ARI Contractor Report 2006-01

Computerized Training in Critical Thinking $(CT)^2$:

A Skill-Based Program for Army Personnel

Susan C. Fischer
Anacapa Sciences, Inc.

V. Alan Spiker
Anacapa Sciences, Inc.

Douglas H. Harris
Anacapa Sciences, Inc.

Evan R. McPeters
Anacapa Sciences, Inc.

November 2006

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

MICHELLE SAMS	**ZITA M. SIMUTIS**
Technical Director	**Director**

Research accomplished under contract
for the Department of the Army

Human Resources Research Organization

Technical review by

Sharon Riedel, U.S. Army Research Institute
XXXX, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Research Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPC-ARI-MS, 2511 Jefferson Davis highway, Arlington, Virginia 22202-3926.

FINAL DISPOSITION: This Research Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Research Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) November 2006	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) February 2004 – November 2006

4. TITLE AND SUBTITLE Web Based Training for Theoretically Derived Critical Thinking Skills	5a. CONTRACT OR GRANT NUMBER W74V8H-04-C-0007
	5b. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Susan C. Fischer, Anacapa Sciences, Inc., V. Alan Spiker, Anacapa Sciences, Inc., Douglas H. Harris, Anacapa Sciences, Inc., Evan R. McPeters, Anacapa	5c. PROJECT NUMBER
	5d. TASK NUMBER
	5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Anacapa Sciences, Inc. 301 E. Carrillo St. Santa Barbara, CA 93101	8. PERFORMING ORGANIZATION REPORT NUMBER 1491-Final
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 851 McClellan Ave. Ft Leavenworth, KS 66027-1360	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Contracting Officer's Representative and Subject Matter POC: Dr. Sharon Riedel

14. ABSTRACT *(Maximum 200 words)*: This report constitutes the final deliverable of a seven-year research and developmen effort to create distance training in critical thinking (CT) for Army personnel. The primary product of this research program was easily distributed and accessed training in CT skills. The development of the training was supported by research that was conducted over three broad phases. The third and final phase of the research program, which is th primary focus of the present report, was devoted to developing, and conducting the necessary research to evaluate, a complete Internet-based training package that provides educational and assessment experiences for eight CT skills. . user-centered design process was used to ensure that the resulting training and assessment products were usable, useful, and well accepted by potential training populations. Consistent with the user-centered design philosophy, four usability investigations were conducted. The effectiveness of the resulting training system was evaluated in an additic two investigations. The results of both evaluation investigations indicated that the training increased a participant's sk on the two critical thinking skills evaluated.

15. SUBJECT TERMS
Critical thinking, computer-based training, web-based training, SCORM, critical thinking skill

SECURITY CLASSIFICATION OF	19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON

16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	38	Ellen Kinzer Technical Publications Speci 703-602-8047

ACKNOWLEDGEMENT

We are very grateful for having had the opportunity to study critical thinking over the past seven years. The main product of this research, a web-based training program that develops eight critical thinking skills, could not have been developed without the support of the Army Research Institute at Fort Leavenworth Kansas. In particular, we are deeply grateful for the continued support of Dr. Sharon Riedel, the project's technical director. Sometime around 1999, she recognized that empirically guided research was greatly needed on critical thinking, and that the Army might benefit from that research. We were fortunate to, first, be given the opportunity to work with Dr. Riedel in developing a model of critical thinking, and, second, to continue to be supported by her to develop an effective training system. Her contributions to the project were many and varied, and very appreciated.

Several employees of Anacapa Sciences, Inc. contributed to the development of Computerized Training in Critical Thinking (CT)^2. Bill Campsey, in particular, was instrumental in shaping the approach we took to training critical thinking. His ideas and words can be found on nearly every page of the training. Even his voice can be heard in the audio portions as we learned that he had significant theatrical talent. Having served in the Army for many years and retired, Bill also served as our subject matter expert and kept the training relevant, interesting, and engaging. He is largely responsible for creating almost all of the scenarios contained in (CT)2 and authored major components of several modules. The training system could not have been created as such without his contributions.

We are also grateful to Dr. Eric Holder and Dr. Patricia Mautone, who contributed greatly to the creation of several of the modules. Eric's industry and creativity produced a large amount of material that was incorporated into the training. He seemed to always be able to find the right graphics and put the package together, even under time constraints. Similarly, we are grateful to Tricia for authoring one of the skills. More than that, though, we are deeply indebted to her for her careful review of all the material and her expertise in multi-media presentation. She is largely responsible for setting standards for the pretests.

Finally, we would like to express our gratitude for the contributions of Anacapa's colleague, Jim Kornell, who authored several of the training components. His ability to see the core of a critical thinking skill and derive its components was instrumental to our approach and to the product. His ideas and words can also be found throughout the training components.

COMPUTERIZED TRAINING IN CRITICAL THINKING (CT)2

A SKILL-BASED PROGRAM FOR ARMY PERSONNEL

EXECUTIVE SUMMARY

Research Requirement:

This report constitutes the final deliverable of a seven-year research and development effort to create distance training in critical thinking (CT) for Army personnel. The impetus to develop easily accessed and distributed training in critical thinking was a perceived need to improve CT skills in Army personnel. The United States military has served a leadership role in promoting thinking skills for good reason. Today's situations demand novel, yet insightful, solutions, which can only be derived from sharply honed thinking skills. Army educational systems, however, have historically devoted far fewer resources to the training of thinking processes than to other important skills. Traditional methods of training tactical decision-making offer a prescriptive model that corresponds to doctrine, and focuses on the *products* of decision-making. The procedural nature of the *doctrinal methods may actually discourage the application or development of thinking skills*, inhibiting the creation of novel solutions that might be the result of CT.

Procedure:

The primary product of this research program was easily distributed and accessed training in CT skills. The development of the training was supported by research that was conducted over three broad phases. The third and final phase of the research program, which is the primary focus of the present report, was devoted to developing, and conducting the necessary research to evaluate, a complete Internet-based training package that provides educational and assessment experiences for eight CT skills. A user-centered design process was used to ensure that the resulting training and assessment products were usable, useful, and well accepted by potential training populations. Consistent with the user-centered design philosophy, four usability investigations were conducted. Feedback obtained in the usability investigations was used to inform and make modifications to the training system. The effectiveness of the resulting training system was evaluated in an additional two investigations. Samples of potential users were asked to complete portions of the training. Their abilities were tested before and after the training to determine if the training had increased critical thinking skill. One of the evaluation investigations also assessed whether an extended version of the training provided any greater learning benefit than the standard version.

Findings:

The results of both evaluation investigations indicated that the training increased a participant's skill on the two critical thinking skills evaluated. These findings indicate that the training program is effective at increasing learning. Because the training and testing components of each of the eight skills follow a similar pedagogy and presentation format, one might expect that each training module would produce a similar learning effect. The results of the second evaluation investigation suggest that the standard training is as effective as the extended training in teaching critical thinking skills. Given the limited time often available for extra training these results suggest that the standard version is sufficient to produce a desired learning effect. The extended version, then, could be used if a student wanted greater explanation or needed some question resolved. The extended version could serve, then as supplemental material, which could be made available for those trainees who desire additional training.

Utilization of Findings:

The main product of the effort was a web-based interactive training system, herein referred to as **C**omputerized **T**raining in **C**ritical **T**hinking, or $(CT)^2$, which was demonstrated to be effective at increasing critical thinking skills. $(CT)^2$ is a modular training system that is accessed from the Internet and can be self-administered. It comprises a number of training and assessment components including pretests, training modules, and posttests for each of eight CT skills. The pretests and training modules are highly interactive, include multiple exercises with corresponding feedback, and utilize multi-media presentation. All of the components are designed to not only increase skill, but also to increase self-awareness of one's thinking. $(CT)^2$ is based on an empirically tested model of CT, which was developed during the research effort, and well-founded pedagogical principles. The particular CT skills on which the system focuses were empirically identified in the first phase of the research, and were found to be important and problematic to Army leadership. The $(CT)^2$ software itself is SCORM compliant and ready for implementation with any learning management system (LMS). $(CT)^2$ is available for use by any Army organization. Groups that are experiencing Command and Control issues related to deficiencies in critical thinking will find it particularly useful.

The completion of this seven-year research program should be only the start to empirical research on CT. It brings to the field two critical sources of power, heretofore unavailable to researchers. The first is a testable model of CT that can guide future research efforts on the construct. The second is (CT)2, which provides an off-the-shelf training package and a model on which future training development efforts can be launched. Therefore, this research positions the Army to conduct more advanced research on critical thinking, which may well provide significant benefit to the Army's objectives.

