

# Effectiveness of Critical Thinking Instruction in Higher Education: A Systematic Review of Intervention Studies

Dawit T. Tiruneh<sup>1</sup>, An Verburgh<sup>1</sup> & Jan Elen<sup>1</sup>

<sup>1</sup>Centre for Instructional Psychology and Technology, Katholieke Universiteit Leuven, Leuven, Belgium

Correspondence: Dawit T. Tiruneh, Centre for Instructional Psychology and Technology, KU Leuven, B-3000 Leuven, Belgium. Tel: 32-1632-6284. E-mail: dawittibebe.tiruneh@ppw.kuleuven.be

Received: November 18, 2013 Accepted: December 19, 2013 Online Published: January 23, 2014

doi:10.5539/hes.v4n1p1

URL: <http://dx.doi.org/10.5539/hes.v4n1p1>

## Abstract

Promoting students' critical thinking (CT) has been an essential goal of higher education. However, despite the various attempts to make CT a primary focus of higher education, there is little agreement regarding the conditions under which instruction could result in greater CT outcomes. In this review, we systematically examined current empirical evidence and attempted to explain why some instructional interventions result in greater CT gains than others. Thirty three empirical studies were included in the review and features of the interventions of those individual studies were analyzed. Emphasis was given to the study features related to CT instructional approach, teaching strategy, student and teacher related characteristics, and CT measurement. The findings revealed that effectiveness of CT instruction is influenced by conditions in the instructional environment comprising the instructional variables (teaching strategies and CT instructional approaches), and to some extent by student-related variables (year level and prior academic performance). Moreover, the type of CT measures adopted (standardized vs. non-standardized) appear to influence evaluation of the effectiveness of CT interventions. The findings overall indicated that there is a shift towards embedding CT instruction within academic disciplines, but failed to support effectiveness of particular instructional strategies in fostering acquisition and transfer of CT skills. The main limitation in the current empirical evidence is the lack of systematic design of instructional interventions that are in line with empirically valid instructional design principles.

**Keywords:** critical thinking, intervention, instructional approaches, teaching strategies, higher education

## 1. Introduction

### 1.1 Background

Promoting students' critical thinking (CT), such as the ability to identify central issues and assumptions in an argument, recognize important relationships, deduce conclusions from information or data provided, evaluate evidence or authority, etc., has been an essential goal of higher education (Halpern, 1993; McMillan, 1987; Paul, 1993). Acquisition of CT skills is considered vital for students to face a multitude of challenges of adult life and function effectively in today's increasingly complex world (Nickerson, 1988; Paul, 1993). However, despite the various attempts to make CT a primary focus of higher education, analyses of existing evidence (e.g., Pascarella & Terenzini, 2005; Tsui, 2002; Van Gelder, 2005) indicates that the level of CT displayed by most students is inadequate. It is argued that classroom instruction is mostly inefficient to help students acquire thinking skills that they could apply to solve important problems within disciplinary areas and in everyday life (e.g., Halpern, 1993; Jonassen, 1999; Nickerson, 1988).

Researchers and educators have been responding to the increasing demand of critical thinkers by designing instructional programs that focus on the acquisition and transfer of CT skills (e.g., Ennis, 1989; Halpern, 1998; Perkins & Salomon, 1988). There is some level of optimism about the capacity of students to become critical thinkers through systematic and well-designed instruction (Halpern, 1998; Mayer, 1992; Pascarella & Terenzini, 2005; Perkins & Salomon, 1988). However, the major challenge has been to determine which instructional interventions yield the greatest CT gains (Halpern, 1993; Pascarella & Terenzini, 2005). A suite of instructional strategies, which are suggested as useful to promote students' CT is available (e.g., Beyer, 2008; Ennis, 1989; Halpern, 1998; Paul, 1993; Tsui, 2002). Yet, there is little consistency with regard to empirically supported instructional conditions that could effectively enhance students' CT (e.g., McMillan, 1987; Pascarella &

Terenzini, 2005).

Overall, the growing body of empirical evidence on effectiveness of CT instructional interventions has been inconsistent. In their extensive review, Pascarella and Terenzini (2005) noted that although there is evidence of positive effect of instruction on CT development, the effect is widely variable and not particularly large. Some attributed the inconsistencies to the complex nature of CT as a construct. Variations in conceptions of the construct among educational researchers may have led them to pursue their own visions of CT based in diverse research traditions (Ennis, 1989; McPeck, 1990; Paul, 1993). Others (e.g., McMillan, 1987; Tsui, 1999) pointed out that the inconsistencies and limited impact may be due to limitations in the design of CT intervention studies, including factors as very brief lapse time between pretest and posttest, and broad measurement instruments. The nature of the instructional interventions such as the instructional approach (Ennis, 1989; Resnick, 1987) and specific teaching strategies (Beyer, 2008; Halpern, 1998) are also mentioned as major factors in influencing effectiveness of CT instruction.

Some student and teacher-related characteristics are also regarded as important in influencing the effectiveness of CT interventions. Student-related characteristics that are shown to be influential are, for instance, students' prior knowledge (Kennedy, Fisher, & Ennis, 1991) and educational level (King, Wood, & Mines, 1990). Moreover, although the direction of influence is ambiguous, research does show significant differences in CT outcomes between male and female students (Giancarlo & Facione, 2001; King et al., 1990). Teacher-related characteristics such as previous training and experience in CT instruction (Beyer, 2008; Pithers & Soden, 2000) are mentioned as influential in the effectiveness of CT interventions. Generally, it appears that there are some conditions in the learning environment that influence the effectiveness of CT instructional interventions. The various instructional conditions could be categorized into three: instructional variables, mainly the instructional approaches adopted and teaching strategies employed; student-related variables including gender, educational level, and prior knowledge; and methodological features of CT intervention studies such as the nature of CT measures employed.

The purpose of this study is to systematically review current CT intervention studies in relation to the aforementioned variables and identify the major components of instructional environments that foster students' CT in higher education.

### 1.2 What is CT?

There are many different definitions of CT. Most of the definitions focus on the conceptualization of CT as a set of cognitive skills. Pascarella and Terenzini (2005) summarized the various popular CT definitions and indicated that CT skills refer to an individual's ability to do some or all of the following: identify central issues and assumptions in an argument, recognize important relationships, make correct inferences from data, deduce conclusions from information or data provided, interpret whether conclusions are warranted based on given data, evaluate evidence or authority, make self-corrections, and solve problems. Halpern (1998) also defined CT as the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. Halpern (1998) identified the following components of CT: understanding how cause is determined, recognizing and criticizing assumptions, analyzing means-goals relationships, giving reasons to support a conclusion, assessing degrees of likelihood and uncertainty, incorporating isolated data into a wider framework, and using analogies to solve problems. In addition to the cognitive skills, scholars (e.g., Halpern, 1993; Giancarlo & Facione, 2001; Kennedy et al., 1992) also mention that there is a motivational dimension to CT termed dispositions. Giancarlo and Facione (2001) pointed out that a more comprehensive view of CT must include dispositions, which refers to a person's inclination to use CT skills when faced with problems to solve, ideas to evaluate, or decisions to make. There is now a consensus that CT, as a broad concept, involves both skills and dispositions (Giancarlo & Facione, 2001; Kennedy et al., 1992; Pascarella & Terenzini, 2005). The dispositions dimension includes truth-seeking, open mindedness, systematicity, analyticity, maturity, inquisitiveness, and self-confidence (Giancarlo & Facione, 2001). In the present review, studies that focused on improvement of student CT skills, dispositions, or both are included.

