Journal of Textbook Research
Vol. 18, No. 2 (August, 2025), 83-130
doi: 10.6481/JTR.202508_18(2).03

Higher-Order Thinking in English Education: Content Analysis of Taiwan's Senior High School Textbooks by Using L Book as an Example

Chia-Hsuan Li

Critical thinking (CT) is an essential higher-order thinking skill for students. Although numerous Taiwanese studies have focused on the teaching of CT and its association with English , few have investigated its representation in English textbooks. In the present study, content analysis of CT representation in Taiwans senior high school English textbooks was performed to determine the extent to which secondary school English textbooks support CT teaching. The 593 questions in the reading sections of the L Book series were used for a case study. Drawing on Blooms taxonomy and Taiwans English curriculum guidelines, the researcher developed a coding scheme comprising four CT categoriesanalysis, evaluation, synthesis, and creation and analyzed how CT was presented within and across textbook volumes. The results indicated that one-third of the questions in the L Book series could enhance students CT skills, with analysis and evaluation being the most prevalent categories. Furthermore, a progressive shift from analysis and evaluation to creation was observed across volumes, indicating a cumulative hierarchy in CT development. These findings offer pedagogical insights for senior high school English teachers and textbook publishers regarding the cultivation of CT skills in students and the improvement of English textbook design in Taiwan.

Keywords: critical thinking, critical thinking representation, English textbook, content analysis

Received: January 2, 2024; Revised: July 29, 2024; Accepted: November 5, 2024

Chia-Hsuan Li, Master, Department of English, National Chengchi University, E-mail: 110551008@g.nccu.edu.tw

教科書研究
第十八卷 第二期（2025 年 8 月），83-130
doi: 10.6481/JTR.202508_18(2).03

英語教材中的高階思考：臺灣高中教科書之內容分析——以 L 版本為例

李佳璇

批判性思考是學生必備的高層次思考能力。臺灣已有許多學者討論如何教學，以及其與英語之相關；但批判性思考在英語教科書是如何呈現卻鮮少被注意。爲瞭解中等學校英語教材能否引導批判性思考，本研究以L版本的閱讀爲例，針對593個問句進行內容分析，探究批判性思考在臺灣高中英語教材中的樣貌。研究者延伸Bloom等人（1956）的認知分類，並參照臺灣英語文領域課綱設計分析架構，共包含四類能力：分析、評價、綜合、創造，分析批判性思考在單冊與跨冊的呈現。結果顯示，L版本約有三分之一的問句能促進學生的批判性思考，以分析、評價佔比最多。研究亦發現，批判性思考能力在跨冊間有漸進式轉變，由分析、評價到創造，表示其累積且分層的特性。根據上述，研究者提出培養學生批判思考能力、改善臺灣英語教科書之建議，將對高中英語教師、教科書書商有教學上的啓發。

關鍵字：批判性思考、批判性思考的呈現、英語教科書、內容分析

收件：2024年1月2日；修改：2024年7月29日；接受：2024年11月5日

李佳璇，國立政治大學英國語文學系碩士，E-mail: 110551008@g.nccu.edu.tw

1. Introduction

Critical thinking (CT) is a must-have higher-order thinking (HOTS) ability for students. In Taiwan, CT has been considered a core competence in the national curriculum guidelines for numerous years. CT debuted in Grade 1-9 Curriculum Guidelines in 2003 as a sub-skill and remained in a crucial place during the educational reformation ten years later, in 2014, as a Core Competency in the new Curriculum Guidelines of 12-Year Basic Education. With its significance, CT has drawn close attention in Taiwan's EFL research field.

The EFL research on CT in Taiwan could be divided into two dimensions. On the one hand, a growing number of Taiwanese researchers have been experimenting with the effectiveness of teaching CT with particular pedagogies, including theme-based instructions (Lin, 2005), debate (Liu et al., 2015), literature circle (Liao, 2009), content-based reading and writing (Liaw, 2007), and case-study (Lai, 2022). On the other hand, another group of researchers has been exploring the relationship between CT and English language skills, such as speaking fluency (Chen, 2021) and persuasive writing (Bauman, 2020). These local papers not only highlighted the instruction of CT but also attested to the close relationship between CT and English abilities. Despite the critical role of CT in Taiwan's national curriculum guidelines and its popularity in the EFL research field, few existing local studies shed light on the definitions and representations of CT in Taiwan's English teaching materials.

In the field of EFL textbook analysis, most local teacher-researchers who have carried out content analysis for Taiwan's officially approved English textbooks emphasized language structures, including text difficulty (Cheng & Chang, 2022), speech acts (Huang & Pai, 2009), and vocabulary (Lee, 2011), or particular issues, such as culture (Li, 2022), gender disparity (Chen, 2019), global competence (Chou et al., 2019), and English as a lingua franca (Luo, 2017). Very few of them examine the representation of CT in English textbooks, namely, how it is defined, presented, and organized within or across volumes. Lacking such knowledge about the

definitions, elements, and patterns of CT, on-site English teachers might have difficulty effectively teaching their students the ability of HOTS through textbooks.

The above literature review shows a noticeable absence of the textbook analysis of CT. Given that CT is one of the core competencies in Taiwan, the present study intended to conduct a content analysis of the representation of CT in the English textbooks for Taiwan's senior high schools, using L Book as an example. The specific aims were twofold: the first one was to investigate the extent to which the English textbooks for secondary schools were conducive to CT instruction, and the second one was to explore any patterns of how CT was presented across different volumes of English textbooks. The following two research questions guided this study:

1. What elements of CT are presented in the English textbooks for Taiwan's senior high schools?
2. What are the patterns of CT representation across volumes of the English textbooks for Taiwan's senior high schools?

It was hoped that this study would discern any problems with CT representation in Taiwan's English textbooks and propose corresponding suggestions. The findings might additionally serve as a guide for English teachers to improve their instruction of CT and as a template for textbook editors to revise their publications.

2. Literature Review

2.1 Critical Thinking

The definitions and categorizations of CT vary remarkably among the theories of philosophical, psychological, and educational domains. Philosophically speaking, "critical thinking is the 'educational cognate' of rationality," and the central task of the cultivation of CT is to foster one's capability of "assess[ing] the probative strength of reasons" (Bailin & Siegel, 2003, p. 182). Ennis (1962) defined CT as "reasonable reflective thinking focused on deciding what to believe or do" (cited in Ennis, 1993, p. 180), whereas Lipman (1988) referred to CT as "thinking that is conducive to judgment" and depends on standards (cited in Paul et al., 1997, p. 4).

For philosophers, CT is the interaction between two sets of dispositions and abilities (Ennis, 1991). Regarding dispositions, a critical thinker must take an inquiring attitude, be independent-minded, stay fair-minded, and respect others (Bailin et al., 1999). They must also (1) understand intended meanings, (2) determine the focus of questions, (3) take an entire situation into consideration, (4) seek and proffer reasons, (5) be well-informed, (6) search for alternatives, (7) quest precision, (8) be reflective on one's beliefs, (9) stay open-minded, (10) withhold judgments without sufficient evidence, (11) take a position, and (12) adopt one's critical thinking abilities (Ennis, 1991). As for abilities, Ennis (1991) created a constitutive list of abilities of ideal critical thinkers, which consists of five abilities related to clarification, two to decision making, three to inference, three to metacognition, and four auxiliary abilities. Clarifying includes: (1) identification of the focus of questions, (2) analysis of arguments, (3) inquiry into and response to challenges, (4) judgment of definitions and handling equivocation, and (5) recognition of underlying assumptions. Deciding involves (6) judgment of source credibility and (7) observation of judgments. Inferencing refers to (8) deductive thinking and judgments of deduction, (9) inductive thinking and judgments of induction, and (10) value judgments. Metacognitive abilities are (11) "suppositional thinking" and (12) integrating the rest of dispositions and abilities to make and defend decisions. The auxiliary abilities are (13) to solve problems, monitor one's thinking, and use a CT checklist, (14) to be sensitive to others' feelings and knowledge level, (15) to adopt "rhetorical strategies," and (16) to be reactive to fallacy.

The psychological tradition of CT differentiates from the philosophical theories in that psychologists emphasize the performances of humans' actual behaviors and disclose our CT processes (Sternberg, 1986). The American Psychological Association (APA, n.d.) defines CT as "a form of directed, problem-focused thinking" with which an individual tests his ideas or solutions to errors. Psychologists often view CT as "a form of problem solving" and are interested in the components or sets of the behaviors of CT (Moon, 2007, p. 37). In Sternberg's (1986, p. 3) "componential" definition, CT consists of "the mental processes, strategies and representations people use to solve problems, make decisions, and learn new concepts." There are three componential skills in CT (Sternberg, 1985; cited in Sternberg, 1986):

metacomponents, performance components, and knowledge-acquisition components. Metacomponents are "higher order executive processes" an individual uses to plan what to do, monitor while doing, and evaluate after things are done. Performance components are "lower order, nonexecutive processes" a person utilizes to implement the metacomponents' instructions and offer feedback. These components vary across domains of performance, such as inductive and deductive reasoning. Under knowledge-acquisition components are "selective encoding" of (ir)relevant information, "selective combination" of related information in an organized form, and "selective comparison" between background knowledge and new information. They are the processes one applies to learn conceptual knowledge or procedures.

As a mixture of philosophical and psychological theories, the educational approach sheds light on the skills students need in classrooms to solve problems, make decisions, and grasp new concepts (Sternberg, 1986). Such a close tie to observations and experiences in classrooms is its benefits. In educational contexts, the definition of CT is more like a "programmatic definition" that expresses practical educational goals (Scheffler 1960; cited in Hitchcock, 2018). Dewey (1910, p. 6), for instance, coined the term "reflective thought" to describe the "active, persistent and careful consideration" of one's beliefs, and took such CT as an ultimate goal of education. Likewise, in Bloom's term, CT is labeled as "intellectual abilities and skills" (Bloom et al., 1956, p. 38) that learners can transfer information and techniques from their past experiences to new problems and situations. Moreover, Bloom et al. (1956) developed a six-level taxonomy of cognitive skills: (1) knowledge, (2) comprehension, (3) application, (4) analysis, (5) synthesis, and (6) evaluation. These skills are sequenced from simple and concrete to complicated and abstract and put in "a cumulative hierarchy" (Krathwohl, 2002, p. 212). That is, the capability of lower skills is the prerequisite for mastering higher skills. Later on, Anderson and Krathwohl (2001), considering that learning objectives and results are what students can do, renamed the six nouns in the form of verbs: from knowledge to "remember," comprehension to "understand," application to "apply," analysis to "analyze," synthesis to "create," and evaluation to "evaluate" (Krathwohl, 2002, p. 214). Another change was the replacement of the top two skills, with evaluating listed as the fifth and creating as the sixth.

The revised taxonomy is also presented hierarchically, and under the six categories are 19 subcategories of cognitive processes (Krathwohl, 2002).