COMPUTERIZED TRAINING IN CRITICAL THINKING $(CT)^2$
A SKILL-BASED PROGRAM FOR ARMY PERSONNEL

CONTENTS

LIST OF TABLES

LIST OF FIGURES

COMPUTERIZED TRAINING IN CRITICAL THINKING $(CT)^2$
A SKILL-BASED PROGRAM FOR ARMY PERSONNEL

This report constitutes the final deliverable of a seven-year research and development effort to create distance training in critical thinking (CT) for Army personnel. The main product of the effort was a web-based interactive training system, herein referred to as **C**omputerized **T**raining in **C**ritical **T**hinking, or $(CT)^2$. $(CT)^2$ is a modular training system that is accessed from the Internet and can be self-administered. It comprises a number of training and assessment components including pretests, training modules, and posttests for each of eight CT skills. The pretests and training modules are highly interactive, include multiple exercises with corresponding feedback, and utilize multi-media presentation. All of the components are designed to not only increase skill, but also to increase self-awareness of one's thinking. $(CT)^2$ is based on an empirically tested model of CT, which was developed during the research effort, and well-founded pedagogical principles. The particular CT skills on which the system focuses were empirically identified in the first phase of the research, and were found to be important and problematic to Army leadership. The $(CT)^2$ software itself is SCORM compliant and ready for implementation with any learning management system (LMS).

The impetus to develop easily accessed and distributed training in critical thinking was a perceived need to improve CT skills in Army personnel. The United States military has served a leadership role in promoting thinking skills for good reason. Military leaders have always had to make tactical decisions in complex and stressful situations where knowledge is incomplete and uncertain, and when sound thinking can make the difference between success and failure. But today's asymmetrical warfare is very unlike the warfare that Army personnel have previously experienced or studied, and therefore requires innovative thinking. Today's situations demand novel, yet insightful, solutions, which can only be derived from sharply honed thinking skills.

Army educational systems, however, have historically devoted far fewer resources to the training of thinking processes than to other important skills. Officers and enlisted personnel are thoroughly educated in the practice and art of warfare. However, *relatively* little training time has been spent on improving the *process* of thinking and decision-making[1]. Traditional methods of training tactical decision-making offer a prescriptive model that corresponds to doctrine, and focuses on the *products* of decision-making (Fallesen, Michel, Lussier, & Pounds, 1996). The procedural nature of the *doctrinal methods may actually discourage the application or development of thinking skills*, inhibiting the creation of novel solutions that might be the result of CT.

If relatively few resources are devoted to developing good thinking habits, Army personnel must develop and hone *their own* methods of thinking to support decision-making. Without explicit training, whatever thinking skills a Soldier possesses are gained through on-the-job experience, fortuitous experiences in training exercises, individual disposition, or other idiosyncratic means such as self-study. Establishing an integrated training program to address the development of thinking skills is clearly preferable to hoping that these skills will somehow

[1] A recent revision to the Command General Staff Officer's Course (CGSOC) curriculum at the Command General Staff College at Fort Leavenworth, Kansas is a notable exception. The proposed curriculum for the CGSOC includes training in critical thinking that is integrated into every unit.

develop on their own. Therefore, the primary purpose of this research program was to provide easily distributed and accessed training in CT skills. The development of the training was supported by research that was conducted over three broad phases. To provide a historical context to the present report, these three phases are briefly reviewed.

Historical Overview of Research Program

The first phase in the development of $(CT)^2$ constituted preparatory research, which investigated CT, CT in the military context, and pedagogical techniques used to teach CT. In the second phase, a prototype CT training program was developed for two CT skills. The final phase, which is the primary focus of the present report, was devoted to developing a complete Internet-based training package that provides educational and assessment experiences for eight CT skills. The major accomplishments of the first two phases are discussed below.

Phase I: Preparatory Research

The purpose of the Phase I effort was to conduct the necessary preliminary research that would support development of a model of CT. The major tasks achieved in Phase I were (1) review existing literature on CT, (2) inventory CT skills identified in literature, (3) develop an empirically testable model of CT, (4) specify the role of CT within the Army, (5) select a set of CT skills for which to develop training, (6) establish lessons learned from the literature on training CT and from existing efforts to teach CT to Army officers, and (7) establish functional requirements for the to-be-developed instructional material.

A thorough review of existing literature on CT concluded that the field is highly fragmented and lacks consensus. A notable lack of empirical study of CT as a psychological process and/or individual difference variable is a major problem with this field, and undoubtedly the reason for its lack of coherence. The Phase I work revealed that research on CT clearly needed an empirically testable model of CT that would serve to ground the construct in objectively observable behavior. The literature review also produced an inventory of CT skills, and provided input to development of a falsifiable model. One of the most important contributions of the preparatory research was a nascent, yet testable, model of CT. Important modifications and clarifications were made to this model later in the second phase of the research.

The preparatory research also examined the role of CT within the Army. A survey of eighteen Army officers at Ft. Hood, Texas was conducted. Participants completed a five-page survey that assessed their opinions and experiences concerning CT skills, predisposing attitudes, and situational conditions within the Command and Control (C2) domain. The officers were asked to rate how important each of thirteen broad categories of CT skills were to C2. They were also asked to indicate if, in their experience, they had observed problems in the execution of each broad skill category. The survey results were then used to select several categories of skills that are key to effective C2. Selecting a subset of key CT skills from the set of 13 skill categories evaluated by the officers was accomplished by applying two criteria to the Fort Hood survey data: the mean importance rating and the reported incidence of problems in executing the CT skills. By jointly considering the two criteria in our selection process, ten[2] CT skills were selected and became the target skills to-be-trained in $(CT)^2$.

---

[2] This list of 10 key CT skills was later reduced to 8 when similarity and overlap among several skills were noted.

Before developing the prototype to $(CT)^2$ however, lessons learned were gathered from (1) the civilian research literature on training CT and (2) current efforts to teach CT to Army officers. The literature review found that educators have used a wide variety of methods to teach CT, which were inventoried in Fischer, Spiker & Berkman (2001). However, only a small proportion of those methods have been empirically validated. Lessons learned from educators at the Command General Staff College (CGSC) at Fort Leavenworth, Kansas State University, the Army War College, and the Army Research Institute revealed a number of issues. Respondents reported that the biggest difficulty instructors experienced came from the Army itself. Specifically, Army culture tends to *discourage* thinking to some extent. Because of the Army's hierarchical nature and the requirement for unit cohesiveness, thinking "out of the box" is often discouraged by leaders and by doctrine. For this reason, courses in CT that are electives typically have low enrollments, and students sometimes balk at the content in CT courses.

A final effort of the preparatory research was to integrate the information that had been gathered to establish functional requirements for $(CT)^2$. The research concluded that a CT training system would need to be (1) practically integrated with current Army training methods, (2) administered at multiple levels of training, and (3) readily accessed by instructors and students in a variety of settings such as the schoolhouse or field unit. With these functional requirements in mind, the second phase of the research program was initiated. A complete discussion of the research and results conducted in the preparatory phase can be found in Fischer and Spiker (2000), and Fischer, Spiker, and Berkman (2001).

Phase II: Prototype Development and Evaluation

Model Development and Validation. The nascent model developed in the preparatory phase was refined in the second phase, and two studies were conducted that served to validate the model. The resulting model incorporates many ideas about CT offered by leading thinkers in philosophy and education. It embodies many of the CT skills and predisposing attitudes discussed in the CT literature. It also specifies the relationships among a variety of variables that previous researchers have discussed, such as the influence of experience and knowledge, and the relationship of CT to cognitive tasks (e.g., judgment and problem solving). The model, however, goes far beyond the largely *rational/analytic* work conducted to date by providing a framework in which CT can be *empirically* investigated as a cognitive process. For the reader's convenience a summary discussion of the model is contained in Appendix A.

The first validation study provided corroborating evidence that the ten CT skills previously identified in the preparatory research were both important and problematic to C2. The results also indicated that the CT model largely captures the skills, situational conditions, and predisposing factors significant to Army leadership.

The second validation experiment tested four of the model's predictions involving the effects of situation (type of material to be processed), type of task, predisposing individual differences, and experience level on CT. It also examined the model's proposition that CT can have negative affective consequences. The model's predictions concerning the type of material to be processed were generally supported. Substantive material took longer and was more effortful to process than low substance material. Response times were longer for degraded substantive material than non-degraded substantive material. Moreover, these effects occurred only when the task was to understand or judge the material, which is predicted by the model. When the task is not one of the predicted purposeful meta-tasks, substance of the material has little effect on time or effort. Tasks in which participants had to make judgments or had to understand material took longer and were more effortful than tasks that required simple identifications, as predicted.

The results of this study failed to support the notion that predisposing individual difference factors affect the tendency to engage in CT skills. A "Need for Cognition" (NFC) scale failed to correlate with any measure of CT. However, potential restriction of range problems on the NFC scale may have produced this null effect. Hence, at this point, we believe that it is premature to conclude that predisposing factors do not affect the application of CT skills.

The study further found that experience level does affect the application of CT skills. Highly experienced participants expressed more questions of disbelief about certain types of material than did less experienced participants. However, negative affect was not a simple, direct outcome of CT in this study. Instead, it appears that other factors such as type of task and type of material determined the level of enjoyment experienced by participants.

In summary, the model of CT successfully generated a host of predictions that, heretofore, had not been empirically tested. The model was sufficiently specified to permit falsification of many of its assertions, which other musings on CT had not provided. The present study took advantage of the specification and tested five of its central predictions. As a result, current knowledge of CT was significantly increased.

