



## Giving Class Averages: Is It Worth It?

Adrian Leis <sup>a \*</sup>, Matthew Wilson <sup>b</sup>

<sup>a</sup> *Miyagi University of Education, Japan*

<sup>b</sup> *Miyagi University, Japan*

Received 10 May 2016 | Received in revised form 26 January 2017 | Accepted 4 February 2017

---

### Abstract

The aim of this study was to test a scoring method for regular testing proposed by the authors which provides feedback on test results to each student based upon his or her increases or decreases with the previous test score (hereafter, Idio-comparative marking). The authors compared the Idio-comparative scoring system with a regular ‘individual result together with class average’ method in an eight-week long quasi-experimental study conducted with 52 first-year Japanese university students. The authors aimed to gain an understanding of whether there were any salient statistical advantages in giving class averages to students to compare with their own result, thus nullifying the Idio-comparative idea. The study used a Pretest-Posttest design, with the results of an analysis of variance suggesting there to be no statistically significant difference ( $p = .77$ ) between the Idio-comparative Group and Give-average Group. The authors suggest the Idio-comparative approach as one to be considered by language teachers in order to assist with building and preserving the confidence of the students participating in the class.

© 2017 EJAL & the Authors. Published by *Eurasian Journal of Applied Linguistics (EJAL)*. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** L2 Learning Motivation; Self-worth Theory; Assessment; Japanese university students; vocabulary acquisition

---

### 1. Introduction

It is no secret that testing is a major part of education in Japan, and English education is no exception. Students are faced with regular listening tests, vocabulary tests, mid-term tests, end-of-term tests - the list goes on. This does not even include the National Center Test for University Admissions or university entrance tests, which decide the fate of thousands of Japanese students every year. Results in these tests, regardless of how major or minor they may seem, have significant washback effects both on the approach students take to their studies and the confidence students have in their scholastic ability. School is the place where students’ egos are built (Covington, 1992), and those getting good scores will often have increased self-confidence, and bad scores may see the confidence of those students deflated and an increase in anxiety (Zimmermann, Schütte, Taskinen, & Köller, 2013). Often when returning tests, teachers give students class averages, which may result in some

---

\* Adrian Leis. Tel.: +61-22-214-3488  
E-mail address: [adrian@staff.miyakyo-u.ac.jp](mailto:adrian@staff.miyakyo-u.ac.jp)  
<http://dx.doi.org/>.....

students, regardless of how much effort they may have put into their studies, seeing themselves as being successful (i.e., those above the average score) and others viewing themselves negatively (i.e., those below the average). Even if, for example, students study hard and achieve a better score compared to previous tests, they may still rank in the lower half of the class, thus see themselves as failures and give up studying (Covington, 1992).

In the present paper, we describe an experiment conducted to analyze whether telling students the average score of the class brings about clear advantages for those students and, as an alternative, propose an Idio-comparative approach to marking tests. The Idio-comparative method does not compare students with others in the class, but the individual scores of each student's previous test. Because it is possible for all students to improve their scores from previous tests, through this approach, every student in the class may be able to view their test results as successful, regardless of where they rank in the class. Therefore, based on the recommendations of the Self-worth Theory (Covington, 1992), we discourage teachers from providing the group average test score to students along with individual scores and instead take upon the Idio-comparative approach in order to create a personalized learning atmosphere in the classroom, in which each student's effort is valued and used to motivate them further in their studies.

## 2. Literature review

### 2.1. *Self-worth Theory in psychology*

About a century ago, a laboratory assistant in Germany, Ferdinand Hoppe (1930), conducted an experiment with his supervisor, Professor Kurt Lewin. In the experiment, Hoppe used a circular conveyer belt, to which several pegs were attached. Hoppe asked several subjects to throw rings over the pegs, with one point being given for successful attempts and zero points for unsuccessful throws. Once the subjects were used to the activity, Hoppe asked them to try to beat their own previous scores. When given this challenge, Hoppe noticed the subjects becoming highly motivated, leaning forward and seemingly concentrating harder in order to beat themselves. However, when Hoppe asked subjects to get better scores than the others participating in the experiment, they seemed to lose motivation, some even giving up. This experiment served as a catalyst upon which Martin Covington (1992, 1998) presented the Self-worth Theory late in the 20th century.

