discriminant and convergent methods to determine whether the *AM* Test can measure associative memory ability. Convergent validity testing can be analyzed by correlating the *AM* Test scores with general information test scores, because good memory ability is needed to remember some information (Schneider & McGrew, 2012). Meanwhile, discriminant analysis testing could be carried out with a test measuring fluid ability because this ability can help an individual solve new problems and have little interrelation with the learning process (ref). In addition, factor analysis methods can also be used to determine the interrelation between the *AM* Test and other subtests in the *TISS*.

References

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). United States: Pearson Prentice Hall.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment: An introduction to tests and measurement* (8th ed.). New York: McGraw-Hill.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21.

Drozdick, L. W., Holdnack, J. A., & Hilsabeck, R. C. (2011). *Essentials of WMS-IV assessment* (Vol. 85). John Wiley & Sons.

Eysenck, M. W.,& Keane, M. T. (2005). *Cognitive psychology: A strudent's handbook.* United State of America: Taylor & Francis.

Flanagan, D. P., & Dixon, S. G. (2013). In C. R. Reynold., Vanest., K. J., Janzen, E. F (eds). *Encyclopedia of special education.* John Wiley & Sons.

Harel, B. T., Pietrzak, R. H., Snyder, P. J., Thomas, E., Mayes, L. C., & Maruff, P. (2014). The development of associate learning in school age children., *9*(7).

Kementrian Pendidikan dan Kebudayaan Republik Indonesia (11 Juli 2014) *Nomor 64 tahun 2014 tentang peminatan pada pendidikan menengah*

Legault, L., Pelletier, L., Green-Demers, I. (2006) Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of educational psychology. (*98)3. 567-582

Lippman, M. Z., & Shanahan, M. W. (1973). Pictorial facilitation of paired-associate learning: Implications for vocabulary training. *Journal of Educational Psychology*, *64*(2), 216.

Lutz, K. A., & Lutz, R. J. (1977). Effects of interactive imagery on learning: Application to advertising. *Journal of Applied Psychology*, *62*(4), 493.

McGrew, K.S. (2005). *The Cattell-Horn- Carroll Theory of Cognitive Abilities: Past, Present, and Future.* In D.P. Flanagan & P.L. Harrison (Ed.),Contemporary intellectual assessment: Theories, tests, and, issues (2nd Ed.), pp. 136-175. New York, NY: Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence: A multidisciplinary Journal, 37*(1),1-10

McGrew, K. S (2014).*Cattell-Horn-Caroll (CHC) theory of cognotitve abilities definition*. Retrieved from http://www.iapsych.com/chcdefsbrief.pdf

McGrew, K. S., LaForte, E. M., & Schrank, F. A (2014). *Woodcock johnson iv technical manual.* Illinois: Houghton Mifflin Harcourt

Nunally, J. C., & Bernstein. (1994). *Psychometric theory.* (3rd)Eds. United State of America. McGraw-Hill

O'Donnell, A. M., Reeve, J.,& Smith, J. K. (2012). *Educational psychology* (3rd)Eds. United State of America: John Willey & Sons, Inc

Ormrord, J. E. (2003). *Educational psychology: developing learners* (4th) Eds. New Jersey: Merril Prentice Hall

Paivo, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 32-38.

Stacy, A. W., Ames, S. L., & Grenard, J. L. (2006). Word association tests of associative memory and implicit processes: Theoretical and assessment issues. *Handbook of implicit cognition and addiction*, 75-90.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn- Carroll model of intelligence. In D.P. Flanagan & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and, issues* (3rd), pp. 99-144. New York, NY: Guilford Press.

Solso, R. L. (2005). *Cognitive psychology* (7th)Eds. United State of America: Pearson Education Inc

Urbina, S. (2004). *Essentials of psychological testing*. United State of America: John Wiley & Sons.

Watkins, M. W., Lei, P. W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, *35*(1), 59-68.

Wilhite, S. C. (1990). Self-efficacy, locus of control, self-assessment of memory ability, and study activities as predictors of college course achievement. *Journal of Educational Psychology*, *82*(4).