Prototype Training System for Two Skills. The newly validated CT model provided the basis on which to design and develop a prototype training system. The prototype system targeted two of the ten CT skills previously identified. The two skills were "Frame the Message" and "Recognize Gist in Material". Because the model proposes that CT is a deliberate, systematic awareness of the process and products of one's own thinking, the prototype training focused on increasing self-awareness. It also targeted common – and potentially serious – errors that people make when they fail to apply the two CT skills. The prototype program highlighted awareness of common errors and taught specific techniques that can be used to overcome them. It presented the student with real-world situations and asked them to complete numerous thinking exercises that require the practice and application of CT skills in a variety of realistic settings.

Evaluation of Prototype Training. To evaluate the effectiveness, usability, and student acceptance of the prototype training system, a third empirical study was conducted during the second phase of the research program. The central objective of the study was to determine whether the prototype CT training system effectively increases measurable indicators of CT compared to two other learning conditions. The study also assessed participating students' attitudes and subjective evaluations of the training as indicators of acceptance and usability. The effectiveness of the prototype system was evaluated in a study in which participants from the 85th Reserve Training Division worked through parts of the training. The results of this study indicate that military students found the training highly acceptable. Although the sample of participants who used the training was small, it was uniformly positively rated. Users found it interesting and well worth their time. The prototype training also appeared to be generally effective at encouraging critical thinking, at least about messages Army personnel must evaluate.

The study clearly showed that the web-based prototype enhanced memory for messages, probably because it encourages greater depth of processing.

The evaluation study also showed that the training inhibited the production of (potentially incorrect) inferences that go well beyond what is explicitly given in the message. Participants who took the training made far fewer unjustified inferences than participants assigned to the other two training conditions. Examination of the responses reveals that training participants did make inferences; however, they justified them by pointing out explicit information given in the message that supported their inferences. Therefore, the prototype system appeared to encourage discrimination of what is "known" or "given" from what might be added (i.e., inferred) by the perceiver.

Phase III: Development of (CT)2

The third and final phase of the research program, which is the primary focus of the present report, was devoted to developing a complete Internet-based training package that provides educational and assessment experiences for eight CT skills. As noted previously, the original list of ten CT skills was reduced to eight after overlap and similarity were taken into account. The complete training package, Computerized Training in Critical Thinking, or $(CT)^2$, incorporates both training and assessment components. For four of the skills, the training component is supplemented with an extended training version.

A user-centered design process was used to ensure that the resulting training and assessment products were usable, useful, and well accepted by potential training populations. Early in the development of $(CT)^2$, a panel of Army training experts was convened to review early conceptions of its design. A systematic design process was also used, starting with the development of technical specifications for each training component, proceeding to storyboarding and content development, and then moving to implementing the content on the web. Consistent with the user-centered design philosophy, four usability studies were also conducted. In each study, Army Soldiers or reservists worked through and evaluated one or more of the training or assessment components. The development process and resulting curriculum is described in the next section.

Overview of $(CT)^2$ Curriculum

With the above discussion as an overview, let us now consider the specific principles on which $(CT)^2$ is based. Here, we briefly describe 18 core principles that were used to guide the development of $(CT)^2$ training materials (Fischer & Spiker, 2004b; Spiker, 2004.) To structure the presentation, these principles have been organized around four themes or meta-principles: (1) CT is a skill, (2) focus CT training on selected aspects of a task, (3) use proven methods of adult learning, and (4) a computer-based infrastructure is needed for training delivery and instructional support. This organization scheme is depicted in Figure 1. The principles associated with each theme are discussed in turn.

Critical Thinking is a Skill

Principle 1: Critical thinking skills can be learned, trained, and transferred. While cognitive skill performance may be affected by inherited traits or abilities, it is generally assumed that,

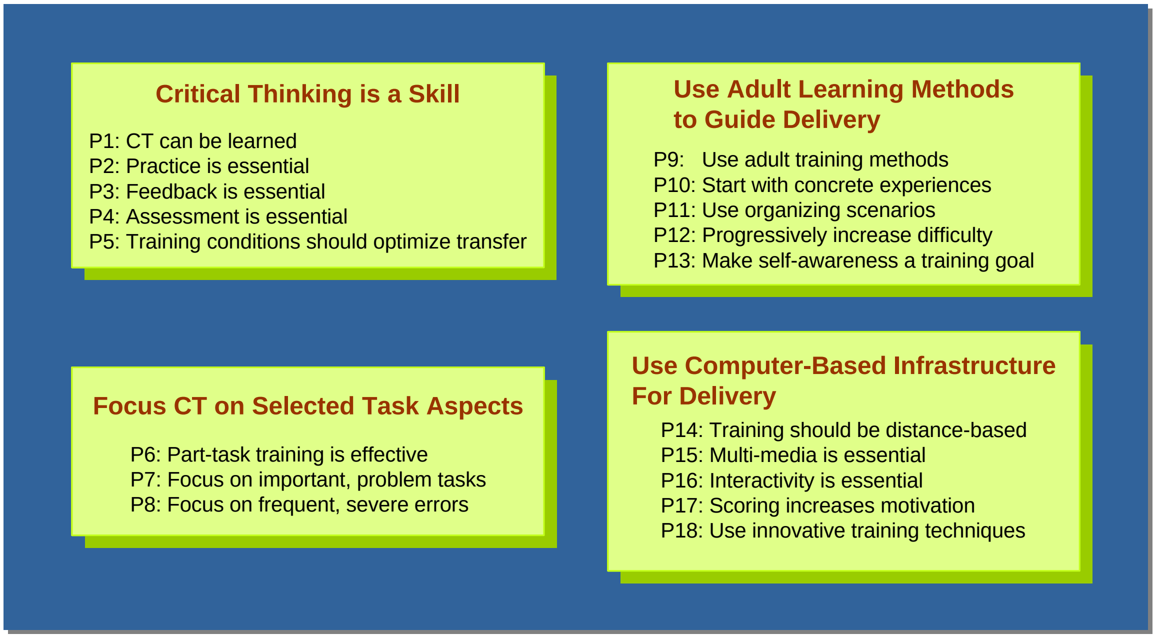

Figure 1. Overview of the 18 principles that structured the (CT^2) program.

with appropriate training and instruction, anyone with normal cognitive capacity can develop the requisite skill proficiency.

Principle 2: Practice is essential. Skill performance does not improve simply with an increase in content knowledge alone. Practice is critical to skill development whether the skill in question is psychomotor or cognitive.

Principle 3: Feedback is essential. Feedback about performance is also required. Learners cannot modify their performance without knowing how well they did. Indeed, learning will not occur in the absence of feedback.

Principle 4: Assessment is essential. Because it is impossible to give feedback in the absence of assessment, this fourth principle is essentially a corollary of the previous principle. Thus, an evaluation of student performance during training is essential so that informative feedback *about* that performance can be provided as students practice a skill.

Principle 5: Training conditions should optimize transfer. Research has shown that leading students toward perfect performance during training does not ensure that such high-performance will transfer to real-world situations. In fact, just the opposite seems to be true—transfer is decreased when mastery learning is required during training.

Focus Critical Thinking Training on Selected Task Aspects

Principle 6: Part-task training methods are most effective. \(CT)^2 provides students with instruction and practice on only those *parts* of real world C^2 tasks that demand CT resources rather than the entire task. For example, rather than having the student analyze an entire OPORD, the instructional materials focus only on *part* of the lengthy document (e.g., enemy intent).

Principle 7: Focus on important and problematic cognitive tasks. To assume that officers need to be trained on all possible C^2 tasks is a mistake. As a practical matter, then, the training focuses on CT skills that have been empirically documented to be important and problematic to C^2.

Principle 8: Focus on frequent and consequential errors. Training the enabling subtasks of a complex CT skill would be inefficient and unnecessary, as it is likely that students already perform them well. A better solution is to evaluate where training is most needed, and then focus training on those problematic areas. One clear indicator of where training is needed is on those components of a task (subtasks) where errors are frequent and where the consequences of errors are severe.

Use Adult Learning Methods to Guide Instructional Delivery

Principle 9: Use training methods appropriate for adults. The adult population of Army personnel is the targeted student audience for $(CT)^2$. Research in instructional design has shown that, to be successful, the methods of instructional delivery must be radically different from that used with younger students in primary and secondary education.

Principle 10: Use concrete experiences to start each training module. Adult learners have to be convinced at the outset of training that it is worth their effort to participate. Ensuring their engagement in the training can be accomplished in a variety of ways. The Experiential Learning Model (ELM), which is used by several military training programs, recommends presentation of a concrete experience at the beginning of the training to "bring the adult student to the table," peak interest, and convince the student that the training may be useful.

Principle 11: Training should be scenario-based. To convince adult students that the training is relevant and important to their work, CT exercises should be presented within a familiar context.

Principle 12: Training should be made increasingly complex and difficult. To maintain the motivation of adult learners throughout the training, it is important to keep them challenged while at the same time avoiding unnecessary frustration. Instruction should thus begin at a level at which the student can successfully build and reinforce key foundational skills, then move on to refine and build upon those skills so they can be applied in increasingly more complex situations.