Based on Bloom's taxonomy and its revision, several different verb lists have been created for learning assessments. To reach a consensus on the verbs' position in the taxonomy, Newton et al. (2020) and Stanny (2016) revisited all the lists collected from Google Search and counted the frequency of verbs within and across categories. The former researcher identified 104 verbs (Stanny, 2016), while the latter enlisted 61 verbs (Newton et al., 2020). In terms of the ambiguity and overlap of verb categorization, Stanny (2016) explained that many words have multiple meanings and that "context modifies meanings" (p. 8). An example was the verb "'recognize' the definition of technical terms" as lower-level thinking but "'recognize' professional situations that produce a conflict of interest" as higher-level thinking.

Different as the above three approaches appear, it is worth noticing that the CT skills each proposes strikingly overlap (Sternberg, 1986). Gubbins (1985), collecting several similar CT taxonomies by numerous researchers, developed a list of thinking skills (cited in Sternberg, 1986). His list included six primary skills and multiple subskills: (1) problem-solving, (2) decision-making, (3) inferencing, (4) divergent thinking, (5) evaluative thinking, and (6) reasoning.

2.2 Textbook Analysis of CT

In the existing studies on CT representation in textbooks of various disciplines, researchers have different preferences in using philosophical, psychological, or educational categorizations. Among the three categories, the most frequently used taxonomy is that of the educator Bloom. In his hierarchical taxonomy of cognitive thinking, the first three are lower-order thinking (LOTS), while the last three are HOTS and are often viewed as CT skills (Ennis, 1993).

With the use of Bloom's taxonomy (Bloom et al., 1956) or its revised version by Anderson and Krathwohl (2001), numerous foreign researchers have tried to probe how levels of thinking skills are presented in English textbooks and to what extent these teaching materials could enhance students' HOTS or CT, including Al-Qahtani (2019), Laila and Fitriyah

(2022), Peyró et al. (2020), Qasrawi and BeniAndelrahman (2020), and Tayyeh (2021). Laila and Fitriyah (2022) adapted a checklist table from Pratiwi (2014), while Tayyeh (2021) conducted descriptive content analysis to explain numerical findings, both focusing on reading comprehension questions in the English textbooks respectively for Indonesian and Iraqi twelfth graders. Similarly, Al-Qahtani (2019) and Qasrawi and BeniAndelrahman (2020) analyzed reading and/or writing activities and tasks in the English textbooks respectively for Palestinian and Saudi Arabian college students. The former researchers utilized a category of verbs under the six levels by OPAR (2012), whereas the latter also adopted descriptive analysis. All of their results, nevertheless, showed that very few HOTS were presented in the English teaching materials; in other words, students were not encouraged to think critically. Unlike these four studies, which used Bloom's HOTS, Peyró et al. (2020) adopted the framework of Anderson and Krathwohl's (2001) revised two-dimension classification to examine the actions and knowledge in the CLIL science textbooks for Spanish sixth graders. Their findings also revealed that more than 60% of the materials did not help promote the CT skills required for the CLIL approach.

Most researchers used Bloom's taxonomy and/or its revised version as a research framework; nevertheless, such adoption might not be as convincing as it seemed. Regarding Bloom's taxonomy, the following two aspects might need to be considered. Firstly, according to Ennis (1993, p. 179), Bloom's taxonomy (Bloom et al., 1956) somehow appears problematic. The six cognitive thinking levels are not hierarchical but "interdependent." Ennis (1981) gives an example: while synthesis and evaluation require analysis, analysis also requires these two skills. Hence, marking more than one CT skill for each question or activity might be possible and essential during the content analysis. One question or activity might encourage two or more kinds of CT skills. Secondly, Ennis (1993, p. 179) also warned that the three higher levels, namely the HOTS, seem to be "too vague." This might be because Bloom's taxonomy was initially designed to be a vastly general principle——"a common language"——so that it could be adapted or modified in accordance with different educational needs (Bloom et al., 1956, p. 10; Krathwohl, 2002, p. 212). In light of its generality, what could be and might have to be done is to

expand the HOTS in Bloom's taxonomy (Bloom et al., 1956) or its revision by Anderson and Krathwohl (2001) and make them as relevant as possible to the Taiwan context. By doing so, the research findings could be more convincing and localized to Taiwan's education of CT.

2.3 CT in the Context of Taiwan's Secondary Education

Scholars in Taiwan have also tried to define CT, and most local researchers take a philosophical perspective. According to Wang (2000), CT is a type of judgmental thinking based on the understanding of narratives and questions. More specifically, Yeh (2003) defined CT as "a complicated cognitive process" involving five kinds of thinking skills: "identification of hypotheses, inferences, deductions, explanations, and evaluations" (cited in Kuo et al., 2008, pp. 171, 174). Wen (2012, p. 3) referred to CT as "the process and result of dialectical thinking carried out by autonomous and self-disciplined people, including questioning, introspection, liberation, and reconstruction to pursue a more reasonable life." In addition, Cheng (2012), reviewing both philosophical and psychological approaches, listed six natures of CT: (1) the thorough and in-depth understanding of a problem or question, (2) the transcendence of self-centeredness and personal experiences, (3) the practice of reason, (4) the connection to and display of language quality, (5) the concern about oneself and his life, and (6) the pursuit of meanings.

When it comes to the education of CT, Taiwan's curriculum guidelines adopted a mixture of philosophical, psychological, and educational approaches. In Taiwan's educational history, CT debuted in Grade 1-9 Curriculum Guidelines in 2003 as a sub-skill of the tenth Core Competence, "Independent Thinking and Problem Solving." "Independent Thinking and Problem Solving," consisting of fundamental thinking, CT, and creative thinking (Curriculum and Instruction Resources Network, CIRN, n.d., abstract), refers to one's capability and habit of thinking independently and reflectively, making judgments systematically, and resolving problems effectively. As a sub-ability, CT, analytic and reflective in nature, is defined as being able to (1) examine, relate, and evaluate a context or a problem, (2) put emphasis on the foci of the problem, (3) collect, organize, analyze, and digest information, (4) decide whether solutions are reasonable with

previously learned experiences, and (5) make practical conclusions (CIRN, n.d., definition). These five sets of action verbs in the CT definition correspond to the different concepts of CT. In the statement, verbs 1 and 5 first see CT as both the capability and disposition of "reasonably reflective thinking" that involves value judgment and asks one to "remain relevant to the main point," which echoes with Ennis' (1993, p. 180) philosophy. Its second focus on problem-solving (e.g., verb 2) and the decoding, comparison, and combination of information (e.g., verb 3) is a psychological point of view of Sternberg (1986). The third focal point of reflective thinking and the ability to transfer background knowledge to new contexts, namely verb 4, are what Dewey and Bloom propose as ultimate educational goals.

During the educational reformation ten years later, in 2014, CT still held a special place in Taiwan's curriculum. In the new Curriculum Guidelines of 12-Year Basic Education (2014), CT can be found in A2, one of the nine Core Competencies. A2 "Logical Thinking and Problem Solving" pertains to the systematic thinking of understanding problems, analysis, CT, and metacognition, with all of which one can take action, reflect oneself, and effectively resolve daily-life problems. More specifically, in Curriculum Guidelines of 12-Year Basic Education: English Domain for Elementary, Junior High School and Upper Secondary School Education (2018), the ENG-A2 for higher-level English learners (i.e., senior high students) requires them to adopt various strategies to gain a deeper understanding of texts, to discern the essence of messages, and to distinguish truth from falsity. Furthermore, under Core Competency is Essential Learning Focus, a combination of nine Leaner Performances and four categories of Learning Contents. Amid the English learner performances for senior high students, the ninth "Logical Reasoning, Judgement and Creativity" includes eight thinking levels (i.e., 9-V-1 to 9-V-8):[1] (1) compare, classify, and sequence, (2) determine the relationship between information, (3) distinguish facts from opinions, (4) generalize and conclude, (5) transfer acquired knowledge to solve problems, (6) integrate information, (7) evaluate and judge, and (8) formulate plans with creativity. As for the learning content for senior

[1] The first code, "9" means the ninth Learner Performance, whereas "D" represents the fourth Learning Content in the Curriculum Guideline for Subject English. "V" refers to senior high students being in the fifth learner stage. The other number, "1," is the numbered skills.

high students, the fourth "Cognitive Capability" (i.e., D-V-1 to D-V-8) also contains the same eight thinking levels. The above verbs in the definitions of (ENG-)A2 and the descriptions of learner performances and learning contents are also a mixture of philosophical, psychological, and educational concepts of CT. Philosophically, it calls for logical, systematic, and reflective thinking, as well as truth-telling. Psychologically, the emphasis is further put on problem-solving and Sternberg's (1986) meta-components, namely planning, monitoring, and evaluating. Educationally, the utmost objective is for students to be able to transfer what they already know to their daily life situations.

The attempts to adopt philosophical, psychological, and educational traditions make the CT definitions and descriptions of its skills more comprehensive. With such comprehensiveness, English teachers and textbook designers might be well persuaded, no matter their stance.

3. Methodology

3.1 Data Sets

The data sets were collected from the English textbooks published by L Book Co., Ltd. L Book Co., Ltd. is a leading publisher of English textbooks for senior high schools, holding about 80% of the market share. The rationale behind its textbook design is to "educate students" instead of merely "proffering knowledge." Speaking of educating a person, CT is not knowledge but a disposition, representation, and ability that one must possess. Considering its high market share and core value on "education," L Book was selected as the research subject in the present study. The L Book's five volumes of English textbooks students used from the fall of 2022 to spring 2023 were analyzed. The first four volumes contain nine units and three reviews, and the fifth has only six units. The reviews were excluded from the data set because of their different designs from other units. Except for the reviews, all units consist of seven sections: "Thinking Aloud," "Reading Strategy," "Reading Selection," "Vocabulary & Phrases," "Sentence Pattern," "Language in Use," and "Listening Strategy." Only the sections related to reading, the first three, were gathered as data, for reading covers most parts of one unit.

Below is an introduction of each (sub)section under the "pre-, during, and postreading" model (Grabe, 1991, p. 396), a reading process where readers first activate their schemata (Carrell, 1988), adopt reading strategies and acquire linguistic knowledge (Güzel, 2022, p. 30), and demonstrate their comprehension through extended production (Petrosky, 1982, p. 24; cited in Zamel, 1992, p. 468). Every unit in L Book starts with two sections for pre-reading. One is "Thinking Ahead (TA)," where an activity and a question serve as a warm-up to increase students' learning motivations. The other, "Reading Strategy (RS)," containing two subparts, reading strategy (RS-RS) and reading comprehension (RS-RC), introduces one reading strategy and includes a short passage and two to three activities or questions for students to practice. The topic of the short passage is closely related to the topic of the main reading text for the next stage, while-reading, which includes one main part and four subparts. "Reading Selection (R)" involves a reading text with activities and questions presented in the subsections of "Language Highlight (R-LH)," "Reading Strategy (R-RS)," "Note the details (R-ND)," and "Reading Comprehension (R-RC)." R-LH focuses on linguistic knowledge, such as transition words and grammatical sentence patterns. R-RS is a similar practice of particular reading strategy to that in the previous section RS-RS. R-ND targets the details of the main reading text and invites students to find out detailed information. R-RC is literacy-based and scenario-situated, as in its design. Behind the reading text, for post-reading, are three other subparts: "Graphic Organizer (R-GO)" with an activity to visualize the text, another "Comprehension Practice (R-RC)" with three questions to check students' understanding of the text, and "Think and Reflect (TR)" with two to four questions to encourage students' critical and reflective thinking about the themes or issues arising from the texts.

