An Investigation of the Effect of Task Design on the Development of Critical Thinking Skills by Engineering Students

Geneviève Légaré

A Thesis

in

The Department

of

Education

Presented in Partial Fulfillment of the Requirements
For the Degree of doctor of Philosophy (Educational Technology) at
Concordia University
Montréal, Québec, Canada

March 2002

0-612-68199-8

Signature page

Abstract

An Investigation of the Effect of Task Design on the Development of Critical Thinking Skills by Engineering Students.

Geneviève Légaré, Ph.D.
Concordia University, 2002

There is much literature on the theory, the practice and the research on critical thinking skills. Several theoretical definitions have been proposed for these skills which are also referred to as "higher order thinking skills". Although Bloom's taxonomy of cognitive skills still serves as the groundwork for the study of critical thinking skills, categories such as "analysis", "synthesis" and "evaluation" have been refined over the years by subsequent scholars and researchers. Hence, more current models tend to include aspects such as metacognitive skills, handling multiple perspectives, decision making skills and critical thinking dispositions.

The rationale for teaching critical thinking skills at all levels of the educational process is not questioned: it is generally assumed that these skills are desirable for any students to function adequately in society. The challenge, however, becomes one of teaching methods. It is assumed that constructivist approaches provide more opportunities to engage critical thinking skills than regular teacher-centered methods because they usually offer novel problems or perplexing situations.

However, research on teaching critical thinking skills, especially in higher education, remains sparse. Most studies, exploratory in nature, are conducted with very small groups of participants, mainly in social science domains. Content of online

collaborative discussions, usually facilitated by an instructor who uses a scaffolding strategy, tends to be the preferred source of data for research on critical thinking skills.

This study investigated the effect of task design on the development of critical thinking skills by engineering students. It was hypothesized that the introduction of complexity in instructional activities would increase the incidence of critical thinking skills. Specifically, a "complexity" variable was incorporated in the instructional tasks carried out individually by the participants. Participants (N = 65) were fourth year engineering students registered in the course "Impacts of Technology on Society" (Engr 492) at Concordia University.

A quasi-experimental time-series design was used to carry out this study, which was conducted over a period of six weeks. Five instructional tasks constituted the treatments; students' written productions provided the main source of data. Individual feedback and instructor's interventions during the data collection were controlled. Students were debriefed upon completion of Task 4. Performance on the first two tasks provided the baseline for the pre-treatment level of students' critical thinking skills. The baseline of the treatment group was compared with that of the non-treatment group to establish group equivalence.

A modified version of Herrington and Oliver's synthesis of critical thinking skills instrument was used to analyze the data. A trend analysis confirmed that the overall treatment had a positive effect on the incidence of observed critical thinking skills. Results indicated that the first level of the complexity variable embedded in the case study (Task 3) triggered a higher incidence of critical thinking, as well as a different set of critical thinking skills. On the second level of complexity, students in Tasks 4 and 5

showed fewer instances of critical thinking skills than in the earlier Tasks. An exploration of the outcome on those tasks suggests that simulation, as an instructional approach, provides too much complexity for individual learners, and so does not trigger a higher incidence of observable critical thinking skills. The results are discussed both in terms of implications for instructional design and for conducting research about critical thinking.

Pour Catherine et Wilfrid, Éva et Alfred

"*Et il leva sur moi un regard de fin de carrière*" --Monsieur Malaussène
Daniel Pennac

Entreprendre des études de troisième cycle est une entreprise dont je ne soupçonnais l'ampleur des étapes à franchir et le niveau de détail à régler avant de parvenir à la ligne d'arrivée. Cette expédition intellectuelle a entraîné dans son sillage de nombreuses personnes. Le soutien de ces personnes tout au long de ce parcours, m'a été des plus précieux. Je leur exprime ma plus grande gratitude.

J'aimerais d'abord remercier Réal Leblanc pour son support inconditionnel, sa compréhension, sa générosité et les sacrifices qu'il a dû faire afin que j'accomplisse ce projet. Sans lui, mon parcours aurait été plus ardu encore.

Je remercie aussi ma famille et mes amis que je ne peux tous nommer ici sans en oublier. Leur support moral et leur encouragement m'ont été des plus précieux.

Ce projet de recherche n'aurait pas été possible sans l'intervention de Mme Bernice Goldsmith. Grâce à elle, j'ai pu réaliser l'étude auprès des étudiants de génie. De plus, c'est grâce à elle que j'ai fait mes premières armes en tant que chargée de cours et assistante à la recherche. Sa détermination, son énergie, son dynamisme et ses idées pour de nouveaux projets sont des sources d'inspiration pour moi.

Je remercie le Dr. Michael Nisbett et le Dr. Roger Balk pour avoir accepté de participer à cette étude. Leur ouverture d'esprit, leur disponibilité à accepter les modifications suggérées et leur tolérance face à mes demandes ont fait d'eux de précieux collaborateurs. Je remercie aussi M. Ray Charron pour m'avoir donné accès à ses étudiants.

Je remercie chaleureusement Mme Anne Browne pour son support, sa compréhension, ses conseils et son amitié qui m'ont été des plus précieux tout au long de mes études supérieures.

De la même manière, je remercie Stef Rucco pour son amitié, son empathie et son souci de faciliter le travail de tous les membres de département, incluant les étudiants. Plusieurs étapes de mon travail ont été rendues plus faciles grâce à ses interventions. Je remercie aussi Denise Brennan pour son support et sa diligence à régler des problèmes techniques.

L'amitié de Laura McEwen ainsi que sa disponibilité m'ont été des plus salutaires. Ses questions, toujours à point nommé, m'ont permis d'éviter plusieurs écueils associés aux étapes du projet. Je l'en remercie.

De la même manière, je remercie Christina Dehler pour son amitié et pour ses généreux conseils par rapport au déroulement du projet de recherche. Sa clairvoyance, son sens de l'organisation, sa confiance et sa sagesse sont toutes des qualités qui m'ont permis de franchir les étapes du doctorat de façon plus sereine.

Je salue l'intégrité et la franchise de Julia Petrov qui m'ont permis de tester sans craintes la pertinence de certaines idées. Son amitié m'a été des plus précieuses au fil des ans.

Je remercie aussi Richard Labonté, Martine Pagé, Jean-François Henri et Vivek Venkatesh pour leur aide apporté à propos des modèles statistiques à utiliser.

Je remercie mon superviseur, le Dr. Dennis Dicks et les membres de mon comité, le Dr. Gary Boyd et le Dr. Richard Schmid, d'abord pour m'avoir laissé la liberté de choisir un projet de recherche qui m'était pertinent et ensuite, pour leurs conseils et leurs suggestions tout au long de mes études doctorales.

De plus, malgré leur occupation du temps chargée, la porte de ces professeurs m'a toujours été ouverte pour obtenir des conseils ou des questions de dernières minutes. Je pense entre autre à M. David Wells avec qui j'ai pu discuter du devis de recherche, au Dr. Allyson Hadwin qui m'a donné le judicieux conseil d'utiliser Nud*ist pour l'analyse de données et aux professeurs Robert Bernard, Steven Shaw et Phil Abrami qui m'ont éclairé quant au processus d'analyses statistiques.

En dernier lieu, je crois humblement que que je n'aurais pas eu le privilège de compléter un doctorat sans le soutien constant de mes parents tout au long de mes études. Ils ont toujours crû que je pouvais y parvenir, toujours prêts à me venir en aide au moindre petit signe de défaillance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The ultimate goal of the educational process is to prepare children, teenagers and young adults for the "real" world. The development of critical thinking (C.T.) skills might be one of the most important parts of that process. Indeed, it is commonly believed that a learner who is able to apply critical thinking skills in a variety of situations will be better equipped not only to survive in the real world but also to take an active and productive part in it. Critical thinking skills are cognitive skills that are supposed to be triggered when an individual faces a novel problem or an unexpected situation. Those skills are believed to be generalizable, that is transferable across domains of knowledge and across situations. The educational specialist must prepare the learner in the expectation that he or she will have to apply those skills at any given time.

