Published every February and August ISSN 1979-7877

EDUCARE
International Journal for Educational Studies

TRISTAN M. ESEO, ARMINA B. MANGAOIL & BETTINA M. SEDILLA

Confidence Judgment and Test Scores: Basis for Proposed Confidence Enhancement

***ABSTRACT**: One of the structured experiences of high school students is taking test. This experience is usually prepared as a regular academic activity in school. It aims to measure students' achievement, which can be gleaned from their test score. How the students provide their impressions or predictive response before they see actual test material and after taking the test are implicit valuations. Thus, implicit valuations are the predictions that answer the question, "what will be their actual test score in mind?". The study investigates the confidence judgment and test scores of graduating high school students by determining their differences, which are the bases for developing a program to enhance student confidence in test-taking. Its participants consist of five hundred seventy five (575) graduating students from San Pablo City National High School in the Province of Laguna, Philippines. The researchers used two instruments, which are the test prediction forms and teacher-made test. The findings showed that (1) there is a significant difference between the students' pre-test prediction and post-test prediction; (2) there is a significant difference between the students' pre-test prediction and actual test score; and (3) there is a significant difference between the students' post-test prediction and actual test score. It is concluded that students' are generally overconfident. Furthermore, it is recommended that there is a need to identify factors that affect students' overconfidence. A module was proposed based on the findings of research.*

***KEY WORDS**: Confidence Judgment; Test Scores; Test Prediction Forms; Teacher-Made Test; Confidence Enhancement.*

INTRODUCTION

One of the structured experiences of high school students is taking test. This experience is usually prepared as a regular academic activity in school. It aims to measure students' achievement which can be gleaned from their test score.

According to Zachary Stein, Theo Dawson & Kurt W. Fischer (2010), "Every year, across the globe, tens of millions of children, adolescents, and adults from all walks of life take tests" (Stein, Dawson & Fischer, 2010:207). This statement affirms the important use of test in a global scale. The use of test plays a pivotal role in shaping our educational institution and society (Bonaccio & Reeve, 2010). Thus, the results of tests in this social activity are not without meanings.

Teachers are active test-users in school. They use test score to measure their students'

About the Authors: **Tristan M. Eseo** is a Special Education Teacher at the *Paaralang Pag-ibig at Pag-asa* at San Pablo City, Laguna, Philippines. **Prof. Armina B. Mangaoil** and **Bettina M. Sedilla, Ph.D.** are the Lecturers at the Faculty of Behavioral and Social Sciences PNU (Philippine Normal University), Taft Avenue, Manila 1000, Philippines. For academic interests, the authors can be contacted via their e-mails at: tmseo77@firstasia.edu.ph, mangaoil.ab@pnu.edu.ph, and bettinasedilla@yahoo.com

How to cite this article? Eseo, Tristan M., Armina B. Mangaoil & Bettina M. Sedilla. (2017). "Confidence Judgment and Test Scores: Basis for Proposed Confidence Enhancement" in *EDUCARE: International Journal for Educational Studies*, Vol.9(2), February, pp.121-128. Bandung, Indonesia: Minda Masagi Press owned by ASPENSI, ISSN 1979-7877.

Chronicle of the article: Accepted (October 9, 2016); Revised (January 20, 2017); and Published (February 27, 2017).

learning ability and skills to help them judge, whether their students have gained the learning competencies that they need to master. They make decisions based on the result of tests (Kane, 2013).

The changes and challenges in the use of assessment tools can provide greater opportunity for teachers to be creative in the way they conduct educational assessment. In doing this, they must not compromise the integrity of the tool that they are to use. They should properly measure the competencies taught by the experts. If this will be achieved, it will make their critical participation in structured experiences of test taking more meaningful and helpful.

There are two things that can be observed on these valuations. These are the explicit valuations and implicit valuations (Damodaran, 2006). Explicit valuations are direct valuations in a test score, i.e. actual test scores. It is a direct quantitative measure for the purpose of the test, while the implicit valuations of tests scores are indirect valuations usually associated with the actual test score and have no manifestations as far as the evaluation of test score is concerned.

