

The Effectiveness of Assessment Module in Taiwan Adaptive Learning Platform (TALP) for Remedial Instruction

Bor-Chen Kuo, Ph.D.¹, Tao-Yi Chang, Ph.D.², and Li-Wei Chang, M. Ed.³

¹National Taichung University of Education, Taiwan, R.O.C. kbc@mail.ntcu.edu.tw

²National Taichung University of Education, Taiwan, R.O.C. snoopyfred@kimo.com

³National Taichung University of Education, Taiwan, R.O.C. liedward0122@gmail.com

INTRODUCTION

Taiwan Adaptive Learning Platform (TALP), called Yin Tsai Wan (因材施教) in Chinese, is the official learning platform of Ministry of Education (MOE) in Taiwan. TALP has applied artificial technologies to offer personal learning environment for students, and the learning material is designed underlying the framework of 12-year Basic Education Curriculum Guidelines. There are four subjects covered in this platform: Chinese, Mathematics, English, and Science. For its instructional repertoire, TALP provides thousands of videos for micro-learning, adaptive assessments for diagnosing students' weakness, and interactive modules for scaffolding learning through Intelligent tutoring agents. Though some studies suggested that assessment was crucial for an online learning platform (Azevedo & Azevedo, 2018), however, most studies have paid more attention to the implementation of instructional videos and its effectiveness on learning (Zengin, 2017; Khalil & Ebner, 2017). Our study aimed to explore whether using assessment module in TALP could assist students to promote their academic achievement in the context of remedial instruction. Four kinds of tests were included in the assessment module of TALP, they are: exercise, dynamic assessment, unit test, and cross-grade diagnosis test. Our study also examined whether tests applied adaptive feedback mechanism would be effective on remedial instruction.

Relevant theories

An assessment mechanism can monitor progress and improve learning outcome only if it has a good feedback mechanism (Bell, 2013). Due to rating restriction, most online testing adopted multiple choice to assess students. Commonly, two types of feedback mechanisms for online multiple choice testing were identified: verification and elaboration (Shute, 2008). Verification is simply to judge whether an answer is correct; and elaboration is to deliver information which supplies relevant cues to scaffold learners toward a correct answer. As for effectiveness of feedback mechanism in the assessment by the meta-analysis (Bangert Drown et al., 1991), the effect size of verification feedback was expected at .22 and elaboration was .53.

Exercise in TALP provides verification feedback and dynamic assessment in TALP is equipped with elaboration feedback. But dynamic assessment in TALP is available in mathematics. Both exercise and dynamic assessment in TALP are presented after students finishing instructional videos. Unit test and cross-grade diagnosis test in TALP both were in multiple choice form but with adaptive feedback mechanism. When students finish tests, both types of tests may generate a report not only informing whether an answer is correct, but also recommend instructional videos according to mistakes students have made. Compared with unit test in TALP, cross-grade diagnosis is more proactive to detect learning weakness through grades. Cross-grade diagnosis test is exceptional for students who suffer misconception from previous grade learning experience (Yang, Kuo, Lin, 2021). Owing to structure knowledge embedded in

TALP, cross-grade diagnosis test can avoid unnecessary and plethoric items during cross-grade testing and it can find learning weakness with better accuracy, too (Wu, Kuo, & Yang, 2012).

The key for a successful remedial instruction is adaptive learning, it may overcome the problem of huge achievement gap between the vulnerable (Liu et al., 2017). Theoretically, assessment module in TALP, especially with adaptive and cross-grade feedback mechanism, should be very supportive for remedial instruction.

Type of study involved

To evaluate the effectiveness of assessment module in TALP, our study applied the passing result of students in National Progress Test as the dependent variable. Our independent variable were types of tests and subjects (Chinese, Mathematics and English). Our main goal was to examine whether tests in the assessment module of TALP are effective for remedial instruction. Our study also included effectiveness of instructional videos in TALP for estimating the overall mean effectiveness of TALP in remedial instruction.

Methods/Experiments

Our study surveyed 16899 students who failed in National Screening Test in May, 2020. These students were diagnosed as at-riskier learners and recommended to receive remedial instruction for 6 months. During this learning period, they applied TALP for their remedial instruction. In December, all participants received National Progress Test (in December, 2020) in order to review the effectiveness of remedial instruction. Both National Screen Test and National Progress Test were administered by MOE in Taiwan via the platform of PRIORI-tbt (Project for Implementation of Remedial Instruction-Technology-Based testing). Though all participant used TALP for learning, but tests in the assessment module of TALP were optional for students. Our study would compare the passing result of National Progress Test between users and non-users of a specific test in the TALP to evaluate its effectiveness.

Statistical Analysis and conclusion

The effectiveness of instructional video and each test in the assessment module in TALP were estimated in Cohen *d*, and all effect sizes were presented by the forest plot (Fig 1.). According to the confidence interval, all effect sizes were significant in Fig 1., and it indicated that students using TALP could have a better shot to pass National Progress Test. By synthesizing effect sizes, the overall mean effect size of TALP (including assessment module and instructional video across subjects) was .448 (CI:.324~.572), and the overall mean effect size of assessment module in TALP was .485(CI:.333~.638), both effect sizes were estimated in random effect model. It seemed that the most effective assessment in TALP was cross-grade diagnosis test in Mathematics, its effectiveness was 1.110. Among tests in the assessment module of TALP, exercise was the least effective test across subjects. Except cross-grade diagnosis in mathematics, effect sizes of cross grade-diagnosis, unit test and dynamic assessment were ranged .471~.572, of which magnitude was around medium effect (Cohen, 1988).

