TPM Vol. 32, No. R2, 2025
ISSN: 1972-6325
https://www.tpmap.org/

EVALUATING A COMPUTERIZED ADAPTIVE LIFE SKILLS ASSESSMENT: EVIDENCE FROM THAI LOWER SECONDARY STUDENTS

PAKAKUL BUDKEN
KHON KAEN UNIVERSITY, THAILAND

PRAKITTIYA TUKSINO
KHON KAEN UNIVERSITY, THAILAND

Abstract This study aimed to evaluate the effectiveness of a Computerized Adaptive Testing (CAT) system for assessing life skills among Grade 9 students in Northeastern Thailand. The life skills scale used in this study was previously developed based on eight core components—Decision Thinking, Critical and Logical Thinking, Creative and Entrepreneurial Thinking, Emotional Control, Self-Awareness, Interpersonal Relationships, Citizenship, and Communication—with 120 polytomous items calibrated using the Graded Response Model (GRM) in Item Response Theory (IRT). In this phase, the CAT system was implemented with 35 students, and the system automatically selected items based on their ability levels (θ). Results revealed diverse life skill proficiency levels among students, with high measurement precision indicated by low standard errors of measurement (SEM) and a well-distributed θ score range. In addition, feedback from both students and teachers reflected positive attitudes toward the adaptive test's usability, fairness, and time efficiency. These findings support the application of CAT in life skills assessment and highlight its potential for enhancing student-centered assessment, educational equity, and scalable testing in Thai education policy reform.

Keywords: Computerized adaptive testing; life skills assessment; secondary education; non-cognitive measurement; Item Response Theory; Thai education; educational technology; psychometric analysis

Correspondence concerning this article should be addressed to Prakittiya Tuksino Email: praktu@kku.ac.th

[Space for the TPM staff**]**

Introduction

In the context of rapidly changing societies, the development of life skills has become a central concern in educational systems worldwide. Life skills, as defined by the World Health Organization (WHO), refer to a set of psychosocial competencies and interpersonal skills that help individuals make informed decisions, solve problems, think critically and creatively, communicate effectively, build healthy relationships, and cope with the demands and challenges of everyday life (WHO, 1999). These competencies are essential for promoting mental well-being, academic achievement, and future employability (UNESCO, 2015). Recognizing this importance, global educational frameworks such as the OECD Learning Compass 2030 and the Sustainable Development Goals (SDG 4.7) have emphasized the integration of life skills into school curricula.

In Thailand, the importance of life skills education is acknowledged through national education policies and curriculum standards. The Basic Education Core Curriculum B.E. 2551 (A.D. 2008), revised in 2017, includes life skills as part of student competencies to be developed throughout schooling. Furthermore, the 13th National Economic and Social Development Plan (2023–2027) underscores the need for youth to acquire 21st-century skills—including emotional regulation, communication, resilience, and self-awareness—as a means to drive national competitiveness and social harmony. However, despite policy commitments, the practical assessment of life skills in Thai schools remains limited. Conventional paper-based or teacher-scored evaluations often suffer from subjectivity, inefficiency, and a lack of diagnostic precision (Junruang et al., 2021).

To address these limitations, Computerized Adaptive Testing (CAT) has emerged as a promising approach. CAT dynamically adjusts item selection based on a student's ability level using item response theory (IRT), thereby improving measurement accuracy while reducing test length and respondent fatigue (Weiss, 2011). While CAT has been widely applied in cognitive assessments and large-scale testing (e.g., GRE, GMAT),

its adoption for non-cognitive constructs such as life skills is still underexplored—particularly in Southeast Asian educational settings. Recent research has begun to demonstrate that CAT systems can be tailored for non-academic domains, enabling personalized and scalable assessments of complex skills (Ebenbeck & Gebhardt, 2022; Papadakis et al., 2022).