Through the Self-worth Theory, Covington (1998) asserts that "the protection of a sense of ability is the student's highest priority" (p. 17), and students will do anything they can, including not trying on purpose in order to preserve their self-esteem when faced with possibilities of failure. This attitude among students comes from the thought that if they study hard but still do not do well on tests, they may be viewed as unintelligent. On the other hand, if students do not study hard on purpose, such as not doing homework or consciously not paying attention in class, their low test scores

may be attributed to this lack of effort, rather than their lack of ability. Although the Self-worth Theory includes some aspects similar to the Attribution Theory (see Weiner, 1974), it also includes actions taken by people to protect their feelings of worth when faced with the possibility of failure, or when they see themselves in a helpless situation in which they feel that they have no control.

The notion of learned helplessness was first introduced by Seligman, Maier, and Geer (1968), who concluded that animals, including humans, tend to become passive when faced with situations in which they feel they have no control over the outcome. In a study by Dweck and Reppucci (1973), however, it was discovered that not all humans give up when faced with failure. Instead, there are those who thrive on the challenge. Further studies (e.g., Dweck, 1975; Kamins & Dweck, 1999) strengthened this idea that some humans approach situations in which they feel they may fail as opportunities to learn and improve their ability in their chosen fields. Mueller and Dweck (1998) suggest that this attitude of not giving up in helpless situations is not necessarily innate. Furthermore, this attitude can be educated in children through praise, not for intelligence, but for the effort put forth.

Developing on the findings and ideas presented in previous studies in the Self-worth Theory, such as those mentioned above, Dweck (2006) argued that humans' ways of thinking and approaches to situations in which they may fail, fall into two main categories: a fixed mindset and a growth mindset. The fixed mindset is a belief that humans' abilities are innate and cannot be changed regardless of the efforts they may make. Those who fall into the fixed mindset category constantly find themselves trying to prove to others that they are smart and therefore hesitate to challenge situations in which they may fail: failure is thought of as a sign of being unintelligent. As a result, although those with a fixed mindset may seem smart at first, because they are only taking on relatively simple problems, they tend to make less progress in their studies: "It is not always the people who start out the smartest who end up the smartest" (Dweck, 2006, p. 5). On the other hand, humans with a growth mindset are more willing to take on challenges, often working at slightly difficult levels, what Krashen (1985) refers to as  $i + 1$ . Learners with a growth mindset see failures as opportunities to better themselves and expand their knowledge and skills, and furthermore, understand proficiency and ability as a result of one's efforts, not something fixed and unchangeable from birth.

Although this positive attitude towards approaching challenges has been well documented in psychology, the amount of literature in second language acquisition (SLA) research is quite limited.

## *2.2. Self-worth Theory in SLA*

Covington's (1992) Self-worth Theory maintains that humans are driven to protect their feelings of competence, and it is through schooling, especially testing, where this competence is challenged. When faced with situations in which students feel they may fail, there is "the fear that one may be judged incompetent, hence unworthy, then the

pursuit of grades becomes an ordeal and the virtually assured result is defensiveness and excuse making” (Covington, 2004, p. 96). The implications of the Self-worth Theory for SLA have not been explored to a great degree, despite anxiety being a well-researched factor for language learners being successful or not (see, for example, Horwitz, 2001; and MacIntyre & Gregersen, 2012). Wiggins (1989) argues that competition with others for better grades can be formidable for students who tend to be anxious in learning. Furthermore, although failure to achieve a goal based on one's own performance can motivate learners to work harder, failure to achieve a goal when competing with others “implies falling short as a person” (Covington, 1998, p. 31).