A lot of literature covers the theory, practice and research on the teaching of critical thinking skills. Several guidelines, based on the work of theorists such as Ennis, Paul, McPeck, and to a certain extent Bloom in the field of training, are available to inform K-12 practitioners on how to integrate critical thinking skills into the curriculum (for example, Baron & Sternberg, (1987); Halpern, (1997); and how to evaluate those skills in classroom settings (Ennis, 1993; Norris, 1989; Baron, 1987). In addition, several attempts, some by "schools of critical thinking", have been made in past decades to promote critical thinking skills in various domains and to guide their integration into curricula. However, research on the subject, especially in higher education, is scant. On one hand, we still know very little about the conditions in which critical thinking happens

(McMillan, 1987). On the other hand, most studies, exploratory in nature, are conducted with small groups in social science domains. Such studies tend to focus either on ways to improve the teaching of critical thinking skills or on the presence of the skills within a constructivist or collaborative framework.

While there are several models providing guidelines for the teaching of critical thinking in the classroom (for example, Baron and Sternberg,1987; Halpern, 1997) this study focuses on the effect of task design on the incidence of critical thinking skills. The reason for not addressing the role of the instructor was primarily based on the fact that as proposed by McMillan (1987) and Norris (1985), we still know very little about the specific characteristics and the individual components of instruction, such as learning context, teaching approaches, types of exercises or tasks and prior knowledge that affect the incidence of critical thinking skills. In this study, one component of instruction, namely types of instructional tasks, has been selected as a strategy to partly address the question as to how to foster critical thinking skills. In short, focusing on one individual component of instruction was a means to isolate and to simplify the problem under study.

The students carried out five written assignments or tasks of increasing complexity. Based on the assumption that a novel situation could trigger the usage of critical thinking skills (Halpern 1998), complexity variables, characterized by increasing difficulty, were embedded in the instructional tasks. Leenders and Erskine's (1987) model of case difficulty was used to establish the levels of complexity in cases. More specifically, the analytical dimension (nature of the problem given) and the presentation dimension (organisation and amount of information given) were used as guidelines to design the instructional tasks. Since the role of the instructor was "controlled", the conceptual dimension of the case difficulty model was left out of the research design. Finally, three types of written assignments were used in this study: essay, case study (complexity level 1) and simulation (complexity level 2). The literature on simulation, case-based and problem-based learning has been searched for general guidelines to design the tasks.

A modified version of Herrington and Oliver's synthesis of critical thinking skills has been used to analyze the data. Frequencies of "Judgement and Interpretation", "Multiple Perspectives", "Imposing Meaning" and "Metacognition" have been recorded. However, given the nature of the tasks, two categories of skills were left out: "Dealing with Uncertainty" and "Decifing of a Path of Action". Indeed the instructional design entailed in the written assignments was not appropriate for the observation of those two cognitive behaviors. In fact, an approach such as verbal protocol analysis for example, would be a better way to record mental processes implied in the skills "dealing with uncertainty" and "deciding of a path of action".

From a practical point of view, the results of this study inform us about the type of instructional method that triggers critical thinking. Furthermore, we obtain information about the type of skills triggered by a specific mode of instruction. From a theoretical standpoint, we gain information about the nature and the incidence of skills given the instructional tasks. While "Judgement and Interpretation" skills are more common, "Multiple Perspectives" and "Imposing Meaning" skills tend to be less frequent. This finding has implications for further research. One recommendation for future research is that small frequencies and triangulation of measures be used to identify more precisely the nature of critical thinking.

Chapter 2 is organized in two sections. First, an overview of the literature on critical thinking and current research in the field is presented. Secondly, the framework of the study is presented, including the rationale, the problem statement, and the research hypotheses.

Chapter 3 describes the context of the study, the methodology, and the procedures entailed in the process. These procedures include the description of the instructional design approach, data collection and coding procedures, training of the coders and inter-rater reliability procedures.

Chapter 4 presents trend analyses of the overall measures of critical thinking skills first, followed by the findings on the specific categories of critical thinking skills: "Judgement and Interpretation", "Multiple Perspectives" and "Imposing Meaning".

Chapter 5 is organized in two sections. In the first part, the results obtained in Task 3, and in Task 4 and 5 are analyzed and discussed. In the second part, reflections on the process of conducting research in the field of critical thinking are offered. This part of the discussion aims at identifying the challenges posed by conducting content analysis in the study of critical thinking.

Chapter 6 synthesizes the findings and the reflections presented in chapter 5. The limitations of the study are revisited in order to propose new avenues for future research.

CHAPTER 2: LITERATURE REVIEW

In this chapter, a comparison of the terms "critical thinking" and "higher order thinking" is drawn. Then, an overview of current instructional approaches that foster critical thinking skills is presented and is followed by a review of research conducted in higher education settings. In particular, five aspects relevant to this study will be highlighted. They are: 1) domain and sample size; 2) research design in a constructivist environment; 3) content analysis of discussion data; 4) use of frequency data; 5) time and practice. Finally, the rationale for the study and its problem statement conclude this section.

Critical thinking skills or higher order thinking?

The literature tends to confuse critical thinking and higher order thinking (Lewis & Smith, 1993; Halonen, 1995). Indeed, several labels –higher order thinking, rational thought, reasoning, problem-solving -- are used interchangeably to describe an intellectual activity that goes beyond what is generally understood as basic (Lewis & Smith, 1993). Essentially, the term "critical thinking" arises in the field of humanities whereas "higher order thinking" has evolved within a scientific context.

On one hand, the field of psychology, evolving in a scientific paradigm, tends to examine higher order thinking in the context of problem solving (Lewis & Smith, 1993). In this case, the steps in the process of solving a problem can be easily identified, resulting in a compartmentalization of skills to be acquired by the learner. In addition, problem solving skills can be easily classified into either basic or higher order thinking skills, so that the acquisition of the former becomes prerequisite to the latter. This dichotomous categorisation of cognitive skills leads to models that influenced all fields related to education. The work of Bloom (1956) and its impact on the field of instructional design present a good example of this hierarchal approach to cognitive skills.

On the other hand, critical thinking is defined as reasonable reflective thinking that is focused on deciding what to believe or do (Ennis, 1987). While Ennis' definition implies some sort of action, others such as McPeck and Brookfield, both cited in Garrison (1991), emphasize the reflective meaning of the term "critical". Halonen (1995), however, suggests an hybrid definition to encompass both action and reflection. Thus critical thinking is "The propensity and skills to engage in activity with reflective skepticism focused on deciding what to believe or do" (Halonen, 1995, p.77).

This definition draws on the discipline of philosophy, wherein the purpose of acquiring critical thinking skills through the processes of argumentation and Socratic dialogue are used to "guard against the propensities of humans to accept fallacious arguments and draw inappropriate conclusions" (Resnick, 1987). From that point of view, critical thinking connotes an evaluative, judgmental flavor. In practice, learners acquire critical thinking skills by exchanging ideas, debating points or, in exams, by developing their answers with supportive arguments (Lewis & Smith, 1993).

However, it now seems generally accepted, notwithstanding the philosophical or the pscyhological paradigm, that what were thought of as basic and higher order thinking skills are interwoven in the teaching and the learning process (Resnick in Lewis & Smith, 1993), to the degree that it becomes futile to distinguish between them when developing teaching approaches. Furthermore, higher order thinking implies that there is an interaction between "cognitive strategies, metacognition and non-strategic (domain-specific knowledge) during novel problem solving" (Young, 1997). In addition, according to Braten (in Young, 1997), the cognitive and metacognitive components "interact in a complementary way".