The unit of analysis is question. By question, it referred to the statements ending with a question mark and/or starting with wh-words, such as "what" and "how," or auxiliary verbs, like "do" or "would." Each question mark or keyword was marked as one piece of data. The rationale behind the focused selection of questions was twofold. On the one hand, questions help probe CT, for "thinking is question driven" (Paul & Elder, 2002, 2007; cited in Rashid & Qaisar, 2016, p. 155) and is displayed "in the form of questions" (Santoso et al., 2018, p. 1).

On the other hand, most learning activities in L Book's English textbooks are phrased in the form of or followed with questions, and those that end only with a period are rare cases. If the statements involve both a question mark and a period, they would still be considered questions. Based on the criteria, 135 questions were gathered from Volume one, 141 from Volume two, 125 from Volume three, 123 from Volume four, and 69 from five. In total, 593 questions were collected from the selected (sub)sections (see Table 1). The whole data set could be divided into three reading stages and four sections: TA as a warm-up and RS and its subsections (i.e., RS-RS and RS-RC) as a pre-learning for reading strategies in pre-reading, R and its subsections (i.e., R-LH, R-RS, R-ND, and R-RC) as the main part in while-reading, and TR as an extension in post-reading.[2] Considering space limit, sample questions of each (sub)part are presented with coding examples in 3.2 Coding Scheme. Other thinking-related tasks, namely R-GO, were excluded from the data set, not only because they are not question-type but also because graphic organizer is another independent research topic that requests further and separate attention.

Table 1. Data Sets

Stages	Sections	V1	V2	V3	V4	V5	Total
Pre-Reading	TA	20	19	13	17	12	81
	RS						
	RS-RS	5	6	0	0	0	11
	RS-RC	9	9	9	6	0	33
While-Reading	R						
	R-LH	6	7	8	4	0	25
	R-RS	6	6	3	4	0	19
	R-ND	27	28	31	27	20	133
	R-RC[a]	34	35	35	35	22	161
Post-Reading	TR	28	31	26	30	15	130
Total		135	141	125	123	69	593

[a] Reading Comprehension (R-RC) and Comprehension Practice (CP) have identical features, and thus were combined together as Reading Comprehension (R-RC).

[2] "Think and Reflect" is renamed as "Thinking and Reflecting" for parallel presentation.

3.2 Coding Scheme

In the present study, CT's educational approach was adopted. The philosophical perspective was excluded because philosophers focus more on logical systems and reasoning, whereas psychologists pay much closer attention to the mental processes of problem-solving. Contrastively, the educational tradition of CT, as a mixture of the two theories (Sternberg, 1986), is more comprehensive and has played the role of the Bible, guiding textbook design and lesson design. Hence, frameworks from the educators, Bloom's taxonomy (Bloom et al., 1956), and its revised version by Anderson and Krathwohl (2001) were modified to create the localized coding scheme for the content analysis of the representation of CT in Taiwan's English textbooks. The categories of the coding scheme included only "Analyzing," "Evaluating," and "Synthesizing" from Bloom's framework, for according to Ennis (1993), the last three are HOTS and are often viewed as CT skills. Synthesizing was still added to the coding scheme, as Taiwan's English test in the college entrance exam, the General Scholastic Ability Test (GSAT), is now prone to add such a synthesis nature in its test design. Another category was Creating from Anderson and Krathwohl's revised version. Creating and Synthesizing both existed in the coding scheme because they differed in the learning outcome learners were required to perform. For Synthesizing, students need to draw multiple elements from different sources and integrate all available information into new constructions. The synthesized production involved both the original contents of the sources and parts of learners' ideas. As for Creating, students typically need to formulate a plan or complete a project from scratch. The created production thus requests more originality of learners' ideas, imagination, and creativity.

Extended from Wang's (2000) definition of CT (see Section 2.3), in the present study, CT was operationally defined as "a type of thinking that puts forward rational value judgments and creative idea generation after in-depth analysis and comprehensive synthesis of the narrative texts and problems." Based on the operational definition, a localized coding scheme was created. The steps were as follows. The researcher first gathered all the contents related to HOTS, namely the above four categories of CT from the Curriculum Guidelines of 12-Year Basic Education: English Domain for Elementary, Junior High School and Upper Secondary School Education

(2018, pp. 17, 24-25, 32-33), which included the lists of the third and the ninth Learner Performance "Language Proficiency in Reading" and "Logical Reasoning, Judgement and Creativity" as well as the fourth Learning Content "Cognitive Capability." Among all the skills on the lists, 9-V-5, "Can apply acquired knowledge to new situations and solve problems," was excluded because it was considered at the level of applying. Then, extracted statements of CT were separated into multiple pieces by verbs since each original statement in the curriculum guideline covered more than one HOTS. To exemplify, the statement of skill 9-V-8, "Can integrate various information, formulate a reasonable plan and accomplish a task with creativity," was chunked into three separate subskills, S-3, C-1, and C-2. 3-V-15, "Can analyze and critique an article and understand the author's perspective, stance and writing purpose," was also split into two, A-7 and E-2; the former was about the objective analysis of the author's writing purpose, while the latter was leveled at evaluation for the subjectivity natures in "critique," "perspective," and "stance." Lastly, the researcher divided those skills into four HOTS ——"Analyzing," "Evaluating," "Synthesizing," and "Creating" (see Table 2). The four categories of HOTS were hierarchical and interdependent, meaning that higher-level thinking skills include lower-level ones as prerequisites. Their corresponding CT skills were also levelized, as they originally were in the curriculum guideline. For instance, A-4 (9-V-2, D-V-2) is above A-1, A-2, and A-3 (9-V-1, D-V-1), the necessary CT skills to master A-4.

During the content analysis, the researcher, considering contexts (Newton et al., 2020; Stanny, 2016), first identified the required thinking skills in each question, such as differentiating and comparing. Then, the marked thinking skills, or action verbs, were categorized according to the present coding scheme and the verb lists of Bloom's taxonomy by Newton et al. (2020) and Stanny (2016). Moreover, based on the abovementioned assumptions of CT skills, instead of marking only one code for each piece of data, the researcher would identify which categories and CT skills a question invites students to perform and label one or more codes. Also, to present the results of CT representations more comprehensively, each identified code would be counted in proportion, regardless of their hierarchy. To exemplify, if a question required A-4 and A-1, both codes would be marked

Table 2. Coding Scheme

Category	CT Skills
Creating	C-1 Can formulate a reasonable plan.[a] (9-V-8, D-V-8) C-2 Can accomplish a task with creativity. (9-V-8, D-V-8)
Synthesizing	S-1 Can generalize multiple pieces of information and arrive at a conclusion. (9-V-4, D-V-4) S-2 Can integrate all the information available to predict potential development. (9-V-6, D-V-6) S-3 Can integrate various information and reconstruct an original text to produce a new interpretation. (9-V-8, D-V-8)
Evaluating	E-1 Can evaluate various pieces of information and make reasonable judgment or suggestions. (9-V-7, D-V-7) E-2 Can critique an article and understand the author's perspective and stance. (3-V-15)
Analyzing	A-1 Can compare multiple pieces of information. (9-V-1, D-V-1) A-2 Can classify multiple pieces of information. (9-V-1, D-V-1) A-3 Can sequence multiple pieces of information. (9-V-1, D-V-1) A-4 Can determine the relationship between two pieces of information based on the context. (9-V-2, D-V-2) A-5 Can distinguish objective facts from subjective opinions based on the context and textual clues. (9-V-3, D-V-3) A-6 Can analyze multiple pieces of information and identify the similarities. (9-V-4, D-V-4) A-7 Can analyze an article and understand the author's writing purpose. (3-V-15)

[a] The codes in the brackets were the original codes from Curriculum Guidelines of 12-Year Basic Education: English Domain for Elementary, Junior High School and Upper Secondary School Education (2018), added to exemplify how the coding scheme was adapted.

and counted. For the calculation of CT proportion, the total counts of CT skills identified in L Book would be the denominator, and the amount of each CT (sub)skill was the numerator (i.e., CT proportion = N of CT (sub) skill / Total of identified CT skills). Both of the portions within and across the five volumes were calculated. In contrast, if any cognitive thinking skills required in the questions were not closely related to HOTS, such as locating information or demonstrating comprehension, they would not be coded into any categories or skills but marked as not applicable (NA). The portion of

NA was calculated separately by dividing the non-applicable questions by the sum of collected data (i.e., NA proportion = N of NA / Total of the data set).

To exemplify the coding procedure, Volume two unit four (V2U4) "Malala: Stronger Than Violence" was used as a demonstration. In the pre-reading stage, the TA question asked students to look at pictures of teenagers and ponder "What can teenagers do to make our world a better place?" This question was coded as S-2 because students had to integrate what they viewed in the pictures and make possible predictions about what teenagers could do for a better world. For another RS-RC question, students had to read a news article and check their understanding of the features of the texts by answering the question, "According to the text, which of the following statements is true?" Such questions only tested their linguistic knowledge and thus were labeled as NA.

In the while-reading stage, firstly, the R-LH question, "What does the phrase 'a few' in Paragraph 1 refer to?" assessed if learners understood what certain words or phrases in the reading text indicated. The question only required students to demonstrate their understanding of the pronoun "a few" by making references within the reading text, and therefore, was marked as NA. Secondly, the R-RS question, "…which kind of article is this passage?" asked students to identify the features of the reading text and determine whether it was a news report or a feature article. The question was also an NA, for it merely checked students' application of the linguistic knowledge of text genres. Thirdly, two of the three R-ND questions: "When the Taliban ruled her hometown, what did Malala's father encourage her to do?" and "Why did the Pakistani government launch a new policy ensuring gender equality in education?" were also labeled as NA in that they only requested students to locate specific information in the reading text. In contrast, to answer the other R-ND question, "Why did the Pakistani government launch a new policy ensuring gender equality in education?" learners had to identify the cause-and-effect relationship for the new policy launched by the Pakistani government mentioned in the reading text. This question was thus marked A-4. Fourthly, among the three R-RC questions, the first one, "What's the main idea of the passage?" only confirmed students' comprehension, considered an NA. The second one, "Why did Malala fight for girls' educational rights?" was also labeled NA, even though

there seemed to be a causal relationship in the question. This was because the text relation had been pointed out so clearly with a discourse marker that learners only needed to locate information——"Malala believed that no one could take away others' right to education. She thus became an advocate of equal rights for female students" (L Book, 2022, p. 83). The last one, "What is the author's main purpose in writing this passage?" where students were asked to analyze the author's purpose of writing the article was coded as A-7.

In the post-reading stage, TR included three questions. The first, "What kind of girl do you think Malala is?" was labeled as both E-1 and S-3, for learners not only had to make reasonable judgments with their own criteria to appraise Malala but also generalize what she has done in her life to conclude her personality or beliefs. The second, "Do you agree that education is the most powerful weapon we can use to change the world?" merely invited learners to share their opinions on whether education is a powerful tool; therefore, it was marked as NA. However, if followed up with the third question, "Why or why not?" the combination of the two would be labeled as E-1, for it required students to evaluate the role of education. A summary of the coding example is displayed in Table 3.