Building on this momentum, our research team previously developed a validated life skills scale for Thai Grade 9 students using a polytomous IRT model with 120 items across eight core components (Junruang et al., 2024). That work focused on item construction, model calibration, and psychometric validation. However, the next critical step is to implement the scale in a real-world CAT environment and evaluate its utility. This study aims to evaluate the implementation of a Computerized Adaptive Life Skills Assessment (CALSA) among Grade 9 students in Thailand by:

1. Assessing students' levels of life skills based on CAT-generated theta (θ) scores;
2. Examining the accuracy and efficiency of the CAT system through standard error (SEM), item usage, and test duration;
3. Exploring students' and teachers' perceptions of the CAT-based assessment system.

LITERATURE REVIEW

Life Skills in Global and Thai Contexts

Life skills are defined by the WHO (1999) as essential psychosocial competencies that enable individuals to deal effectively with the demands of daily life. These include decision-making, critical thinking, creative thinking, effective communication, interpersonal relationship skills, self-awareness, coping with emotions, and coping with stress. International frameworks such as the OECD Learning Compass 2030 and UNESCO's Education 2030 Agenda have emphasized the integration of life skills as foundational for 21st-century learning (OECD, 2020; UNESCO, 2015).

In Thailand, life skills education is integrated into the national curriculum, particularly through the **Basic Education Core Curriculum B.E. 2551 (A.D. 2008)** and its revised standards (2017). The curriculum mandates the development of competencies such as problem-solving, communication, self-regulation, and emotional intelligence. However, the implementation and assessment of these skills remain inconsistent, with educators citing limited resources, subjective evaluation methods, and a lack of standardized tools (Saenboonsong & Poonsawad, 2024).

Assessment Challenges in Life Skills Education

Unlike academic subjects with clearly defined content domains, life skills involve complex behavioral traits that are difficult to assess objectively. Traditional pen-and-paper methods or rubric-based teacher assessments often result in inconsistencies and may not fully capture the dynamic nature of student competencies (Junruang et al., 2021). Moreover, such methods are labor-intensive and limited in terms of scalability and diagnostic feedback, especially when applied to large student populations or in under-resourced school contexts.

Computerized Adaptive Testing (CAT): Concepts and Benefits

Computerized Adaptive Testing (CAT) is grounded in Item Response Theory (IRT) and provides an efficient method for assessing individual performance by tailoring test items to a respondent's ability level (Weiss, 2011). Instead of presenting a fixed set of questions, CAT dynamically selects items based on real-time responses, enhancing measurement precision while reducing test length and respondent fatigue (Van der Linden & Glas, 2010).

In educational contexts, CAT has proven successful in large-scale assessments (e.g., GRE, TOEFL) and is increasingly explored in classroom-based diagnostic tools. Key benefits include:

- Enhanced accuracy through adaptive item selection
- Shorter testing time with fewer but more informative items
- Personalized feedback based on ability estimates (θ)

Although most CAT applications have focused on cognitive domains, recent studies demonstrate that CAT can also be applied to measure non-cognitive constructs such as emotional intelligence, self-concept, social skills, and life skills. Ebenbeck and Gebhardt (2022) successfully applied a CAT approach to assess inclusive progress in special education. Likewise, Papadakis et al. (2022) identified over 150 educational applications using adaptive mechanisms to assess problem-solving and creativity. The challenge lies in constructing robust item banks that align with behavioral indicators and ensuring valid parameter estimation under polytomous IRT models such as the Graded Response Model (GRM). While the feasibility of CAT for life skills has been demonstrated in pilot studies, large-scale implementation in Southeast Asia remains limited.

Gaps in Thailand's Life Skills Evaluation

Despite a strong policy focus on 21st-century competencies, Thailand still faces significant challenges in the effective assessment of life skills among school-aged children. Existing tools used in classrooms tend to be paper-based checklists, self-reports, or observation protocols with limited standardization or diagnostic precision. Furthermore, many schools lack psychometrically validated instruments that can provide formative feedback or support personalized learning paths. The use of item response theory (IRT) models in Thailand's educational measurement is still emerging, and the application of computerized adaptive testing (CAT) in the context of non-cognitive domains—especially life skills—has not been widely adopted. Most prior research has focused on cognitive outcomes, leaving a gap in terms of adaptive measurement for behavioral and socio-emotional competencies. Additionally, few studies have explored user perspectives—such as student and teacher attitudes—towards CAT-based life skills assessments in Thai schools. Consequently, there is a pressing need to pilot and evaluate systems that integrate CAT with validated life skills frameworks, allowing for more efficient, fair, and personalized assessment. This study responds to that gap by implementing a CAT-based life skills test within real school settings, focusing on both technical measurement outcomes and user feedback.