Principle 13: Self-awareness of one's own critical thinking should be a goal of training. Since CT is a System 2 process (see discussion of CT model), we are consciously aware when we are engaged in an "episode" where CT skills are utilized. Self-awareness of one's own CT is thus an important aspect of becoming a more effective critical thinker, and as such, it is a key step in helping further an Army officer's intellectual development.

Computer-Based Infrastructure is Necessary for Instructional Delivery

Principle 14: Training should be distance-based. The need for engaging stimuli, timed responses, interactive experiences, and real-time feedback necessitates the use of a computer-based infrastructure to deliver the instruction and archive the data. To maximize the distribution of training across students, a distance learning or web-based instruction using the public Internet is the most efficient option.

Principle 15: Multi-media is essential. The technological advantages of web-based instruction are soon lost if the method of delivery is nothing more than “electronic page-turning.” Since we rarely know which preferred learning style a student has, it makes the most sense to take a multi-media approach for all students, where the stimulus materials comprise a mix of text, graphics, tables, audio, video, animation, direct observation, and so forth.

Principle 16: Interactivity is essential. Requiring students to frequently interact with the training materials, rather than passively read information on a computer screen, will increase motivation, decrease drop out rates (a chronic problem with distance learning courses), and promote active learning.

Principle 17: Scoring increases motivation. One of the core tenets of adult education is that instruction will be effective to the extent that it promotes positive self-esteem. Adults respond well to, and are intrinsically motivated by, feedback about their performance. One of the most direct methods of feedback is to provide a score.

Principle 18: Use innovative training techniques. The Socratic method of guided questioning is a particularly effective pedagogical approach for CT. Unfortunately, web-based training cannot begin to duplicate this complex, highly personalized student-instructor encounter. Nevertheless, it is possible, with a variety of innovative training techniques, to lead students through a series of scripted exercises that require them to reveal their reasoning as they critically think through a problem.

(CT)^2 Purpose

The ultimate focus of (CT)^2 is to arm students with sufficient knowledge, skills, and attitudes in CT so they may avoid common C^2 errors. Consequently, the pedagogical approach is designed to address practical problems and deficiencies in CT that arise during the performance of real-world C^2 tasks.

Intended Population. (CT)^2 was originally designed for Army personnel in leadership positions, although even the most junior personnel can understand and benefit from the system. The training has been successfully applied and tested on ROTC students as well as on junior and mid-level officers. In order to make the training as meaningful as possible, realistic Army messages and situations are employed; therefore, a prerequisite to its use is that students should have some Army knowledge. They should be familiar with Army terminology, acronyms, and standard report formats. Students should, for example, be able to understand standard Army messages such as task organization tables, operation orders, mission statements, and battlefield sketches. Although $(CT)^2$ assumes students already have some domain knowledge, it gives them practice in applying strategies to help them use that knowledge more effectively.

Scope and Breadth. CT encompasses a wide array of component skills. Since it is neither practical nor desirable to teach *every* critical thinking skill that an Army officer might require, which ones should be the focus of training? To determine the program's training content, a systematic selection process was employed using a mix of literature review, content analysis, field surveys, and statistical analysis. The eight high-payoff CT skills are summarized in Table 1. The left-hand column gives the name of the skill, while the second column offers a working definition of the skill in practical terms. The third column lists the C^2 tasks most directly supported by the skill, with the right-most column specifying the errors and problems most frequently reported with that task. The skills are listed in no particular order.

Table 1. Overview of the Eight High-Payoff Critical Thinking Skills

Skill	Definition	Primary C^2 Task	C^2 Errors and Deficiencies
1. Frame the Message	The ability to identify the essential elements of a message, understand their relationships, and describe a high fidelity representation of the message.	Clarify the intent of the commanders 1 and 2 levels up	Difficulty in establishing clear and accurate understanding of commander's intent Difficulty in conveying clear commander's intent
2. Recognize Gist in Material	The ability to sort through the details in a message (written, graphical, visual, auditory, and/or tabular) and extract the gist therein.	Restate mission objectives provided by upper echelon to write own mission statement	Too much detail in operations orders (OPORDs) that must be filtered to establish gist that supports writing of own mission statement Too little time at lower echelons to accurately extract essence of mission
3 Develop an Explanation that Ties Information Elements Together in a Plausible Way	The ability to: ? Arrange evidence logically ? Highlight the gaps in knowledge. ? Develop an explanation or multiple explanations based on evidence ? Evaluate explanation(s) for plausibility	Interpret reports of recent enemy activities in area of interest to estimate enemy intent and predict enemy actions	Overlook seemingly unrelated facts Fail to assess the quality of information Difficulty in filtering excessive information Tendency to embellish enemy activity reports—over-reports of enemy contact and movement Tendency to discount initial reports
4. Generalize from Specific Instances to Broader Classes	The ability to recognize and then classify specific facts/incidents/events as part of a general category.	Interpret reports of enemy disposition	Fail to accurately induce patterns of overall movement based on report instances Tendency to disregard reports that do not match expectations Tendency to inflate information in reports
5. Use Mental Imagery	The ability to accurately create mental images in one's mind of how resources will be applied and events will unfold within a situation.	Develop scheme of maneuver War game courses of action (COAs)	Fail to visualize events Fail to include sufficient detail in COAs Fail to consider contingencies Fail to consider how plans could go wrong Generate only one COA Fail to consider combat multipliers Difficulty in keeping track of mobile forces
6. Challenge One's Bias	The ability to consistently reevaluate one's current view of the situation for prejudice or bias as new information is received.	Change own-unit plans based on new tactical input	Tendency to "fight the plan" General reluctance to change plans
7. Examine Other Peoples'	The ability to view and interpret a set of circumstances from the perspectives of different individuals, different cultures/religions, and differ-	Interpret reports of recent enemy activities in area of interest	Failure to accurately estimate enemy intent

Perspectives	ent timeframes (historical perspective).		
8. Decide When to Seek Information Based on its Value and Cost	The ability to evaluate the need for new information in terms of its cost in: time, resources, risk	Assess current situation	Tendency to spend too much time planning and gathering information Tendency to make quick decisions without gathering more information

Curriculum Elements

To facilitate flexibility of use, $(CT)^2$ is modularized into several components for the eight separate, but complementary, skills. The training modules and associated tests can be selected individually from a learning management system. Figure 2 shows that pretest, training module and posttest are available for each of the eight skills. In addition, an extended training module is available for Skills One, Two, Three and Four.

Figure 2. Course elements of $(CT)^2$.

Pretests. A pretest is provided for each skill for the purposes of 1) enabling student self-assessment of proficiency in application of the skill, 2) increasing student sensitivity to the errors and deficiencies associated with lack of or inappropriate skill application, 3) increasing student self awareness of CT capability, and 4) presenting feedback to initiate the process of learning skill application. While the pretest for each skill requires about the same amount of time (20 minutes or so), pretests differ somewhat in the format employed to meet pretest objectives. Brief descriptions of the pretests for Skill 2 (Frame the Message) and Skill 4 (Generalize from Specific Instances to Broader Classes) are provided here to illustrate the types of formats employed for the pretests.

The Skill 2 pretest is organized around three short passages that represent the kinds of messages typically encountered by an Army officer--such as intent statements or operational plans. Following each passage, the student is required to make four types of responses designed to assess ability to recognize the gist of a message: 1) writing a one sentence summary to capture the gist of the passage, 2) selecting, from a list, the points that should be included to capture the gist, 3) selecting from a list the statement that best captures the gist, and 4) ranking a set of statements on how closely each is related to the gist of the passage. After the test is completed, feedback is provided on the student's ability to capture the gist of a message.

The Skill 4 pretest is organized around three scenes or scenarios, each involving some type of stability and support operation. Two are described via photographs in which features and situations must be viewed for a limited time by the student and then classified by responding to questions. In the third, a scenario described in a situation report is classified by answering a set of questions. In each case, after the scenes or scenarios are removed from sight, the student is asked to 1) make a classification based on memory of the defining features that were present, 2) report whether certain critical features or elements were present, and 3) identify the critical features that might alter the classification if they were known by the student to be present or absent. After the test is completed, feedback is provided on the student's ability to generalize from specific instances to broader classes.

Training Modules. A training module is provided for each of the eight high-payoff critical thinking skills listed and defined earlier in Table 1. The modules were developed to meet Army training needs and to adhere to the four themes and 18 principles listed in Figure 1 and described and discussed in the previous section of this manual. Each module is designed to require about two hours to complete, on average; however, individual differences among students will probably result in a sizable distribution of times around this two-hour target. A description of each training module is provided in the next major section of the manual.

Posttests. A posttest for each training module is provided for use in assessing student comprehension of the critical thinking concepts presented in the module, and for evaluating student capability for applying the skills learned. Posttests vary in format from module to module, depending on what testing techniques were considered most effective in meeting the assessment objectives for the module. Training administrators may employ posttests to evaluate student progress and to identify training weaknesses. Students may employ posttests for self-evaluation of their progress and to identify weaknesses in their critical thinking knowledge and skills.