Furthermore, Ennis (1987) suggests that the ability to think critically is not enough to qualify "good" thinking: One needs not only to possess the required skills, but should also be willing to act upon and to apply them at the appropriate moment (Ennis, 1987; Halpern, 1998). In essence, we are speaking about critical thinking skills and critical thinking dispositions (Halpern, 1998). According to Tishman, Jay and Perkins (1993), dispositions are composed of three elements: 1) the ability to carry through a behaviour; 2) the sensitivity or alertness to appropriate occasions to display the behaviour and 3) the inclination to behave in a certain way. Seven dispositions or tendencies have to be considered while fostering critical thinking in the classroom (Tishman & Andrade, 1993). They are:

1) disposition to be open-minded and adventurous
2) disposition toward sustained intellectual activities
3) disposition to clarify and seek understanding
4) disposition to be planful and stategic
5) disposition to be intellectually careful
6) disposition to seek and evaluate reasons
7) disposition to be metacognitive.

While some dispositions are easier to observe in practice, those attitudinal components are best taught by an enculturation process (Tishman, Jay, & Perkins, 1993). According to the authors, the teacher, instead of focusing solely on the critical thinking skills, could use cultural exemplars, encourage class interactions and use direct teaching strategies to cultivate, to nurture the dispositions to think critically.

Halonen (1995) also addresses the issue of dispositions but she adopts a broader perspective. Speaking of "propensities" rather than dispositions, she extends Passmore's definition of "critical spirit". She suggests we consider four propensity elements to critical thinking : the first propensity is affective which is, according to Brookfield (in Garrison, 1991), essential to motivate the learner. An element of surprise in the instruction, for example, would trigger the use of critical thinking skills. Then, there is the attitudinal aspect, which as Halonen explains, borrows on Paul's conception of intellectual and moral virtues. The virtues include intellectual humility and courage, integrity, empathy, perseverance, fairmindedness and confidence in reason. The third element is physiological readiness (fatigue, well-being etc.) which, according to Halonen, is rarely addressed in the critical thinking literature. Finally, metacognition is also considered a propensity and is defined as a "kind of scrutiny given to the process, the product and the changes in the thinker that result from critical thinking activities". (Halonen, 1995; p.77).

In summary, three key features relevant for the purpose of the current research can be highlighted. First, it is no longer useful to isolate thinking skills as high or low. Instead, "higher order" thinking skills take many forms, including metacognitive skills, and interact in a complementary fashion. Second, all definitions imply that critical thinking skills or higher order thinking occur when a novel problem is being solved (Young, 1997) or in a "perplexing situation" (Lewis & Smith, 1993; Halpern, 1998). Remarkably, this key feature is generally taken for granted, even though the usage of thinking skills depends on the individual's history (Lewis & Smith, 1993). In other words, a particular problem may require higher thinking skills for person A, yet person B, because of previous experience with a similar case, will find the same problem relatively easy. Person B would then be solving a routine problem, and therefore, would not be using critical thinking skills. Finally, being good thinker requires not only to ability to think critically, but also a readiness to think critically in appropriate circumstances.

For educators, the challenge then becomes one of instructional design: How does one create a learning situation whereby most individuals will have to use critical thinking skills to solve a problem? How do you incorporate the enculturation process for the teaching of critical thinking dispositions? What are the conditions and the characteristics of instruction that are most likely to trigger higher order thinking? More specifically, when does critical thinking occur during the learning process? Are there identifiable conditions which stimulate the use of critical thinking skills? In order to answer some of these questions, it is necessary to look at the instructional approaches that are used to foster the development of critical thinking in the learner.

Instructional approaches

According to Halpern (1998), the teaching of critical thinking is based on two assumptions. First critical thinking skills are clearly identifiable and definable so they can be taught to be applied appropriately. Secondly, since the skills are recognizable, and teachable, the students can be made effective thinkers. (Halpern, 1998, p.451). These assumptions lead to two branches of the literature concerning the teaching of critical thinking skills. On one side there are provisions focusing mainly on the role of the teacher and offers strategies that can be taken to foster critical thinking. In that body of literature, instructional models or strategies for integrating a specific set of critical thinking skills (for example, developing arguments), using different approaches (inquiry, problem-based etc.) are proposed. The other branch of the literature is more technology driven. In this area, interactive or collaborative environments are created, while drawing on constructivist principles to foster critical thinking skills. Both paths are explored in the following paragraphs.

Teaching strategies

If we assume that critical thinking encompass both the cognitive skills and the dispositions to think critically, the role of the teacher as a facilitator becomes paramount. Indeed, Underbakke and his colleagues suggest that for learning to happen, the second most powerful predictor of how much students learn is the teacher. The most important predictor is students prior knowledge. The authors pursue this point by suggesting that "since we cannot control "students' prior knowledge" the only remaining variable is the teacher's performance" (Underbakke, Borg, & Perterson, 1993; p.138).

General models of teacher training for the integration of critical thinking in their course outlines tend to be organized around three components. Fogarthy and McTighe (1993) for example, propose the three story model. The first level focuses on the acquisition of critical thinking skills. The second level focuses on the design of opportunities to involve the students, that is the instructional strategies to foster critical thinking attitudes. The last level targets lifelong learning of critical thinking skills, so the teachers are presented with strategies to increase the chance of transfer from learning in the classroom to real-life situations.

Halpern (1998), on the other hand, suggests more specifically what should be taught to teachers. She proposes that the instruction of critical thinking should encompass four elements: the disposition and attitudinal element; instruction in and practice with critical thinking; the structure of the instructional activities aimed at tansfer across contexts; and a metacognitive component used to assess and direct learning. In essence, teaching students to become good thinkers requires more than the acquisition and practice of the skills per se and so should include the modelling of good thinking as a means to address the question of disposition.

As mentioned earlier, there are in the literature several instructional strategies and approaches proposed to teach critical thinking in general , or in a specific set of skills. Given the purpose of this study, only two main strategies—discussion and writing exercises—will be addressed here. An overview of other strategies and the skills they address is presented for the purpose of illustration in Appendix I.

Good thinking is either modelled or facilitated by the teacher (Halpern, 1998) or demonstrated through an enculturation process (Tishman et al., 1993). While some propose that critical thinking is more a dialectic process (Paul, 1987; Kuhn, 1999), others suggests that writing is more appropriate for fostering individual thinking skills (Wade, 1995).

The main advantage of classroom discussions as a strategy to foster critical thinking is the possibility for the student to benefit from peer-to-peer interaction. As mentioned earlier, exchange of ideas allows one's peers to be challenged. Verbal exchanges might also constitute an opportunity for weak writers to express themselves more easily (Baron, 1987). In addition, the teacher can model good thinking or scaffold the instruction. King (1995) for example, uses a questioning strategy to foster both the acquisition of critical thinking skills and dispositions. Before class, each student prepares two or three "thoughtful" questions, and in class, the teacher selects one to discuss as a group or in pairs. Although King recognizes that such an approach requires more time than lecturing, she suggests that the strategy is centered on students' needs. The main disadvantage of class discussions is that not every student gets to participate (Wade, 1995).

Writing is also another activity that is frequently used to foster critical thinking skills. Wade (1995) for example, contends that writing exercises have the advantages of getting everyone to practice taking an active part in the exercise. Writing also fosters dialectic reasoning, that is, the students have time to look at the other side of an argument and allow for the refinement of ideas. Finally, and perhaps more importantly, much of the time spent writing focuses on reflection, not on immediate reaction about what is said in class (Wade, 1995).

