

## Stanford-Binet Intelligence Scale Form L-M Predictive Power on Academic Achievement

Eugenius Tintus Reinaldi, Rahmat Hidayat

Faculty of Psychology, Gadjah Mada University, Indonesia

eugeniusreinaldi@gmail.com

### Abstract

Stanford-Binet Intelligence Scale Form L-M test is widely used in Indonesia to assess the academic capacity of elementary school students. However, its predictive power upon academic achievements has not been examined. This research represents a preliminary attempt at closing this gap. Stanford-Binet scores obtained 1 to 3 years earlier were used to explain variations in three subject marks of 156 an Elementary School from the 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> Grades. Simple regression analysis shows that 4.3% to 25.4% of the variance can be explained by Stanford-Binet scores, indicating a low to moderate predictive power. The results suggest a limited predictive power of the Stanford-Binet Form L-M test for applications in the assessment of the academic capacity of elementary school children.

**Keywords:** Academic achievement, intelligence test, predictive validity, Stanford-Binet form L-M test.

### Abstrak

*Tes Stanford-Binet dipergunakan secara luas di Indonesia dalam asesmen kapasitas akademik siswa sekolah dasar. Namun, daya prediksi tes ini terhadap prestasi akademik belum pernah diperiksa, setidaknya dalam beberapa dekade terakhir ini. Penelitian ini berusaha menjadi awalan untuk mengurangi kesenjangan ini. Skor Stanford-Binet yang diperoleh 1 hingga 3 tahun sebelumnya dipergunakan untuk menerangkan variasi dalam nilai-nilai tiga mata pelajaran dari 156 siswa sebuah Sekolah Dasar dari Kelas 3, 4 dan 5. Analisis regresi sederhana menunjukkan bahwa 4,3% sampai 25,4% dari variasi nilai dapat diterangkan oleh skor Stanford-Binet, mengindikasikan adanya daya prediksi yang rendah sampai menengah. Hasil ini menunjukkan daya prediktif yang terbatas dari Tes Stanford-Binet untuk dipergunakan dalam asesmen kapasitas akademik siswa sekolah dasar.*

**Kata Kunci:** Prestasi akademik, tes inteligensi, validitas prediktif, tes Stanford-Binet form L-M.

## Introduction

As a profession that has an evidence-based character, the psychology professionals should prioritize evidence from research in making decisions related to psychological services, one of which is in the form of psychological tests. Every psychological test given to clients ideally has evidence showing that the test tool has good validity and reliability, and is effective in revealing certain psychological constructs. The problem is that until now there are still psychological testing tools that are still being used, even though these tools are not supported by the latest research data.

In the education setting, psychological tests are used to select students to enter school, provide scholarships, placement in certain programs, measure student learning progress, classify students based on their ability to absorb instructions in class, identify fast learners and slow learners in class, identify student's problem, and identify student's career interests (Anastasi & Urbina, 2007; Miller, Lovler, & McIntire, 2013). Various meta-analysis studies have shown that intelligence measurement can predict a person's academic performance (Roberts, Markham, Matthews, & Zeidner, 2005).

One of the intelligence testing tools that are still often used in Indonesia is Stanford-Binet. Ackerman & Beier (2005) stated that the Binet-Simon test and its revisions (including Stanford-Binet) are still used for academic placement tests and identify symptoms of mental retardation in various countries because these tests have proven to be very effective in predicting academic success in children and adolescent in the last 100 years. However, this testing tool in Indonesia is not supported by the latest research data which states that this testing tool can predict students' academic achievement.

The Stanford-Binet test used in Indonesia is the Stanford-Binet Intelligence Scale Form L-M, compiled in 1960 and is the third revised edition. This testing tool has been adapted and used in Indonesia since the 1970s (Wulan, 1995). This testing tool is a combination of Form L and Form M, which is a version of the Binet test from 1937 (Becker, 2003). Stanford-Binet Intelligence Scale Form L-M consisted of 20 age levels, starting from year II to Superior Adult III. These low age levels are mostly served for children and adults who have mental retardation, while Superior Adults levels I, II, and III are mostly presented for intelligent children. This test is mostly used to test children aged 6 to 10 years old (Wulan, 1995). Becker (2003) stated that this version of the Stanford-Binet test only measures one intelligence factor, namely general intelligence (g).