For example, one can determine if (i.e. using actual test score) a student passed or failed the test following a criterion set in the manual of the test. These are the explicit valuations. However, implicit valuations are subjective perceptions of the test users in test scores. The meanings associated with are not found in the test or its direct intended purpose. These are test users' interpretation of the test scores, which are beyond what the test scoring protocols or manual provides.

How the students provide their impressions or predictive response (e.g. predictions of test scores) before they see actual test material and after taking the test are implicit valuations. Thus, implicit valuations are the predictions that answer the question, "what will be their actual test score in mind?". UNESCO (United Nations Educational, Scientific, and Cultural Organization), in 2011, advocates that:

> [...] learning in education could be more effective and inclusive when students are engaged in thinking about their own and others' thinking, thereby developing a metacognitive awareness of the basis for their own present thinking and of the development of their thinking as they learn (UNESCO, 2011:30).

This advocacy to develop student metacognition and self-regulated learning could inform the challenges of learning and education in the 21st century. The research and understanding of cognitive processes had grown far since the beginning of the cognitive revolution in the 1950's up to the present. This paradigm shift from the behaviorism has led researchers to break new grounds on the nature of human behavior (Miller, 2003).

J. Flavell (2012), on his seminal work, contributed to this vast development in cognitive psychology. It continued to progress and had been expanding with its marriage to other scientific discipline (Flavell, 2012). Emerging theories and the use of metacognitive constructs to further analyse cognition and human behavior included new terms like confidence judgment, judgment of learning, tip of the tongue states, feeling-of-knowing, and source judgment (Dunlosky& Metcalfe, 2011).

In relation with its emergence, it can also be noted in literature that these new terms used in the metacognitive research did not deviate essentially with other terms used in the field of psychological science such as self-regulation, self-regulated behavior, and self-regulated learning (Efklides & Misailidi, 2010; and McIntire, Miller & Lovler, 2011). Certain patterns were observed regarding the derivation of terms used on metacognitive research.

These include uses of tests have a purpose of: (1) defining metacognition; (2) identify students level of metacognitive abilities; and (3) improving student learning (Efklides & Misailidi, 2010). On the other hand, factors that affect test and performance outcome include: personal characteristics, behavioral approaches to test taking, and test material (Hassanbeigi *et al.*, 2011).

Most of the studies on metacognition from its definition, measurement, and practical application typically used standardized tests (Bajar, 2013). But, the limited purpose of standardized test does not usually addressed

ISSN 1979-7877 and www.mindamas-journals.com/index.php/educare

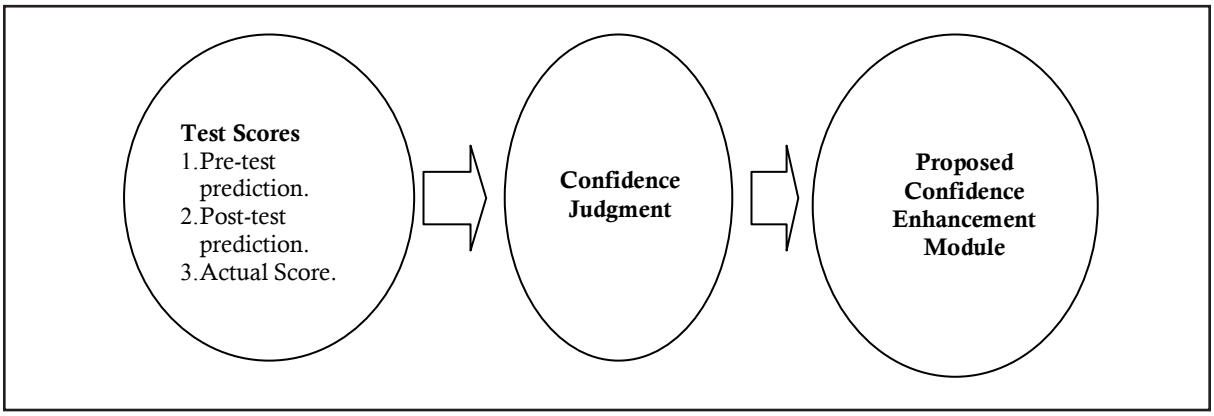

Figure 1:
Conceptual Framework of the Study

the issues on ecological validity (Hacker, Bol & Keener, 2013), processes that must link instruction. However, by using test material in a natural or classroom setting, the concern for these issues can be addressed (Rosenthal *et al.,* 2010; and Hacker, Bol & Keener, 2013).