3.3 Inter-Rater Reliability

To carry out the coding more objectively and reach more convincing results, another researcher was invited to conduct the coding of content analysis together with the present researcher. The partner researcher was one of the researcher's classmates in graduate school, and both of whose master's thesis were instructed by the same professor. The partner researcher agreed that the HOTS (i.e., Analyzing, Evaluating, Synthesizing, and Creating) are CT skills. She also considered the categorization to be much more meaningful because it corresponds to the teaching site, where the textbook design should fulfill the expectations of the curriculum guideline.

Table 3. A Coding Example of V2U4

Stages	Sections	Questions	CT skills
Pre-reading	TA	Look at the pictures and discuss the following question with your partner. What can teenagers do to make our world a better place?	S-2 (make possible prediction)
	RS-RC	According to the text, which of the following statements is true? (A) The text provides a complete analysis of the event. (B) The text covers the author's opinion about this attack. (C) The text includes the most essential information in the first paragraph.	NA (acquire linguistic knowledge)
While-reading	R-LH	What does the phrase 'a few' in Paragraph 1 refer to?	NA (understand text contents)
	R-RS	According to the statements you checked, which kind of article is this passage? ☐ News report. ☐ Feature article.	NA (acquire linguistic knowledge)
	R-ND	1. When the Taliban ruled her hometown, what did Malala's father encourage her to do? 2. Why did the Pakistani government launch a new policy ensuring gender equality in education? 3. What message did Malala send to the world in her UN speech?	1. NA (locate information) 2. A-4 (identify cause-and-effect) 3. NA (locate information)
	R-RC	1. What's the main idea of the passage? 2. Why did Malala fight for girls' educational rights? 3. What is the author's main purpose in writing this passage?	1. NA (understand text contents) 2. NA (locate information) 3. A-7 (analyze the author's purpose)
Post-reading	TR	1. What kind of girl do you think Malala is? 2. Do you agree that education is the most powerful weapon we can use to change the world? Why or why not?	1. E-1 (make reasonable judgment), S-1 (generalize, conclude) 2. (1) NA (personal opinion); (2) E-1 (make reasonable judgment)

The co-coding steps were as follows. First, the researcher introduced the coding scheme to the partner rater and trained her on how to analyze questions. V2U4 was used as an example to check her understanding. Once ready, we conducted the coding separately and dealt with the five randomly selected units: V1U5, V2U6, V3U3, V4U2, and V5U1. Following the individual coding was an online discussion to identify the problems and reach a consensus if there were disagreements between the two raters. After the analysis, the (in)consistency between the two rates on each item was marked as 1 for agreement and 0 for disagreement. Next, the researcher calculated the percentage of consensus, drawing a matrix where the columns were different raters, and the rows represented the variables collected from the data. This is the traditional measurement of inter-rater reliability named "percent agreement," calculated by dividing the scores of agreements by the total scores (McHugh, 2012, p. 277). This study used such a method because its calculation is simple and allows analyzers to discern any problematic variables. In the present study, the percentage of agreement between the researcher and the partner was about 84%, which meant the inter-rater reliability was well-established.

4. Results

4.1 CT Representation Within Volumes

Table 4 shows the sum of NA questions found in each volume of L Book. About two-thirds (69.3%, 411 out of 593) of questions seemed not conducive to developing learners' HOTS. In contrast, of the remaining 182 CT-provocative questions, the CT skills were counted 214 times, as some of the questions required more than one skill. The proportion of CT skills in the five volumes is presented in Table 5. In an overview, the four categories of CT skills all appeared in the different units of the different volumes of English textbooks. Of all the identified data (i.e., 214 counts), Analyzing (41.6%) was the most dominant CT skill, with 17.8% of A-4 and 12.6% of A-7. Evaluating (28.5%) was the second dominant CT skill, with E-1 (25.7%) accounting for the most. On the other hand, Synthesizing (19.2%) and Creating (10.7%) appeared rather less frequently, respectively including 11.2% of S-1 and 7.5% of S-2, and 4.7% of C-1 and 6% of C-2.

Table 4. Proportion of NA in L Book

	N of NA Questions	Proportion of NA (%)
V1	89	64.0
V2	103	70.1
V3	91	72.8
V4	81	61.8
V5	47	63.5
Total	411	69.3

Table 5. Proportion of CT in L Book

	N of CT	Proportion of CT (%)
CT		
Creating	23	10.7
C-1	11	4.7
C-2	13	6.0
Synthesizing	41	19.2
S-1	24	11.2
S-2	16	7.5
S-3	1	0.5
Evaluating	61	28.5
E-1	55	25.7
E-2	6	2.8
Analyzing	89	41.6
A-1	10	4.7
A-2	2	0.9
A-3	4	1.9
A-4	38	17.8
A-5	0	0
A-6	8	3.7
A-7	27	12.6
Total	214	100

More specifically, Table 6 points out the portion of the NA question in each (sub)section. On the other hand, for the identified CT-provocative questions, Table 7 shows the proportion of each CT skill in the different (sub)sections of L Book. The detailed findings of each (sub)section are presented below in three stages: pre-reading (i.e., TA, R-RS, R-RC), while-reading (i.e., R-LH, R-RS, R-ND, R-RC), and post-reading (i.e., TR).

Table 6. Proportion of NA in Each Section of L Book

	N of NA Questions	Proportion of NA (%)
Pre-Reading (TA, RS)	100	80.0
While-Reading (R)	251	74.3
Post-Reading (TA)	60	46.2
Total	411	69.3

Table 7. Proportion of CT within Volumes (%)

	Pre-Reading (TA, RS)	While-Reading (R)	Post-Reading (TR)
CT			
Creating	3.3	0	25.9
C-1	0	0	11.8
C-2	3.3	0	14.1
Synthesizing	40.1	22.2	8.2
S-1	3.3	20.2	3.5
S-2	36.7	2.0	3.5
S-3	0	0	1.2
Evaluating	23.3	6.1	56.5
E-1	23.3	0	56.5
E-2	0	6.1	0
Analyzing	33.3	71.7	9.4
A-1	10.0	5.1	2.4
A-2	0	2.0	0
A-3	0	4.0	0
A-4	10.0	35.4	0
A-5	0	0	0
A-6	10	3.0	2.4
A-7	3.3	22.2	4.7
Total	100	100	100

4.1.1 Pre-Reading

In the first two sections, TA and RS, most questions (80%) seemed unable to engage students in CT but targeted their background knowledge or personal experiences (see Table 6). Only one-fifth of the questions focused on HOTs, with 33.3% of Analyzing, 23.3% of Evaluating, 40.1% of Synthesizing, and 3.3% of Creating (see Table 7). More specifically, only four Analyzing skills, A-1 (10%), A-4 (10%), A-6 (10%), and A-7 (3.3%), were found. For Evaluating, none of E-2 (0%) could be found but mostly E-1 (23.3%). As for Synthesizing, the relatively dominant CT skill here, 3.3% of S-1 and 36.7% of S-2 were identified. Below is an example of S2, where students needed to integrate the information in the picture and make predictions. By sharp comparison, Creating merely accounted for a small portion (3.3% of C-2).

> V5U6 (see Figure 1): Look at the picture. Who are these people, and what are they doing? Why are some of the people chained up, and what will happen to them? [3] (L Book, 2022, p. 147, TA) (both S-2, make possible prediction)

> V1U9: Does the graffiti on the wall make the city ugly or beautiful? Why? (L Book, 2022, p. 193, TA) (E-1, make reasonable judgment)

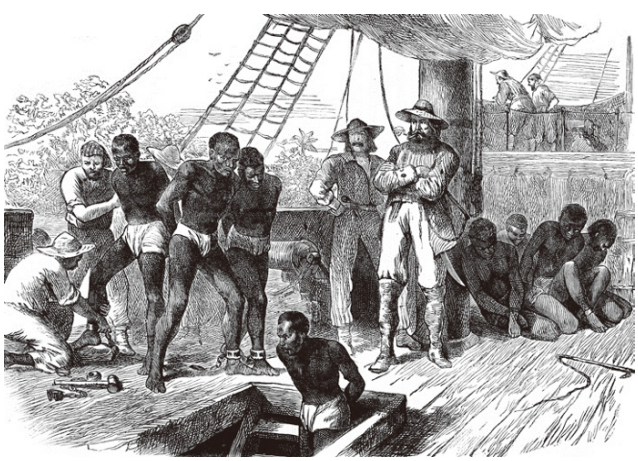

Figure 1. Example of S-2 from V5U6
Source: L Book (2022, p. 147).

[3] CT-provocative questions were underlined to differentiate them from other descriptions in the examples. This explanation is consistently applied to all underlined elements that follow.

4.1.2 While-Reading

In the third section, R, a significant number of the questions (74.3%) in its subsections were lower-level thinking (see Table 6). Among the scarce CT skills, Analyzing was primarily required in while-reading, taking up 71.7%. All of its six subskills, except for A-5, were found, with A-4 (35.4%) and A-7 (22.2%) appearing the most. The rest identified CT skills were 6.1 % of Evaluating (E-2) and 22.2% of Synthesizing (20.2% of S-1 and 2% of S-2) (see Table 7).

Below are four examples of Analyzing from V4U6, V5U5, V5U1, and V3U5. In the first example from V4U6, an introduction to Mumbai's unique food delivery system, Dabbawalas, its R-RC question asked learners to categorize the features of and identify the similarities between Mumbai's Dabbawalas and Taiwan's Foodpanda and Uber Eats. The second example, V5U5, introduced the threats and reasons for the rising water level in Venice, Italy. After reading the text, learners were invited to analyze how the author organized the essay by reordering the ideas in the question. For V5U1, it was a summary and an analysis of the British writer Rudyard Kipling's poem "If –." The reading text explained the qualities of a successful life mentioned in the poem, one of which is willpower. During their reading process, learners had to ponder why willpower was important. The cause-and-effect relationship was not clearly presented in the text; hence, the ability to analyze the text critically was a must. In the last example from V3U5, students were asked to read each sentence carefully in the first paragraph of the main reading text to analyze how different sentences supported one idea. Students had to pay attention to transition words, such as "In spring…," "In summer…," "… in autumn…," and "Then, when winter comes…," and multiple adjectives, both of which were used by the textbook author to delineate how Kyoto looked "throughout the year." Through such analysis, students would learn that the writer aimed to describe something logically and visually.