### 1.3 CT Instructional Approaches

Ennis (1989) categorized the various approaches to CT instruction as *general*, *infusion*, *immersion*, and *mixed*. In the general approach, CT is taught separately from the presentation of the content of existing subject matter. The infusion approach attempts to integrate CT instruction in standard subject matter instruction and makes general principles of CT explicit to the students. In this approach, students are encouraged to acquire and explicitly practice CT skills through deep and well-structured subject matter instruction. The immersion approach also tries to incorporate CT within standard subject matter instruction. However, general CT principles and procedures are

not made explicit to students with the assumption that they will acquire the thinking skills as a consequence of engaging in the subject matter instruction. The mixed approach consists of a combination of the general approach with either the infusion or immersion approach together. In the mixed approach, there is a separate thread or course aimed at teaching general principles of CT, but students are also involved in subject-specific CT instruction where either the objectives of CT are explicit or implicit (Ennis, 1989).

Regardless of the approach, CT instruction is mainly based on the assumption that there are clearly identifiable and definable thinking skills which are domain-independent, and can be taught to students to recognize and apply them appropriately in daily life situations and future careers (Halpern, 1988; Nickerson, 1988; Perkins & Salomon, 1988). The goal of CT instruction is, therefore, to help students acquire and transfer those domain-independent thinking skills to solve problems faced in everyday life (e.g., Ennis, 1989; Glaser, 1984; Halpern, 1998). In achieving this goal, a question that has been topic of debate is whether CT is best taught as a separate course adjunct to the standard curriculum, or embedded in academic disciplines (Ennis, 1989; Nickerson, 1988; Resnick, 1987). There is no complete agreement thus far on whether we should teach CT skills in domain-independent (general) courses or integrate CT instruction within existing subject-matter courses (Ennis, 1989; Mayer, 1992; McPeck, 1990; Nickerson, 1988). In this review, effectiveness of the CT instructional approaches adopted in various intervention studies (general, infusion, immersion and mixed) will be examined.

#### *1.4 Teaching Strategies for CT Development*

It is argued that student improvement in CT skills and dispositions hardly occur simply as an incidental outcome of subject matter classroom learning (e.g., Beyer, 2008; Nickerson, 1988). Many of the students, being novices or having little experience, are less capable of acquiring and transferring thinking skills to out-of-classroom contexts. Studies have suggested that empirically supported teaching strategies that encourage, stimulate, and facilitate students' acquisition and transfer of thinking skills are essential for CT development (e.g., Beyer, 2008; Halpern, 1993). Scholars (e.g., Halpern, 1998; Nickerson, 1988; Perkins & Salmon, 1988) asserted that the issue of transfer must be addressed either CT is taught using the general or discipline-embedded approach. Nickerson (1988) clearly elaborated the risks involved in both the general and discipline-embedded approach: "A risk of teaching a specific aspect of thinking only in a "content-free" way is that the student will acquire some understanding of that aspect but fail to connect that knowledge to the many situations in life in which it could be useful. A risk of teaching the same aspect of thinking only within the context of a [standard subject matter ] course is that the student will fail to abstract from the situation what is really context independent and again will not transfer what has been learned to other contexts" (p. 34). In this review, attempts are made to identify and examine effectiveness of the various teaching strategies employed in selected intervention studies.

#### *1.5 CT Measures*

One of the challenges in evaluating the effectiveness of CT instruction is related to the measurement of CT (e.g., Halpern, 1993). Researchers employ various kinds of CT measures that cover a broad range of formats, scope, and psychometric characteristics and, thus, evaluation of student acquisition and transfer of CT skills is problematic. It is argued that the different CT measures vary in their conceptualization of CT and nature of items involved in measuring transfer of acquired thinking skills (e.g., Halpern, 1993; McMillan, 1987). Therefore, it is likely that our evaluation of instructional interventions in enhancing CT depends on the CT measure employed. In this review, we will examine CT outcomes in relation to CT measures employed.

#### *1.6 Rationale for and Objectives of the Present Review*

Despite the large body of research in relation to the teaching of CT in higher education, it appears that there is a considerable gap in our knowledge about the conditions under which instruction could result in greater CT outcomes. There are a few relatively recent systematic reviews which attempted to analyze the evidence on CT instruction in colleges and universities (e.g., Abrami et al., 2008; Behar-Horenstein & Niu, 2011; Ten-Dam & Volman, 2004). However, these reviews are subject to some limitations. For example, Ten Dam and Volman reviewed studies which are mainly theoretical and those which do not employ specific instructional interventions. The design of the review and quality criteria applied to the studies reviewed appears to be fairly minimal, and it is likely that studies with significant weaknesses were included. For example, in most of the studies reviewed, students' self-reports were used in measuring CT outcome.

Abrami et al. (2008) conducted an extensive meta-analysis on the effect of instructional interventions in the development of CT skills and dispositions targeting all educational levels. An explicit search strategy, including a wide range of sources and criteria for inclusion and exclusion of studies was employed. However, the review included studies at all levels of education that made it difficult to understand the nature of instructional

interventions in the context of higher education. In addition, their meta-analysis does not focus on how student-related characteristics such as academic performance, gender, and educational levels influence the effectiveness of CT instruction.

The other recent review is the one conducted by Behar-Horenstein and Niu (2011), which examined empirical evidence on the teaching of CT in higher education. However, the review included only studies which employed a few of the standardized CT measures (namely, the Cornell Critical Thinking Test, the Watson-Glaser Critical Thinking Appraisal, and the California Critical Thinking Skills Test). It is likely that relevant studies that adopted various other CT measures (either standardized or non-standardized) are excluded in their review, which potentially minimizes the representativeness of the included studies. In addition, the review does not focus on how student and instructor-related characteristics could influence the effectiveness of CT instruction in higher education.

It is clear that previous reviews provide only limited information about the conditions under which instruction could enhance students' CT in higher education. None of the above reviews have systematically examined the effectiveness of instructional interventions in relation to student and teacher-related characteristics. In addition, the impacts of different teaching strategies employed in the various intervention studies were not adequately analyzed in previous reviews.

In this contribution, attempts are made to systematically examine current empirical evidence on the effectiveness of instructional interventions in fostering university students' CT. Specifically, features of instructional interventions of individual studies in relation to CT instructional approach, teaching strategy, student characteristics, and CT measurement are analyzed.

The following are the research questions that we will answer in this review study:

- What instructional interventions have an effect on the development of CT skills and dispositions among university students?
- To what extent do student-related characteristics including gender, academic performance, and educational level influence the impact of CT interventions?
- Are there differences in effectiveness of instructional interventions in relation to specific teacher characteristics (particularly previous experience in CT instruction and formal CT training)?
- Are there differences in effectiveness of instructional interventions in relation to CT measures employed?