In summary, both approaches have their disadvantages and advantages. The selection of one particular strategy might very well depend on practical constraints of the learning situation. In any case, these are but two approaches to teach critical thinking. Perhaps the wisest approach would be to use a combination of both class discussion and writing exercises in order to reach a broader spectrum of students' needs (Baron, 1987).

Interactive encvironments

Researchers and developers of interactive computer-based environments generally assume that the application of broad constructivist principles in the design of instruction should trigger the use of critical thinking skills by the learners. For instance, Young (1997) contends that the key features of situated cognition are crucial when higher order thinking is the "target of instruction". Hence, the learning environment should integrate collaborative social interaction, authenticity, scaffolding, as well as reflection about the metacognitive processes.

Following the same logic, Herrington and Oliver (1999) have created a multimedia environment that incorporates nine characteristics of situated learning, as a means to investigate students' thinking abilities. These characteristics are somewhat more precise than what Young (1997) proposes. They include: 1) authentic context; 2) complex authentic activities; 3) multiple perspectives; 4) expert performances; 5) coaching and scaffolding; 6) opportunity for collaboration; 7) reflection; 8) articulation; and 9) authentic assessment (p.5).

In another study, McLoughlin & Oliver (1998) and McLoughlin (1997) put a stronger emphasis on social interaction as a means to foster higher order thinking in a distance learning environment. Vygotsky's concept of proximal development as well as reciprocal teaching strategies have been used because "when learners have to explain ideas to each other ..." a more explicit and organised understanding can result, often referred to as higher order thinking (Mercer in McLoughlin and Oliver, 1998).

Bullen (1998) created a collaborative computer-conferencing environment with the purpose of exploring student-student interactions as well as student-facilitator interactions. Cognitive dissonance was believed to arise from the collaboration, thus increasing opportunities to create conditions favorable for the use of critical thinking skills. The content of the telediscussions was analyzed using Norris and Ennis critical thinking skills criteria. The work of Newman and his colleagues represents another attempt to explore the incidence of critical thinking in computer supported cooperative work (Newman, Webb, & Clive, 1995).

Adelsköld, Aleklett, Axelsson, & Blomgren (1999) designed a genuine problem-based approach for graduate students from all domains studying at a distance. The research group used figures, data and charts as a basis for their instructional strategy to trigger discussions that might lead to the development of critical thinking skills. As part of the learning process, students had to exchange ideas and thoughts as well as determine their learning goals.

In summary, both branches of the literature represent attempts to foster good critical thinking. As we have seen, several approaches in both classroom education and in computer-based interactive environment have been proposed as means to enhance the learning of critical thinking skills. The next issue then is to explore the extent to which the teaching and the learning of such skills is efficient. Issues pertaining to the measurement of critical thinking skills, both for the purpose of evaluation and research are covered in the next section.

Evaluation of critical thinking skills: State of research

The field of critical thinking is alive and kicking, yet research on the subject, at least in higher education settings, remains scarce. On one hand, the number of experimental studies available is somewhat limited. In addition, the results from critical thinking tests are disappointing (Norris, 1985), to the point that some have been questioning not only the validity of the instruments (Norris, 1985; Coffman, 1987), but also the reason for using such tools to assess the efficiency of instruction (McMillan, 1987; Norris, 1989). On the other hand, case studies lead to results that are difficult to replicate or to generalize to the population. Some researchers conduct content analyses of on-line discussions, classroom interactions and written assignments to attest of the incidence of critical thinking. The challenges of measuring and assessing critical thinking are explored in following sections.

Two authors have conducted reviews of empricial evidence on critical thinking. McMillan (1987), used the term "critical thinking" and found 21 studies conducted between 1950 and 1985. Seven of those studies were by the same authors (Dressel and Mayhew). Gibbs (1985) on the other hand, used more stringent selection criteria based on experimental assumptions. He retained only nine studies conducted over approximately the same time span. An overview of their conclusions is presented in the following paragraphs.

Perhaps the most striking observation is the use of weak experimental designs in most studies. Only one study (Bailey, 1979) used a true experimental design reported by McMillan (1987). All the other studies used a quasi-experimental approach. Gibbs (1985) recognized that practical constraints prevented the researchers from using randomization procedures, which would reinforce the generalizability of the results.

Another common point made by both reviewers is that, although the evidence is weak, college education does seem to contribute to gains in critical thinking. In addition, senior students tend to do better than freshman. College students who do well on critical thinking tests upon entry do not improve as much as the ones who did poorly. This might suggest that college education is more efficient for weaker than stronger students. Although McMillan warns about the danger of such hasty conclusion, he fails to recognize that such phenomenon might be attributable to rival hypotheses such as a "ceiling effect or tendency to regress towards the mean" (Borg & Gall, 1989). McMillan (1987), however, underlines that these results have to be taken with caution because of extraneous variables such as maturation and mortality.

Instructional variables such as peer-to-peer interaction, teaching approaches, teacher effectiveness, level of student participation in class, student-teacher interaction and learner characteristics have been explored, and again, inconclusive results are obtained. There might be gain in critical thinking in general but "association evidence is weak" (Gibbs, 1985; p.145). Although the evidence suggests that critical thinking can be taught efficiently (Halpern, 1998), the current instruments are not precise enough to reveal what exactly is contributing to accrued critical thinking (McMillan, 1987). However, two recent studies, Tsui (1999) and Anderson, Howe, Soden, Halliday, & Low (2001), have demonstrated that it is possible to associate critical thinking with more than one instructional factor. What is interesting about these studies is that they used an alternative modes of evaluation. Tsui used a self-reported measure, Anderson et al. triangulated content analysis and critical thinking tests. These studies will be examined subsequently.

Another criticism of critical thinking tests concerns the issue of time. In order to think critically, students need time to reflect (McMillan, 1987; Coffman, 1987). Testing in a contrived environment is probably not conducive to applying critical thinking skills, but also to applying them properly. Furthermore, the time lapse between the teaching and the practice of critical thinking - for example over one term - might be too short to prove any long term impacts of the instruction (McMillan, 1987). If longitidunal studies would represent a better alternative, other factors such as maturation, mortality, and college experience factors might constitute rival hypotheses (McMillan, 1987). In short, there is no simple recipe for designing a research study on critical thinking skills.

Hence the results of research on critical thinking have been disappointing. In view of the constraints outlined earlier, it seems that the limited results might be attributable to the instruments used in evaluating or conducting research on critical thinking skills (Gibbs, 1985; Norris, 1985; Coffman, 1987; McMillan, 1987; Norris, 1989). Most of the studies reviewed by McMillan and Gibbs used standardized tests as a means to measure critical thinking in pre-test posttest studies. While those authors surreptitiously suggest that such tests are not appropriate for research projects, others contend that we can improve them (Norris, 1989; Ennis in Norris, 1989).

There is a wide choice of standardized instruments available on the market (see for example, Baron, 1987). The tests most frequently used in both research and evaluation studies are the Watson-Glaiser Critical Thinking Appraisal and the Cornell Test of Critical Thinking (Gibbs, 1985; McMillan, 1987; Norris, 1989). Those instruments are composed of multiple choice normative test items and cover various sets of critical thinking skills such as inductive and deductive reasonning, judgement and so on (Norris, 1989). The main criticism, as we have seen earlier, is that those tests are not designed to measure the specific impacts of instructional methods, teacher effectiveness, and course design on critical thinking. The lack of evidence might lead to faulty a conclusion, mainly, that the instruction was ineffective. A good example of this can be observed in Anderson et al. (2001). They found some evidence of critical thinking gains when they analyzed peer interaction and students' written assignments, but the participants performed poorly on the Smith-Wetton Critical Reasoning Test. The outcome of the test led the authors to conclude that critical thinking is not generalized to another situation. Given the results obtained using other measuring methods, the authors could have legitimately challenged the validity of the instrument.