In one example of research conducted in Japan that focused on the Self-worth Theory and modeled on Dweck and Repucci's (1973) study, Leis (2014) distributed English crossword puzzles to junior high school students and timed how long it took them to complete the puzzles. Half of the students were given puzzles that could easily be solved, and half were given puzzles that were impossible to solve (i.e., there were insufficient squares for the words to be entered). When the students were able to complete the puzzles, they were asked to indicate their success by raising their hands and saying, “Finished!” The purpose of having the students indicate that they had been successful was to create a classroom atmosphere in which those who failed to complete the puzzles would compare their ability to those who could complete them. Even though both groups were able to complete the first crossword puzzle within a statistically similar length of time, when it came to the sixth crossword puzzle, which was the same as the first puzzle but presented slightly differently, the students who had been given possible puzzles were able to complete the puzzle significantly quicker than the group of students who had been given impossible to solve puzzles. The repeated experience of failure, which was beyond the control of the students with impossible to solve puzzles, together with the decrease in self-confidence as a result of noticing the other students being able to solve puzzles that they were not, resulted not only in their lack of improvement in the final puzzle but also several students cheating (e.g., adding extra squares or using incorrect spelling) in order to solve the impossible to complete puzzles.

In his conclusion, Leis (2014) encouraged instructors to create learning environments in which students would not compare their abilities to others, but to their own previous performances. One feasible way of achieving this comparison to oneself would be in the way tests are returned. Whereas many teachers appear to give the class averages to students when returning tests, Leis discouraged this, instead recommending a scoring system in which students' test results were compared individually. It was argued that putting the students' focus on their own increased or decreased scores would be effective indications of the efforts they had made for the test being returned. In this paper, we coin this approach to scoring “Idio-comparative Scoring” which will be explained in further detail in the following section.

### *2.3. Idio-comparative scoring*

The idea of Idio-comparative scoring stems from Covington's (1998) conviction that when students feel their self-worth may be threatened, they will take steps to put the blame on lack of effort rather than on insufficient ability. By using the Idio-comparative scoring method, the authors believe students will be less likely to compare themselves with others, thus decreasing the possibility of students' feelings of worth being threatened. In the Idio-comparative scoring approach, teachers do not give the class averages to students when returning tests. Instead, we propose that teachers give a plus or minus score along with the test result to each student, comparing his or her current test score to that of the previous test. By doing so, teachers are able to acknowledge the effort (or lack thereof) of all students, even those whose test scores are below the average of their peers. The authors believe that this appreciation of effort on behalf of teachers will reduce the feeling of helplessness and increase the drive to improve one's own score in the following test.

Table 1. Examples of Idio-comparative Scoring

Student	Previous Test	Current Test	Idio-comparative Score
Albert	95	80	- 15
Betty	55	70	+15
Charlie	60	60	± 0
Daisy	65	50	- 15
Ed	25	40	+ 15

*Note.* Idio-comparative scores are calculated by measuring the difference between the current test result and the previous test result.

Table 1 displays examples of Idio-comparative scoring for five students in a hypothetical class. In this situation, if the instructor were only to give the class average (i.e., 60), Albert, Betty, and Charlie would be satisfied with their performances because their scores are above or equal to the average, whereas Daisy and Ed would be disheartened at being below the average. However, using the Idio-comparative scoring method, Albert and Daisy would realize that their scores are not good enough, as they had decreased by 15 points in comparison to their previous test scores, suggesting a lack of effort on their behalf. On the other hand, Betty and Ed would see the benefits of their efforts in class and study with increases of 15 points compared to their previous test scores. Seeing the benefits of effort is especially important for Ed, who seemingly has studied hard for this test. If the average is given to the students, Ed may see his efforts as fruitless because regardless of how hard he studies, he is still unable to reach standards set by the others in his class.

Although various other theories surrounding second language learning motivation have received quite a lot of attention in research conducted in Japan, almost none have focused on describing the motivation of Japanese students from the perspective of the Self-worth Theory. This is rather surprising as Covington (1992) suggests

school to be the place where students' feelings of worth are developed, especially through tests. With English education in Japan having a reputation for being centered on tests (e.g., see Sasaki, 2008), it may be beneficial to conduct more studies on the second language learning motivation of Japanese students based on the Self-worth Theory.

### 3. The Study

#### 3.1. Research question

The idea of an Idio-comparative scoring system is new, and statistical backing is yet to be provided. Therefore, in the early stages of research regarding this approach, the authors wish to investigate whether giving the class averages brings any particular advantages to learning, thus nullifying the hypothesized benefits of the Idio-comparative scoring approach. In the present study, the authors aim to answer the following research question (RQ):

RQ: Does using the Idio-comparative scoring approach result in statistically significant improvements in students' performances on tests?