## 2. Method

A systematic literature search was conducted to identify and retrieve empirical studies relevant to this review. Three databases were searched: Web of Science, ERIC and PsycInfo. The reference sections of previous review articles were also scanned for relevant articles. We used the following set of keywords (or possible synonyms) to search the relevant articles: *critical thinking* in relation to (*higher, postsecondary, tertiary*) *education, university, college, intervention, instruction, teach\*, learn\*, influence, effect\*, develop\**. We limited our search to empirical studies published in peer-reviewed/refereed journals between 1995 and 2012 (November). This search resulted in a total of 88 articles.

Next, the abstracts of all the 88 articles were read by the first and third author to decide whether the full text of an article should be retrieved or not. The two authors initially agreed on the following criteria for inclusion and exclusion of articles. In order to be included, the study should present some kind of instructional intervention; should involve teacher-led classroom instruction or computer-based instruction, or some sort of instruction by the researcher/teacher/tutor; should compare the CT outcomes of participants (e.g., control group with experimental group, pretest with posttest) through either standardized or non-standardized generic CT measures; and should focus on students in higher education. If any of the above criteria were not met, the study was rejected. For example, if a study employed only student self-reports as CT measure, or if there was no pre-test, or if the CT measure employed was only domain-specific, the study was rejected. Based on these inclusion/exclusion criteria, the abstracts of all the articles were read and coded independently by the two authors. The following ratings were used during coding: 1 (the article is not suitable); 2 (the article is possibly suitable); and 3 (the article is suitable). Articles which were rated "1" by the two authors were immediately rejected. In more than 80% of the cases, the authors selected similar articles as either suitable or not suitable for the review. In a few cases, adequate information was not found from the abstract and the two authors read the full text either to include or exclude the article. For the articles which were coded contrastingly by the two authors, discussion was held until full agreement was reached either to include or exclude the article. The outcome of this procedure was a total of 33 articles from which data were extracted.

A data extraction sheet was used to record all the required information from the individual study reports. The extraction of data involved the CT instructional approach, specific teaching strategies employed, student characteristics (gender, year level, and academic performance/GPA), and delivery of CT instruction (teacher training, previous experience in CT instruction). Extraction of methodological features included mainly research design (experimental, quasi-experimental and one group pretest-posttest) and type of CT measure employed (standardized and non-standardized tests).

### 3. Results

Overall descriptive information about each of the included studies is presented in Appendix A.

#### 3.1 CT Instructional Approaches

We primarily categorized the 33 studies included depending on the CT instructional approach (general, infusion, immersion and mixed). Five (15%) studies applied the general approach, nine (27%) studies applied the infusion approach, 16 (49%) studies applied the immersion approach, and three (9%) studies applied the mixed approach (see Appendix B). When evaluating the overall effectiveness of studies merely based on the instructional approach adopted, we found that the majority of the studies which employed the general approach reported significant student CT improvement ( $n=4$ , 80%). In the second and third place are studies that adopted the mixed ( $n=2$ , 67%) and the infusion ( $n=5$ , 56%) approaches, respectively. Half of the studies which adopted the immersion approach reported significant CT gains, which is the least proportion compared to the other three approaches. However, this finding should be interpreted with caution as studies which adopted the general and mixed approaches are relatively limited in number compared to the infusion and immersion approaches.

#### 3.2 Teaching Strategies

We categorized all the teaching strategies employed in the studies into two: direct and implicit (see Appendix C). Studies within the direct teaching category employed explicit explanation of CT procedures at the early phase of instruction followed by a combination of instructional activities such as teacher modeling, scaffolding, role playing, and small group discussion. Studies within the implicit teaching category employed various teaching strategies that embed CT without any explicit emphasis on CT skills.

The evidence we found regarding the effectiveness of those implicit teaching strategies is inconclusive. For example, Garside (1996) examined the effect of lecture and group discussion methods of teaching on CT development. Students participated in the study as part of the scheduled curriculum in the regular class hours in which classes were randomly assigned to 6 conditions: 3 groups exposed to the lecture method and 3 other groups to discussion method of teaching for 18 weeks (36 hours). There was no explicit instruction on aspects of CT skills to either of the conditions, and it was hypothesized that student involvement in various group activities and assignments could result in significant CT growth. The results revealed no significant difference on the CT performance of the two instructional groups. In another pre-experimental study (Stark, 2012), sophomore students were taught a research methods course in which they were required to critique subject related scientific examples and discuss their evaluation with other students in the classroom. No significant pretest-posttest CT improvement was reported.

Three studies (Semerci, 2006; Sendag & Odabasi, 2009; Yuan, Kunaviktikul, Klunklin, & Williams, 2008) examined the effect of problem-based learning (PBL) environments on CT improvement. In those three studies, problems were presented to the students which required them to give solutions through small group discussion and other collaborative activities without explicit instruction on CT principles and procedures. Sendag and Odabasi (2009), for example, examined the effect of an online PBL environment involving ill-structured problem scenarios on students' CT skills. They found that students in the online PBL group significantly outperformed those in the online instructor-led group. The other two studies (Semerci, 2006; Yuan et al., 2008) similarly compared the CT skills of two groups of students in which one group was taught in a PBL environment and the other through lecture method. The PBL group in both studies significantly outperformed the lecture group.

Two studies (Chen, Liang, Lee, & Liao, 2011; Wheeler & Collins, 2003) which examined the effect of concept mapping on CT skills have reported varied findings. Chen and his colleagues taught a course where one group of first year nursing students used concept maps which required students to discuss in small groups and graphically represent and organize their ideas; whereas the control group was taught using lecture method. It was found that the concept mapping group significantly outperformed the lecture group in their CT outcomes (Chen et al., 2011). On the other hand, the use of concept mapping of patient information to prepare junior nursing students for clinical experiences brought no significant improvement on the students' CT compared to lecture method (Wheeler & Collins, 2003).

Higher-order questioning is the other implicit teaching strategy which was employed in three studies (Barnet & Francis, 2012; Renaud & Murray, 2008; Williams, Oliver, & Stockdale, 2004). Renaud and Murray (2008) examined the effect of repeated practice in higher-order questioning on CT improvement among first year students. Renaud and Murray reported that repeated practice in higher-order thinking questions did not result in significant CT improvement compared to students who practiced answering lower-level thinking questions. In the same way, Williams and his colleagues did not find significant differences in CT scores between two treatment groups in which one group received higher-order and the other lower-order questions during a semester long instruction. However, both studies by Renaud and Murray (2008) and Williams et al. (2004) reported that the higher-order question group significantly outperformed the lower-order question group on domain-specific CT measures. Similarly, Barnet and Francis (2012) examined whether repeated quizzes containing higher order thinking questions could improve students' CT outcome. They assigned 3 sections of students into one of the three conditions: quizzes containing factual multiple-choice items; factual essay items; and essay items requiring higher-order thinking. The quiz items for the three sections were prepared from the same content areas and the only difference was their level of difficulty. No significant CT score difference was detected among the three sections.