> V4U6: Dabbawalas play an important role in the lives of many Mumbai residents, and so do food delivery services such as Foodpanda and Uber Eats in that of many Taiwanese. According to the passage, which of the following is the proper content for the intersection? (L Book, 2022, p. 141, R-RC) (A-1 compare; A-2, classify; A-6, identify dis/similarities)

V5U5: How does the author organize the passage? Put the following ideas in the correct order… (L Book, 2022, p. 125, R-RC) (A-3, sequence)

V5U1: Why is willpower an important ability that one needs to have? (L Book, 2022, p. 6, R-ND) (A-4, identify cause-and-effect)

V3U5: Circle the keywords that signal the development of Paragraph 1. How is the idea "throughout the year" supported? (L Book, 2022, p. 106, R-LH) (A-7, analyze the author's writing purpose)

Examples of Synthesizing were from V1U2, V1U5, and V2U3. V1U2 was a break-up letter a teenage girl wrote to say goodbye to her cellphone. In the reading text, the pros and cons of smartphones were mentioned. Its R-RC question required students to generalize what had been mentioned in different paragraphs to summarize the reason for writing this letter. Another similar example was V1U5, which told the life story of a plastic bag with personification. In V1U5, its R-RC asked students to take a concluding stance on plastic bags to respond to other personified marine animals' concerns. V2U3 was adapted from American writer O. Henry's short story *Mammon and the Archer*. In the narration, different characters hold different opinions on true love and money, as indicated in different paragraphs. To answer the R-RC question, learners also needed to generalize the descriptions and plots to make conclusions about the characters' perspectives.

V1U2: Why does Sally want to leave John? (A) John cheats on her. (B) Sally's friends dislike him. (C) She has changed for the worse. (D) John seldom pays attention to her. (L Book, 2022, p. 33, R-RC) (S-1, generalize, conclude)

V1U5 (see Figure 2): According to the passage, which of the following is most likely the plastic bag's response? (A) We never meant to hurt anyone. I think it's fair to say we plastic bags are also victims. (B) It's all your fault! Your friends should have seen it clearly before swallowing a plastic bag. (C) It's unfair to blame us. According to a study, less than half of all sea turtles have eaten plastic. (D) Since we were created to

make humans' lives easier, some negative effects are unavoidable. (L Book, 2022, p. 107) (S-1, generalize, conclude)

V2U3 (see Figure 3): Based on the passage, which of the following statements best describes Old Anthony Rockwall's concept of money? (L Book, 2022, p. 53, R-RC) (S-1, generalize, conclude)

Figure 2. Example of S-1 from V1U5
Source: L Book (2022, p. 107).

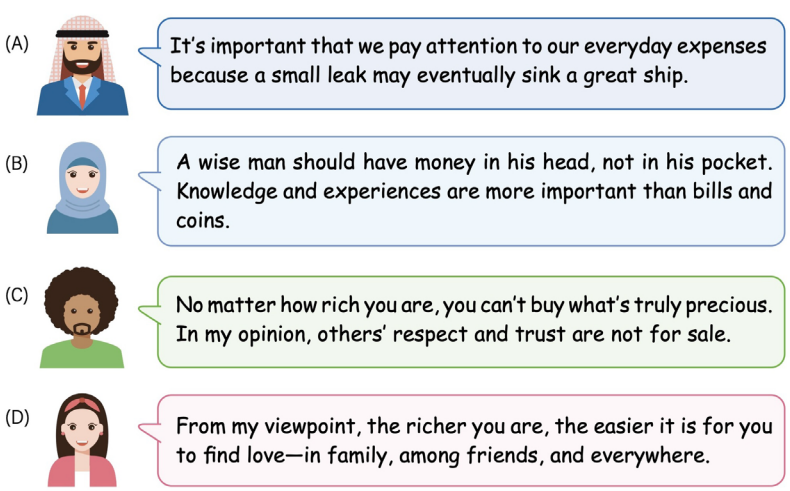

Figure 3. Example of S-1 from V2U3
Source: L Book (2022, p. 53).

4.1.3 Post-Reading

In the final section, TR, more than half of the questions targeted students' higher-level skills (see Table 6). TR involved all of the four categories of CT, including 9.4% of Analyzing (A-1, A-6, A-7), 56.5% of Evaluating (E-1), 8.2% of Synthesizing (S-1, S-2, S-3), and 25.9% of Creating (C-1, C-2) (see Table 7). Below are some instances from different units and volumes. Synthesizing could be found in V4U4 and V4U9. S-2 existed in V4U4, where the Tao people's culture of flying fish season on Orchid Island, Taiwan, was described. The article involved several aspects of flying fish, such as origin, traditional ceremonies, and restricted rules. To answer the TR question, students had to generalize all the information mentioned in the text and summarize what they could learn from the Tao people. S-3 and C-2 were in V4U9, the rewritten version of the novel *To Kill a Mockingbird*. Since it was a story with a sad ending, its TR question invited learners to synthesize what they had read from the narration and then come up with a new and different ending for the story on their own. Two other instances of Creating could be seen in V3U6 and V2U9. V4U7 was a history of Iceland's road to gender equality. Extended from the description, its TR question encouraged students to think of their experiences of being treated unfairly and then formulate feasible solutions to deal with the discrimination they had faced. As in V2U9, after reading about examples of eponyms, learners were asked to come up with more examples and the reasons behind them on their own.

> V4U4: What can we learn from the Tao people's attitude toward the flying fish season? (L Book, 2022, p. 91, TR) (S-1, generalize, conclude)

> V4U9: The author ends the story with Tom Robinson's death and Atticus Finch's comforts. If you could change the story, how would you end it? (L Book, 2022, p. 227, TR) (S-3, reconstruct; C-2, be creative)

> V4U7: Have you been treated unfairly because of your gender? For example, are you expected to do more household chores at home or perform better in certain school subjects? If you are faced with such discrimination, how will you deal with it? (L Book, 2022, p. 175, TR) (C-1, formulate plans)

V2U9: Think of a popular brand name. What do you think this name would mean if it were turned into an eponym? Try to make a sentence as an example. (L Book, 2022, p. 207, TR) (C-2, be creative)

4.2 Patterns of CT Representation across Volumes

Table 8 displays the CT representation separately in each volume. The overall findings showed that all four categories of CT skills had a similar proportion in each volume, with Analyzing (40.5% on average) and Evaluating (28.9% on average) being the most dominant two skills and Synthesizing (19% on average) and Creating (11.4% on average) the least. If viewed carefully, evidence of the features of CT in each (sub)section and the enhancement of CT skills could be found. Such patterns across volumes are discussed below by the pre-while-post reading model.

Table 8. Proportion of CT in Each Volume (%)

	V1	V2	V3	V4	V5
CT					
Creating	7.7	8.8	10.8	11.5	18.5
C-1	1.9	4.4	5.4	3.8	11.1
C-2	5.8	4.4	5.4	7.7	7.4
Synthesizing	23.1	13.0	27.1	17.3	14.8
S-1	15.4	8.7	19.0	9.6	0
S-2	7.7	4.3	8.1	5.8	14.8
S-3	0	0	0	1.9	0
Evaluating	26.9	21.7	37.8	28.8	29.6
E-1	26.9	21.7	32.4	25.0	22.2
E-2	0	0	5.4	3.8	7.4
Analyzing	42.3	56.5	24.3	42.4	37.0
A-1	3.9	8.7	2.7	3.9	3.7
A-2	0	2.2	0	1.9	0
A-3	1.9	2.2	0	1.9	3.7
A-4	26.9	13.0	5.4	23.1	14.8
A-5	0	0	0	0	0
A-6	1.9	6.5	2.7	3.9	3.7
A-7	7.7	23.9	13.5	7.7	11.1
Total	100	100	100	100	100

4.2.1 Patterns of CT Representation in Pre-Reading

In Pre-Reading, particularly the TA section, as displayed in Figure 4, there were two nuances of the four categories of CT skills across volumes. On the one hand, every volume started its units with Synthesizing and Analyzing, especially S-2 making predictions and A-1 comparing. Learners were invited to observe pictures or watch videos and make general guesses about the upcoming topic. On the other hand, almost each of the five volumes motivated students to perform Evaluating at the beginning of each unit, except for V2. Most of V2's pre-reading questions involved few CT skills but invited learners to share their personal experiences, which might serve as a connection to the coming topic of the main reading text since V2 tried to convey messages about personal growth.

4.2.2 Patterns of CT Representation in While-Reading

In the (sub)sections of while-reading, three patterns could be found. To begin with, Analyzing was the most frequently required CT (sub)skill, particularly A-4 determining text relationship (16.6% on average) and A-7 analyzing the author's purposes (12.7% on average). The proportion of this CT skill was an up-and-down flow – the ups in V1, V2, and V4 and the downs in V3 and V5 (see Figure 5). Possible reasons might be the composition of sentences in the main reading text. Even though almost

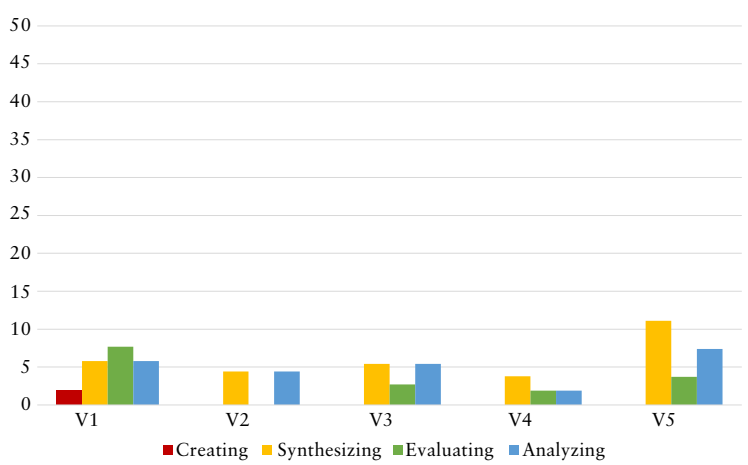

Figure 4. Patterns of CT Representation in Pre-Reading

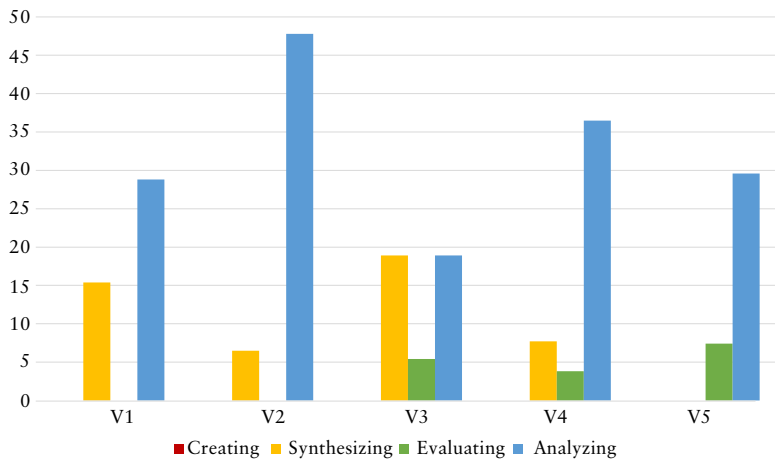

Figure 5. Patterns of CT Representation in While-Reading

every unit in each volume inquired students about the relationship between pieces of information, especially cause-and-effect, some answers were clearly presented in the main reading text with discourse markers, such as "because," "so," "thus," or "therefore." In such cases, the CT skill A-4 was unnecessary because students only had to locate the discourse markers. In addition, there was much more Analyzing in V2, especially A-7. The potential reason might be that the units in V2 tried to convey various vital messages about personal growth to senior high students. By analyzing the writer's purpose, students could understand the topic or central message of a passage and why it was essential to target readers. That is to say, given the chance to perform A-7, students could uncover the hidden messages in V2 units. Take U4 and U5 for example. V2U4 described Malala's life story and aimed to inspire learners that "teenagers can make a difference in society" as long as they firmly stuck to their beliefs. V2U5 was about Walt Disney's career history and Disney animations. Its reading text motivated students to overcome difficulties and pursue their dreams. Lastly, Evaluating, specifically E-2 criticizing the author's perspectives, started to appear from V3 onward, though with a slight portion (5.4%) (see Table 8). The R-RC questions in V3, V4, and V5 focused more on E-2 because the target students were second- and third-graders preparing for the English writing test in Taiwan's college entrance exam, the GSAT. Analyzing the author's

attitudes and perspectives benefited their English writing ability. Through Evaluating, students could not only realize that English writings were not entirely neutral but also learn how to express their opinions in their essays.