In general, the Stanford-Binet test has been tested for validity, so it can be used as a measure of general intellectual ability, has a good correlation with achievement, and can distinguish between gifted, mentally retarded, and neurologically impaired people (Roid & Barram, 2004; Kaplan & Sacuzzo, 2005; Anastasi & Urbina, 2007). The third edition of this test also can attract young children to do the task given because it uses toys as media. Varied test administration can also keep testees interested throughout the test (Becker, 2003). These are the reasons why Stanford-Binet Intelligence Scale Form L-M is still used in Indonesia.

Academic achievement is defined as the level of education completed and the ability to achieve success in learning. Academic achievement is the result of education and can be measured by continuous examinations and assessments (Kalojya, Basu, & Basu, 2017). Kalojya, et al. (2017) also states that to measure academic achievement, standardized and developed achievement tests can be used for each subject in school, because academic achievement is measured through the achievement of each student at the end of the semester. In general, intelligence still is a strong factor in predicting academic achievement (von Stumm, Hell, & Chamorro-Premuzic, 2011, Downey et al., 2014).

A psychological testing tool must have a predictive function if the test is designed to predict future performance (Azwar, 2016). Predictive validity measures the accuracy of the test function in predicting variations in criterion score changes, not referring to the accuracy of the test score in describing the construct it measures (Allen & Yen, 1979). Azwar (2008) states that predictive validity is estimated if the test

performed serves as a predictor of future performance. Nunnally (1981) states that predictive validity is important to use as a basis for making decisions about education, especially to measure children's readiness.

Since the evaluation of the Stanford-Binet Intelligence Test Form L-M conducted by Wulan in 1995, until now this test has not been tested for its predictive validity on academic achievement. Although there is no actual research on the predictive power of this test on student academic achievement, this test is still used as a basis for making decisions at school. This has become a problem because Messick (in Azwar, 2016) stated that validation is a continuous process. The purpose of this study is to test the predictive power of the Stanford-Binet Intelligence Scale Form L-M towards students' academic achievement.

## Methods

This study is testing the predictive validity of the Stanford-Binet Intelligence Scale Form L-M on the academic achievement of elementary school students. This testing tool is still widely used in educational settings, even though there has been no research about its predictive validity on students' academic achievement in Indonesia. Elementary students were selected as subjects on this research because as Wulan (1995) stated the Stanford-Binet Intelligence Scale Form L-M mostly tested children aged 6 to 10 years old. Students' academic achievement is obtained from the documentation of final exams scores from both odd and even semesters of 2016/2017 academic year students in 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade for Mathematics, Indonesian Language, and Civics subjects.

These three subjects were chosen as the criterion for academic achievement because based on the document "Kurikulum 2013: Kompetensi Dasar untuk Sekolah Dasar/Madrasah Ibtidaiyah" compiled by the Ministry of Education and Culture (2013), these three subjects are grouped in Group A; subjects whose content is developed by the central government. Other subjects in group A are science, social studies, and religion and character. Students in 3<sup>rd</sup> grade have not received science and social studies subjects. That's why only Mathematics, Indonesian Language, and Civics subjects were included in this study.

On the other hand, the selection of students in 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade was based on the timing of the psychological test and the time of final exams. The academic achievement data used is final exams score data for odd and even semesters of 2016/2017 academic year. Astuti et al. (2013) show there is a significant relationship between final exams scores and report cards in math subjects for 3<sup>rd</sup> grade students in an even semester, so final exams score data is quite representative as student academic achievement data.

This study's purpose is to test the predictive validity, so the predictor score (Stanford-Binet Intelligence Scale Form L-M score) must already be available before students have a criterion value (odd and even semester final exams). Students in the school where the data were collected took psychological tests in two periods. The first period is before they enter 1<sup>st</sup> grade, for the selection of acceleration programs. The second period is when students are in 2<sup>nd</sup> grade, ahead of their promotion to 3<sup>rd</sup> grade, also as an acceleration program selection for students who have good achievements during school. This is the reason why 1<sup>st</sup> and 2<sup>nd</sup> graders were not included in this study because they did not have psychological test results yet during odd semester final exams.