The majority of the studies included in this review employed explicit and direct teaching of CT procedures and we categorized them under the direct instruction strategy (see Appendix C). Common to all the studies in this category is that instruction begins with teacher explanation of the thinking procedures, rules, guidelines and followed by instructional activities that involve students in more extensive discussion and practice of the thinking skills. For example, in three studies (Bensley & Haynes, 1996; Bensley, Crowe, Bernhardt, Buckner, & Allman, 2010; Solon, 2007) the CT instruction begun with teacher explanation of the CT principles, and followed by teacher modeling and coaching through practical exercises. The direct instruction group in all the three studies significantly outperformed the other group of students in which CT was not explicitly taught. The other study (Yeh, 2009) which emphasized teacher modeling and scaffolding of CT activities reported significant CT improvement compared to students which did not receive direct CT instruction. Allegretii and Frederick (1995) reported that an intervention that targeted the teaching of argument analysis using teacher modeling in an interdisciplinary ethical reflection course resulted in a significant pretest-posttest CT improvement. Similarly, other studies (Hitchcock, 2004; Mazer, Hunt, & Kuznekoff, 2007; Plath, English, Connors, & Beveridge, 1999; Reed & Kromrey, 2001) which employed some variations of direct instruction strategy yielded significant CT improvement (see Appendix C for details). On the other hand, two studies (McLean & Miller, 2010; Nieto & Saiz, 2008) which employed some variations of direct instruction strategies reported no significant CT improvement.

The large majority of studies included in the review focused on the teaching of CT skills rather than CT dispositions or both. Only two studies were found that focused on fostering students' CT dispositions (Reed & Kromrey, 2001; Toy & Ok, 2012). Both studies adopted direct instruction in which CT strategies were taught explicitly combined with teacher modeling and continuous feedback. However, both studies reported no significant improvement on students' CT dispositions.

Overall, the evidence we found regarding the effect of implicit instruction on CT improvement is inconsistent. Although PBL among the implicit teaching strategies appear to consistently result in greater CT improvement, the rest of the teaching strategies demonstrated inconsistent findings. On the other hand, direct instruction in thinking skills combined with other teaching strategies including teacher modeling, scaffolding, coaching, and small group discussion appear to consistently result in greater CT improvement compared to implicit teaching strategies.

### *3.3 Effectiveness of CT Instruction across Various Student Characteristics*

We attempted to examine the extent to which some student-related characteristics influenced the impact of instructional interventions. However, a major challenge we faced was insufficiency of student-related information in the reports of the reviewed studies. Most of the studies do not provide the requisite information related to student variables. For example, no study was found in which gender is considered in evaluating the effectiveness of CT instructional interventions. We found 12 (36%) studies which clearly mentioned the number of male and female participants, but none of the studies reported separate analysis of pretest-posttest CT gains for male and female participants. The CT scores in all the studies were not categorized by gender and, therefore, we do not know whether a particular CT intervention was more effective to the male or female students. Using the limited information provided in some of the studies, we examined effectiveness of instructional interventions in relation to academic performance and educational level.

### 3.3.1 Academic Performance

Only one study (Williams et al., 2004) examined the effect of an intervention with respect to students' exam performance on the course. They found that high exam performers significantly outperformed the low exam performers in their posttest CT scores. No study compared students' CT outcomes with respect to their previous academic performance (GPA). A great majority of the studies did not even report the participants overall GPA. Only five (15%) studies reported that participants in the control and experimental conditions had similar GPA before the intervention, and all the five studies reported significant improvement on students' CT skills after the intervention (see Appendix A). Although the evidence is limited, this may indicate that the effect of instructional interventions could be masked when students in the experimental and control groups have diverse GPA.

### 3.3.2 Difference in Educational Level (Year Level)

Twenty three (70%) studies clearly reported the year level of the study participants. We observed that the large majority of the studies (n=10) targeted first year university students, followed by second year students (n=8) and least number of studies (n=5) focused on senior students (third and final year students). Among the ten studies in year one, seven (70%) of them reported significant CT improvement. Four (50%) studies in year two and three (60%) studies in senior level also reported significant CT improvement (see Appendix A).

In two studies (Chau et al., 2001; Hitchcock, 2004), participants were students from different year levels. In the one group pretest-posttest design by Chau and her colleagues, a total of 83 students were involved (38 students were from year one and 45 students were from year two). The students' posttest CT score overall did not significantly improve after the instructional intervention compared to the pretest score. When the CT scores of the students were compared by year level, again the difference was not significant (Chau et al., 2001). Similarly, a study by Hitchcock (2004) in which the research participants were from different year levels (second, third and fourth year) indicated that differences in mean gain by year level were not statistically significant.

We also examined the effectiveness of instructional interventions in relation to year level and CT instructional approach. When CT instructional approach employed is either general or infusion and targeted at first year students, all of the studies reported significant CT improvement. Among the six (75%) immersion studies targeted at second year students, only half of them reported significant CT improvement. We could not find significant variation in relation to effectiveness of CT approach and year level for senior students, but there appears to be adequate evidence that first year students benefited much from CT instruction when the approach adopted involves either teaching of CT principles as a separate course or when CT principles are explicitly taught within the subject matter. In addition, the immersion approach appears to be relatively more effective when employed to second year students (50% success) than first year students (33% success).

### 3.4 Teacher-Related Characteristics and CT Outcomes

We analyzed differences in CT outcomes when the teachers involved in the interventions were either trained or experienced, or when one or more of the research authors implemented the intervention, or when regular classroom teachers with no previous experience or training were assigned to teach. Twenty seven (82%) studies clearly revealed in their research report the individuals involved in implementing the CT intervention. Among the 27 studies, the instructional intervention for the 15 (56%) studies was implemented by one or more of the research authors from which nine (60%) of the studies yielded significant CT improvement. In 12 (44%) studies, the regular classroom teachers were involved in implementing the instructional intervention and five (41%) of the studies reported significant CT improvements. No information is available in those studies whether the classroom teacher(s) had previous training on CT instruction, except in one study by Mazer et al., 2007. Although the evidence is limited, it seems that significant CT improvement was observed when either the research authors or trained classroom teachers were involved in implementing the CT instructional intervention.

We examined whether there is association between the CT approach adopted and the teachers involved in implementing the instructional intervention. There is a little evidence that when the CT approach is that of general or mixed type, it does not seem to matter whether the intervention is implemented by the regular classroom teacher or the researcher(s). The success or failure rate of the CT interventions is similar across those two approaches. However, in the case of immersion and infusion approaches, we found that the instructional intervention resulted significant CT improvement mostly when either the intervention is implemented by the researcher(s) or a trained classroom teacher (see Appendix A).

### 3.5 CT Measurement

The most commonly employed standardized CT measures were the Cornell Critical Thinking Test (CCTT), California Critical Thinking Skills Test (CCTST), Watson-Glaser Critical Thinking Appraisal (WGCTA), and

Ennis-Weir CT Essay Test. The non-standardized measures were those developed by the research authors (researcher-made tests).

We compared the outcomes of CT instruction when either standardized or non-standardized measures were employed (see Appendix D). Studies which employed the non-standardized CT measures reported significant improvement on the posttest or between the experimental and control conditions more frequently (93% of the studies) than the standardized CT measures (55% of the studies). For example, two infusion approach studies (Anderson et al., 2001; Bensley & Haynes, 1995) that employed similar teaching strategies and research design but different CT measures reported differing CT outcomes. Anderson and his colleagues employed a standardized CT measure and found no significant CT improvement but the study by Bensley and Haynes that employed non-standardized CT measure reported significant CT improvement. The CT measure may be one reason for the difference in the CT outcomes of the two studies. We also noted that some variations on CT outcomes could be explained depending on the task/item format of the standardized CT measure employed. For example, in the study by Plath et al. (1999) in which two CT measures were employed together, significant CT improvement was revealed on the CT measure that required students to respond to the open-ended item than the multiple choice format.