4.2.3 Patterns of CT Representation in Post-Reading

In TR, three main patterns of the four categories of CT skills could be seen (see Figure 6). Firstly, almost each of the five volumes motivated senior high students to perform Evaluating at the end of each unit. The proportion of Evaluating was much more significant than the other three CT skills, and E-1——making reasonable judgment or suggestion——was the most dominant CT skill in TR (56.5%) (see Table 7). Secondly, less and less Analyzing could be found from V1 to V5. In V1 and V2, most questions requested students to analyze the author's purpose (i.e., A-7). The reason might be that learners who read V1 and V2 were freshmen; therefore, the training of CT skills might have to start with easier ones, namely Analyzing. Thirdly, the portion of Creating gradually increased from V1 to V5. The percentage of Creating was once one-third as much as that of Evaluating in V1, V2, and V3, from which Creating grew from half as much to nearly equivalent to the other CT skill. This might indicate a change from the CT skills of Analyzing and Evaluating to Creating. It was also likely that L Book attempted to motivate students to be more creative.

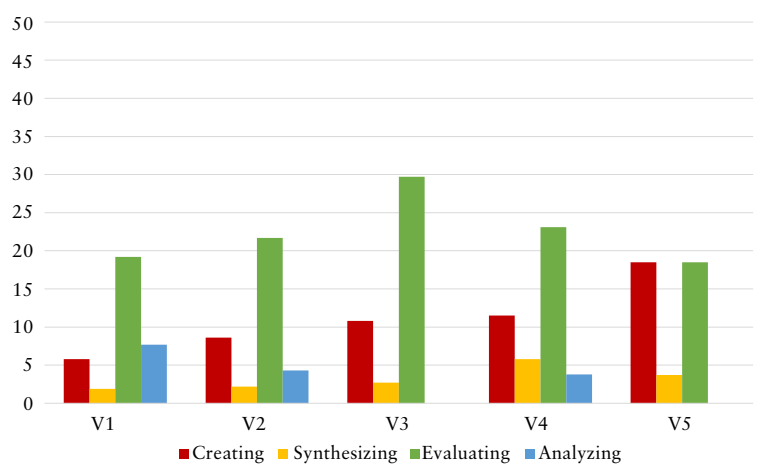

Figure 6. Patterns of CT Representation in Post-Reading

5. Discussion

5.1 CT Representation Within Volumes

5.1.1 The W-Shaped Flow of Cognitive Level

To better understand the CT representations within volumes, the researcher visualized and conceptualized the cognitive flows respectively in pre-while-post reading stages (see Figure 7). The conceptual diagram showed that the required CT skills varied according to the different purposes of the (sub) sections. In each (sub)sections, the unique features of CT skills formed a W-shaped flow of cognitive level within the five volumes of L Book's English textbooks ——from "Evaluating" and "Synthesizing," to "Analyzing" and "Synthesizing," and to "Evaluating" and "Creating". The essential CT skills in each section of the reading parts in L Book's English textbooks—— E-1 and S-2 in pre-reading (i.e., TA, RS), A-4, A-7, and S-1 in while-reading (i.e., R), E-1, C-1, and C2 in post-reading (i.e., TR) ——demonstrated ups and downs among the four hierarchical categories of CT skills in the coding scheme (see Section 3.2), with creating being the highest one. This result was similar to Qasrawi and BeniAndelrahman's (2020) study. In their calculation of the frequency of HOTS in each unit of the English textbooks, though not specified directly, it could be found that their textbooks also started with [Evaluation] in the warm-up videos (motivating students to make predictions), continued with Analysis and Synthesis in the reading (asking learners to make inferences and summarize), and ended with

Figure 7. W-Shaped Flow of Cognitive Level

Note. The figure is a conceptual diagram, not created using the aforementioned quantitative data.

Evaluation in CT (encouraging students to evaluate) and Synthesis [Creating] in writing (inviting learners to make plans and compose writing).

The pre-while-post model for reading might explain the rationale behind the W-shaped flow of cognition. According to Grabe (1991), reading lessons should be divided into the three-stage framework of "pre-, during, and postreading" (p. 396). In the first stage, pre-reading aims to activate learners' schemata, namely their background knowledge and personal experiences, of a topic to motivate them to read and prepare them for reading. In schema theory, pre-reading is of utmost significance, for the lack of "schema activation" may lead to reading difficulties for language learners (Carrell, 1988; cited in Grabe, 1991, p. 390). In the case of L Book, TA was a warm-up in pre-reading. Its questions invited students to perform Evaluating (i.e., E-1, evaluating an issue) and Synthesizing (i.e., S-2, integrating information to make predictions). Both subskills required their schemata. In the next stage, while-reading targets both learners' reading strategies (Grabe, 1991) and linguistic knowledge (Güzel, 2022). During the reading process, students not only need to decode texts and understand text relationships (i.e., A-4) but also have to analyze pieces of information (i.e., A-1, A-2, A-3, A-6, A-7) and synthesize (i.e., S-1, S-3) them as a whole conclusion. In L Book, R and its subsections were where students perform these Analyzing and Synthesizing skills. In the final stage, post-reading engages learners in both comprehension check and deeper analysis of texts. From the viewpoints of reading theorists, "the only way to demonstrate comprehension is through extended discourse where readers become writers" (Petrosky, 1982, p. 24; cited in Zamel, 1992, p. 468). This means learners can express their understanding of texts and connections to the topics and reconstruct their own meanings. As an extension of reading, the TR of L Book was a suitable place for students to discuss critical questions or issues, share their values, and make judgments (i.e., E-1). Through the performance of Evaluating, learners become connected to what they have acquired and will reach an advanced level——being capable of "taking a critical stance," one of the goals of English learning (Hedge, 2003; cited in Alyousef, 2006, p. 67). Also, TR questions challenged students to construct problem-solving plans with creativity (i.e., C-1, C-2). Through the performance of Creating, they rebuild what they've read in the texts with their thoughts to create new meanings.

5.1.2 The Most Frequently Required CT Skills

In particular, four CT skills were apparently of high proportion in the questions of the reading parts in L Book's English textbooks (see Table 5): A-4 and A-7 from Analyzing, E-1 from Evaluating, and S-1 from Synthesizing. A-4 and A-7 appeared frequently. The rationale behind the high numbers of A-4 and A-7 might be related to the design of Taiwan's curriculum guidelines for English learning. According to the Curriculum Guidelines of 12-Year Basic Education: English Domain for Elementary, Junior High School and Upper Secondary School Education (2018), the deep analysis of text contents (e.g., A-7) and the clarification of crucial information (i.e., A-4) are two essential capabilities of systematic thinking that Taiwan's senior high students must have, clearly mentioned in the A2 core competency. Another reason might concern the washback effect. The design of Taiwan's English tests in the college entrance exam GSAT also promotes the CT skills of A-4 and A-7. In the latest *GSAT English Test Preparation Guide* provided by the College Entrance Examination Center (CEEC, 2019), one of the test objectives is to evaluate students' ability to extract important information in a reading text to further analyze, compare, and infer the information. Analysis includes understanding the author's writing purpose (i.e., A-7), whereas inference requires clarifying pieces of information (i.e., A-4). Examples could be found in the archive of test papers on the CEEC website (e.g., analysis of writer's purpose in the year 2022; inference in the year 2023).

E-1 was also a high-portion CT skill in L Book. It is a must-have characteristic of a critical thinker. According to philosophers, critical thinkers must be well-informed, open-minded, independent-minded, and judgmental (Bailin et al., 1999). Nonetheless, the higher percentage of E-1 seemed to subordinate other CT skills and lower their proportions, such as S-1, S-3, C-1, and C-2. This might, in turn, make the CT representation in the questions of L Book's reading parts, especially those in TR, somehow homogenous rather than diverse. In pursuit of the diversity of the CT skills, L Book might need to make some changes to the TR questions to engage learners in multiple kinds of CT skills. More Creating could be added by asking students to solve the issues mentioned in the reading texts. For

instance, in V1U5, whose topic is environmental issues, its TR questions may motivate students to devise plans to reduce plastic products. Also, more Synthesizing and Creating could be involved by inviting learners to reconstruct the ending or beginning of texts, as such creation also required a synthesized understanding of the origin text. V2U6, adapted from American author R. J. Palacio's novel *Wonder*, or V3U3, the Greek mythology of Prometheus, could be good sources. By doing so, Taiwan's senior high students might have more opportunities to perform different CT skills in English learning. Other possible training for S-3 might include questions asking students to compose a short summary for a passage or fill in the blanks of a synthesized summary.

S-1 was the other most frequently required CT skill. It appeared almost in every unit of the R-RC——some accompanied graphic organizers, while others offered contextualized scenarios——to check students' overall understanding of a reading text and evaluate if they could transfer their knowledge in different contexts. The synthesized and contextualized design of this might be a product under Taiwan's curriculum guidelines for English learning. According to the Curriculum Guidelines of 12-Year Basic Education: English Domain for Elementary, Junior High School and Upper Secondary School Education (2018), one of the principles for the edition of English teaching materials is to provide "examples of real-life situations and [diverse] communicative exercises that are related to the topic of the unit" (p. 51). To provide real-life situations, most questions in textbooks were contextualized. Dialogues were also added in the R-RC to engage learners in diverse linguistic exercises. Examples could be found in the RC of V1U5, V2U3, and V3U4 (see Section 4.1.2). Another reason might concern the washback effect. The design of Taiwan's English tests in the college entrance exam GSAT also encourages S-1 and the diversity of questions. In the latest *GSAT English Test Preparation Guide* provided by the CEEC(2019), one of the test objectives is to evaluate students' ability to understand text contents and extract critical messages (i.e., S-1) (CEEC, 2019). The latest mixed item type is devised specifically to evaluate such a HOTS skill. The forms of questions in the mixed item type include multiple choices, short-answer questions, fill-in-the-blanks, etc.

5.1.3 The Missing CT Skills

In contrast to the most frequently required CT skills, four CT skills were found scarce or missing in the questions of the reading parts in L Book's English textbooks: A-2, A-3, and A-5 from Analyzing, and S-3 from Synthesizing.