Significantly, students as a rule do poorly on standardized tests. Norris (1985) for example, reports that the median scores of undergraduate students on the Watson-Glaser CTA tests range between 52 and 60 out of a possible 80. The best performances are recorded for MBA (66) and medical students (68). The results are even lower for high school students. These results led Norris (1985) to two conclusions that should be taken with caution. First, the exercise of critical thinking as measured with standardized multiple choice tests is an uncommon phenonmenonm amongst learners. Secondly, critical thinking is uncommon amongst the general population (Norris, 1989), even to the extent that less than half adults think critically (Garrison, 1991). Again, critical thinking may be an uncommon phenomenon that occurs in complex and messy situations (Halpern, 1998), but it would be too reductionist to base our understanding on the sole evidence of these tests.

In a later article, Norris (1989), without rejecting all together the use of the main standardized tests, challenges the validity of such instruments. One one hand he contends that even if a test item is well designed and tested to measure effectively the presence of critical thinking, students during the test may not respond as expected due to time or dispositions, as mentioned earlier. This does not mean that the student is not apt to think critically. On the other hand, he contends the tests are designed to measure the product of critical thinking, not the process of thinking critically. The student thus may come to the correct conclusion, using faulty or inapropriate reasonning.

The distinction between evaluating the product and the process of critical thinking is an important one, especially considering how they are defined. If we consider the disposition to think critically, that is a willingness to apply critical thinking, it would be paramount to develop at least some sort of measure that factors in the thinking process, including ability to think, sensitivity to the occasion, and disposition to act (Ennis, 1987; Tishman et al., 1993).

Several suggestions have been made to improve the instruments. The most common one is that the tests should have some kind of open form (Coffman, 1987), to provide an opportunity to justify answers. The Ennis-Weir Essay Test caters to that need, requiring multiple choice resopnse, but also a justification for the choice. While those tests are more costly to use because of the time required in marking them, they represent a viable alternative to generic standardized tests.

Another suggestion is to develop in-house tests (Coffman, 1987; McMillan, 1987). The advantage would be that the test developers are more in tune with what the curriculum addresses and as a result, the test items would be more closely related to what the students have learned. The test would be more enlightening as to what was effective in terms of instruction or teaching approaches (McMillan, 1987). They would also be more sensitive to the context of learning and the subject matter that was taught.

If some suggest solutions to improve the validity of standardized tests, others take a more radical stance and challenge the basic rationale of those tests. Coffman (1987) for example, suggests that the modern approach to developing tests be questionned. He raises the issue of the norm, that is, the use experts' understanding of what consitutes critical thinking skills and the development of test items that seek to elicit that outcome. He revisits the work of Tyler. Tyler first observed participants' behaviors in situ and then developed items to match these behaviors -in other words, developing a situation-specific "norm". A recent study on decision making in social science domains, (Osana, Bennett, & Tucker, in press) also raised the issue of using the "expert" norm to explore adolescents' underlying reasons for making decisions. Using a norm as the basis for creating assessment models or test items poses interesting challenges, with regard to time and money constraints as well as intellectual issues.

Notwithstanding the issues raised above, the reliablity of standardized tests is not seriously challenged in the literature. Gibbs reports that the reliability value on multiple choice and essay tests are usually .70. The Glaser-Watson CTA reliability values range from .69 to .89 whereas the averaged coefficient values for the Cornell Test is slightly higher (Level X: .80 and Level Y: .71).

In more recent studies however, it seems that the use of standardized instruments as a means to measure gains in critical thinking has been put aside. The preferred mode of inquiry tends to be content analysis. We have seen earlier that some are conducting content analyses of student's contribution in on-line environments (Henri, 1992; Newman, Webb, & Clive 1995; Bullen, 1998; McLoughlin & Oliver, 1998; Herrington & Oliver, 1999; Winnips, Collis, & Moonen, 2000); Others are observing peer interaction in the classroom (McLoughlin, 1997; McLoughlin, Baird, Pigdon, & Woolley, 2000; Anderson, Howe, Soden, Halliday, & Low, 2001). Some are combining content analyses and testing strategies to assess critical thinking (see for example Anderson et ai., 2001).

Finally, Tsui (1999), used a completely different approach. The author surveyed fourth year college students about the impact of their education on critical thinking skills. She used a self-reported measure to question the students about their relative growth in their perceived ability to think critically. Using a block regression analysis, she explored the relationship between two independent variables and their impact on critical thinking. The variables were "Course" which comprised information about the number of courses by domain (writing, science etc.); and "Instruction" which included items such as "paper criticized by teacher" "multiple choice exam" etc. She found that student perceptions related to both variables, but one cannot consider either of them independently. Obviously, the use of a survey as a tool and the design of the study (one self-reported measure of critical thinking) to collect data might be criticized. However, what is interesting here is the idea of using self-reported measure for the assessment of critical thinking. We may envisage a research design that would incorporate a carefully developed self-assessment tool that could be triangulated with other intruments or methods to analyze the impact of instruction on critical thinking.

In conclusion, there are some limitations entailed with these modes of measurement, difficulties that arise in conducting a content analysis. Nevertheless, if standardized tests might not be ideal to assess the incidence of critical thinking skills nor to explore the specific contributions of instructional stategies on gains in critical thinking, content analyses represent a viable choice, at least for the purpose of research.

In this section, we explored some of the challenges and the difficulties entailed in the teaching and the evaluation of critical thinking. In the next section, some elements of theoretical interest are revisited, in order to establish a framework and rationale for the study.

Framework for the study

In this section, five features in the literature have been highlighted to create a framework for this study. These features, sometimes descriptive, sometimes reflective, synthesize the state of research on critical thinking skills and lead to a rationale for the study. They are 1) domain and size of sample; 2) research design in a constructivist environment; 3) use of content analysis of discussion data to assess critical thinking; 4) use of frequency data to assess critical thinking; and 5) time and practice.

Domain and size of sample

The first noticeable aspect of the current state of research concerning critical thinking skills in higher education settings is that the sample sizes are usually relatively small. In case studies, such as McLoughlin (1997), Bullen (1998), McLoughlin & Oliver (1998), as well as in evaluative experiments (Adelsköld, Aleklett, Axelsson, & Blomgren,1999; Herrington & Oliver, 1999), samples tend to include less than 20 participants, and most often less than 10. In addition, most studies tend to be conducted in social sciences domains such as ethics (Bullen, 1998; Calkins, & Armstrong, 2000; teacher education (McLoughlin, 1997), and educational technology (Herrington & Oliver, 1999; Winnips, Collis, & Moonen, 2000). To my knowledge, no study has been done investigating critical thinking skills of engineering students. In contrast with the studies listed here, the participants will apply critical thinking skills to a domain of study (social studies) different from their main academic programme (science) (Quellmalz, 1987).

Research design in a constructivist environment

With the exception of Herrington and Oliver's work (1999), most studies do not investigate the specific characteristics of constructivism as independent variables. In fact, it is generally assumed that situated learning, for example, as supported by Young (1997), will automatically create an environment favorable for the development of critical thinking skills, as well as trigger the use of those skills. From a pedagogical point of view, constructivist principles might appear to create favorable conditions. However, from a research point of view, it is difficult to attribute actual critical thinking outcomes to a specific characteristic of constructivism.

For instance, one might question the specific contribution of a collaborative strategy in the development of critical thinking skills, that are, after all, an individual manifestation. Several factors within collaboration might trigger higher order thinking: team composition, interest in subject, willingness to work with others, nature of the task, experience with a group and of course, individuals' a priori ability and disposition to think critically.