The eight posttests employ a variety of testing techniques. Multiple-choice questions are used extensively for testing knowledge acquisition and, combined with other techniques, for testing skill application. Many posttests require the application of critical thinking skills to operational scenarios and sample materials such as messages and photographs. Formats for student responses include classification, matching, ranking and rating by means of dragging-and-dropping items. Analytical responses are also made by selecting items from lists by checking boxes, selecting YES or NO, or selecting TRUE or FALSE.

Extended Training Modules. Extended training modules are provided for the skills considered most fundamental to critical thinking: Skill 1 (Frame the Message), Skill 2 (Recognize Gist in Material), Skill 3 (Develop an Explanation) and Skill 4 (Generalize from Specific Instances to Broader Classes). Each extended training module consists of an introduction and from two to five training elements. The organization of each extended module is illustrated here with a closer examination of the module for Skill 1, which consists of an introduction and five elements. The other extended training modules follow essentially the same approach and format.

The introductory tutorial provides answers to the following questions: What is framing? What is a frame? How does framing help you think? What are the benefits of framing a message? What does a message frame look like? What is the plan for training this skill?

In Element 1 the student learns about structured messages and their frames, and practices identifying correct frames for messages of different structures and matching parts of structured messages to parts of the frame.

Element 2 provides instruction and practice in analyzing unstructured messages. A categorization system is introduced and employed to categorize message components. Practice is then provided in applying the system to messages and, also, in identifying the relationships that exist among message components.

The objective of Element 3 is to teach the student how to identify weak spots in a message. The student is instructed in the six different ways that a message might contain weak spots, and given practice in identifying these weak spots in sample messages. The student is also instructed and given practice in distinguishing between message components that are weak because they are unclear and components that are weak because the information is uncertain.

Element 4 addresses the resolution of weak spots in a message. The student is instructed in and given practice in procedures that are effective in resolving message weaknesses due to lack of clarity and uncertainty. The student is also instructed and given practice in gauging the strength of inferences that may be required for resolving weak spots in messages.

The final element, Element 5, provides instruction in the difference between evidence and inference, and provides practice in evaluating inferences.

Flexible Use and Progression of Course Elements.

The time required to work through all of the training modules in the $(CT)^2$ program may exceed the training time available for a group of students. On average, it will take approximately two hours to work through each of the eight CT skills, for a total of 16 hours. If a unit has only 8 hours available for CT training, which 4 modules should be selected? If only 4 hours can be spared for training, which 2 should be selected?

As discussed below, we recommend two possible courses of action for selecting modules/skills for training when time is limited. On the one hand, one can examine the types of errors that each training module is designed to address, and compare those to known areas of weakness within one's operational or training unit. Alternatively, one can arrange the skills in order of difficulty, and concentrate on those CT skills that match the unit's present CT skill level.

Because the training materials within each skill were developed as standalone items, none of the training modules are prerequisites for any other. Consequently, either method is valid for selecting a subset of modules for conducting CT training. Below, we briefly describe how each method could be employed.

Recommended Use - Matching CT Skills to Errors.

To use this CT skill selection strategy, the analyst or trainer would first construct a list of the most troublesome or vexing errors/breakdowns in the C^2 domain observed in his/her unit. Then, depending on the time available for training, pick the CT skills whose targeted errors best match the ones that have been occurring in the field. Since the listing of C^2 errors in Table 1 is only partial, the analyst will have to use his/her judgment to find the best matches. In Table 2, we provide a set of decision rules that could be used to guide selection.

Recommended Progression – Matching CT Skills to Difficulty Level.

Another way to select the CT skills to be trained is based on estimated difficulty level. Table 3 categorizes the eight CT skills based on an estimated level of challenge. For simplicity, the skills are grouped into low, medium, or high difficulty level. Please note that these are only estimates, as empirical validating data concerning relative skill difficulty have yet to be collected. However, this listing provides a reasonable starting point for making a skill selection decision when training time is limited.

Software Technical Specifications.

$(CT)^2$ is a SCORM compliant web-based training package. The content structure is primarily

Table 2. Decision Rules to Guide Selection of Training Modules

If the unit is having problems understanding, translating, or writing commander's intent statements or following commander guidance	THEN	Use CT Skill 1: Frame the Message
If the unit is having trouble reading through complex OPORDs, is taking too long in writing their own mission statements, or not able to filter out irrelevant information from mission objectives	THEN	Use CT Skill 2: Recognize Gist in Material
If the unit is having trouble interpreting reports of enemy intent, does a poor job in assessing the quality of information, or is inconsistent in reporting enemy contact and movement	THEN	Use CT Skill 3: Development Explanation that Ties Information Together
If the unit is having trouble extracting patterns of movement based on individual reports or is not able to interpret reports of projected movements accurately	THEN	Use CT Skill 4: Generalize from Specific Instances to Broader Classes
If the unit is failing to consider contingencies, include sufficient detail in COAs, or anticipate how time and place affect how the world looks	THEN	Use CT Skill 5: Use Mental Imagery
If the unit is showing a general reluctance to change a failing plan (i.e., fighting the plan)	THEN	Use CT Skill 6: Challenge One's Bias
If the unit is failing to accurately estimate enemy intent or is unable to appreciate the other side's point of view	THEN	Use CT Skill 7: Examine other People's Perspectives
If the unit is spending too much time planning and gathering information and is unable to make quick decisions without the need to gather more information, or if they tend to make quick decisions without enough information	THEN	Use CT Skill 8: Decide When to Seek Information Based on its Value and Cost

made up of HTML pages. However, other media types used in the training include:

? Javascript files (.js)
? CSS Style files (.css)
? Adobe Flash files (.swf)
? image files (.jpg and .gif)
? audio files, (.wav and .mp3)
? video files (.mpeg)
? PDF files (.pdf)

All programming, SCORM and otherwise, is handled on the client side via Javascript. No server side programming (e.g. ASP, PHP, CGI) is used in the training to ensure compatibility with a wide variety of server configurations. All database capabilities of $(CT)^2$ are handled through SCORM; therefore, no specific database vendor is required for its use.

At the SCORM level, $(CT)^2$ comprises eight *skill* directories (e.g. Skill One, Skill Two), as well as directories containing global reference files that are used by all Sharable Content Objects (SCOs) during runtime. Each skill directory is in turn broken down into individual SCOs, housed in separate folders. All SCORM manifest files are included in the root level of the SCOs. As a rule, initiation of each SCO is handled by the *index* page of its directory.

Consistent with SCORM compliancy standards, the training package is designed to run on an LMS. $(CT)^2$ is currently being tested on *Moodle*, which is an open source LMS. However, any 2004 SCORM compliant LMS should be compatible with each of the training and assessment components of $(CT)^2$.

All administrative features of the training can be accessed through an LMS. The following features can be tracked through the LMS database:

? SCO initiation
? SCO page to page movement
? SCO test scores
? SCO time on test

Development Process

i. Tech Specs (our outlines)
ii. Storyboarding
iii. Panel Review
iv. Usability studies
    1. Los Alamitos
    2. Ft. Riley
    3. Ft. Lewis
    4. Ft. Hood

Evaluation Investigations

To evaluate whether (CT)2 increases CT skill, two investigations were conducted. The first investigation was designed to assess the learning achieved through one training module compared to a control group. This study was conducted at an umbrella week provided by Fort Riley, Kansas. The second investigation assessed whether the extended version of the training provides additional learning over and beyond the gains in skill achieved through the standard training module. This study was conducted with volunteer ROTC students from the University of California, Santa Barbara.

Experiment 1

The central objective of this study was to determine the learning effect, if any, of training provided by $(CT)^2$ on a particular CT skill. To meet this objective, a comparison was made between Soldiers that received $(CT)^2$ training and a control group that did not receive $(CT)^2$ training. The training component of one CT skill, Skill 8: *Decide When to Seek Information Based on Its Value and Cost*, was used in the study. The central hypothesis tested was that participants who complete $(CT)^2$ training on one particular CT skill will display superior performance on that skill. Post and pre training assessments were taken on both the experimental and control groups to control for potential variation in pre-training ability.

Method

Participants. Twenty-one male Soldiers participated in the evaluation. Their ranks included 1st Lieutenant (n = 5) , 2nd Lieutenant (n = 4), Captain (n = 11), and Major (n = 1). Eleven participants were randomly assigned to the no-training condition and 10 were randomly assigned to the training condition. Participants were originally tasked to attend the session as part of umbrella week activities at Fort Riley, Kansas. However, upon arrival participants were informed that their participation was voluntary. All participants provided informed consent.

Materials. The materials used in this study were a subset of the (CT)2 training materials. Specifically, the materials were taken from the training module for Skill 8: *Decide When to Seek Information Based on Its Value and Cost*. Participants were first given a pretest to assess their ability on this skill, which was developed specifically for this investigation. This pretest was essentially a parallel version of the post-test component of the module. Participants assigned to the experimental condition also completed the training module for Skill 8. Finally, participants were given the post-test for the module, assessing the participants' ability to decide when to seek information based on its value and cost. A brief demographic questionnaire was also used.