A-2 and A-3 were two similar CT skills that were of low frequency. A-2 and A-3 seldom appeared alone in the questions of the reading parts in L Book but mostly along with other CT skills. For example, A-2 was required together with A-1 and A-6 in the R-RC questions of V4U6. The reason for the scarcity of A-2 and A-3 might be the limited unit analysis of questions in the present study. That is to say, A-2 and A-3 might not appear in the form of questions but, in fact, might exist in other forms and in other subsections, such as Graphic Organizer (R-GO), which encourages learners to outline a graphic for each reading text systematically. In R-GO activities, students need to categorize pieces of information from the main reading text or sequence them into a logical order.

A-5 was the only CT skill that did not exist. The capability of A-5 is a new trend in literacy-oriented test design, as shown in the example test items on CIRN (n.d.). A-5 is also a critical reading subskill, for it involves an evaluative process (Schell, 1967). According to Gubbins (1985), "distinguishing between facts and opinions" is one of the subskills of "Evaluative Thinking Skills" in his Matrix of [six] Thinking Skills (cited in Sternberg, 1986, p. 34). Significant as it is, A-5 was somehow missing in the questions of the reading parts in L Book's English textbooks. This might be because about half of the units in each volume are expository essays, a genre that contains primarily facts. To engage students in A-5, maybe L Book could extract essays whose contents refute the public's ideas. In this way, both subjective opinions and objective facts could be presented simultaneously, and learners would have chances to discern the differences between the two.

S-3 was another least frequently required CT skill that only appeared once in V4U9's TR (see Section 4.1.3). As a HOTS skill, Synthesizing helps learners build an in-depth understanding of the information in the texts (Cooper, 2006). To synthesize a text, students have to draw elements from different sources and put them into new and clearer structures. According

to Bloom et al. (1956, p. 162), Synthesizing will require "all of the previous categories to some extent." As stated, the scarcity of S-3 was probably due to its prerequisite of idea originality and the coverage of other CT skills. It might be difficult for textbook publishers to design S-3 relevant tasks, for English teachers to instruct S-3, and for students to perform such a CT skill. Hence, it is suggested that L Book could add more questions similar to the one in V4U9 to develop learners' S-3. Also, English teachers should guide their students in practicing writing synthesized summaries of the main reading text in the classroom to improve their S-3.

5.1.4 The Causes of NAs

In addition to the scarce or missing CT skills, another issue was the high proportion of NAs. About two-thirds of the questions in the reading-related (sub)sections of L Book's English textbooks were NAs, suggesting that the textbook had more LOTS, lower-level thinking, than HOTS, higher-level thinking or CT. Such results correspond to the prior literature that about 60% to 80% of the questions in the reading parts of their English textbooks could not promote students' HOTs (Laila & Fitriyah, 2022; Peyró et al., 2020; Tayyeh, 2021). The low proportion of CT skills in the questions of reading-related parts in L Book's English textbooks can be explained by exploring the nature of each (sub)section.

In the reading parts of L Book, a significant number of NAs were marked in RS-RS, RS-RC, R-LH, R-RS, and R-ND. Most of the questions from these subsections were lower-level thinking. The cause of their lack of CT skills might be their original design and their teaching and learning purposes. RS-RS, RS-RC, and R-RS, were devised to instruct students in specific reading strategies and provide opportunities for them to apply what they have learned. R-LH focused mainly on the knowledge of the English language, such as grammatical knowledge and the reference of pronouns. R-ND was designed to assist learners in comprehending the texts with several guided detail-oriented questions, such as those inquiring about what, where, or when. To answer R-ND questions, students had to scan for certain information throughout the text or skim for the main idea. Most of the above questions either requested students to recall or understand linguistic knowledge or required them to adopt reading strategies. In other

words, these subsections were both linguistic and meta-cognitive oriented in their nature and thus were only at the Remembering, Understanding, and Applying levels.

LOTS as the questions might be, their existence in English textbooks is indispensable for EFL learners. They play an essential role in drawing students' close attention to what they need to know to understand the reading texts and instructing them how to practice different reading strategies in their reading process. Still, the high percentage of the LOTS needed to be taken into consideration. Other than the LOTS, English textbooks might add one or two more questions at Analyzing or Synthesizing levels to these subsections to train students in both lower- and higher-level thinking.

5.2 Patterns of CT Representation Across Volumes

5.2.1 Patterns of CT Representation in While-Reading

In while-reading, especially the subsection of R-RC, there existed a movement from Analyzing to Evaluating, or from A-7 to E-2. The transition from analyzing the author's purpose to evaluating the author's attitudes indicated an enhancement of CT skills from V1 to V5. Given the slight difference between the two CT skills, both A-7 and E-2 are beneficial to learners not only in their English reading but also in English writing. Through A-7 and E-2, senior high students could grasp how good writers compose their English essays to convey meaningful messages and express personal opinions. Gradually, they would be more capable of writing English essays, which is a needed language ability for GSAT. The rationale behind such training seems similar to the first part of Halliday's systematic functional theory of language development. In his theory, there are three teaching and learning cycles of English writing: modeling of texts, joint construction of texts, and independent composition of texts (cited in Firkins et al., 2007). In the case of L Book, the reading texts in the English textbooks play the role of writing models, and CT skills A-7 and E-2 serve as guidance for students in analyzing and evaluating the model texts. For imitating writing, L Book also provides students with handbooks, including guided writing instructions to help students compose essays similar to the model texts. As for the other two stages, it is English teachers' responsibility

to use L Book's teaching materials well to design lessons inviting students to compose English essays collaboratively and individually.

Compared with Analysis, Evaluating took up a relatively slight portion, only with E-2. However, E-1, criticizing information and making reasonable judgments, was missing in while-reading. Evaluative reading is critical reading, defined as "the judgment of veracity, validity, of worth of the ideas read" (Robinson, 1966; cited in Smith, 1972, p. 164). Such judgment and evaluation are based on the criteria readers developed with their previous experiences. It is a crucial skill, particularly in the modern world full of false information. To help learners develop their critical reading ability, various reading materials are essential, and attention needs to be given in the classroom (Smith, 1972). Therefore, it is suggested that textbook publishers and English teachers could provide students with extended reading materials where they would have the opportunity to read information from multiple perspectives and discuss or comment on matters with their own standards. Take V1U9 for instance. The reading text of V1U9 was arguments for and against graffiti. To engage learners in evaluative reading, further reading texts about opinions from different individuals, such as artists, cleaners, or residents, would be helpful.

5.2.2 Patterns of CT Representation in Post-Reading

In post-reading, or the section of TR, there was a move from Analyzing and Evaluating to Synthesizing and Creating, which proved the following two concepts. First, the level of thinking skills is a "cumulative hierarchy" (Krathwohl, 2002, p. 212), meaning the mastery of the previous CT skills is the prerequisite for the mastery of the next CT skills. Learners would need to learn how to analyze and evaluate before they move on to synthesize and create. The questions in the reading parts of L Book's English textbooks first assisted students in becoming proficient in the CT skills of Analyzing and Evaluating in the first three volumes. Then, they challenged learners to perform the CT skills of Synthesizing and Creating. Second, Creating is one of the ultimate educational goals of English learning in Taiwan. According to Taiwan's Curriculum Guidelines of 12-Year Basic Education: English Domain for Elementary, Junior High School and Upper Secondary School Education (2018, p. 2), to cultivate Taiwanese students' core competency in

English learning, our English lessons should “foster learners’… creativity”. Under such an essential setting, L Book’s English textbooks could be viewed as a helpful and well-designed tool for English teachers to reach the final educational goal of developing students’ creativity.

6. Conclusion

This study explored the elements of CT in the English textbooks for Taiwan’s senior high schools and the patterns of CT representation across volumes with a coding scheme of CT skills, including Analyzing, Evaluating, Synthesizing, and Creating. The questions collected from reading-related (sub)sections in L Book’s English textbooks were examined by content analysis. The findings showed that about one-third of the questions are conducive to developing students’ CT skills and revealed that patterns of CT representation existed within and across volumes in each (sub)section.

6.1 Pedagogical Implications

6.1.1 For Textbook Publishers

Undoubtedly, L Book’s English textbooks could cultivate students’ different kinds of CT skills, and the existing patterns across volumes boost its training. Nevertheless, the findings also disclosed three issues in the questions of reading-related sections that might need corresponding adjustments from the English textbooks’ publisher and English teachers. Firstly, E-1 seemed more dominant in post-reading (i.e., TR) but missing in while-reading (i.e., R). L Book is suggested to diversify the TR questions so that Taiwan’s senior high students may have more opportunities to perform multiple kinds of CT skills. Providing various extended reading materials to present numerous perspectives for learners to integrate and criticize information is also advised. Secondly, A-5 was missing in the questions of reading-related sections. To fill the gap, L Book is recommended to involve both opinions and facts in reading texts to allow students to distinguish the differences between the two. Thirdly, S-3 was found scarce in pre-, while-,

and post-reading stages. To engage learners in integrating and reconstructing texts, L Book is advised to add more questions similar to the one in V4U9.

6.1.2 For English Teachers

Not only English textbook publishers but also Taiwan's senior high school English teachers are responsible for the abovementioned issues. What English teachers could do is to modify their pedagogies in their classrooms. Firstly, to evenly engage students in different CT skills, English teachers are suggested to add various HOTS questions in different reading stages. For pre- and while-reading, English teachers could adopt the Directed Reading Thinking Activity (DRTA), which involves three steps: predicting, reading, and proving (Stauffer, 1969). Before reading, students make assumptions about the texts with their schemata. Then, learners, reading the texts carefully, look for reasonable proofs to validate their predictions and generate conclusions (Chang, 2004). During the reading process, English teachers could also help students read meaningfully and critically by instructing them in the following techniques: determining text relationships (e.g., chronological and causal), making inferences from contexts, making conclusions, explaining their understandings (Rosenshine, 1980; cited in Chang, 2004). It is also effective to jot down personal notes in the margin, which could be further used for extended discussion and evaluation in post-reading. English teachers could invite students to share if they (dis) agree with the reading texts and why. Secondly, to engage students in E-1, English teachers could create opportunities for them to lead a discussion or leave comments on certain issues by offering extended materials involving diverse perspectives. Thirdly, in response to the missing A-5, before changes are made to the English textbooks, English teachers are recommended to design supplementary reading materials that can help develop students' A-5 and to add A-5 as a test item type to tests and exams. By doing so, English teachers could train and assess their students' A-5. Fourthly, faced with the scarcity of S-3 in English textbooks, English teachers are advised to devise related lesson plans. For instance, English teachers could guide their students to practice writing a synthesized summary of the main reading text to habitually and gradually cultivate their S-3.

6.2 Limitations and Suggestions for Future Research

The limitations of the present study are twofold. First, the collected data was limited to the questions. If the learning tasks were not in the form of questions, there was no data for analysis and discussion, which might, in turn, affect the overall results of the CT representation. Therefore, future researchers are suggested to involve both questions and activities in their studies. Second, data collection was limited only to the content analysis of English textbooks. However, how English teachers teach and use textbooks might influence the target CT skills. Hence, future researchers are recommended to conduct class observations further to justify whether English teachers engage their students in the target CT skills.