METHOD

The researchers used the descriptive-correlational design (Gravetter & Wallnau, 2010; and Gravetter & Forzano, 2012). The confidence judgment is observed through the post-test prediction, that is the students' re-evaluation of their prediction after taking the test and answering it. The interest of studying how students use their test result is to elaborate the simple notion of relying on input-output process of students' actual performance in answering the teacher-made test. It must be clarified that students test scores are influenced by factors, such as students' capabilities in answering the test, the nature of the test or material itself, and how they perceived actual observation of results (Cole, 2008).

Rather than emphasizing a single notion of a test score, i.e. actual test score, which is to pass or fail the students, an aspect of this research highlighted a reflective value of test scores to measure their confidence judgment. This will help students sustain self-regulatory function in using test scores, because the feedback of information from the teacher-made score can led students to monitor their learning.

Students' confidence judgment is the re-evaluation of pre-test prediction using of post-test-prediction scores (Hadwin & Webster, 2013; and Henneman, 2014). This will be sustained by a proposed module for student active control of learning or self-regulated learning. See figure 1.

Test scores are goals in the mind of students. Test scores are important to better understand student learning in a classroom. According to B.J. Zimmerman (2008) and B.J. Zimmerman & D.H. Schunk (2011), self-regulated learning is defined as the process whereby students activate and sustain cognitions and behavior systematically oriented toward the attainment of their learning goals (Zimmerman, 2008; and Zimmerman & Schunk, 2011).

The participants of this study were graduating high school students of San Pablo City National High School composed of five hundred seventy-five students. The participants belong to the last batch of basic education curriculum for the school year 2014-2015, which is being replaced by the new curriculum known as the "K to 12 Program" through the initiative of the Department of Education in the year 2011 (Okabe, 2013).

In this study, test scores are indicators for the measure of students' cognitive and metacognitive capabilities (Roebers, Schmid & Roderer, 2009; and Efklides & Misailidi, 2010). Students are not typically made aware of their confidence judgment before and after taking test, since they are more concerned

with actual test scores than test predictions. However, by using the actual result of the test with student test predictions, it is certain to clarify and measure the confidence judgment of students in a classroom setting (Duckworth *et al.*, 2011; and Praetorius *et al.*, 2013).

The modules to be implemented can serve as a program in enhancing students' confidence in predicting accurate test result. It can be tested for its effectiveness in moderating the confidence of students' test predictions.

The study focused on the test predictions of graduating high school students' actual test score in the English teacher-made test. The components of test predictions were pre-test prediction and post-test prediction. It was used to measure the confidence judgment of students. The predictions component can be considered as a plausible study, since it can be observed in a highly structured activity in school or classroom setting (Black & Wiliam, 1998; and Schunk & Zimmerman, 2007).

This study is grounded on using student actual test score in an English teacher-made test. This score is basically used for determining student mastery of the subject. Teachers are depending on its practical use in a real classroom setting. Without this score, teachers cannot provide information on how learning takes place in the minds of students (Shepard, 2000).

From this ground, the researchers measure the confidence judgment of students in taking test. It is measured with test predictions on the macro-level of confidence judgement (*cf* Rosenthal *et al.,* 2010; Krebs & Roebers, 2012; and Hickey, 2014).

The first part of the study focused on the monitoring aspect metacognition. It highlights the nature of test scores and confidence judgment. The second part of the study was a proposed module to help students' in the accuracy of prediction in taking test based. This is the control aspect of the self-regulated learning, which can be observed when the module is implemented for its effect to enhance students' confidence.