The effect of scaffolding strategy on critical thinking skills represents another challenge for the study of critical thinking skills within a constructivist framework. Very often scaffolding, at least in projects regarding critical thinking skills, takes the form of individualized feedback (for example, Bullen, 1998) or individual coaching (McLoughlin, 1997). Despite noticeable exceptions (for example, Herrington & Oliver, 1999; McLoughlin, 2000), scaffolding is seldomly integrated into the instructional activities per se. Moreover, even in comparative studies, the effect of scaffolding is often disputed. For example, Winnips, Collis and Moonen (2000) did not find significant results in their study that specifically investigated scaffolding.

To summarize this point, constructivist learning environments tend to be taken for granted when designing a study on the use of critical thinking skills. The potential for pedagogical benefits is not disputed here. Difficulty arises not only in testing, but also in simply describing the effects of specific characteristics such as scaffolding and collaborative learning on the development of critical thinking skills. Since the characteristics of constructivism are numerous, the range of instructional design options expand, thus complicating the interpretation of results.

Content analysis of discussion data

The third, and perhaps the most interesting feature in current research on critical thinking skills is that the majority of investigators use the content of telediscussions or oral performance to assess the presence of those skills. Talk-aloud protocols have been used as a means to investigate critical thinking skills. However, the usefulness of oral performance is currently being challenged in the field (Herrington & Oliver, 1999), for one reason, "because at the point when students are engaged in problem-solving, they become quiet, possibly due to cognitive overload". One solution has been to use less obstrusive observational methods such as videotaping students' interactions in the classroom (Anderson et al., 2001).

On the other hand, the study of telediscussions necessarily entails the recording and the classification of lower order speech events, such as social and procedural utterances (for example, Bullen, 1998; Henri, 1989). High frequencies in such categories might obscure more interesting research data, the less frequent higher order thinking skills. As mentioned earlier, critical thinking tends to happen in novel situations, which are, by definition, exceptional (Halpern,1998).

Furthermore, content analysis is not a straightforward task. Ahuvia (2001) suggests that even in traditional content analysis where raters are coding "denotative" or manifest meaning as opposed to "connotative" or latent content in visual or written texts, the inter-rater agreement might still be relatively low. He cites a study by Gilly (1988) where raters had to identify the type of setting (restaurant, private house, public space) in ads. This relatively simple coding scheme led to an inter-rater agreement of 67%. Ahuvia concludes that achieving reliability when the main object of research is to identify the connotative meaning of a focal text might be an impossible task.

For the purpose of illustration, the proposed standard of inter-rater agreement for content analysis is .90 (Miles & Huberman in Harwell, 1999). If we transpose the difficulty of analyzing content into the field of critical thinking skills, it is not surprising that reliability coefficients even on standardized tests range from .70 to .80. In addition, given the fact that critical thinking skills are complex and ill-defined (Halpern, 1998) and that some skills might be composed of more than one element (Ennis, 1993), achieving an acceptable level of inter-rater agreement when conducting a content analysis might be an elusive goal.

There is no ready-made solution to the challenge of achieving an acceptable level of agreement in content analysis of critical thinking skills. There are, however, some provisions that can be made to improve the inter-rater reliability, such as a comprehensive training sessions for the coders or the negotiation of the meaning of the codes between raters (Ahuvia, 2001). Another indirect strategy consists of reinforcing the internal validity of rating data, by ensuring that the coders are " interchangeable" or that they represent a parallel form of instrument (Harwell, 1999). This strategy entails the control of extraneous variables such as history, mortality, instrumentation, which contains instrumentation decay and practice effect, and finally, the threat of imitation or diffusion (Campbell & Stanley, 1966). Those issues have been adressed in the present study and are explained in the Methods chapter.

Use of frequency data

The fourth feature of current research on critical thinking skills concerns the expectations researchers hold for the frequency results. Indeed, the most common way to assess critical thinking skills is to count their frequencies during a given learning event. High frequencies might make measurement and comparisons easier. However, it is not logical to expect high frequencies for a phenomenon that, by definition, should only happen in novel situations, in non-routine types of problems. In fact, the challenge, as we have seen earlier, is to create a learning situation that will be novel for the majority of the learners.

Frequencies on critical thinking skills are not as enlightening as we would wish. At most, one might obtain a high count on certain indicators; one might be able to show that, given certain conditions, "judgement" for example, will occur more frequently than "self-regulatory activities" within a given group of participants. We are simply observing the incidence of the skills; we are not assessing the quality of thinking nor the disposition to think critically. A student could exhibit a skill in a given situation, but the usage of that skill might be inapropriate given the circumstances (Ennis, 2002). In addition, a high count might be due to personal characteristics of the participants, or their previous history. In short, results on frequency counts are difficult to generalize. One solution, when practical constraints are not too limiting, would be to use several measures of critical thinking skills at various times (McMillan, 1987; see also Anderson et al., 2001).

Time and practice

Thinking critically requires time and reflection (McMillan, 1987). Constrained environments or testing situations where time is limited might lead to inconclusive or misleading results on the performance of critical thinking (McMillan, 1987; Gibbs, 1985; Norris, 1989). On the other hand, demonstrating critical thinking in online environments (synchronous or asynchronous) might be too demanding on the part of the students who have to read, think and type a reply while online. While practical constraints does not allow to conduct a study over more than one term as proposed by McMillan (1987), the issues of allowing time for reflection and practice of the skills are addressed in this study. Students will be writing weekly essays over a period of six weeks.

Rationale for the study

It seems that while research about critical thinking yields mixed or disappointing results, strategies for integrating the teaching of critical thinking in the curriculum are voluminous (Norris, 1985). Yet, we still have limited knowledge as to which specific parameters of the instruction are contributing to the incidence of critical thinking. This study focuses on one parameter of instruction, that is the instructional strategies. The problem statement is as follows:

Problem statement

The proposed research will investigate whether task design can contribute to the development of students' higher order thinking skills. Specifically the level of complexity in task design is manipulated in this study (see Leenders and Erskine's model of Case difficulty presented in Table 3.3).

Research questions and hypotheses

Main hypothesis:

The introduction of "complexity" in instructional activities will increase the use of critical thinking skills.

Secondary hypotheses:

Hypothesis 1: Complexity "Level 1" variable

$H0 = \mu_{dT2-T3} = 0,$

$HA = \mu_{dT2} < \mu_{dT3}$

The complexity Level 1 variable corresponds to the second level of the analytical dimension (a problem is given; the student has to find a solution) and to the second level of the presentation dimension (the case contains an average amount of information with some extraneous data) (see Table 3.3).

Hypothesis 2: Complexity "Level 2" variable

$H0 = \mu_{dT3-T4} = 0,$

$HA = \mu_{dT3} < \mu_{dT4}$

The complexity variable Level 2 corresponds to the third level of the analytical dimension (student has to identify both the problem and the solutions) and the third level of the presentation dimension (the information provided is less organized and contains a large amount of extraneous data) (see Table 3.3).

<u>Hypothesis 3</u>: Evidence of transfer

$H0 = \mu_{dT4\text{-}T5} = 0,$
$HA = \mu_{dT4} < \mu_{dT5}$

<u>Hypothesis 4</u>: Baseline

$H0 = \mu_{dT1\text{-}T2} = 0,$
$HA = \mu_{dT1} \neq \mu_{dT2}$

<u>Hypothesis 5</u>: Group equivalence

$H0 = \mu_{dCtrl\text{-}Trmt\,(T2)} = 0,$
$HA = \mu_{dCTRL} \neq \mu_{TRMT}$

The relationship between writing abilities and critical thinking skills will be explored.