Five studies (Anderson et al., 2001; Reed & Kromrey, 2001; Renaud & Murray, 2008; Stark, 2012; Williams et al., 2004) employed both generic and domain specific CT measures in parallel. All the five studies reported significant improvement on domain specific CT measures. However, only one study (Reed & Kromrey, 2001) reported significant improvement on both generic and domain specific measures. Overall, the evidence seems to indicate that our evaluation of the effectiveness of CT instruction could be influenced by the type of CT measures employed.

#### **4. Discussion**

##### *4.1 CT Instructional Approaches*

Among the four CT instructional approaches, it appears that CT skills are effectively enhanced when either the general or mixed approach is employed. However, this finding should be interpreted with caution as studies which adopted the general and mixed approaches are limited in number compared to the infusion and immersion approach. Perhaps, the very small number of studies which adopted the general and mixed approach compared to the immersion and infusion may indicate that there is a shift towards integrating CT instruction within subject-matter courses. A closer look at the CT instructional approaches adopted by studies included in previous systematic reviews (Abrami et al., 2008; Behar-Horenstein & Niu, 2011) also confirm that the majority (over 75%) of the intervention studies targeted either the immersion or infusion approach. This may indicate that embedding CT instruction within specific subject matter domains rather than teaching in separate courses is being considered as a more promising route to help students become critical thinkers.

However, despite the apparent shift towards discipline-embedded approach, the interventions which were in line with both the infusion and immersion approach seem to result in limited effect in students' abilities to transfer the acquired CT skills during subject-matter instruction to new tasks. Moreover, comparison of the effectiveness of the two approaches resulted in only marginal differences. Only about half of the immersion and infusion approach studies reported significant CT development. This finding does not support the one by Abrami et al. (2008) in which the immersion approach was clearly reported the least effective compared to the infusion approach. This difference may be due to the difference in the scope of the two studies. Abrami and his colleagues targeted intervention studies conducted at all educational levels, while our review targeted interventions only in higher education setting. The other reason could be the enormous variations in the individual studies in issues other than whether the general, infusion, immersion or mixed approaches were employed. The various features of studies across identical CT instructional approach differed in terms of specific teaching strategies employed, student and teacher related variables, and CT measures. Overall, our finding suggests that CT instructional approach alone may not determine effectiveness of instruction in fostering students' CT.

##### *4.2 Teaching Strategies*

There is some evidence to suggest that direct teaching strategies, which are based on explicit and detailed explanation of CT principles, are more effective than the implicit teaching strategies. Previous evidence (Beyer, 2008; Paul, 1993) also showed that such explicit teaching strategies are effective in improving students' CT. However, the evidence on effectiveness of teaching strategies that attempt to implicitly embed CT skills in subject matter instruction is inconsistent. The effect of most of the implicit teaching strategies (higher-order questioning, concept mapping, and small group discussion) was found to be inconclusive. For example, it appears that encouraging students to practice on subject related higher-order questions is less useful in fostering

CT. We noticed that such practice resulted in significant improvement in measures that are closely related to the original task (domain-specific CT measures). However, students were not able to transfer the acquired domain-specific thinking skills to novel contexts (as measured by domain-generic CT measures).

Overall, a closer analysis of most of the studies which attempted to implicitly embed CT skills within subject matter courses reveals that the designed interventions suffer from a few important limitations. Primarily, most of the interventions focused on a particular aspect of the instructional process and did not take into consideration all important design principles suitable for the development of CT. For instance, in the interventions which involved a discussion method of teaching (Garside, 1996; Stark, 2012), it is not clear (a) which components of CT were targeted during the subject-matter instruction, (b) what kind of tasks/problems were designed for discussion (it is not clear, for example, whether the tasks encouraged retention of facts or application to other contexts), (c) what role each of the students had during group discussions, and (d) what type of feedback or coaching was given by the instructor.

In the interventions which involved higher-order questioning, a number of limitations could be identified such as (a) absence of clear emphasis on the CT skills that were targeted in the design, (b) limited information regarding the various instructional activities during the whole teaching learning process apart from the provision of higher-level or lower-level questions at the end of each lesson, and (c) limited information on the nature of questions given to the students (e.g., abstract questions or questions meaningful to the students). Overall, there is evidence to suggest that the interventions did not (adequately) focus on important instructional activities which could have resulted in greater CT improvement. We argue that the inconsistent impact of the instructional interventions may be due to the inadequate application of empirically valid instructional principles in the design process.

#### *4.3 Student and Teacher-Related Characteristics*

Most of the student variables which might explain differential gains in CT skills and dispositions were not adequately reported in the targeted studies. For example, despite the theoretical and empirical arguments on gender differences in CT outcomes (Clinchy, 1996; Erickson & Strommer, 1991; Giancarlo & Facione, 2001; King et al., 1990), the current empirical research surprisingly does not examine effectiveness of instructional interventions in relation to gender. None of the studies included in this review reported the CT outcomes of male and female students separately. We found a little evidence with regard to academic performance (GPA), which may suggest that effect of instruction in fostering CT could be masked when students in the experimental and control groups have diverse prior knowledge (in terms of GPA). It appears that most of the immersion and infusion approach studies considered academic performance (or subject-matter understanding) essentially irrelevant to learners' CT outcomes. Only one study (Williams et al., 2004) examined the effectiveness of CT instruction with respect to students' academic performance, in which high exam performers outperformed the low exam performers in their CT outcomes. Although we do not find adequate evidence from the study reports to link academic performance with CT outcomes, previous review by McMillan (1987) indicated that students' content knowledge might influence the outcomes of instruction which target CT.

Regarding the influences of teacher-related characteristics, we noticed that students improved in their CT skills when either the research author(s) or trained classroom teacher(s) were involved in implementing the CT intervention. However, we want to caution that a large portion of the studies (44%) did not provide information on whether the regular classroom teachers involved in implementing the instructional intervention had either previous experience or formal training in CT instruction. We could not suggest, therefore, a strong association between teacher experience/training and effectiveness of a particular CT intervention. We noticed another interesting finding regarding the relationship between CT instructional approaches and experience of the teacher. In the immersion or infusion approach studies, frequent significant CT improvements were reported mainly when the instructional intervention was implemented by either the research author(s) or trained classroom teacher(s). Such relationships were not observed when studies adopted either the general or mixed approaches. This may suggest that the infusion and immersion approaches require teachers' adequate training and preparation to bring about greater CT improvement.

Overall, it could be generalized that adequate information regarding student and teacher-related variables is hardly reported in the current CT empirical evidence. Thus, definitive conclusions cannot be made regarding most of the student and teacher-related characteristics that significantly influence effectiveness of CT instructional interventions. We suggest that, in appraising effectiveness of CT instructional interventions, future research should emphasize the impact of a particular intervention in relation to various student-related characteristics and report the data. Such data would help practitioners establish clear evidence of the impact (or

lack of impact) of some student characteristics such as gender, year level, academic discipline, and academic performance in CT instruction. For example, if there is consistent evidence that some student characteristics could influence effectiveness of CT instruction, then such evidence would allow educators and researchers make informed decisions on how to design CT instruction tailored to a particular group of students.