METHODOLOGY

Research Design

This study employed a quantitative, survey-based design using a Computerized Adaptive Testing (CAT) platform to assess students' life skills and evaluate the performance and acceptability of the assessment system. The design was exploratory in nature, aimed at gathering theta estimates, test functioning data, and user feedback from a targeted implementation in real school settings.

Participants

The participants were 35 Grade 9 students (aged 14–15 years) from four secondary schools in northeastern Thailand. The schools were selected using purposive sampling based on their willingness to participate in digital assessment pilot projects under the educational innovation network of the Faculty of Education, Khon Kaen University. All students were administered the CAT life skills assessment under standardized conditions. Participation was voluntary, and informed consent was obtained from both students and their guardians. The study received ethical clearance from the Human Research Ethics Committee of Khon Kaen University (HE673128).

TABLE 1 Demographic Characteristics of Participants (n = 35)

Variable	Category	Frequency	Percentage (%)
Gender	Male	18	51.40%
	Female	17	48.60%

Instruments and CAT System

Life Skills Assessment Framework The assessment tool was based on a validated framework comprising eight core components of life skills:

1. Decision Making (DM)
2. Critical and Logical Thinking (CLT)
3. Creative Thinking (CET)
4. Emotional Control (EC)
5. Self-awareness (SA)
6. Interpersonal Relationship (IR)
7. Communication and Expression (CE)
8. Coping with Stress (CS)

Each component was measured using polytomous items aligned with four behavioral levels. The item bank, initially developed and validated in earlier phases of the research, consisted of 120 items, calibrated using the Graded Response Model (GRM) under IRT.

CAT Platform and Configuration The CAT was deployed via a custom Moodle-based system integrated with an adaptive plugin. Key configurations included:

- Starting point: Average $\theta = 0.00$
- Item selection method: Maximum Fisher Information

- Termination rule: Standard Error of Measurement (SEM) ≤ 0.30 or maximum 15 items per student
- θ Estimation: Expected A Posteriori (EAP)

The system logged each test session, recording θ estimates, SEM, item counts, and test duration.
Usability and Feedback Questionnaire To assess user experience, a brief questionnaire was administered to both students and teachers after the CAT session. The instrument consisted of Likert-scale items (1–5) and open-ended questions addressing perceived usefulness, clarity of instructions, test engagement, and suggestions for improvement. Content validity was verified by three experts in educational technology and measurement.

RESULTS

Life Skills Proficiency Levels
The results are organized into three parts, addressing the study objectives: (1) life skills proficiency levels derived from CAT-based θ estimates, (2) technical performance of the CAT system, and (3) user feedback from students and teachers. The CAT system estimated the life skills proficiency (θ values) for each student based on their responses. The θ scores were normally distributed, with values ranging from –1.47 to +1.78. The average θ score was 0.16 (SD = 0.76), indicating a slightly above-average skill level among participants.

TABLE 2 Distribution of θ Scores for Life Skills Proficiency (n = 35)

θ Score Range	Level of Life Skills	Frequency	Percentage (%)
Above 1.00	Very High	6	17.10%
0.51 – 1.00	High	9	25.70%
–0.50 – 0.50	Moderate	13	37.10%
–1.00 – –0.51	Low	5	14.30%
Below –1.00	Very Low	2	5.70%

The computerized adaptive testing system produced life skills proficiency scores (θ) for each of the eight components among 35 Grade 9 students. Table 3 displays the descriptive statistics, including the mean, standard deviation (SD), and score range for each life skill component. Overall, the θ values ranged from approximately −1.34 to +2.20, indicating wide individual differences in students' life skill proficiencies.
The component with the highest average score was Critical and Logical Thinking (M = 0.51, SD = 0.91), followed by Emotional Control (M = 0.47), Interpersonal Relationships (M = 0.45), and Decision Thinking (M = 0.42). On the lower end, Self-Awareness showed the lowest average (M = 0.33), though still within the positive ability range. The distribution of scores suggests that most students possessed moderately developed life skills, with a few high-performers and a few students needing targeted interventions.
A radar chart (Figure 1) visually represents the mean θ values across the eight components. The chart highlights relatively balanced competencies across domains, with a slight emphasis on cognitive-regulatory skills.