Research data used in this study are secondary data obtained from school and UKP UGM (Unit Konsultasi Psikologi UGM/Gadjah Mada University Psychological Consultation Unit). Psychological test data is owned by UKP UGM, while the final exams score is owned by the school. To obtain the needed data, researchers must get permission from these institutions.

After being allowed to conduct research in the school, the researcher was assisted by a teacher at said school to collect final exams scores from other classroom teachers. Then the data is given to the UKP supervisor to retrieve psychological test data.

There were 102 data samples, consisting of 52 3<sup>rd</sup> grade, 30 4<sup>th</sup> graders, and 20 5<sup>th</sup> grade students for 2016/2017 academic year from an elementary school which has MoU with UKP UG Deary et al. (2013) and Anastasi (1976) stated that a person's intelligence tends to be constant and stable throughout his life as long as there is no drastic change in conditions in their life. So, the IQ score gained when the students were tested in 1<sup>st</sup> or 2<sup>nd</sup> grade is estimated to be the same as their IQ when they are in 3<sup>rd</sup>, 4<sup>th</sup>, or 5<sup>th</sup> grade.

IQ scores will be the predictor, while final exams score in even and odd semesters of the 2016/2017 academic year will be the criterion. Predictors arise when a study attempts to predict a person's score on one measurement based on that person's score on another measure. Meanwhile, the criterion shows the measurement results that have special treatment (Murphy & Davidshofer, 2005). The analysis used is regression analysis because it can be used if a study aims to predict the value of a variable based on the value of other variables that have been known previously (Howell, 2013). Simple regression will be used because there is only one predictor used to predict criterion (Field, 2005).

After data analysis is done, constants will be obtained which are useful for formulating the line equation as shown here (Field, 2005).

$$Y_i = (\beta_0 + \beta_1 X_i) + \epsilon_i$$

$Y_i$  is the end result wanted to predict, while  $X_i$  is the value of the subject on the predictor variable.  $\beta_0$  and  $\beta_1$  are the regression coefficient, with  $\beta_1$  is the gradient of the straight line and  $\beta_0$  is the intercept of the line.  $\epsilon_i$  is the residual measurements, which show the difference between the value predicted by the equation for the subject  $i$  and the actual value obtained by the subject  $i$ . The formula for this equation is often written without writing the residual (Field, 2005).

## Results and Discussion

### Results

There are 102 data used from students in 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade. Exam scores used as a criterion are from Mathematics, Indonesian Language, and Civics subjects. Table 1 shows the descriptive statistics of the variables in this study.

Table 1. Descriptive Statistic

	3 <sup>rd</sup> Grade		4 <sup>th</sup> Grade		5 <sup>th</sup> Grade		3, 4, 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Math1	82.94	9.650	80.46	9.615	80.66	13.353	81.35	10.950
Math2	78.08	9.046	80.33	10.454	79.52	13.348	79.32	11.023
Indo1	81.71	8.687	87.30	7.394	84.60	10.272	84.57	9.067
Indo2	81.46	7.790	87.94	7.930	81.76	10.419	83.80	9.217
Civics1	90.08	8.005	89.94	4.901	86.44	8.031	88.87	7.244
Civics2	84.52	7.599	91.39	6.344	80.96	8.164	85.76	8.533
Stanford-Binet Test Score	124.13	12.072	127.97	13.265	120.75	8.608	124.60	12.019
N	52		30		20		102	

Note: Math1: final exam odd semester for Mathematic subject; Math2: final exam even semester for Mathematic subject; Indo1: final exam odd semester for Indonesian Language subject; Indo2: final exam even semester for Indonesian Language subject; Civic1: final exam odd semester for Civics subject; Civic2: final exam even semester for Civics subject.