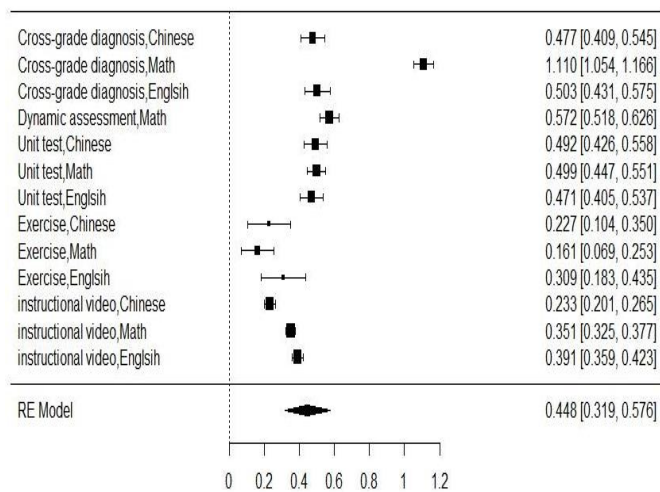


Fig1. Forest plot

Real World Application

The participant in our study were elementary school students receiving remedial instruction. Our data was derived from the real educational practice, and implementing TALP in remedial instruction is also one of the educational policies by MOE in Taiwan. The sample size of our study and the magnitude of effect size in our findings should be persuasive to support the effectiveness of the assessment module in TALP for remedial instruction. The results of cross-grade diagnosis and units test also confirmed that adaptive feedback mechanism was effective for remedial instruction. Instructional video in TALP seemed less effective than dynamic assessment, unit test and cross-grade diagnosis test, but the design of tests and feedback is based on the content of instructional videos in TALP. Our study won't encourage students purely concentrate on the assessment module in TALP due to its good effectiveness, because the effectiveness of tests were not independent but correlated with instructional videos.

Future

Our findings also imply that the cross-grade diagnosis test in TALP is effective on remedial instruction, but the effectiveness of cross-grade diagnosis test is nearly as the same as unit test in Chinese and English. The future study should pay more attention to why cross-grade mechanism works exceptionally well in mathematics, instead of Chinese and English.

SUMMARY

Our findings imply that assessment module in TALP is effective to promote academic achievement on remedial instruction and the magnitude of its effect size close to medium effect (Cohen, 1988). The results of cross-grade diagnosis test and unit test in TALP both support that adaptive feedback mechanism can facilitate remedial instruction. Compared with unit test, cross-grade diagnosis test is only superior in mathematics (Fig1.), it indicates that cross-grade mechanism is more effective on remedial instruction in mathematics than Chinese and English. Our results show that most tests in the assessment module in TALP are more effective than instructional videos, however, instruction is the most important part in learning and assessment doesn't work without instruction. To those who are interested in online learning platform or in remedial instruction, our study suggest that students should try to make use of feedback from assessment after watching instructional videos from online learning platforms.

REFERENCES

Azevedo, A., & Azevedo, J. (2018). Handbook of Research on

E-assessment in Higher Education. IGI Global: Hershey, PA, USA.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *The Journal of Educational Research*, 85(2), 89-99.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Khalil, M., & Ebner, M. (2017). Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories. *Journal of computing in higher education*, 29(1), 114-132.

Liu, M., McKelroy, E., Corliss, S. B., & Carrigan, J. (2017). Investigating the effect of an adaptive learning intervention on students' learning. *Educational technology research and development*, 65(6), 1605-1625.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.

Wu, H.-M., Kuo, B.-C., & Yang, J.-M. (2012). Evaluating Knowledge Structure based Adaptive Testing Algorithms and System Development. *Educational Technology and Society*, 15(2), 73-88.

Yang, S. J., Kuo, B. C., & Lin, C. P. (2021). A study of knowledge structure-based computerized adaptive test system for elementary school students on the effectiveness of remedial math. *International Journal on Digital Learning Technology*, 13(1), 23-48.

Zengin, Y. (2017). Investigating the use of the Khan Academy and mathematics software with a flipped classroom approach in mathematics teaching. *Journal of Educational Technology & Society*, 20(2), 89-100.



Bor-Chen Kuo received the Ph.D. degree in electrical and computer engineering from the Purdue University, West Lafayette, IN, in 2001. He is currently a Director, Department of Information and Technology Education, Ministry of Education, Taiwan, and a Distinguished Professor of the Graduate Institute of Educational Measurement and Statistics,

National Taichung University of Education, Taiwan. Since 2015. He served as the President of Chinese Association of Psychological Testing. He is the Associate Editor of *Educational Psychology* since 2016. Professor Kuo received an Outstanding Research Award from the R.O.C Ministry of Science and Technology (MOST) in 2017. His research interests include computerized adaptive testing, self-regulation learning, cognitive diagnostic modeling, machine learning, and artificial intelligence in education.