The training items were programmed in an html to allow for the intended web-based delivery. The materials were run via an Internet Explorer web-browser from a laptop computer.

Procedure. Upon arrival, participants were greeted, and were asked to select a card that randomly assigned them to the training or no-training group. Then, they were given a brief overview of the project, its purpose, a description of the data collection procedures, and a discussion of how the data would be used. Next, the participants read and completed the consent form and a demographic questionnaire. After all questions were answered, all 21 participants completed the Skill 8 pretest. After completing the pretest, participants assigned to the no-training group took a 15-minute break, and then completed the Skill 8 Posttest. After completing the Skill 8 Pretest, the experimental group completed the Skill training module, followed by the Skill 8 Posttest. They then completed a posttest questionnaire soliciting their opinions about the training. The experimental group was allowed to take breaks, as required, while conducting the training.

Results

Participants' posttest scores were subjected to a One-Way Analysis of Covariance (ANCOVA) that examined differences in post-test scores between the training and no-training groups, using participants' pretest scores as a covariate to control for baseline skill. The training group performed significantly better than the no training group on the posttest $F(1, 18) = 10.121$, $p < .01$, = 59.922. The pretest and posttest means (unadjusted and adjusted for pretest) are listed in Table 2.

Table 2. Pretest and posttest means and standard deviations by group

Group	Pretest Mean (SD)	Posttest Mean (SD) (unadjusted)	Posttest Mean (SD) (adjusted)
Training (N=10)	50.5 (11.2)	64.0 (8.4)	61.4 (8.4)
No Training (N=11)	42.3 (13.8)	47.7 (12.3)	50.1 (12.3)

Discussion

The results indicate that completing the Skill 8 training module increased a participant's skill at deciding when to seek information based on its value and cost. These findings indicate that $(CT)^2$ is effective at increasing learning on this particular skill. Because the training and testing components of each of the eight skills follow a similar pedagogy and presentation format, one might expect that each training module would produce a similar learning effect. However, further research would be necessary to test that assumption.

Experiment 2: Extended versus Standard Training Module

The central objective of the second evaluation investigation was to determine if an extended training module provides any greater learning than the standard training for a particular CT skill. The extended version, of course, requires a greater time commitment than does the standard version. Hence, the study sought to determine if the greater time cost was worth the investment, i.e., paid off in greater learning. This investigation was accomplished with the assistance of the University of California-Santa Barbara (UCSB), Military Sciences Department ROTC Surfrider Battalion. Cadets and cadre from the ROTC department volunteered to participate in the study. The central hypothesis tested was that participants who complete an extended training version of a particular CT skill will display superior performance on that skill compared to participants who complete the standard training module. The study also evaluated participant's subjective preferences between the two versions of the training.

Method

Participants. Twenty-one participants completed the experiment. Eight completed the extended version of the training and 13 completed the standard version. Of these participants, 5 were cadre (instructors) and 16 were cadets (students). Eighteen males and 3 females comprised the participant sample. See table ? for a breakdown of participants by training condition. All participants provided informed consent. For each participant who completed the standard version of the training, a contribution of $25 was made to the UCSB ROTC Surfrider Battalion's Morale and Welfare fund. For each participant who completed the extended version of the training, $75 was contributed to the fund.

Facilities and Equipment. The experiment was conducted at the university's Military Science computer laboratory. Six laboratory computers and 3 laptops were used to display the training and other materials in the experiment.

Training Materials. Four components of the Skill 4 (Generalize from Specific Instances to Broader Classes) training module were used as materials in this study. The pretest component of Skill 4 was used to assess the participants' pre-training ability regarding this CT skill. Th extended training version and the standard training version of the Skill 4 module were also used. Following completion of the either the extended or standard training version, participants were also asked to complete the posttest component of Skill 4 as a post training assessment of their skill level. The training items were programmed in an html to allow for the intended web-based delivery. The materials were run via an Internet Explorer web-browser from a laptop computer.

Procedure. Prior to the investigation, visits were made to the UCSB Military Science department to brief the cadre and cadet participants on the purpose of the study. At that point, interested parties were invited to sign up to participate in the study. Volunteers were randomly assigned to either the standard training condition or the extended training condition. Participants were then assigned time slots to complete their assessment and training obligations.

Upon arrival, participants were greeted and seated at a computer terminal. Participants were then given a brief overview of the project, data collection and use of the data. Then, they were given a brief overview of the project, its purpose, a description of the data collection procedures, and a discussion of how the data would be used. After preliminary questions were answered, participants completed the Skill 4 pretest. After completing the pretest, participants randomly assigned to the extended training condition completed their training, followed by the Skill 4 posttest and a post training questionnaire. The participants who had been assigned to the standard training version followed the same procedure, except they completed the standard training module. After completing the training and questionnaire, both training groups received a short email inquiry once a day for the next 10 work days, asking them to note instances of how the training concepts they learned applied to their everyday affairs. The participants' responses were submitted via email.

Results

Participants' posttest scores were subjected to a One-Way Analysis of Covariance (ANCOVA) that examined differences in post-test scores between the extended and standard training groups, using participants' pretest scores as a covariate to control for baseline skill. There was no significant difference found between the post test performances of the two groups $F(1, 18) = 000$, $p > .05$. The pretest and posttest means (unadjusted and adjusted for pretest) are listed in Table 3. The two groups' averages were almost identical indicating that the extended version did not produce a learning advantage over the 2-hour version as measure by Skill 4 posttest performance.

Group	Pretest Mean (SD)	Posttest Mean (SD) (unadjusted)	Posttest Mean (SD) (adjusted)
Extended (N=8)	120.2 (12.3)	174.5 (34.6)	174.4 (34.6)
Standard (N=13)	120.2 (10.3)	174.5 (29.1)	174.5 (29.1)

Table 3. Pretest and posttest means by group

Participants' subjective ratings of the overall training were analyzed by means of an ANOVA comparing the average rating of the extended version to the average rating of the standard version. The ANOVA showed no significant difference in the average ratings of the two groups (M = 1.84 for standard version; M = 1.75 for extended version. $F(1, 18) = .168$, $p > .05$. All of the participants, except one in the 2-hour version (who rated the training as 3–neutral), rated the training as either very good (1) or good (2), reflecting positively on the training but no differences between the two versions.

Discussion

The results suggest that the standard training was as effective as the extended training in teaching the skill of generalizing from specific instances to broader classes. Given the limited time often available for extra training these results suggest that the standard version is sufficient to produce a desired learning effect. The extended version, then, could be used if a student wanted greater explanation or needed some question resolved. The extended version could serve, then as supplemental material, which could be made available for those trainees who desire additional training.

Applications (VAS)

Future Research

The completion of this seven-year research program should be only the start to empirical research on CT. It brings to the field two critical sources of power, heretofore unavailable to researchers. The first is a testable model of CT that can guide future research efforts on the construct. The second is $(CT)^2$, which provides an off-the-shelf training package and a model on which future training development efforts can be launched. We see many possibilities for the future of CT research. Below, we discuss what we believe to be the most interesting potential topics of future research.

One of the requirements imposed on $(CT)^2$ was to make it maximally *accessible* by Soldiers anywhere in the world. Hence, early in its design, the design decision was made that it would be delivered via the Internet. $(CT)^2$ was to be a web-based program that would provide distance training, perhaps stand alone, on CT skills important to Army concerns. Informational content is truly unlimited with web-based delivery, and computers possess an exceptional capability to create opportunities for practice. However, distance training, and its web-based instantiation, impose constraints on the content and training experiences that can be delivered to students. We have found that the greatest limitation of current web-based training is that it provides only limited capability to deliver feedback to students' responses. For a skill like critical thinking, limitation on the quality and amount of feedback might severely hamper learning. While several evaluation studies have shown that $(CT)^2$ does produce learning and increases CT skill, it is likely that feedback tailored to students' thinking would be more effective. The classic pedagogy that appears to be maximally effective at generating clear thinking is the Socratic method, in which a tutor points out the frailties of a student's thinking by asking repeated and pointed questions. The technology to emulate the Socratic method on a computer is available in the form of intelligent tutors. In recent years, largely due to research efforts funded by the United States military, intelligent tutoring technology has reached impressive capability. Therefore, it is theoretically possible to create computer-based CT training that might meet this highest standard of feedback. Future research should explore the possibility of creating a truly Socratic training system designed to increase CT skills.

$(CT)^2$ rests on the common assumption that increasing self-awareness of one's own CT will recursively improve the same. Indeed, the most widely used textbooks on CT solely use exercises that increase self-awareness of one's thinking. Yet, the effectiveness of increasing self-awareness on improving thinking skills has not been empirically shown. More research is needed to further investigate the relationship between meta-cognition concerning CT and the development of the same.

$(CT)^2$ was developed to improve CT within individuals. Yet, an emergent property of team behavior may be its composite ability to critically think. It is not clear whether the most effective teams are composed of members who are all good critical thinkers, or have only one or two members with this skill. The effect of the distribution of CT skills across a team likely interacts with communication capability, level of cooperation and other team characteristics. Team composition and the distribution of critical thinking skills across the team is a new, potentially important, area of study that should be pursued.