<u>Hypothesis 6</u>:

Participants whose English writing skills are high will demonstrate critical thinking skills more frequently than participants whose skills are low.

Reasons

Issue of transfer

In the proposed study, participants educated within the scientific domain will be asked to apply their thinking skills within a social science domain. The skills, according to Quellmalz (1987), are said to be transferable from one domain to the other (see Table 2.1).

Table 2. 1 Examples of Applications of Higher Order Reasoning Skills in Three Subject Domains.

	Science	Social science	Literature
Analyze	Identify the components of a process or the features of animate or inanimate objects	Identify the components of an argument or the elements of an event	Identify the components of literary, expository, and persuasive discourse
Compare	Compare the properties of objects or events	Compare the causes and effects of separate events and of social, political, economic, cultural, geographic features	Compare meanings, themes, plots, characters, settings, and reasons
Infer	Draw conclusions, make predictions, pose hypotheses, tests, and explanations	Predict, hypothesize, conclude	Explain characters' motivations in terms of cause and effect
Evaluate	Evaluate the soundness and significance of findings	Evaluate the credibility and significance of arguments, decisions, and reports	Evaluate form, believability, significance, completeness, and clarity

Written exercises

Instead of using talk-aloud protocols or discussions held in electronic forums, student written work in a regular credit course will be used as the main source of data. The decision to use written assignments as a strategy to assess the incidence of critical thinking is in keeping with the issues of time and practice outlined above.

Individual work

For reasons outlined earlier, scaffolding strategies and collaborative learning have been deliberately left out of the design of the current study. The point is to neutralize as much as possible extraneous variables such as individual feedback, in order to examine the specific impact of a given strategy on the critical thinking performance.

Complexity of the situation

In keeping with the literature, it is assumed that a novel situation or an ill-defined problem triggers the use of critical thinking skills. Thus, complexity of instructional content becomes an independent variable in the study. To generate critical thinking, instructional tasks have been designed following principles borrowed from the field of simulation (Jones, 1985; Thiagarajan & Stolovitch, 1978), case study in teaching (Leenders & Erskine, 1989) as well as problem-based learning (McBurney, 1995; Adelsköld et al., 1999). Specifically, Jones' (1985) guidelines for designing simulations were followed to establish a correspondance with reality (Thiagarajan & Stolovitch, 1978). The simulations had to contain the descriptions of the role of the participant, the mandate, the context and the tools necessary to carry out the simulation.

Data

Finally, to analyze the incidence of critical thinking, a modified version of Herrington & Oliver's (1999) synthesis of higher order thinking skills, which is essentially based on Resnick's work and comprises most current articulations of critical thinking, will be used as a framework of analysis. Specifically, only the indicators "Judgement and Interpretation", "Multiple Perspectives" and "Imposing Meaning" will be used to analyze student work. The instructional design of the course under study will not generate examples of "Deciding on a Path of Action", "Self-Regulation of Thinking" and "Uncertainty". Herrington and Oliver's synthesis is presented in Table 2.2.

Table 2. 2 Herrington & Oliver's Synthesis of Higher Order Thinking.

Revised characterisation of HO	Corroborating definitions of HO thinking from other theorists	Indicators for the purpose of classification
Uncertainty	• Asking, clarifying questions (Ennis, 1993)	• Any questions or statement seeking clarification of requirements, or uncertainty.
Deciding on a Path of Action	• Deciding what to do (Lewis and Smith, 1993) • Browsing and searching (Duchastel, 1990)	• Any statement referring to a suggested course of action • Any question asking opinion on a course of action.
Judgement and Interpretation	• Identifying conclusions, reasons and assumptions (Ennis, 1993) • Developing and defending a position on an issue (Ennis 1993) • Defining terms in a way appropriate for the context (Ennis, 1993) • Making contributions which are relevant and connected to prior discussion (Newmann, 1990).	• Any statement which seeks to defend a position taken on an issue • Any statement which connects to, and furthers the discussion • Any statement which defines terms in a way appropriate for the context
Multiple Perspectives	• Angling (establishing different perspective) (Duchastel, 1990) • Assuming the role of questioner and critic (Newmann, 1990)	• Any statement which suggests an alternative approach • Any statement which challenges a conclusion or a previously made point by providing an alternative perspective. • Any statement which challenges a perspective given in the MM program.
Imposing Meaning, effortful thinking and multiple solutions	• Drawing conclusions when warranted, but with caution (Ennis, 1993) • Offering explanations for conclusions (Newmann, 1990) • Deciding what to believe (Lewis & Smith, 1993) • Integrating (interrelating conceptual elements) (Duchastel, 1990) • Generating original and unconventional ideas, explanations, hypotheses or solutions to problems (Newmann, 1990) • Creating a new idea, a new object, or an artistic expression (Lewis and Smith, 1993) • Making a prediction (Lewis and Smith, 1993)	• Any statement which states a conclusion • Any statement which offers a summary of the point of view adopted • Any statement which states a belief or original perspective on the subject matter. • Any statement which proposes alternative solutions to problems. • Any statement which recognises that alternative approaches have different costs and benefits
Self-Regulation of Thinking	• Applying metacognitive skills (Vockell & van Deusen, 1989)	• Any statement which expresses an awareness of thinking processes or understanding. • Any statement which acts on awareness of thinking to affect change.

CHAPTER 3: METHODS

Participants

The participants in this study are students in Engineering and Computer Science enrolled in Impacts of Technology on Society (ENGR492) offered by Concordia University in Montreal. It is a mandatory course for fourth year students although a few third year students may decide to take the course earlier in their program. The main goal of the course is "to give an understanding of the issues and complexities of the role and effects of technology on society and the planet". The content is organized using a matrix framework whereby a set of different technologies is presented from different perspectives. Some of the dimensions covered include: environmental, social, ethical. Instructional approaches are varied and include lectures, class discussions, guest speakers, case studies, and the use of WebCT.

There are three sections of ENGR492 offered per term (Fall and Winter), which represents a potential sample size of approximately 300 students. Since this study was conducted in the Winter term, the potential sample was reduced to 150 participants. It should be noted though that for the winter term, enrollment in the Social Aspects courses tends to be lower than in the fall. Participation rate in this study for the three sections is shown in Table 3.1.