#### 4.4 CT Measures

There is evidence to suggest that the CT measures studies employed may influence our evaluation of effectiveness of CT interventions. The finding overall shows that a relatively large number of studies, which employed non-standardized CT measures reported significant improvement in the posttest or between the experimental and control group than those studies which employed standardized measures. It is not clear on what type of items or tasks were included in the non-standardized (researcher-made) CT measures. However, it is more likely that items which were related to the tasks presented during the instructional intervention were administered, and that led to higher student CT outcome.

Moreover, even within studies which employed standardized CT measures, it appears that variations in effectiveness of instructional interventions could be explained by the item format. We found that CT measures that required students to respond to essay items yielded significant CT gains compared to those which require students to respond to multiple choice items. Although suggesting appropriate CT measuring instruments is beyond the scope of the present review, we want to point out that type of CT measure employed may influence our evaluation of the effectiveness of CT instruction and, therefore, attention needs to be given to the measurement of CT.

#### References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, 78(4), 1102-1134. <http://dx.doi.org/10.3102/0034654308326084>
- Allegretti, C. L., & Frederick, J. N. (1995). A model for thinking critically about ethical issues. *Teaching of Psychology*, 22(1), 46. [http://dx.doi.org/10.1207/s15328023top2201\\_14](http://dx.doi.org/10.1207/s15328023top2201_14)
- Alwehaibi, H. U. (2012). Novel program to promote critical thinking among higher education students: Empirical study from Saudi Arabia. *Asian Social Science*, 8(11), 193-204. <http://dx.doi.org/10.5539/ass.v8n11p193>
- Anderson, T., Howe, C., Soden, R., Halliday, J., & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. *Instructional Science*, 29(1), 1-32. <http://dx.doi.org/10.1023/A:1026471702353>
- Angeli, C., & Valanides, N. (2009). Instructional effects on critical thinking: Performance on ill-defined issues. *Learning and Instruction*, 19(4), 322-334. <http://dx.doi.org/10.1016/j.learninstruc.2008.06.010>
- Astin, A. (1993). *What matters in college? Four Critical Years Revisited*. San Francisco: Jossey-Bass.
- Barnett, J. E., & Francis, A. L. (2012). Using higher order thinking questions to foster critical thinking: A classroom study. *Educational Psychology*, 32(2), 201-211. <http://dx.doi.org/10.1080/01443410.2011.638619>
- Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning*, 8(2), 25-42.
- Bensley, D. A., & Haynes, C. (1995). The acquisition of general purpose strategic knowledge for argumentation. *Teaching of Psychology*, 22(1), 41-45. [http://dx.doi.org/10.1207/s15328023top2201\\_13](http://dx.doi.org/10.1207/s15328023top2201_13)
- Bensley, D., Crowe, D., Bernhardt, P., Buckner, C., & Allman, A. (2010). Teaching and Assessing Critical Thinking Skills for Argument Analysis in Psychology. *Teaching of Psychology*, 37, 91-96. [http://dx.doi.org/10.1207/s15328023top2201\\_13](http://dx.doi.org/10.1207/s15328023top2201_13)
- Beyer, B. (2008). How to teach thinking skills in social studies and history. *Social Studies*, 99(5), 196-201. <http://dx.doi.org/10.3200/TSSS.99.5.196-201>
- Chau, J. P. C., Chang, A. M., Lee, I. F. K., Ip, W. Y., Lee, D. T. T., & Wootton, Y. (2001). Effects of using videotaped vignettes on enhancing students' critical thinking ability in a baccalaureate nursing programme. *Journal of Advanced Nursing*, 36(1), 112-119. <http://dx.doi.org/10.1046/j.1365-2648.2001.01948.x>
- Chen, S.-L., Liang, T., Lee, M.-L., & Liao, I.-C. (2011). Effects of concept map teaching on students' critical thinking and approach to learning and studying. *Journal of Nursing Education*, 50(8), 466-469.

- <http://dx.doi.org/10.3928/01484834-20110415-06>.
- Clinchy, B. M. (1996). Connected and separate knowing: Toward a marriage of two minds. In N. R. Goldberg, J. M. Tarule, B. M. Clinchy, & M. F. Belenky (Eds.), *Knowledge, Difference, and Power: Essays Inspired by Women's Ways of Knowing* (pp. 205-247). New York: Basic Books.
- Daud, N. M., & Husin, Z. (2004). Developing critical thinking skills in computer-aided extended reading classes. *British Journal of Educational Technology*, 35(4), 477-487. <http://dx.doi.org/10.1111/j.0007-1013.2004.00405.x>
- Elliott, B., Oty, K., McArthur, J., & Clark, B. (2001). The effect of an interdisciplinary algebra/science course on students' problem solving skills, critical thinking skills and attitudes towards mathematics. *International Journal of Mathematical Education in Science and Technology*, 32(6), 811-816. <http://dx.doi.org/10.1080/00207390110053784>
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, 18(3), 4-10. <http://dx.doi.org/10.3102/0013189X018003004>
- Erickson, B. L., & Strommer, D. W. (1991). *Teaching College Freshmen*. San Francisco: Jossey-Bass Publishers.
- Facione, P., & Facione, N. (1992). *The California Critical Thinking Disposition Inventory*. Millbrae: California Academic Press.
- Garside, C. (1996). Look who's talking: A comparison of lecture and group discussion teaching strategies in developing critical thinking skills. *Communication Education*, 45(3), 212-227. <http://dx.doi.org/10.1080/03634529609379050>
- Giancarlo, C., & Facione, P. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *The Journal of General Education*, 50(1), 29-55. <http://dx.doi.org/10.1353/jge.2001.0004>
- Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39(2), 93-104. <http://dx.doi.org/10.1037/0003-066X.39.2.93>
- Halpern, D. F. (1993). Assessing effectiveness of critical thinking instruction. *The Journal of General Education*, 42(4), 238-254. <http://www.jstor.org/stable/27797197>
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. Dispositions, skills, structure training, and metacognitive monitoring. *The American psychologist*, 53(4), 449-455. <http://dx.doi.org/10.1037/0003-066X.53.4.449>
- Hitchcock, D. (2004). The effectiveness of computer-assisted instruction in critical thinking. *Informal Logic*, 24(3), 183-217.
- Huff, M. T. (2000). A comparison study of live instruction versus interactive television for teaching MSW students critical thinking skills. *Research on Social Work Practice*, 10(4), 400-416.
- Jonassen, D. H. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory* (pp. 215-239). New Jersey: Lawrence Erlbaum Associates.
- Kennedy, M., Fisher, M. B., & Ennis, R. H. (1991). Critical thinking: Literature review and needed research. In L. Idol, & B. F. Jones (Ed.), *Educational Values and Cognitive Instruction: Implications for Reform* (pp. 11-40). Hillsdale, NJ: Lawrence Erlbaum.
- King, P., Wood, P., & Mines, R. (1990). Critical thinking among college and graduate students. *The Review of Higher Education*, 13(2), 167-186.
- Kumta, S., Tsang, P., Hung, L., & Cheng, J. (2003). Fostering critical thinking skills through a web-based tutorial programme for final year medical students-a randomized control study. *Educational Multimedia and Hypermedia*, 12(3), 267-273.
- Magnussen, L., Ishida, D., & Itano, J. (2000). The impact of the use of inquiry-based learning as a teaching methodology on the development of critical thinking. *Journal of Nursing Education*, 39(8), 360-364.
- Mayer, R. E. (1992). Teaching of thinking skills in the sciences and mathematics. In D. F. Halpern (Ed.), *Enhancing Thinking Skills in the Sciences and Mathematics* (pp. 95-116). New Jersey: Lawrence Erlbaum Associates.
- Mazer, J., Hunt, S., & Kuznekoff, J. (2007). Revising general education: Assessing a critical thinking