In order to answer this RQ, the authors compared scores on vocabulary tests of three groups of students (including a control group) and whether also giving students the class averages would result in a higher level of vocabulary acquisition, thus higher scores than the Idio-comparative scoring method at the end of the eight-week study. It was thought that if including the class average on weekly tests along with individual scores resulted in higher vocabulary acquisition, it would nullify the authors' proposal for Idio-comparative scoring to be implemented into EFL classrooms in Japan.

#### 3.2. Participants

A total of 52 first-year students attending a public university in northeast Japan participated in the present study. Their majors were Nursing ( $n = 17$ ) and Business Planning ( $n = 35$ ). Within the sample, there were 36 female students and 16 male students. The students studied English communication in three separate classes, with all Nursing majors in one class, and the Business major students in the other two classes (17 students and 19 students). The class with Nursing majors made up the Idio-comparative Group, and one of the classes of Business majors was used as the Give-average Group. The Control Group was made up of the other class of Business majors. Although past research has shown that motivation to learn English is often higher for those majoring in English and with higher linguistic proficiency, no salient difference was observed among non-English majors (see Leis, Suzuki, & Ando, 2011). Therefore, it was deemed by the authors that comparing two groups with different majors would be satisfactory for this study. The participants indicated in a questionnaire at the beginning of the study that they studied English for an average of 25.29 minutes per week outside of class, suggesting to the authors that their motivation to learn the language was not noticeably high. However, the students'

regular language instructor commented that they were all eager to participate in class, implying a positive attitude to learning English.

### 3.3. Procedure

The present quasi-experimental study was conducted over eight weeks as part of the participants' regular English lessons. The design of the study is displayed in Table 2. In Week 1, the students were given a 60-item vocabulary test based on words randomly chosen from levels 2700 to 3000 of the JACET 8000 word list. The 60 items were presented in a Cloze-style test. Thirty items were multiple choice and 30 required the students to write the answer. The tests were all conducted during class time under the direct supervision of one of the researchers who was the teacher of all the students. The scores of the three classes were averaged and analyzed to confirm there were no statistically significant differences between these pretest scores. The three classes were then randomly designated as the Give-average Group (GA Group), Idio-comparative Group (IC Group), and Control group. Then, the GA Group and IC Group were given booklets containing words from levels 2700 to 3000 of the JACET 8000 word list, divided into six 50-word lists, which would be tested from Week 2 to Week 7 (six weeks) of the study. When returning the tests from the previous week, the GA Group was told the class average along with their own score. On the other hand, the IC Group, in addition to that week's score, were given a plus or minus score in comparison to the previous test score. During this time, members of the Control Group were neither given the word lists nor required to take the weekly tests. In Week 8 of the study, all three groups took the same vocabulary test taken in Week 1. Both the pretest and posttest were conducted using Google Forms, allowing the item order to be randomized, also reducing issues of human error such as missed items and items answered twice, which can occur in paper tests. Unlike the pretest and posttest, the quizzes from Week 2 to Week 7 were conducted in class using a paper format in order to easily provide the test result feedback, idio-comparative scores or class averages, which cannot be accomplished as effectively through online means.

Table 2. Description of the Procedure for the Present Study

Group	Week 1	Weeks 2 - 7	Week 8
IC	Pretest	Weekly tests with students given individual comparisons	Posttest
GA	Pretest	Weekly tests with students given class average	Posttest
Control	Pretest	No testing	Posttest

Note. GA: Give-average; IC: Idio-comparative.

## 4. Results

First, the data were tested for normality and homogeneity of variance. The results suggested that the pretest data (Skewness = -.034; Kurtosis = -.603) were normally distributed. This was confirmed by conducting a Shapiro-Wilk test of normality, which

proved to be non-significant ( $p = .595$ ). The posttest data showed similar results with Skewness (.379) and Kurtosis (-.650), as well as the Shapiro-Wilk test of normality ( $p = .157$ ) suggesting the data were normally distributed and satisfactory for analyses.