The relationship between intelligence, ethics and CT has been pursued since Watson and Glaser (1980) first conceptualized CT in the early 1940s. Having progressive political attitudes, these early researchers hypothesized that liberal viewpoints were related to high levels of CT. Since then, others with conservative political perspectives have also posited that conservatism is an indicator of CT. While the bias is obvious in these positions, the fact remains that we know little about the relationships among intelligence, ethics, and CT. We can create good theories of CT only by knowing, via empirical demonstration, what it is not. The construct validity of CT should be pursued in future research, or the field will continue to follow the fragmented path that has been its course since the 1940s.

References

Fallesen, J.J., Michel, R.R., Lussier, J.W., & Pounds. J. (1996). *Practical thinking: innovation in battle command instruction.* Army Research Institute Technical Report 1037.

Fischer, S.C., & Spiker, V.A. (2000). *Application of a theory of critical thinking to army command and control*. Santa Barbara, CA: Anacapa Sciences.

Fischer, S.C., & Spiker, V.A., Berkman, A. A. (2001). *Functional requirements for web-based training of critical thinking*. Santa Barbara, CA: Anacapa Sciences.

Fischer, S.C. & Spiker, V.A. (2004a, May). *Critical thinking training for Army officers*. *Vol II: A model of critical thinking* (Phase II Final Report). Alexandria, VA: US Army Research for the Behavioral and Social Sciences.

Fischer, S.C. & Spiker, V.A. (2004b, June). *Critical thinking training for Army officers*. *Vol III: A web-based training program to develop the critical thinking skills of Army officers.* (Phase II Final Report). Alexandria, VA: US Army Research for the Behavioral and Social Sciences.

Spiker, V. A. (2004). *Core Concepts in ARI's Critical Thinking Training*. Paper presented at Critical Thinking Workshop, November 2004.

Watson, G., & Glaser, E.M. (1980). *Critical Thinking Appraisal.* New York: Psychological Corporation.

APPENDIX A:

A MODEL OF CRITICAL THINKING

Critical thinking was first conceived in the early 1940's by two psychologists, Goodwin Watson and Edward Glaser. Watson and Glaser also developed the first test of CT, the Watson-Glaser Critical Thinking Appraisal (WGCTA) (Watson & Glaser, 1980), which is still widely used. Since then, almost all of the thinking and theoretical development of CT has been conducted by educators and philosophers. It is unclear why psychologists have played a small role in this work; however, see Halpern (1996) and Baron & Sternberg (1986) for notable exceptions. Because most of the work on CT has been conducted by educators and philosophers, the construct has not endured the kind of empirical inspection typically bestowed upon constructs developed by psychologists. Its relationship to other, well-established psychological constructs such as IQ, working memory, and reasoning, for example, has rarely been studied. It is the authors' admittedly subjective opinion that the lack of empirical study of CT and its relationship to other individual difference dimensions has produced a fractionated view of the construct. Without the grounding of data, theorists have been free to postulate divergent concepts of CT. An effort in philosophy to reach a consensus definition of CT in 1993 had little effect on unifying the field.

To fill this gap, Fischer and Spiker (2004a) developed a model[3] of CT that is sufficiently specified to permit empirical testing and falsification. The model identifies CT's role within the related fields of reasoning and judgment, which have been empirically studied since the 1950s and are better understood. It incorporates many ideas about CT offered by leading thinkers (e.g., Paul & Elder, 2001) in philosophy and education. It also embodies many of the relevant variables discussed in the CT literature (e.g., predisposing attitudes, experience, knowledge, and skills) and specifies the relationships among them. The model can, and has been, used to make testable predictions about the factors that influence CT and about CT's psychological consequences. It also offers practical guidance to the development of CT training. In this manual we offer an overview of the model's main features. But first, it is necessary to briefly review current thinking about reasoning and judgment, on which the model is based.

Dual System Theory of Reasoning and Judgment

Prior to the early 1970's, the dominant theory stated that people made judgments by calculating (1) the probability and (2) the utility of competing options. Although this *rational choice* model took on a variety of forms, all versions posited a rational actor who made calculations of probability and/or utility, and selected the option that had the highest value. In the 1950's, however, researchers began to notice that the model failed to predict *actual* behavior (Meehl, 1954; Simon, 1957). Evidence that falsified the rational choice theory accumulated over the following decade.

[3] The model described in this manual is a revised version of a previous model discussed in earlier reports (Fischer & Spiker, 2000; Fischer, Spiker, & Berkman, 2001).

In the early 1970s, an alternative theory proposed that people use heuristics, as opposed to the rational weighing of relevant factors, to make judgments. The "new" theory was, and continues to be, supported by empirical study. The heuristic theory states that many judgments are based on intuition or rules of thumb. It does not propose that *all* judgments are made intuitively, just that there is a tendency to use such processes to make many judgments. The most recent versions of heuristic theory, in fact, propose that *two* cognitive systems are used to make judgments (Kahneman, 2003). The first system, *intuition*, is a quick, automatic, implicit process that uses associational strengths to arrive at solutions. The other, *reasoning*, is effortful, conscious, and deliberately controlled. Since the 1970's, multiple and similar two-process theories have been proposed to explain judgment. To accommodate the multiple theories, many researchers now refer to the implicit associational type of process as *System 1*, and the conscious deliberate process, *as System 2*. The following example shows how these two processes may lead to different judgments.

> *Suppose a bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?*

Most people's immediate judgment is that the ball costs 10 cents. This is a response derived from *intuition or System 1*, which again, is quick, automatic, and relies on associations. The strong mathematical association between $1.10, $1, and 10 cents leads to this quick, *but wrong*, judgment. The ball can't cost 10 cents because then the bat would have to be $1, which would make it only 90 cents more than the ball. The more effortful deliberately controlled *reasoning, or System 2*, process usually produces a different, and correct, answer. When people spend the time and effort to think about the problem, they usually realize the ball must cost 5 cents and the bat must cost $1.05. Hence, in this example, the two systems produce different judgments. It would be a mistake to conclude that System 1 *always* produces different judgments than System 2, however. Nor does System 1 always produce an incorrect answer, nor one that is poorer than one produced by System 2.

In fact, researchers have shown that expert performance in any field, which is commonly the gold standard, is often driven by intuition derived from extensive experience (e.g., Klein, 1999). That said, expert performance is not without fault, and studies have shown that even experts make errors in judgment when well-learned associations lead them astray. The associational processes used in System 1 that make expert performance so quick and powerful are the same processes that are responsible for systematic errors that experts sometimes make. Additional weaknesses of System 1 are that it depends on the quality and amount of experience an individual possesses, and it can't be used effectively in novel situations. System 2 reasoning also has its strengths and weaknesses. While it is highly useful in novel situations and problems, it is also slow and effortful. It usually cannot be utilized concurrently with other tasks and, like System 1, it can also produce wrong judgments.

Most researchers posit that judgments can be made using either system. Some models posit that System 1 is the preferred system because it demands fewer resources and is less effortful to apply. Most recent theories, however, believe that Systems 1 and 2 *run in parallel* and *work together*, capitalizing on each other's strengths and compensating for weaknesses. For example, many researchers believe that one function of the controlled deliberate process is to monitor the products of the automatic process. System 2 is thought to endorse, make adjustments to,

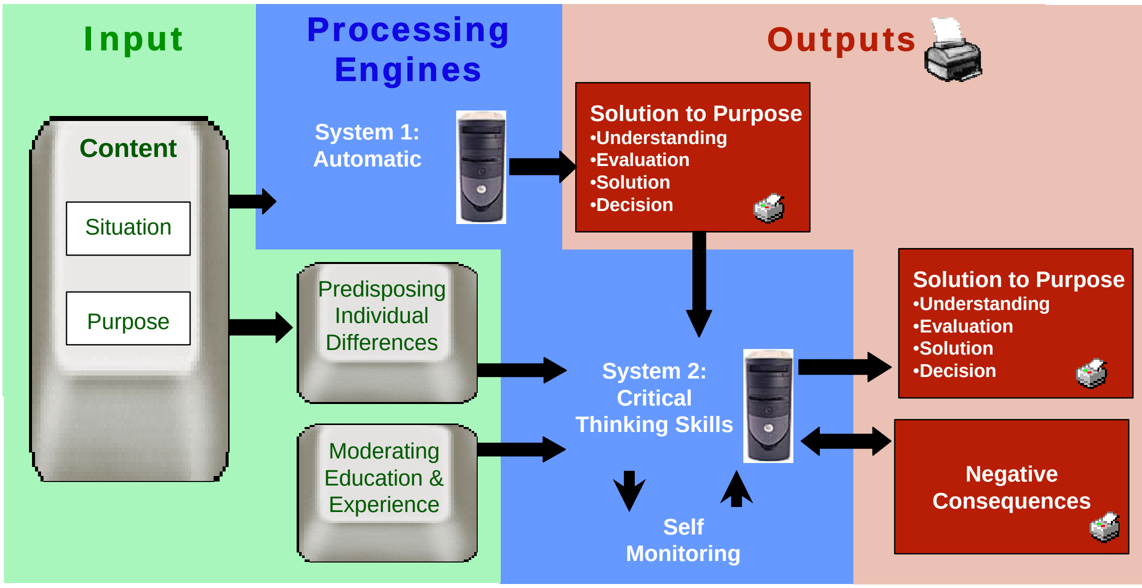

Figure 1. Model of critical thinking.

correct, or block the judgment of System 1. However, if no intuitive response is accessible, System 2 may be the primary processing system used to arrive at judgment. Sloman (2002) states that the systems work hand in hand as "two experts who are working cooperatively to compute sensible answers."