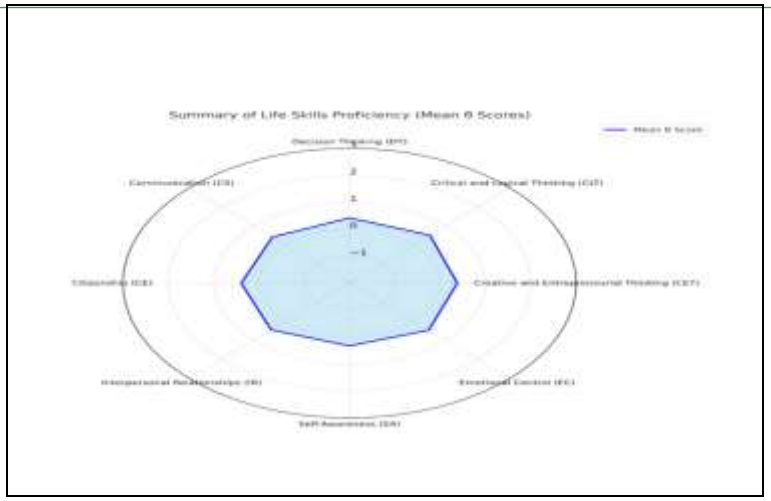


FIGURE 1 A radar chart Life Skills Proficiency Results

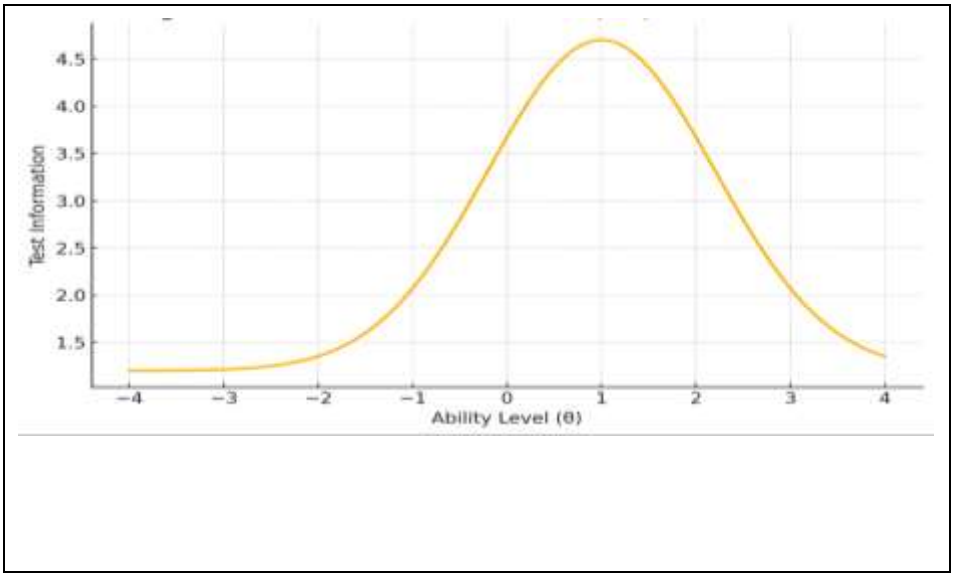
Technical Performance of CAT The efficiency and precision of the CAT were evaluated through three metrics: number of items administered, standard error of measurement (SEM), and test duration.