Next, in order to measure whether any statistically significant differences could be observed between the three groups in the study, in particular the IC and GA Groups, the authors conducted a one-way analysis of variance (ANOVA) using SPSS Version 22. At the pretest stage, no statistically significant differences could be observed between the three groups ( $F(2, 49) = .56, p = .57$ ), and the 95% Confidence intervals (95%CI) contained enough overlapping to suggest the three groups' averages were similar and thus satisfactory for data analyses. Skewness and Kurtosis were close to zero, suggesting a relatively balanced distribution of the scores within each group.

At the posttest stage, a one-way ANOVA was once again conducted to measure statistical differences between the groups. The ANOVA was significant  $F(2, 49) = 19.95, p < .001$ , with very strong effect sizes ( $\eta_p^2 = .45$ ). A follow-up Tukey Test was conducted to evaluate pairwise differences in the means. Results indicated that both the IC and GA Groups were significantly greater than the Control Group ( $p < .001$ ), but no such difference was observed between the IC Group and GA Group ( $p = .77$ ). Table 3 displays the statistics for the present study.

In order to confirm a type II error (reporting a lack of statistically significant difference when one in fact does exist) had not been made in the follow-up Tukey Test of the IC Group and GA Group, a post hoc power analysis was conducted using Faul and Erdfelder's (1992) software package G\*Power. The results revealed there was more than sufficient power ( $1 - \beta = .91$ ) to conclude that no such error existed, and there was indeed no statistically significant difference in the posttest scores of the IC and GA Groups.

Table 3. Statistical Comparisons in the Present Study

Group	n	Test	Mean	SD	Skewness	Kurtosis	95%CI
IC	17	Pre	14.06	3.80	0.31	0.55	[12.41, 16.34]
		Post	23.47	7.07	0.76	0.55	[19.74, 27.50]
GA	19	Pre	13.95	4.17	-0.67	0.03	[12.25, 16.38]
		Post	24.89	6.96	-0.04	0.53	[22.15, 29.72]
Control	16	Pre	12.75	3.79	0.54	-0.66	[10.73, 14.77]
		Post	12.56	3.65	-0.12	0.57	[10.62, 14.51]

## 5. Discussion

The results confirm the authors' suspicions that giving class averages to students does not bring any salient benefits from the perspective of higher scores in vocabulary tests. This also means, however, that the Idio-comparative group did not see a



significant improvement over the group of students given the class averages. Several reasons could be hypothesized for this result: students not being used to the Idio-comparative approach to marking; and the short-length of the study.

First, it is well-known by many teachers in Japan, whether they are teachers of the English language or not, that giving averages and comparing within the group are looked upon with much importance by educators in relation to university entrance examinations. Many students and teachers often refer to the *hensachi*, or standard rank scores, when making decisions about the university they wish to enter. This score is used as a means of comparison to all other students in the country, and the required *hensachi*, to enter every university in Japan, as well as individual *hensachi* are provided to both students and their high school teachers.

For example, the national *hensachi* average based on practice tests is always given a benchmark score of 50, with the majority of students being ranked with *hensachi* scores between 30 (very low compared to the national average) to 70 (very high compared to the national average). A student with a *hensachi* of 75, for example, is ranked in the top 1% of test takers (Nanapi, 2015). The *hensachi* has also been used in research to measure the English language proficiency of subjects in SLA research (see Hirai, 2014).

Although the participants in the study were not asked about their *hensachi*, it is likely that they were aware of this score as high school students before applying to a public university, and may have been more familiar with a scoring method utilizing comparison, rather than a scoring approach only comparing individual progress. Therefore, this new system of marking may have had some influence on the students' approach to study. It is possible that some participants preferred to know the average score and used this as a motivation to study harder. In a study conducted over a longer time span, participants could become used to the Idio-comparative marking method, possibly resulting in higher motivation to study, and better results on vocabulary tests than the other groups. As Wiggins (1989) stresses, when one attempts to beat others and be the best, it is a never-ending battle, unless one quits, of course. However, when one attempts to beat his or her previous ability, it leads to a higher drive to achieve more. Based on this advice, we feel it is imperative for students not to see other students in the class as their rivals but view their only rival as themselves and strive to be better today than they were yesterday.