The similarities between descriptions of CT and System 2 are striking. The words "effortful, controlled, deliberate, purposeful, and conscious" are frequently used to describe both. In the next section we will see that Fischer and Spiker's model proposes that CT's primary function is to monitor and control the judgments produced by System 1 association reasoning. It further proposes that System 2 is the engine that powers CT skills. It follows that errors in reasoning can occur when CT has failed to serve its monitoring and correcting function.

Overview of CT Model

System 1 and 2 Engines. As shown in Figure 1, the CT model assumes that CT skills are executed by System 2.[4] CT skills serve to monitor, evaluate, and control the judgments produced by the System 1 associational process. Hence, Figure 1 shows that System 1 judgments provide input to CT skills. The two processes are thought to *run in parallel* and *interact* to produce judgments. Because System 1 is truly an automatic and uncontrolled process it cannot be initiated or stopped. For this reason, CT monitors only the *products*, and not the *process*, of System 1. Because System 1 is quick, it often comes to judgment before System 2. However, CT skills executed by System 2 may override, or confirm, that judgment. Therefore, System 2 has the potential for controlling judgment, although it may not always utilize that potential.

CT skills can provide *thorough examination* of the problem at hand. While System 1 typically derives only one solution

[4] We do *not* claim that all processing executed in System 2 involves CT skills, but that all critical thinking requires the kind of analytic, rational processing that only System 2 can provide. For example, System 2 may drive other deliberate processes such as controlled visualization. However, controlled visualization is not considered a CT skill.

(Klein, 1999), CT skills can provide multiple potential solutions. System 1 works to narrow possible action paths, which is often highly effective when the task must be accomplished quickly and when the problem space is limited. However, when the problem space is novel or complex or when solutions must be innovative, CT skills are more powerful. CT skills have the meta-cognitive capability to monitor the progress of their own processing. In this sense, CT is *recursive*, as represented by the curved arrow leading out and back into the "System 2: CIT Skills" processor in Figure 1.

Figure 1 also shows how the processing engines of CT and System 1 interact with environmental and individual factors. Both systems receive initial input from the environment in the form of information about a situation or problem that requires judgment. Part of that input is a meta-task that defines the general *purpose* of judgment. The other part of the input is information about the situation. System 1 immediately and automatically begins processing of the input by searching through its associational network for potential solutions that will satisfy the purpose. CT skills, motored by the System 2 processing engine, receive the same input, filtered through predisposing individual difference factors, which are discussed in greater detail below. CT skills may, or may not, also begin processing at that time. If CT skills are engaged, they will begin to evaluate solutions offered by System 1 or they will apply deliberate reasoning to the problem.

Whether or not CT skills are utilized depends on a variety of factors, including individual predisposition and situational variables. The sum value of these factors provide the impetus to engage in effortful CT skills, but that motivation must exceed some threshold value for CT to be initiated. Let's examine each component of the model in more detail.

The Context: Situation and Meta-Task. As noted above, the opportunities for judgment are set in motion by contextual factors, which include (1) the situation and (2) purpose (meta-tasks). While the automatic System 1 will engage in all conditions, two characteristics of the situation *must* be present to elicit CT activity. First, stimulus material *must* contain substantive information. Second, there must be sufficient time available to engage in an appropriate CT skill. Other characteristics of the situation make it more likely that CT will be initiated. For example, the presence of conflicting information, disordered or unorganized material, uncertain information, and complex material all make it more likely that CT will be engaged.

CT is not an end in itself, but serves other objectives specified by purpose (meta-tasks). The purpose also dictates the specific response that will be required to successfully end the execution of a CT skill. For example, the situation may include a meta-task to *understand, make an evaluation, make a decision, or solve a complex problem*. Whether final judgment is based on CT or System 1 processing, System 2 determines when the requirements of the purpose have been met. Hence, successful completion of the meta-tasks as determined by System 2 can also provide input that terminates a CT episode, as depicted in the bidirectional arrow in Figure 1.

Predisposing Factors. Predisposing factors influence the likelihood of a person using, or persisting in using, a CT skill. Like features of the situation, they serve as input conditions to CT skill, and as a filter through which the situation and purpose are evaluated. Some may be key factors that strongly affect an individual's use of a CT skill. Other factors may have a weaker relationship to CT, perhaps increasing the likelihood of engaging in CT by a marginal amount. In summary, predispositions are *measurable* ways in which people differ, whether *fixed* or *modifiable*, that influence the *use* or *persistence of use* of a CT skill.

Moderating Variables, Education, and experience Moderating variables are individual variables that influence how, and how well, CT skills are performed. For example, domain expertise, recent experience, and education influence the quality of the reasoning produced by CT. They do not, however, influence whether one executes a CT skill, as do predisposing factors.

Negative Experiential Consequences. There is a general consensus in the literature that individuals are reluctant to engage in CT. This is based on widespread observation of incoherent reasoning, nonsensical beliefs, lack of respect for evidence, poor CT test scores, and unsupported decision-making in various American populations. Indeed, much of the CT literature is devoted to a movement to increase the application of CT in various populations. One of the central topics has been the question of why the public seems disinclined to use CT. Some theorists posit that individual characteristics, such as intellectual laziness, arrogance and cowardice (which are represented in the model as predisposing individual differences), are the reasons why CT is avoided. The model of CT discussed here, however, posits that negative affective consequences associated with the application of CT are the primary inhibitory sources.

The model posits that individuals who engage in CT for any substantive length of time are likely to experience negative affective reactions. For example, CT can produce mental fatigue, increased effort, increased anxiety, cognitive dissonance, and decreased self-esteem. Negative affect experienced during a CT episode might be countered by positive affect that is the result of a positive outcome (e.g., solving a difficult problem) that, in turn, is a direct result of CT. Therefore, the application of CT may be positively rewarded and hence, increases in its use may be realized. Some individuals, then, may not experience negative affect associated with CT. But at the very least, by definition, CT requires more effort than System 1 processing, and is therefore a less desirable means to achieve judgment in that limited sense.

Figure 1 shows that negative experiential consequences serve as both a byproduct of CT and as input to the decision to maintain a CT episode, as depicted by the bidirectional arrow in Figure 1. When the affective consequences of applying a CT skill become too negative, the motivation to maintain the episode is decreased. If the negative consequences are sufficiently strong, they may result in a cessation of the episode.

System 1 and 2 Products. As shown in Figure 1, Systems 1 and 2 have the capacity to generate products in the form of judgments. The double arrow leading to and away from System 2 solutions indicates that this deliberate process is capable of evaluating its own products. In contrast, the single arrow leading from System 1 to its judgment product shows that this associational process does not have that same metacognitive capability. Note also that the System 1 judgments also serve as input to System 2 monitoring and evaluation roles.

The Quality of Judgment. The *task* posed by a particular situation should not be confused with the system that is used to solve it. For example, one may have the task of understanding a commander's intent statement that could be achieved using associational processes of System 1 or controlled CT skills powered by System 2. Therefore, an individual who is trying to understand an intent statement may or may not be using a CT skill to do so. Even more important, the application of CT skills driven by System 2 *does not always* produce the best solution to a task. It would be a mistake to encourage consistent use of CT because that strategy would deny the power and effectiveness of System 1. Similarly, it is not advisable to only develop associational processes because controlled deliberate reasoning can (1) produce superior solutions, and (2) provide necessary checks on the products of System 1. Moreover, the issue of which system is most effective is practically irrelevant because most theorists believe that both are almost always used in conjunction to produce a solution. Hence, the real issue that determines the quality of a solution is probably *how well the two systems interact*.

The quality of a solution produced by the application of a CT skill may also be affected by how well the skill is executed. Decrements in performance may be produced by *failing to apply* a component of the CT skill (e.g., failing to clarify ambiguous information in a message or failing to consider alternative explanations for a pattern of data), *failing to accurately perform* a component of the skill, or by *lacking sufficient knowledge* that can be processed by the CT skill. Therefore, one could apply a CT skill and still produce inferior solutions to a task. Moreover, it is not possible to determine whether System 1 or System 2 was applied to derive a solution based on the solution alone. The quality of a solution may also be affected by moderating variables such as educational level and experience. These issues are important to the design of training that seeks to improve CT skills. As discussed in the following section, the model of CT presented here was used to identify a pedagogical approach and principles of training CT. This approach was then used to design $(CT)^2$.