There are two instruments used in this study. These are teacher-made test in English and test prediction forms. The teacher-made test was validated by the English teachers, who also helped in the study. To establish the reliability and validity of the instrument, the researchers used the table of specification for content validation and conducted a test of reliability from the data that was provided by the English teachers. It is composed of sixty items.

The test prediction forms were validated by research and educational experts in the field. Its face validation was also conducted with another group of graduating high school students. The procedure for the data gathering highlights the necessary steps in the collection of data for interpretation.

Weeks before the actual schedule of the test in English, the researchers went through the following procedures: (1) approval from the leaders in the school community; (2) orientation of English teachers and test proctors on the nature of the study; (3) gathering of pertinent information before the actual test administration; (4) preparation of material; and (5) gathering of test result after the test. The data collected were used to compute for the t-test of correlated means in answering the hypotheses of the study.

RESULTS AND DISCUSSION

This study aimed to determine the following hypotheses and develop a module based on these findings: (1) Is there a significant difference between students' post-test prediction and actual test score?; (2) Is there a significant difference between pre-test prediction and post-test prediction?; (3) Is there a significant difference between students' pre-test prediction and actual test score?; and (4) What module can be develop to enhance students' confidence in test-taking?

Hypothesis 1. Table 1 shows that there is a significant difference between the students' pre-test prediction and post-test prediction.

It can be said that students are overconfident in their pre-test predictions based from their post-test predictions. The overconfidence on pre-test predictions or its inaccuracy can be calibrated through their accurate post-test predictions to get a proximate value of the pre-test predictions. The lesson on confidence judgment can addressed this need.

ISSN 1979-7877 and www.mindamas-journals.com/index.php/educare

Table 1:
Computed Values of Pre-Test Prediction and Post-Test Prediction

n = 575	Pre-Test Prediction	Post-Test
Mean	28.23	30.92
Standard Deviation	7.12	8.80
Pearson Correlation	0.51	-
Computed t Stat	8.03	-
t Critical two-tailed	1.96	-

Table 2:
Computed Values of Pre-Test Prediction and Actual Test Scores

n = 575	Pre-Test Prediction	Actual Score
Mean	28.23	23.74
Standard Deviation	7.12	5.34
Pearson Correlation	0.30	-
Computed t Stat	14.37	-
t Critical two-tailed	1.96	-

Table 3:
Computed Values of Post-Test Prediction and Actual Test Scores

N = 575	Post-Test Prediction	Actual Score
Mean	30.92	23.74
Variance	77.37	28.52
Standard Deviation	8.80	5.34
Pearson Correlation	0.23	
Computed t Stat	18.78	
t Critical two-tailed	1.96	

Hypothesis 2. Table 2 shows that there is a significant difference between the students' pre-test prediction and actual test score.

This result indicates students' low performance on mastery of the English subject as observed in the actual test scores. Considering that the pre-test prediction are much higher than it. Students' low performance can proximate this difference by providing the following lessons in the proposed module.

Hypothesis 3. Table 3 shows that there is a significant difference between the students' post-test prediction and actual score.

This result can be interpreted that students are overconfident on their post-test predictions compared with the actual test scores. This overconfidence is expected to be calibrated with the following lessons, which are focusing on Study Skills and Habits and Test-Taking Strategies. The four lessons found in the module are designed to help students in their academic work, i.e. taking test.

The whole activities found in the module are made in such way that students can simulate the test taking situation. In particular, this will make students confidence judgment be accurate as their go through the learning process. These include lesson on confidence judgment, goal setting, developing study skills and habits, and using test-strategies.

CONCLUSION

Based on the findings that determined, students' are generally overconfident in predicting their actual score in a teacher-made test. However, one thing needs important mentioning in this general observation. That is when the students were to provide their post-test prediction, it did not help them accurately adjust predictions in reference to their target actual or pre-test predictions which is quite unexpected.

Instead, the students' overconfidence was exacerbated as observed in the post-test

prediction. This observed overconfidence predicts an actual score that was observed on the computed mean of test predictions, which are above the computed mean of actual scores. There is a need to conduct research and activities that will help students, who are overconfident to elaborate further its inconsistent result for students to have accurate predictions of actual test scores and for an enhanced confidence judgment.