In addition, the length of the study may not have been long enough to result in a statistically significant difference between the two groups at the posttest stage. Studies investigating the effects of various lengths of studying abroad on motivational changes in participants (e.g., Sasaki, 2011), have suggested that a minimum of one university semester (i.e., three to four months) is required to see any clear changes in students' drive to study and consequently changes in language proficiency. It will be beneficial to conduct future studies over a longer timespan in order to gain a deeper understanding of whether using the Idio-comparative approach when returning tests

can be more beneficial for increasing students' learning motivation than giving the class averages.

Finally, the large number of female subjects in comparison to male subjects in the present study may have been influential on the results. In Dweck and Reppucci's (1973) study, which focused on younger learners, it was discovered that male students tend to keep on trying more in comparison with female students when faced with situations in which they may fail. More recently, however, the opposite pattern was apparent in a study conducted with subjects of a similar age to those in the present study, in which male participants tended to give up easily, but females did not (Leis, 2013). More consideration of the gender of subjects, as well as their motives for learning may bring about more salient results in future studies.

The imbedded belief among Japanese students that knowing the group average is important, the relatively short length of the study, and the imbalance of gender may have affected the present study to the extent that no significant difference was observed between the Give-Average Group and Idio-comparative Group. However, the results in this study have suggested that there are no advantages in giving class averages to students. The comparison with one's own score may bring about a stronger feeling of responsibility as has been discussed and recommended in classic papers in psychology, such as Crandall, Katkozsky, and Crandall (1965) and Dweck and Reppucci (1973). With the backing of theoretical studies of the Self-worth Theory in psychology research, and the lack of distinct statistical advantage for giving the class averages to students in the study, the authors come to the conclusion that the benefits of not giving students the class averages may outweigh giving class averages to students.

Admittedly, the idea of an Idio-comparative approach to test marking is fairly unfamiliar and more research is necessary, the authors would like to encourage teachers to abstain from giving the class averages, and instead use this Idio-comparative approach in order to create classroom atmospheres in which the effort and proficiency of each individual student is compared not with those of others but only with their own.

## **6. Conclusion**

Competition with others is overrated and should be eliminated as much as possible from schooling (Covington, 1998). Despite some studies suggesting that competition with other students' results in worsened academic performance (e.g., Boggiano & Pittman, 1992), many schools and teachers appear to encourage the idea that one must be better than all the others. In a discussion on the importance of personalized instruction in the classroom, Keefe (2007) argues, "personalization is the inevitable design and fabric of schooling" (p. 223). It is the responsibility, Keefe opines, of schools and teachers to provide a learning environment in which each individual student is given instruction based upon his or her own learning style and progress. "If grades are given, they are assigned on an absolute basis and certify what a student has or has

not learned, not where the student stands in relation to classmates” (Keefe, 2007, p. 222).

In the present paper, the authors have introduced the idea of Idio-comparative scoring, an approach, which encourages students to compare their test scores to their previous scores rather than the scores of others in the class. Because the Idio-comparative approach did not appear to show any salient differences in the scores of a class receiving class averages, the authors are satisfied that evidence has been provided to show that this new approach can be useful for the language classroom. Because research surrounding the Self-worth Theory in SLA is still very minimal, there is a need for further studies in this area, especially in the EFL environment in Japan. We hope more research can be conducted to investigate the benefits of the Idio-comparative scoring approach and whether this way of providing personalized feedback holds any distinct advantages or disadvantages in comparison to giving scores along with class averages.

### Acknowledgements

The authors wish to thank the participants in the present study. Many thanks also go to Sachiko Nakao for her comments on a previous version of this paper. Much appreciation also goes to the two anonymous reviewers for their valuable suggestions. All remaining errors are the responsibility of the authors.