The condition for a predictor to predict the criterion is that the predictor has a correlation with the criterion ( $p < 0.05$ ). If there is no correlation between the predictor and the criterion ( $p > 0.05$ ), then the predictor can't predict the criterion. Table 2 shows the overall correlation and regression test results between Stanford-Binet test score and final exams score for the odd and even semester of 2016/2017 academic year for Mathematics, Indonesian Language, and Civics subjects. The analysis was carried out as a whole (3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade at once) and per class (3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade are analyzed separately).

**Table 2.** Correlation and Determination Coefficient of Stanford-Binet Intelligence Scale Form L-M with Final Exams Score for Odd and Even Semester

Grade	Subjects	Stanford-Binet Intelligence Scale			
		Odd Semester		Even Semester	
		r	R <sup>2</sup>	r	R <sup>2</sup>
3.4.5	Mathematics	0.207*	0.043	0.339**	0.115
	Indonesian L.	0.329**	0.109	0.257**	0.066
	Civics	0.286**	0.082	0.341**	0.116
3	Mathematics	0.164	-	0.381**	0.145
	Indonesian L.	0.289*	0.083	0.183	-
	Civics	0.235	-	0.224	-
4	Mathematics	0.480**	0.231	0.440*	0.194
	Indonesian L.	0.503**	0.253	0.421*	0.177
	Civics	0.383*	0.147	0.321	-
5	Mathematics	0.090	-	0.114	-
	Indonesian L.	0.294	-	-0.024	-
	Civics	0.452*	0.205	0.504*	0.254

\* $p < 0.05$ ; \*\* $p < 0.01$

Overall, the Stanford-Binet Intelligence Scale Form L-M test can predict students' academic achievement, which is indicated by their final exams score, although for each subject in each semester, the predictive power is different. The intelligence test can predict a mathematic final exam score in an odd semester by 4.3%. This means that the mathematic final exam score in the odd semester can be predicted by other factors up to 95.7%. Meanwhile, for even semester mathematics, the Stanford-Binet test can predict final exam scores up to 11.5%, leaving 88.5% for other factors to be the predictors of mathematic final exam score in an even semester. The intelligence test also can predict the Indonesian Language subject in the odd semester at 10.9%, while in the even semester its predictive power is 6.6%. For the Civics subject, the intelligence test can predict the odd semester final exam score at 8.2% and 11.6% in the even semester.

For 3<sup>rd</sup> grade students, Stanford-Binet Intelligence Scale Form L-M can predict the odd semester Indonesian Language final exam up to 8.3% and the even semester mathematic up to 14.5%. For 4<sup>th</sup> grade students, the intelligence test can predict the odd semester final exam for mathematic, Indonesian Language, and Civics, as well as mathematic and Indonesian Language in the even semester. Its predictive power moves between 14.7% to 25.3%. While for 5<sup>th</sup> grade students, the intelligence test can only predict final exam score for Civics subject for both odd and even semesters (20.5% in odd semester; 25.4% in even semester).

Table 3 shows the  $\beta$  coefficient of each criterion, which is needed to create an equation formula for predicting final exams scores for each subject in odd and even semesters. The use of the formula can predict students' final exams score from their Stanford-Binet Intelligence Scale Form L-M score, by multiplying the test score with the beta coefficient, then adding it up with the constant.