The recommendation for research aimed to gain a better understanding of why the students turned out to be overconfident. What are the factors that affect students' overconfidence in a teacher-made test? How do these variables relate with each other? What are the practical applications of this knowledge to help students predict scores accurately?

Specifically, to identify what are the factors that affect students' in pre-test prediction and post-test prediction; pre-test prediction and actual test score; post-test prediction and actual test-score; prior to taking a teacher-made test, such as (i.e. age, test anxiety, and motivation); behavioral approaches in taking a teacher-made test (i.e. use test strategies, goal setting skills, and study habits); level of emotions during, before and after taking the test; perception of the test material (before and after) taking the teacher-made test with regards to its level of difficulty and test expectancy; and, lastly, test the module for its effectiveness.[1]

References

Bajar, P.A. (2013). *Metacognition as Correlates of Achievement in Chemistry Using Predict-Explain-Observe-Explain Approach.* Manila: PNU [Philippine Normal University] Press.

Black, Paul & Dylan Wiliam. (1998). "Assessment and Classroom Learning" in *Assessment in Education: Principles, Policy & Practice*, Vol.5(1), March.

Bonaccio, S. & C. Reeve. (2010). "The Nature and Relative Importance of Students: Perception of the Sources of Test Anxiety" in *Learning and Individual Differences*, pp.617-625.

Cole, Robert W. (2008). "Educating Everybody's Children: Diverse Teaching Strategies for Diverse Learners". Available online at: http://www.ascd.org/publications/books/107003/chapters/Educating-Everybody%27s-Children [accessed in Manila, Philippines: August 15, 2016].

Damodaran, Aswath. (2006). "Valuation Approaches and Metrics: A Survey of the Theory and Evidence". Available online at: http://people.stern.nyu.edu/adamodar/pdfiles/papers/valuesurvey.pdf [accessed in Manila, Philippines: August 15, 2016].

Duckworth, A. *et al*. (2011). "Role of Test Motivation in Intelligence Testing". Available online at: http://www.pnas.org [accessed in Manila, Philippines: August 15, 2016].

Dunlosky, J. & J. Metcalfe. (2011). *Metacognition.* London: SAGE Publication.

Efklides, A. & P. Misailidi. (2010). *Trends and Prospects in Metacognition Research.* New York: Springer Science.

Flavell, J. (2012). "Metacognition and Cognition Monitoring: A New Era of Cognitive-Developmental Inquiry" in *American Psychologist*, pp.906-911.

Gravetter, F.J. & L.B. Wallnau. (2010). *Essentials of Statistics for the Behavioral Sciences*. Belmont, CA: Wadsworth Cengage Learning.

Gravetter, F.J. & L.B. Forzano. (2012). *Research Methods for the Behavioral Sciences*. Belmont, CA: Wadsworth Cengage Learning, 4th edition.

Hacker, D.J., L. Bol & M.C. Keener. (2013). "Metacognition in Education: A Focus on Calibration" in J. Dunlosky & R. Bjork [eds]. *Handbook of Memory and Metacognition.* New Jersey: Lawrence Erlbaum Associates, pp.429-455.

Hadwin, A.F. & E.A. Webster. (2013). "Calibration in Goal Setting: Examining the Nature of Judgment Confidence" in *Learning and Instruction*, pp.37-47.

Hassanbeigi, A. *et al*. (2011). "The Relationship between Study Skills and Academic Performance of University Students" in *Procedia Social and Behavioral Sciences*, pp.1416-1424.

Henneman, G. (2014). "Suggestions for Building Students Confidence". www.statemanjournal.com [accessed in Manila, Philippines: December 1, 2016].

Hickey, C. (2014). "Hundreds of High School Seniors Refuse to Take New State Standardized Exam". Available online at: Retrieved December 10, 2014, from www.kwgn.com/2014/11/14/hundreds-of-high-school-seniors-refused-to-take-new-state-standardized-exam/ [accessed in Manila, Philippines: December 25, 2016].