TABLE 3 CAT Test Statistics

Statistic	Mean	SD	Min	Max
Items Administered	10.46	2.78	7	15
SEM (Standard Error)	0.28	0.07	0.17	0.3
Test Duration (minutes)	9.6	3.1	5.2	15.8

The average SEM of 0.28 confirmed the test's measurement precision, and most students required fewer than 12 items to complete the test, indicating the CAT's strong efficiency and adaptivity.

FIGURE 2 Test Information Function (TIF) for the Full Scale
This figure illustrates that the CAT provides optimal information and measurement precision for individuals with θ values



between –1.00 and +1.50, aligning with the majority of student ability levels.
User Feedback from Students and Teachers Post-assessment feedback was gathered from 35 students and 4 teachers. Key findings include: Students reported high satisfaction with the CAT process, noting that the test

felt "short but challenging," and 91.4% agreed it helped them reflect on their personal strengths and weaknesses. Teachers appreciated the system's automatic scoring, interpretability of θ scores, and time efficiency. All teachers indicated interest in using the platform again in broader settings.

TABLE 4 Summary of User Satisfaction **(**Likert**-**scale**)**

Item	Mean (out of 5)	SD
The test was easy to understand	4.4	0.68
The questions matched real-life skills	4.31	0.73
I prefer CAT over paper-based tests	4.57	0.55
The results were clear and useful	4.49	0.61

DISCUSSION

The findings of this study provide important insights into the feasibility and utility of implementing a Computerized Adaptive Testing (CAT) approach to assess life skills among Grade 9 students in Thailand. The discussion below is organized according to the study's three main objectives: measurement outcomes, technical performance of the CAT system, and user experience

Interpretation of Life Skills Proficiency Levels

The θ distribution indicated that the majority of students possessed moderate to high levels of life skills, with only a small proportion falling below average. This finding aligns with previous national reports suggesting that Thai adolescents possess foundational interpersonal and decision-making skills, but often lack advanced competencies in areas such as emotional regulation and stress management (Office of the Basic Education Commission [OBEC], 2022).

This supports the idea that life skills education in Thailand is gaining traction at the school level, yet highlights a continued need for personalized interventions to address skill gaps—especially in students identified in the "low" or "very low" range by the CAT system.

Effectiveness and Precision of the CAT System

The CAT algorithm demonstrated a high level of measurement efficiency, with students completing the test using an average of just over 10 items, compared to the full item pool of 120. The average standard error of measurement (SEM) of 0.28 confirms that the adaptive testing mechanism succeeded in balancing precision with test length.

This result parallels findings by Weiss (2011) and van der Linden (2016), who emphasized the value of adaptive testing in reducing test fatigue and improving student motivation. In the Thai context, this efficiency is particularly relevant due to time limitations in the school day and the increasing need to implement short yet meaningful assessments.

Moreover, the Test Information Function (TIF) showed the highest precision between θ = –1.0 to +1.5, aligning well with the ability levels of the tested population. This ensures the CAT provides valid and reliable scores for the majority of students, although future refinements could add items targeting extreme ability levels to improve test accuracy at the tails of the distribution.

User Experience and Implementation Implications

Feedback from students and teachers was overwhelmingly positive. Students noted that the adaptive format made the test feel more engaging and less stressful, while teachers valued the immediate feedback and detailed reports. This aligns with previous research by Ebenbeck & Gebhardt (2022), who found that adaptive assessments promote positive testing attitudes, especially in formative contexts. The strong acceptability also suggests that integrating CAT into Thailand's school assessment infrastructure is feasible, especially when paired with appropriate teacher training and policy support. Notably, the transparency of θ scores and their interpretation could support **data-driven decisions**, including tailored learning plans or school-based interventions.

Educational and Policy Relevance

Thes results offer meaningful implications for **education reform and digital transformation** in Thailand. As the Ministry of Education continues to promote competency-based education, the integration of smart, adaptive assessment tools becomes a crucial enabler. The CAT platform used in this study—customized within a Moodle-based system—demonstrates that such innovations can be **locally developed, scalable, and user-centered**.

More broadly, the study contributes to the **global discourse on non-cognitive assessment**, adding empirical evidence from Southeast Asia to support the feasibility of using CAT beyond academic subjects.