**Table 3.** Beta Coefficient and Prediction Formula

Grade	Subjects	$\beta_0$	$\beta_1$	Prediction Formula
3.4.5	Math, odd semester	64.350	0.154	$\widehat{Mat1} = 64.350 + 0.154 (Binet)$
	Math, even semester	47.251	0.266	$\widehat{Mat2} = 47.251 + 0.266 (Binet)$
	Indonesian, odd semester	55.867	0.234	$\widehat{BI1} = 55.867 + 0.234 (Binet)$
	Indonesian, even semester	62.581	0.176	$\widehat{BI2} = 62.581 + 0.176 (Binet)$
	Civics, odd semester	70.289	0.156	$\widehat{PKn1} = 70.289 + 0.156 (Binet)$
	Civics, even semester	58.405	0.224	$\widehat{PKn2} = 58.405 + 0.224 (Binet)$
3	Math, even semester	42.660	0.285	$\widehat{Mat2} = 42.660 + 0.285 (Binet)$
	Indonesian, odd semester	55.906	0.208	$\widehat{BI1} = 55.906 + 0.208 (Binet)$
4	Math, odd semester	40.734	0.322	$\widehat{Mat1} = 40.734 + 0.322 (Binet)$
	Math, even semester	44.714	0.295	$\widehat{Mat2} = 44.714 + 0.295 (Binet)$
	Indonesian, odd semester	52.751	0.278	$\widehat{BI1} = 52.752 + 0.278 (Binet)$
	Indonesian, even semester	58.455	0.241	$\widehat{BI2} = 58.455 + 0.241 (Binet)$
	Civics, odd semester	73.033	0.132	$\widehat{PKn1} = 73.033 + 0.132 (Binet)$
5	Civics, odd semester	57.807	0.256	$\widehat{PKn1} = 57.807 + 0.256 (Binet)$
	Civics, even semester	38.626	0.362	$\widehat{PKn2} = 38.626 + 0.362 (Binet)$

Note:  $\beta_0$ : constant;  $\beta_1$ : Stanford-Binet Intelligence Scale Form L-M beta coefficient; Binet: Stanford-Binet Intelligence Scale Form L-M test score; Mat1: final exam score for odd semester mathematic; Mat2: final exam score for even semester mathematic; BI1: final exam score for odd semester Indonesian Language; BI2: final exam score for even semester Indonesian Language; PKn1: final exam score for odd semester Civics; PKn2: final exam score for even semester Civics.

## Discussion

This study aims to test the predictive validity of the Stanford-Binet Intelligence Scale Form L-M test towards students' academic achievement. In this study, students' achievement was seen from the final exams score for odd and even semester for three subjects: Mathematics, Indonesian Language, and Civics. The results of the analysis show that this test can predict students' academic achievement.

The correlation coefficient shown in table 2 is the validity coefficient of the Binet test against each criterion. Murphy & Davidshofer (2005) stated that although in theory, the correlation coefficient moves between 0.0 to 1.0, in practice the validity coefficient tends to be small. Even a very well-structured test rarely gets a correlation of more than 0.5. Cronbach (in Azwar, 2016) states that coefficients ranging from 0.3 to 0.5 can make a good contribution. Azwar (2016) adds that what must be taken into consideration in interpreting the validity coefficient is the usefulness of the test score in decision making, so it is possible that a test that has a low validity coefficient can still be used to help make decisions.

Stanford-Binet Intelligence Scale Form L-M test can predict final exams scores for Mathematics, Indonesian Language, and Civics in odd and even semesters for students in 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade, with various contributions ranging from 4.3% to 11.6%. For 3<sup>rd</sup> grade students, this test can predict the odd semester Indonesian Language final exam score with a contribution of 8.3% and the even semester math final exam score with a 14.5% contribution. In 4<sup>th</sup> grade students, the test only failed to predict the even semester Civics final exam score. The rest can be predicted with a contribution between 14.7% to 25.3%. In 5<sup>th</sup> grade students, the test can predict the Civics final exam score for odd and even semesters, with contributions of 20.5% and 25.4%, respectively.

This result is in line with Goleman (in Devi, 2012) and Downey, et al. (2014) which states that IQ contributes 20% to success in life. In addition to intelligence factors, success in academics is also associated

with emotional management and control as well as high self-awareness, and low extraversion. In students aged 8 to 10 years old, adaptability is one of the factors that contribute to academic achievement, and in students aged 11 to 13 years old, the factors that contribute to academic achievement are stress management ability, intrapersonal emotional intelligence, and adaptability (Brouzos, Misailidi, & Hadjimatheou 2014). Other factors that can affect academic achievement are socioeconomic status, learning style, self-confidence, adaptability, and social competence (Chen, Huang, Chang, Wang, & Li 2010; von Stumm, et al., 2011; Brouzos, et al., 2014; Gordon & Cui, 2016).