- instructional model in the basic communication course. *The Journal of General Education*, 56(3-4), 173-199. <http://dx.doi.org/10.1353/jge.0.0000>
- McLean, C., & Miller, N. A. (2010). Changes in critical thinking skills following a course on science and pseudoscience: A quasi-experimental study. *Teaching of Psychology*, 37, 85-90. <http://dx.doi.org/10.1080/00986281003626714>
- McMillan, J. H. (1987). Enhancing college students' critical thinking: A review of studies. *Research in Higher Education*, 26(1), 3-29. <http://dx.doi.org/10.1007/BF00991931>
- McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational Researcher*, 19(4), 10-12. <http://dx.doi.org/10.3102/0013189X019004010>
- Nickerson, R. (1988). On improving thinking through instruction. *Review of Research in Education*, 15(3), 3-57. <http://www.jstor.org/stable/1167360>
- Nieto, A., & Saiz, C. (2008). Evaluation of Halpern's "structural component" for improving critical thinking. *The Spanish Journal of Psychology*, 11(1), 266-274.
- Pascarella, E. T., & Terenzini, P. T. (2005). *How College Affects Students. A third decade of research, Vol. 2*. San Francisco: Jossey-Bass.
- Paul, R. (1993). *Critical thinking: What every person needs to survive in a rapidly changing world*. Santa Rosa, CA: Foundation for Critical Thinking.
- Perkins, D., & Salomon, G. (1988). Teaching for transfer. *Educational Leadership*, 46(1), 22-32. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=afh&AN=8524829&site=ehost-live&scope=site>
- Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational Research*, 42(3), 237-249. <http://dx.doi.org/10.1080/001318800440579>
- Plath, D., English, B., Connors, L., & Beveridge, A. (1999). Evaluating the outcomes of intensive critical thinking instruction for social work students. *Social Work Education: The International Journal*, 18(2), 207-217. <http://dx.doi.org/10.1080/02615479911220201>
- Reed, J., & Kromrey, J. (2001). Teaching critical thinking in a community college history course: Empirical evidence from infusing Paul's model. *College Student Journal*, 35(2), 201-216.
- Renaud, R. D., & Murray, H. G. (2008). A comparison of a subject-specific and a general measure of critical thinking. *Thinking Skills and Creativity*, 3(2), 85-93. <http://dx.doi.org/10.1016/j.tsc.2008.03.005>
- Resnick, L. B. (1987). *Education and Learning to Think*. Washington DC: National Academy Press.
- Semerci, N. (2006). The effect of problem-based learning on the critical thinking of students in the intellectual and ethical development unit. *Social Behavior and Personality*, 34(9), 1127-1136. <http://dx.doi.org/10.2224/sbp.2006.34.9.1127>
- Sendag, S., & Odabasi, H. F. (2009). Effects of an online problem based learning course on content knowledge acquisition and critical thinking skills. *Computers & Education*, 53(1), 132-141. <http://dx.doi.org/10.1016/j.compedu.2009.01.008>
- Solon, M. (2007). Generic critical thinking infusion and course content learning in introductory psychology. *Journal of Instructional Psychology*, 34(2).
- Stark, E. (2010). Enhancing and assessing critical thinking in a psychological research methods course. *Teaching of Psychology*, 39(2), 107-112. <http://dx.doi.org/10.1177/0098628312437725>
- Szabo, Z., & Schwartz, J. (2011). Learning methods for teacher education: The use of online discussions to improve critical thinking. *Technology Pedagogy and Education*, 20(1), 79-94. <http://dx.doi.org/10.1080/1475939X.2010.534866>
- Ten-Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: Teaching strategies. *Learning and Instruction*, 14, 359-379. <http://dx.doi.org/10.1016/j.learninstruc.2004.01.005>
- Toy, B., & Ok, A. (2012). Incorporating critical thinking in the pedagogical content of a teacher education programme: Does it make a difference? *European Journal of Teacher Education*, 35(1), 39-56. <http://dx.doi.org/10.1080/02619768.2011.634902>
- Tsui, L. (1999). Courses and instruction affecting critical thinking. *Research in Higher Education*, 40(2). <http://dx.doi.org/10.1023/A:1018734630124>

- Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *The Journal of Higher Education*, 73(6), 740-763. <http://dx.doi.org/10.1353/jhe.2002.0056>
- Van-Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College Teaching*, 53(1), 41-46. <http://www.jstor.org/stable/27559216>
- Wheeler, L., & Collins, S. (2003). The influence of concept mapping on critical thinking in baccalaureate nursing students. *Journal of Professional Nursing*, 19(6), 339-346. [http://dx.doi.org/10.1016/S8755-7223\(03\)00134-0](http://dx.doi.org/10.1016/S8755-7223(03)00134-0)
- Williams, R. L., & Worth, S. L. (2001). The relationship of critical thinking to success in college. *Inquiry: Critical Thinking across the Disciplines*, 21, 5-16.
- Williams, R. L., Oliver, R., & Stockdale, S. (2004). Psychological versus generic critical thinking as predictors and outcome measures in a large undergraduate human development course. *The Journal of General Education*, 53(1), 37-58. <http://dx.doi.org/10.1353/jge.2004.0022>
- Yang, Y., Newby, T., & Bill, R. (2008). Facilitating interactions through structured web-based bulletin boards: A quasi-experimental study on promoting learners' critical thinking skills. *Computers & Education*, 50(4), 1572-1585. <http://dx.doi.org/10.1016/j.compedu.2007.04.006>
- Yeh, Y. C. (2009). Integrating e-learning into the Direct-instruction Model to enhance the effectiveness of critical-thinking instruction. *Instructional Science*, 37(2), 185-203. <http://dx.doi.org/10.1007/s11251-007-9048-z>
- Yuan, H., Kunaviktikul, W., Klunklin, A., & Williams, B. (2008). Improvement of nursing students' critical thinking skills through problem-based learning in the People's Republic of China: A quasi-experimental study. *Nursing & Health Sciences*, 10(1), 70-76. <http://dx.doi.org/10.1111/j.1442-2018.2007.00373.x>

## Appendix A

Selected study features of articles included in the review

Study	CT Approach	Teaching strategy	Educational Level	Delivery of CT Instruction (RT, RA)	CT Measure	Effect(+, 0)
Allegretti and Frederick (1995)	General	Direct instruction (DI)-teacher modeling & small group discussion	Senior	RA	CCTT	+
Alwehaibi (2012)	General	DI (teacher modeling & small group discussion using real-world examples)	2 <sup>nd</sup>	RA	CTAI	+
Anderson et al. (2001)	Infusion	DI (teacher modeling & peer-based critique exercises)		RA	DS CRT	+
AngeliandValanides (2009) <sup>a</sup>	Infusion	Peer discussion & reflection			CCTT	+
Barnet and Francis (2012)	Immersion	Higher-order questioning	2 <sup>nd</sup>	RT	WGCTA	0
Bensley and Haynes (1995)	Infusion	DI (teacher modeling & coaching)	1 <sup>st</sup>	RT	NS	+
Bensley et al.	Infusion	DI (teacher modeling &		RA	NS	+