CONCLUSION AND IMPLICATIONS

The results offer actionable insights at both the classroom and system levels:

- **For Educators and Schools:**
  - CAT platforms can be used to **support formative assessment** practices, especially in areas that are traditionally difficult to measure, such as emotional and interpersonal skills.
  - Real-time feedback allows for **individualized intervention**, helping teachers differentiate support based on student profiles.
  - Teachers' positive response to the system indicates **high scalability** with appropriate training.
- **For Policy Makers and Assessment Authorities:**
  - The study supports **policy alignment with Thailand's Basic Education Core Curriculum (BECC) and the 13th National Economic and Social Development Plan**, which emphasize life skills and learner-centered evaluation.
  - It also aligns with international movements such as **OECD's Learning Compass 2030**, which calls for personalized, competency-based assessment systems.
  - The CAT model used in this study can serve as a **blueprint for future national-level implementation**, including integration into Thailand's digital learning platforms or national testing frameworks.
- **For Educational Technology Developers:**
  - The success of a locally built CAT system (via Moodle) suggests the **potential for sustainable, cost-effective development** of smart assessment platforms in developing countries.
  - Open-source frameworks could be customized to serve diverse educational contexts and linked with AI-driven learner analytics in the future.

LIMITATIONS

While the findings present valuable insights into life skills assessment using adaptive testing, certain limitations should be noted. First, the sample size of 35 students, although adequate for demonstrating system functionality, limits generalizability. Second, the participants were from a single geographical region, which may not reflect the diversity of student populations across Thailand. Third, although IRT provides precise estimation, CAT's adaptability was tested only in a simulated single-session environment. Future studies should involve longitudinal testing, larger multi-site samples, and integration into national educational platforms to enhance scalability and impact.

REFERENCES

1. Ajibade, A. T., & Fashiku, C. O. (2019). Effectiveness of life skills education on adolescents' social
2. behavior. *Journal of Education and Practice, 10*(24), 23–30. https://doi.org/10.7176/JEP
3. Ebenbeck, N., & Gebhardt, M. (2022). Simulating computerized adaptive testing in special educa
4. tion based on inclusive progress monitoring data. *Frontiers in Education, 7*, 878009.
5. https://doi.org/10.3389/feduc.2022.878009
6. Khamreang, B., & Siribanpitak, P. (2020). Development of a life skills test for Thai secondary school
7. students. *Kasetsart Journal of Social Sciences, 41*(3), 592–598.
8. https://doi.org/10.1016/j.kjss.2020.02.008
9. Ministry of Education, Thailand. (2017). *Basic Education Core Curriculum B.E. 2551 (A.D. 2008): R*
10. *evised 2017 Edition*. Bangkok: MOE Publishing.
11. Office of the Basic Education Commission (OBEC). (2022). *Thailand Life Skills Development Frame*
12. *work*. Bangkok: Ministry of Education.

13. Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2021). Computational thinking and life skills: A
14. systematic review of empirical studies. *Education and Information Technologies, 26*(6), 7377–7401.
15. https://doi.org/10.1007/s10639-021-10579-2
16. Roid, G. H., & Haladyna, T. M. (2018). *Measurement and assessment in teaching* (11th ed.). Pearson.
17. Saenboonsong, W., & Poonsawad, P. (2024). Gamification for enhancing creative thinking skills in early childhood education. *Early Child Development and Care, Advance online publication*.
18. https://doi.org/10.1080/03004430.2024.2298275
19. van der Linden, W. J. (2016). *Handbook of item response theory: Volume one—Models*. CRC Press.
20. Weiss, D. J. (2011). Better data from better measurements using CAT. *Journal of Methods and Meas*
21. *urement in the Social Sciences, 2*(1), 1–27. https://doi.org/10.2458/v2i1.12359
22. World Health Organization. (1999). *Partners in Life Skills Education: Conclusions from a United Na*
23. *tions Inter-Agency Meeting*. Geneva: WHO. https://www.who.int/mental_health/media/en/30.pdf