### References

- Boggiano, A. K., & Pittman, T. S. (1992). *Achievement and motivation: A social-developmental perspective*. Cambridge, UK: Cambridge University Press.
- Covington, M. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. New York, NY: Cambridge University Press.
- Covington, M. (1998). *The will to learn*. New York, NY: Cambridge University Press.
- Covington, M. (2004). Self-worth theory: Goes to college or do our motivation theories motivate? In D. M. McInerney, S. V. Etten (Eds.) *Big theories revisited*. (pp. 91-114). Charlotte, NC: Information Age Publishing.
- Crandall, V. C., Katkozy, W., & Crandall, V. J. (1965). Childrens' beliefs in their own control of reinforcement in intellectual-academic situations. *Child Development*, 36, 91-109. <http://doi.org/b6ff76>
- Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology*, 31(4), 674-685. <http://doi.org/c2zrzk>
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Ballantine Books.
- Dweck, C. S., & Reppucci, N. D. (1973). Learned helplessness and reinforcement responsibility in children. *Journal of Personality and Social Psychology*, 25(1), 109-116. <http://doi.org/dtz6bq>
- Faul, F., & Erdfelder, E. (1992). GPOWER: *A priori, post-hoc, and compromise power analyses for MS-DOS* [Computer program]. Bonn, Germany: Bonn University, Department of Psychology.
- Hirai, A. (2014). A Review of Four Studies on Measuring Vocabulary Knowledge. *Vocabulary Learning and Instruction*, 3 (2), 85\_92. doi: 10.7820/vli.v03.2.hirai
- Hoppe, F. (1930). Untersuchungen zur Handlungs – und Affektpsychologie IV [Psychological studies of action and affect IV: Success and failure]. *Erfolg und Misserfolg Psychologische Forschung*, 14, 1-63. <http://doi.org/fsxt3p>

- Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, 21, 112-126. <http://doi.org/dbv3tc>
- Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: implications for contingent self-worth and coping. *Developmental Psychology*, 35(3), 835-847. <http://doi.org/cvfs7j>
- Keefe, J. W. (2007). What is personalization? *Phi Delta Kappan*, 89(3), 217-223. <http://doi.org/b4v2>
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. London, UK: Longman.
- Leis, A. (November, 2013). *The effects of classroom English on proficiency and self-confidence: A self-worth perspective*. Presented at The Tohoku English Language Education Society Miyagi Conference, Sendai, Japan.
- Leis, A. (2014). The self-confidence and performance of young learners in an EFL environment: A self-worth perspective. *JES Journal*, 14, 84-99. Retrieved from: [https://www.academia.edu/8481565/The\\_Self-confidence\\_and\\_Performance\\_of\\_Young\\_Learners\\_in\\_an\\_EFL\\_Environment\\_A\\_Self-worth\\_Perspective](https://www.academia.edu/8481565/The_Self-confidence_and_Performance_of_Young_Learners_in_an_EFL_Environment_A_Self-worth_Perspective)
- Leis, A., Suzuki, W., & Ando, A. (2011). English ability and the L2 motivational self among Japanese university students. *Journal of the Tohoku English Language Education Society*, 31, 69-80. Retrieved from: [https://www.academia.edu/26059533/English\\_ability\\_and\\_the\\_L2\\_motivational\\_self\\_among\\_Japanese\\_university\\_students](https://www.academia.edu/26059533/English_ability_and_the_L2_motivational_self_among_Japanese_university_students)
- MacIntyre, P., & Gregersen, T. (2012). Affect: The role of language anxiety and other emotions in language learning. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for language learning: Insights from research, theory & practice*. (pp. 103-118). Basingstoke, UK: Palgrave MacMillan.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33-52. <http://doi.org/c3bnbh>
- Nanapi. (2015, April 18). Imasara hito ni kikenai...hansachi tte nani? [It's too late to ask, but what is a *hansachi*?] [Web log post]. Retrieved from <http://news.livedoor.com/article/detail/10020777/>
- Sasaki, M. (2008). The 150-year history of English language assessment in Japanese education. *Language Testing*, 25(1), 63-83. <http://doi.org/c2d5jd>
- Sasaki, M. (2011). Effects of various lengths of study-abroad experience on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45(1), 81-105. <http://doi.org/fq7vd4>
- Seligman, M. E., Maier, S. F., & Geer, J. H. (1968). Alleviation of learned helplessness in the dog. *Journal of Abnormal Psychology*, 73(3), 256-262. <http://doi.org/fmkw5z>
- Weiner, B. (1974). *Achievement motivation and attribution theory*. Morristown, NJ: General Learning Press.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 48, 18-25. <http://doi.org/b4v4>
- Zimmermann, F., Schütte, K., Taskinen, P., & Köller, O. (2013). Reciprocal effects between adolescent externalizing problems and measures of achievement. *Journal of Educational Psychology*, 105(3), 747-761. <http://doi.org/b4v5>

---

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the Journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).