The third edition of the Stanford-Binet test, although it has been used since the 1970s in Indonesia, is still able to predict student academic achievement. This is inseparable from the structure of the Stanford-Binet Intelligence Scale Form L-M, which is 40% of this test measures knowledge (Becker, 2003). While academic achievement is learning outcomes, and achievement tests (final exam) aim to measure educational outcomes in schools rather than measuring patterns of knowledge or skills (Simpson & Weiner, in Kaloiya, et al., 2017). The Ministry of Education and Culture (2013) explains that four core competencies must be possessed by students, namely religious attitudes (Core Competency 1), social attitudes (Core Competencies 2), knowledge (Core Competencies 3), and application of knowledge (Core Competencies 4). One of these four core competencies (knowledge) is what the Stanford-Binet Intelligence Scale Form L-M test measures. This is why this test can still predict students' academic achievement even though this test has been designed for a long time.

The formula obtained from regression analysis can be used to calculate the minimum IQ score that can be received by the school if they want to select students for some sort of special class. For example, if the school wants students who enter the special class can obtain scores at least 80 in mathematics, the school can use this formula:

$$(\text{Mat1}) = 64.35 + 0.154 (\text{Binet})$$

to calculate final exam scores in the odd semester.

$$(\text{Mat2}) = 47.251 + 0.266 (\text{Binet})$$

to calculate the final exam score in the even semester. By entering value 80 into Mat1 and Mat2, the minimum Stanford-Binet Intelligence Scale Form L-M test score required is 123.

The limitation of this study lies in the usage of secondary data. The documentation data obtained by the researcher can not be changed or added according to the wishes of the researcher because these are actual data. The researcher also can not control the conditions when the data is taken, both during the psychological test and during final exams. The use of documentation also causes the criterion to be limited. The researcher can not use report cards as criteria in this study. On the other hand, the use of data documentation also has its advantages. The data obtained are actual data, following real conditions in the school, and free from research bias.

## Conclusion

This research is an early stage of research in testing the predictive validity of the Stanford-Binet Intelligence Scale Form L-M test on students' academic achievement. This study shows that this test can predict students' academic achievement. This is because one of the constructs measured by the test is knowledge. Meanwhile, academic achievement, which is measured by the final exams score, is the result of learning. One of the four core competencies of the 2013 Curriculum is knowledge. This knowledge factor allows the test to still predict students' academic achievement.

Based on the findings of this research, it is recommended for educational purposes to prioritize the use of the Stanford-Binet Intelligence Scale Form L-M test. The line equation formula obtained from this study can be used as a basis if the school wants to make a selection for a special class. For example, if the school wants a student entering a special class to have a math score of at least 80, then the minimum Binet Test score accepted is 123.

As for the limitations of this study due to the use of secondary data, further research can use primary data. Researchers can administer tests on research subjects, thus researchers have the freedom to choose research subjects, predictor variables, and criteria variables.