Kane, M. (2013). "Validating the Interpretations and Uses of Test Scores" in *Journal of Educational Measurement*, pp.1-73.

Krebs, S.S. & C.M. Roebers. (2012). "The Impact of Retrieval Processes, Age, General Achievement Level, and Test Scoring Scheme for Children's Metacognitive Monitoring and Controlling" in *Metacognition & Learning*, 7, pp.75-90. DOI:10.1007/s11409-011-9079-3.

McIntire, S.A., L. Miller & R. Lovler. (2011). *Foundations of Psychological Testing: A Practical Approach.* London: SAGE Publication, 4th edition.

Miller, George A. (2003). "The Cognitive Revolution:

[1] ***Statement*:** Herewith, we declare that this paper is our original work; it is not product of plagiarism and not reviewed or published by other scholarly journals elsewhere.

ISSN 1979-7877 and www.mindamas-journals.com/index.php/educare

A Historical Perspective" in *TRENDS in Cognitive Sciences*, Vol.7, No.3 [March]. Available online also at: http://www.cs.princeton.edu/~rit/geo/Miller.pdf [accessed in Manila, Philippines: December 25, 2016].

Okabe, Masayoshi. (2013). "Where Does Philippine Education Go? The 'K to 12' Program and Reform of Philippine Basic Education". Available online at: https://ir.ide.go.jp/dspace/bitstream/2344/1267/3/ARRIDE_Discussion_No.425_okabe.pdf [accessed in Manila, Philippines: December 25, 2016].

Praetorius, A.K. *et al*. (2013). "Judgment Confidence and Judgment Accuracy of Teachers in Judging Self-Concepts of Students" in *The Journal of Educational Research*, pp.64-76.

Roebers, C.M., C. Schmid & T. Roderer. (2009). "Metacognitive Monitoring and Control Processes Involved in Primary School Children's Test Performance" in *British Journal of Educational Psychology*, pp.749-767.

Rosenthal, G. *et al.* (2010). "Do Students Know if They Answered Particular Questions Correctly on a Psychology Exam" in *Journal of Instructional Psychology*, pp.57-62.

Schunk, D.H. & B.J. Zimmerman. (2007). "Influencing Children's Self-Efficacy and Self-Regulation of Reading and Writing through Modeling" in *Reading & Writing Quarterly*, 23(1), pp.7-25.

Shepard, Lorrie A. (2000). "The Role of Classroom Assessment in Teaching and Learning". Available online at: http://cresst.org/wp-content/uploads/TECH517.pdf [accessed in Manila, Philippines: December 25, 2016].

Stein, Zachary. Theo Dawson & Kurt W. Fischer. (2010). "Redesigning Testing: Operationalizing the New Science of Learning" in Myint Swe Khine & Issa M. Saleh [eds]. *New Science of Learning: Cognition, Computers, and Collaboration in Education*. New York: Springer, pp.207-224.

UNESCO [United Nations Educational, Scientific, and Cultural Organization]. (2011). *Current Challenges in Basic Science Education.* Paris: UNESCO Education Sector.

Zimmerman, B.J. (2008). "Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects" in *American Educational Research Journal*, 45(1), pp.166-183.

Zimmerman, B.J. & D.H. Schunk. (2011). "Self-Regulated Learning and Performance: An Introduction and an Overview" in B.J. Zimmerman & D.H. Schunk [eds]. *Handbook of Self-Regulation of Learning and Performance.* New York: Routledge, pp.1-12.

ISSN 1979-7877 *and* ***www.mindamas-journals.com/index.php/educare***

Students of San Pablo City National High School in the Philippines
(Source: http://www.ugnayan.com/ph/Laguna/SanPablo, 15/1/2017)

The participants of this study were graduating high school students of San Pablo City National High School composed of five hundred seventy-five students. The participants belong to the last batch of basic education curriculum for the school year 2014-2015, which is being replaced by the new curriculum known as the "K to 12 Program" through the initiative of the Department of Education in the year 2011.

ISSN 1979-7877 and www.mindamas-journals.com/index.php/educare