Last but not least, it is of the utmost significance that English teachers, textbook publishers, and future researchers understand that CT is not one single thinking skill but involves different kinds of HOTS such as Analyzing, Evaluating, Synthesizing, and Creating, just as Bailin and Siegel (Bailin & Siegel, 2003, p. 188) stated, "Critical thinking does not describe one type of thinking among others, but is an umbrella term that refers to the quality of thinking, whatever the context or activity." Only when educators and researchers are aware of the diversity of CT could they cultivate their students' CT skills with a comprehensive focus and provide them with multiple contexts to practice their thinking skills.

Acknowledgement

This paper is the author's unpublished MA thesis at National Chengchi University. The author is grateful to Professor Yi-Ping Huang for her constructive instruction and two anonymous reviewers for their insightful and invaluable comments and suggestions.

References

Al-Qahtani, E. M. (2019). Critical thinking pedagogy: Using textbooks evaluation and content analysis techniques for Saudi university students. *International Journal of Linguistics, Literature and Translation*, *2*(5), 239-244. https://doi.org/10.32996/ijllt.2019.2.5.26

Alyousef, H. S. (2006). Teaching reading comprehension to ESL/EFL learners. *Journal of Language and Learning*, *5*(1), 63-73.

American Psychological Association. (n.d.). Critical thinking. In *APA dictionary of psychology*. Retrieved August 17, 2022, from https://dictionary.apa.org/critical-thinking

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). Longman.

Bailin, S., & Siegel, H. (2003). Critical thinking. In N. Blake, P. Smeyers, R. Smith, & P. Standish (Eds.), *The Blackwell guide to the philosophy of education* (pp. 181-193). Blackwell.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies*, *31*(3), 285-302.

Bauman, C. (2020). *Writing acts as a vehicle for the development of critical thinking in an adolescent L2 environment: An exploratory study* [Unpublished master's thesis]. Wenzao Ursuline University.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longmans.

Carrell, P. L. (1988). Some causes of text-boundedness and schema-interference in ESL reading. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 101-113). Cambridge University Press.

Chang, Y.-Y. (2004). Critical thinking on language instruction. *Curriculum & Instruction Quarterly*, *7*(2), 41-55. https://doi.org/10.6384/CIQ.200404.0041

Chen, T.-P. (2019). Gender representation in English textbooks of senior high school: A content analysis. *Journal of National Taichung University: Humanities & Arts*, *33*(2), 25-44.

Chen, Y.-J. (2021). *The influence of critical thinking skill trainings on English fluency*

and coherence [Unpublished master's thesis]. National Taipei University of Education.

Cheng, W.-M. (2012). The analysis of the essence of critical thinking and its implications of teaching. *Journal of Education & Psychology*, *35*(3), 95-113.

Cheng, Y.-S., & Chang, S.-C. (2022). Do senior high school English textbooks sufficiently prepare students for the high-stakes college entrance examinations? A corpus-based analysis of text difficulty. *Journal of Textbook Research*, *15*(1), 43-80. https://doi.org/10.6481/JTR.202204_15(1).02

Chou, P.-I., Wang, Y.-T., & Wu, S.-T. (2019). Content analysis of core concepts of global competence in junior high school English textbooks in Taiwan: Taking the content domains of the cognitive tests in the PISA 2018 global competence assessment as a framework. *Journal of Textbook Research*, *12*(1), 1-38. https://doi.org/10.6481/JTR.201904_12(1).01

College Entrance Examination Center. (2019, September). *111 GSAT English test: Preparation guide*. https://reurl.cc/qNOL33

Cooper, J. (2006). *Literacy: Helping children construct meaning* (6th ed.). Houghton Mifflin.

Curriculum and Instruction Resources Network. (n.d.). *Grade 1-9 curriculum guidelines core competence: Independent thinking and problem solving.* https://cirn.moe.edu.tw/WebContent/index.aspx?sid=9&mid=221

Curriculum guidelines of 12-year basic education. (2014).

Curriculum guidelines of 12-year basic education: English domain for elementary, junior high school and upper secondary school education. (2018).

Dewey, J. (1910). *How we think*. D. C. Heath.

Ennis, R. H. (1991). Critical thinking: A streamlined conception. *Teaching Philosophy*, *14*(1), 5-24.

Ennis, R. H. (1981). Eight fallacies in Bloom's taxonomy. In C. J. B. Macmillan (Ed.), *Philosophy of education 1980* (pp. 269-273). Philosophy of Education Society.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice*, *32*(3), 179-186.

Firkins, A., Forey, G., & Sengupta, S. (2007). Teaching writing to low proficiency EFL students. *ELT Journal*, *61*(4), 341-352. https://doi.org/10.1093/elt/ccm043

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, *25*(3), 375-406. https://doi.org/10.2307/3586977

Güzel, S. (2022). The importance of reading activities in reading comprehension. *Disiplinlerarası Yenilik Araştırmaları Dergisi*, *2*(1), 29-34.

Hitchcock, D. (2018). Critical thinking. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2018 ed.). Stanford University. https://plato.stanford.edu/entries/critical-thinking/

Huang, M.-C., & Pai, I.-F. (2009). Speech acts in elementary school English instruction: Textbook analysis and teacher perceptions. *SPECTRUM: NCUE Studies in Language, Literature, Translation*, *5*, 117-140. https://doi.org/10.29601/YYWYLL.200907.0007

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212-218.

Kuo, J., Lee, J.-P., & Yu, C.-J. (2008). Validity and revising for Cornell critical thinking test level I. *Journal of Educational Research and Development*, *4*(4), 169-200.

Lai, H.-Y. (2022). Integrating case teaching into business English to enhance students' critical thinking and English presentation skills: A case study of a department of applied foreign languages in northern Taiwan. *Journal of National Taipei University of Business*, *40*, 1-22.

Laila, I., & Fitriyah, I. (2022). An analysis of reading comprehension questions in English textbook based on revised Bloom's taxonomy. *Journal of English Teaching*, *8*(1), 71-83. https://doi.org/10.33541/jet.v8i1.3394

Lee, M. Y.-F. (2011). How words inform Taiwanese junior high school students: A textbook corpus analysis. *English Teaching and Culture*, *7*, 113-122. https://doi.org/10.29721/ETC.201106.0010

Li, L. (2022). Analysis on Chinese culture in junior middle school English textbooks: Take PEP as an example. *International Journal of Social Science and Education Research*, *5*(8), 715-723. https://doi.org/10.6918/IJOSSER.202208_5(8).0102

Liao, M.-H. (2009). Cultivating critical thinking through literature circles in EFL context. *SPECTRUM: NCUE Studies in Language, Literature, Translation*, *5*, 89-115. https://doi.org/10.29601/YYWYLL.200907.0006

Liaw, M.-L. (2007). Content-based reading and writing for critical thinking skills in an EFL context. *English Teaching & Learning*, *31*(2), 45-87. https://doi.org/10.6330/ETL.2007.31.2.02

Lin, C.-Y. (2005). Incorporating critical thinking into English language teaching in

the film-based classroom. *Journal of Education and Foreign Language and Literature*, *2*, 53-69. https://doi.org/10.6372/JEFLL.200512.0053

Liu, E. P.-H., Wu, V. W.-C., & Shieh, R.-S. (2015). Enhancing EFL students' critical thinking and writing: An asynchronous debate instructional design. *English Teaching & Learning*, *39*(3), 33-59. https://doi.org/10.6330/ETL.2015.39.3.02

Luo, W.-H. (2017). A study of internationally-published English textbooks from the perspective of English as a lingua franca concerning content and pedagogy. *Journal of Textbook Research*, *10*(2), 133-159. https://doi.org/10.6481/JTR.201708_10(2).05

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276-282.

Moon, J. (2007). *Critical thinking: An exploration of theory and practice*. Routledge.

Newton, P. M., Da Silva, A., & Peters, L. G. (2020). A pragmatic master list of action verbs for Bloom's taxonomy. *Frontiers in Education*, *5*, Article 107. https://doi.org/ 10.3389/feduc.2020.00107

Paul, R. W., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations*. California Commission on Teacher Credentialing.

Paul, R., & Elder, L, (2002). *Critical thinking: Tools for taking charge of your professional and personal life* (2nd ed.). Sage.

Paul, R., & Elder, L. (2007). Critical thinking: The art of Socratic questioning. *Journal of Developmental Education*, *31*(1), 36.

Peyró, M. C. R., Herrero, E. C., & Llamas, E. (2020). Thinking skills in primary education: An analysis of CLIL textbooks in Spain. *Porta Linguarum*, *33,* 183-200. https://doi.org/10.30827/portalin.vi33.26647

Pratiwi, N. (2014). *Higher order thinking skill in reading exercise (An analysis of reading exercises in pathway to English textbook for the eleventh grade of senior high school students)* [Bachelor's thesis, The Syarif Hidayatullah State Islamic University]. Institutional Repository UIN Syarif Hidayatullah Jakarta. https://repository.uinjkt.ac.id/dspace/handle/123456789/25375

Qasrawi, R., & BeniAndelrahman, A. (2020). The higher and lower-order thinking skills (HOTS and LOTS) in ulnlock English textbooks (1st and 2nd editions) based on Bloom's taxonomy: An analysis study. *International Online Journal of Education and Teaching*, *7*(3), 744-758.

OPAR. (2012). *Bloom's taxonomy 1956*. https://humanities.utm.my/languageacademy/wp-content/uploads/sites/7/2016/10/Bloom-Taxonomy.pdf

Rashid, S., & Qaisar, S. (2016). Developing critical thinking through questioning strategy among fourth grade students. *Bulletin of Education and Research*, *38*(2), 153-168.

Santoso, T., Yuanita, L., & Erman, E. (2018). The role of student's critical asking question in developing student's critical thinking skills. *Journal of Physics: Conference Series*, *953*, Article 012180. https://doi.org/10.1088/1742-6596/953/1/012180

Schell, L. M. (1967). Distinguishing fact from opinion. *Journal of Reading*, *11*(1), 5-9.

Smith, H. K. (1972). Critical reading. In R. Karlin (Ed.), *Reading for all* (pp. 163-172). International Reading Association.

Stanny, C. J. (2016). Reevaluating Bloom's taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences*, *6*(4), Article 37. https://doi.org/10.3390/educsci6040037

Stauffer, R. G. (1969). *Directing reading maturity as a cognitive process*. Harper & Row.

Sternberg, R. J. (1986). *Critical thinking: Its nature, measurement, and improvement*. National Institute of Education. https://files.eric.ed.gov/fulltext/ED272882.pdf

Tayyeh, M. N. (2021). An analysis of reading comprehension questions in English textbook "English for Iraq" according to revised Bloom's taxonomy. *Turkish Journal of Computer and Mathematics Education*, *12*(7), 2868-2874.

Wang, K.-H. (2000). Critical thinking. In *National Academy for Educational Research web of words*. Retrieved March 27, 2024, from https://terms.naer.edu.tw/detail/df1e290c456e8e62a9609216d068aaa8/

Wen, M.-L. (2012). Critical thinking and its pedagogy: Dialogue, emancipation and reconstruction. *Taiwan Education Review*, *675*, 2-8. https://doi.org/10.6395/TER.201206.0002

Zamel, V. (1992). Writing one's way into reading. *TESOL Quarterly*, *26*(3), 463-485. https://doi.org/10.2307/3587174