(2010) <sup>b</sup>		coaching)				
Chau et al. (2001)	Mixed	Role playing and discussion using prompts	1 <sup>st</sup> , 2 <sup>nd</sup>	RT	CCTT	0
Chen et al. (2011)	Immersion	Concept maps & small group discussion	1 <sup>st</sup>		NS	+
Daud and Husin, (2004) <sup>b</sup>	Immersion	Discussion method		RA	CCTT	+
Elliot et al. 2001	Immersion	Discussion using demonstration & repeated exercises	1 <sup>st</sup>	RA	WGCTA	0
Garside (1996)	Immersion	Small group discussion	1 <sup>st</sup>	RT	NS	0
Hitchcock (2004)	General	Independent learning with computer-based tutorial	2 <sup>nd</sup> , 3 <sup>rd</sup> , 4 <sup>th</sup>		CCTST	+
Huff (2000)	Immersion	Small group teacher-led discussion	1 <sup>st</sup>		CCTST	0
Kumta et al. (2003)	Immersion	Peer discussion	Senior	RT	NS	+
Magnussen, Ishida, and Itano (2000)	Immersion	Inquiry-based learning (small group discussion & tutor guidance)	1st, 4th	RT	WGCTA	0
Mazer et al. (2007) <sup>b</sup>	Infusion	DI (explicit teaching of the CT strategies, exercises, peer & self-evaluation)	1st	RA	NS	+
McLean and Miller (2010)	Infusion	DI(explicit teaching of CT skills followed by reflection & class presentations)	Senior	RT	WGCTA	0
					NS	0
Nieto and Saiz (2008)	General	DI (Modeling, real-world exercises & feedback)	Senior	RA	CCTT	0
Plath et al. (1999)	Mixed	DI (Arguments & debates with teacher feedback)	Senior	RA	Ennis-Weir	+
					CCTT	0
Reed and Kromrey (2001)	Infusion	DI (modeling, small group discussion &	1st	RA	Ennis-Weir	+

feedback)						
					DS	+
					CCTDI	0
Renaud and Murray (2008)	Immersion	Higher-order questioning	1st	RA	WGCTA	0
					DS	+
Semerci (2006)b	Immersion	PBL with small group discussion	2nd	RA	NS	+
Sendag and Odabasi (2009)	Immersion	PBL (using ill-structured problem scenarios & small group discussion)	2nd		WGCTA	+
Solon (2007)b	Mixed	DI (coaching)	1st	RA	CCTT	+
Stark (2012)	Immersion	Discussion method (critiquing scientific examples)	2nd		CCTT	0
					DS	+
Szabo and Schwartz (2011)	Immersion	Discussion method		RT	Ennis-Weir	+
Toy and Ok (2012)	Infusion	DI (through questioning, role playing & case study)	2nd	RA	CCTDI	0
Wheeler and Collins (2003)	Immersion	Concept maps	2nd	RA	CCTST	0
Williams et al. (2004)	Infusion	DI (coaching through CT practice questions)	1st, 2nd, 3rd	RA	WGCTA	0
					DS	+
Yang et al. (2008)	Immersion	Group discussion		RA	CCTST	+
Yeh (2009)	General	DI (Modeling, scaffolding & small group discussion)			NS	+
Yuan et al. (2008)	Immersion	PBL through small group discussion	2nd	RT	CCTT	+

*Note.* + denotes significant CT improvement from the pretest to posttest or between the experimental and control conditions; 0 denotes no significant CT improvement. RT=CT instruction delivered by the Regular classroom Teacher; RA=CT instruction delivered by the Research Author(s); DS=Domain Specific; NS=Non-standardized; RT=CT instruction is delivered by the Regular classroom Teacher; RA=CT instruction delivered by the Research Author(s).

<sup>a</sup>The study focused on comparison of the general, infusion and immersion approaches, in which infusion was found superior.

<sup>b</sup>Studies in which students in the experimental and control group had similar GPA.

## Appendix B

Summary of the study findings in relation to CT instructional approach

CT approach	N	Effect	
		+ (n)	0 (n)
General	5	4	1
Infusion	9	5	4
Immersion	16	8	8
Mixed	3	2	1
Total	33	19 (58%)	14 (42%)

*Note.* + denotes significant CT improvement from the pretest to posttest or between the experimental and control conditions; 0 denotes no significant CT improvement.

## Appendix C

Summary of the teaching strategies

Category	Teaching strategies
Direct instruction	Teacher modeling & small group discussion (Allegretti & Frederick, 1995; Anderson et al., 2001; Hitchcock, 2004)
	Explanation of thinking skills, teacher modeling, scaffolding, coaching & feedback (Alwehaibi, 2012; Bensley & Haynes, 1995; Bensley et al., 2010; Nieto & Saiz, 2008; Reed & Kromrey, 2001; Solon, 2007; Toy & Ok, 2012; Yeh, 2009)
	Explanation of thinking skills, exercises, reflection, peer evaluation (Angel i& Valanides, 2009; Mazer et al., 2007; McLean & Miller, 2010)
Implicit instruction	Explanation of CT guidelines & role playing (Chau et al., 2001; Plath et al., 1999)
	Higher-order questioning (Barnet & Francis, 2012; Renaud & Murray, 2008; Williams et al., 2004)
	Concept maps & small group discussion (Chen et al., 2011; Wheeler & Collins, 2003)
	Discussion method (Daud & Husin, 2004; Elliot et al., 2001; Garside, 1996; Huff, 2000; Kumta et al., 2003; Magnussen, Ishida, & Itano, 2000; Stark, 2012; Szabo & Schwartz, 2011; Yang et al., 2008)
	Problem-based Learning (PBL) (Semerci, 2006; Sendag & Odabasi, 2009; Yuan et al., 2008)

**Appendix D**

## Instructional approach, CT measurement and research design

CT approach	CT measure				Research design		
	Standardized measures		Non-standardized measures		E (n)	QE (n)	OG (n)
	n	Effect (+,0)	n	Effect (+,0)	Effect (+,0)	Effect (+,0)	Effect (+,0)
General (N=5)	4 (80%)	3(+) 1(0)	1(20%)	1(+)		1(+) 1(0)	3(+)
Infusion (N=9)	6 (66%)	2(+) 4(0)	3(34%)	3(+)	1(+)	4(+) 4(0)	
Immersion (N=16)	11 (68%)	5(+) 6(0)	5(32%)	4(+) 1(0)	2(+) 1(0)	6(+) 5(0)	2(0)
Mixed (N=3)	3 (100%)	2(+) 1(0)	0 (0%)			1(+)	1(+) 1(0)
Total	24 (73%)		9 (27%)		4 (12%)	22 (67%)	7 (21%)

*Note.* + denotes significant CT improvement from the pretest to posttest or between the experimental and control conditions; 0 denotes no significant CT improvement. E=Experimental; QE=Quasi Experimental; OG=One group pretest-posttest.

**Copyrights**

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

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).