## References

- Ackerman, P., & Beier, M. (2005). Knowledge and intelligence. In O. Wilhelm, & R. Engle, *Handbook of understanding and measuring intelligence*. California: Sage Publications.
- Allen, & Yen. (1979). *Introduction to measurement theory*. Monterey: Brooks/Cole. Publishing Company.
- Anastasi, A. (1976). *Psychological testing (4th edition)*. New York: Macmillan Publishing Company.
- Anastasi, A., & Urbina, S. (2007). *Tes psikologi (edisi ketujuh)*. Jakarta: PT Indeks.
- Astuti, E., Tarini, K., & Budiyono. (2013). Ketuntasan nilai UAS dan nilai rapor serta hubungan nilai UAS dengan nilai rapor siswa kelas III semester II mata pelajaran matematika SD Negeri se-Kecamatan Bonorowo Kabupaten Kebumen tahun ajaran 2011/2012. *Ekuivalen*, 4(1). Diunduh tanggal 26 Desember 2017, dari <http://ejournal.umpwr.ac.id/index.php/ekuivalen/article/view/596>
- Azwar, S. (2008). *Dasar-dasar psikometri*. Yogyakarta: Pustaka Pelajar Offset.
- Azwar, S. (2016). *Reliabilitas dan validitas (edisi 4)*. Yogyakarta: Pustaka Pelajar.
- Becker, K. (2003). *History of the Stanford-Binet intelligence scales: Content and psychometrics (Stanford-Binet Intelligence Scales, Fifth Edition Assessment Service Bulletin No. 1)*. Itasca: Riverside Publishing.
- Brouzos, A., Misailidi, P., & Hadjimattheou, A. (2014). Associations between emotional intelligence, socio-emotional adjustment, and academic achievement in childhood: The influence of age. *Canadian Journal of School Psychology*, 29(2), 83-99.
- Chen, X., Huang, X., Chang, L., Wang, L., & Li, D. (2010). Aggression, social competence, and academic achievement in Chinese children: A 5-year longitudinal study. *Development and Psychopathology*, 22, 583-592.
- Deary, I., Pattie, A., & Starr, J. (2013). The stability of intelligence from age 11 to age 90 years: The Lothian birth cohort of 1921. *Psychological Science*, 24(12), 2361-2368.
- Devi, A. (2012). Assessing emotional intelligence: Academic performance & overall personality development. *Indian Journal of Industrial Relations*, 48(2), 354-367.
- Downey, L., Lomas, J., Billings, C., Hansen, K., & Stough, C. (2014). Scholastic success: Fluid intelligence, personality, and emotional intelligence. *Canadian Journal of School Psychology*, 29(1), 40-53.
- Field, A. (2005). *Discovering statistic using SPSS (2nd edition)*. London: Sage Publication.
- Gordon, M., & Cui, M. (2016). The intersection of race and community poverty and its effect on adolescents' academic achievement. *Youth & Society*, 1-19.
- Howell, D. (2013). *Statistical methods for psychology*. Wadsworth: Cengage Learning.
- Kaloiya, G., Basu, S., & Basu, S. (2017). Academic achievement, behavioral and emotional problem among marginalised children. *The International Journal of Indian Psychology*, 4(3), 36-49.
- Kaplan, R., & Saccuzzo, D. (2005). *Psychological testing: Principles, applications, and issues (6th edition)*. Toronto: Wadsworth Cengage Learning.



- Kementerian Pendidikan dan Kebudayaan. (2013). *Kurikulum 2013: Kompetensi dasar sekolah dasar (SD)/madrasah ibtidaiyah (MI)*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Miller, L., Lovler, R., & McIntire, S. (2013). *Foundations of psychological testing. A practical approach (4th edition)*. California: Sage Publication.
- Murphy, K., & Davidshofer, C. (2005). *Psychological testing: Principles and applications (6th edition)*. New Jersey: Pearson Education, Inc.
- Nunnally, J. (1981). *Psychometric theory (2nd edition)*. New Delhi: Tata McGraw-Hill Publishing Co Ltd.
- Roberts, R., Markham, P., Matthews, G., & Zeidner, M. (2005). Assessing intelligence: Past, present, and future. In O. Wilhelm, & R. Engle, *Handbook of understanding and measuring intelligence*. California: Sage Publication.
- Roid, G., & Barram, R. (2004). *Essentials of stanford-binet intelligence scales (SB5) assessment*. New Jersey: John Wiley & Sons, Inc.
- Shipstead, Z., Harrison, T., & Engle, R. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, 11(6), 771-799.
- Soysal, S., Tan, S., & Aldemir, S. (2012). Accuracy of WISC-R and Raven Standard Progressive Matrices tests in mathematical success of children with ADHD. *Dusunen Adam The Journal of Psychiatry and Neurological Sciences*, 25, 17-26.
- Stankov, L. (2004). Complexity in human intelligence. In R. Stenberg, J. Lautrey, & T. Lubart, *Models of intelligence: International perspectives*. Washington, DC: American Psychological Association.
- Styles, I. (1999). The study of intelligence - the interplay between theory and measurement. In M. Anderson, *The development of intelligence*. East Sussex: Psychology Press.
- von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science*, 6(6), 574-588.
- Wulan, R. (1995). Evaluasi penggunaan tes binet. *Buletin Psikologi, Tahun III, Nomor 2*, 49.