Table 3. 1 Total Number of Participants in the Study

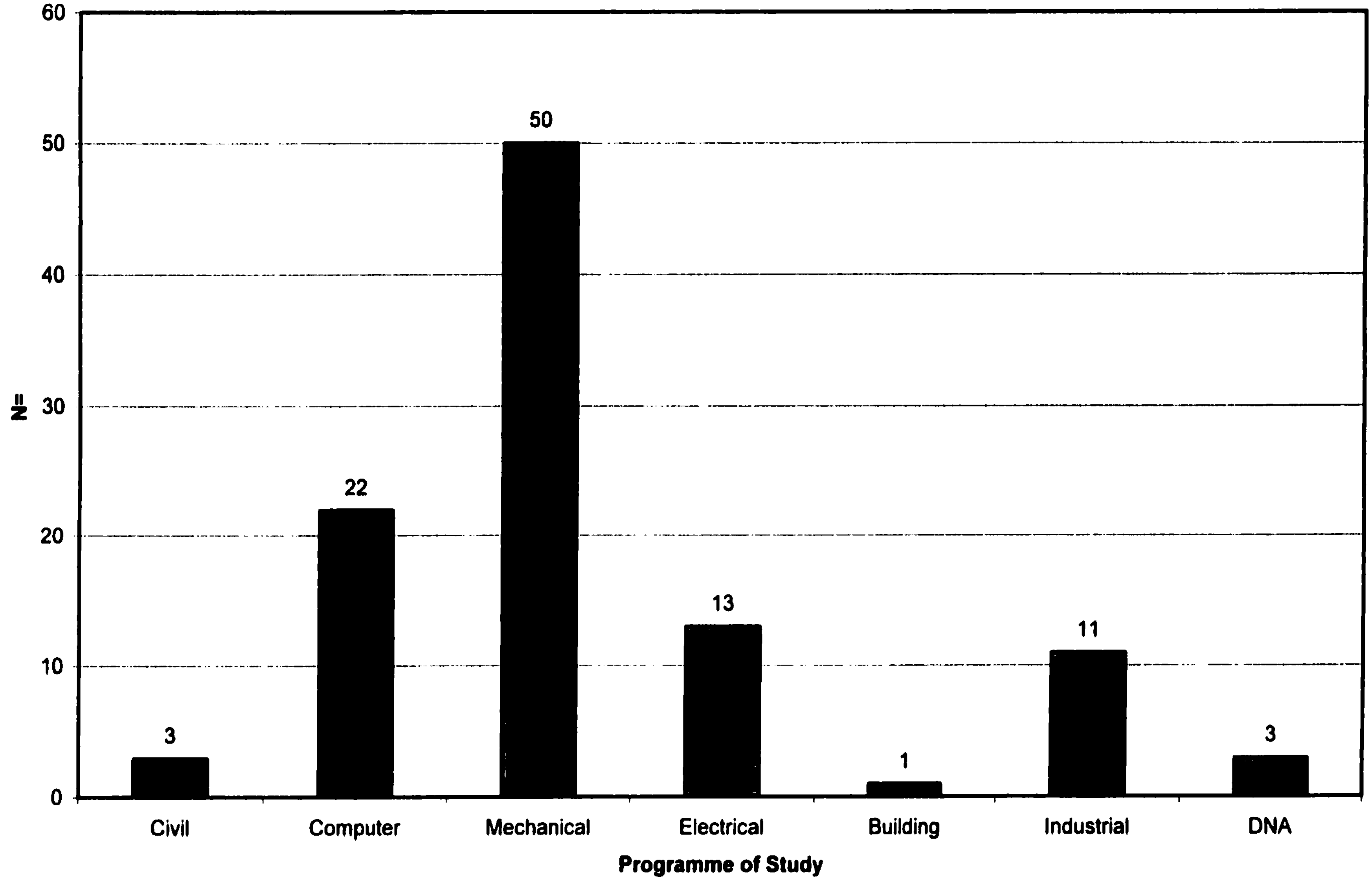
Sections	Condition	Enrollment Winter 2001	Agreed to participate in the study
Section A	Treatment	51	44
Section B	Treatment	33	30
Section C	Non-Treatment (Baseline)	32	29
TOTAL		116	103

Here it should be added that although the enrollment in section C for the winter 2001 was officially 51 students, only 32 students were present when I presented the project in class, thus affecting the potential size of the sample.

Out of 116 students, 110 filled out the consent form and the short survey (see Appendices II and III). Among those who filled out the questionnaire, 103 agreed to participate in the study, with female students constituting 13.6% of the sample (N = 14). This represents a participation rate of 93.6%. Participation in the study was voluntary. The students were told that their non-participation would not affect their grade. Nevertheless, if a student chose not to participate, he or she still had to carry out the tasks.

Unlike other engineering courses, ENGR492 sections are heterogeneous. Students come from all programmes offered by the University: civil, industrial, mechanical, electric, computer and building engineering. Figure 3.1 shows the distribution of the students per programme.

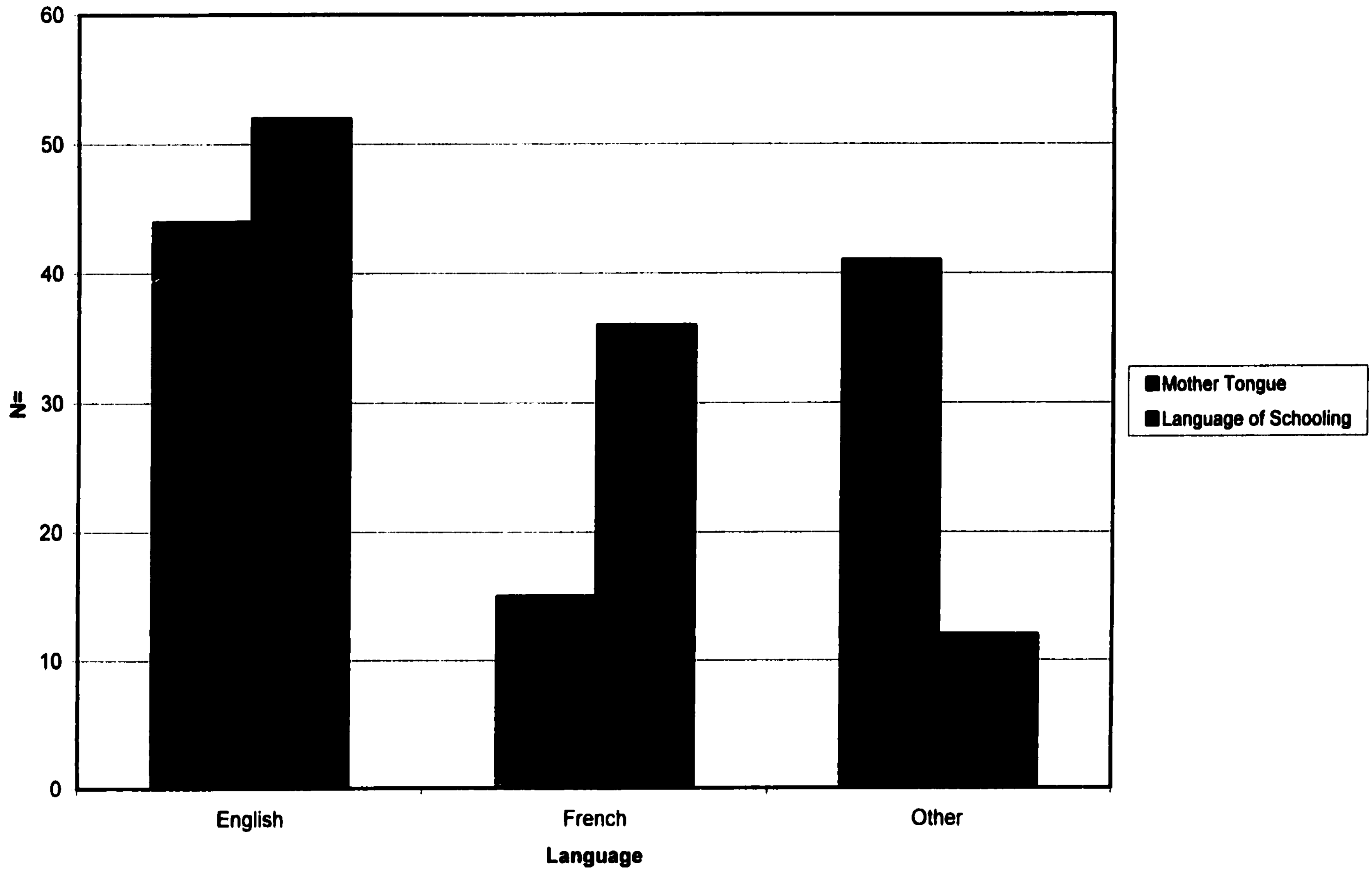
Figure 3.1 Distribution of Students per Programme of Study (N=103)



The majority of the students is registered in mechanical engineering (N = 50). Only three students come from civil engineering and one from building studies.

In addition, one of the most striking characteristics of the Engineering and Computer Science clientele is the diversity of ethnic origins. The distribution of the participants' mother tongue and language of schooling is presented in Figure 3.2.

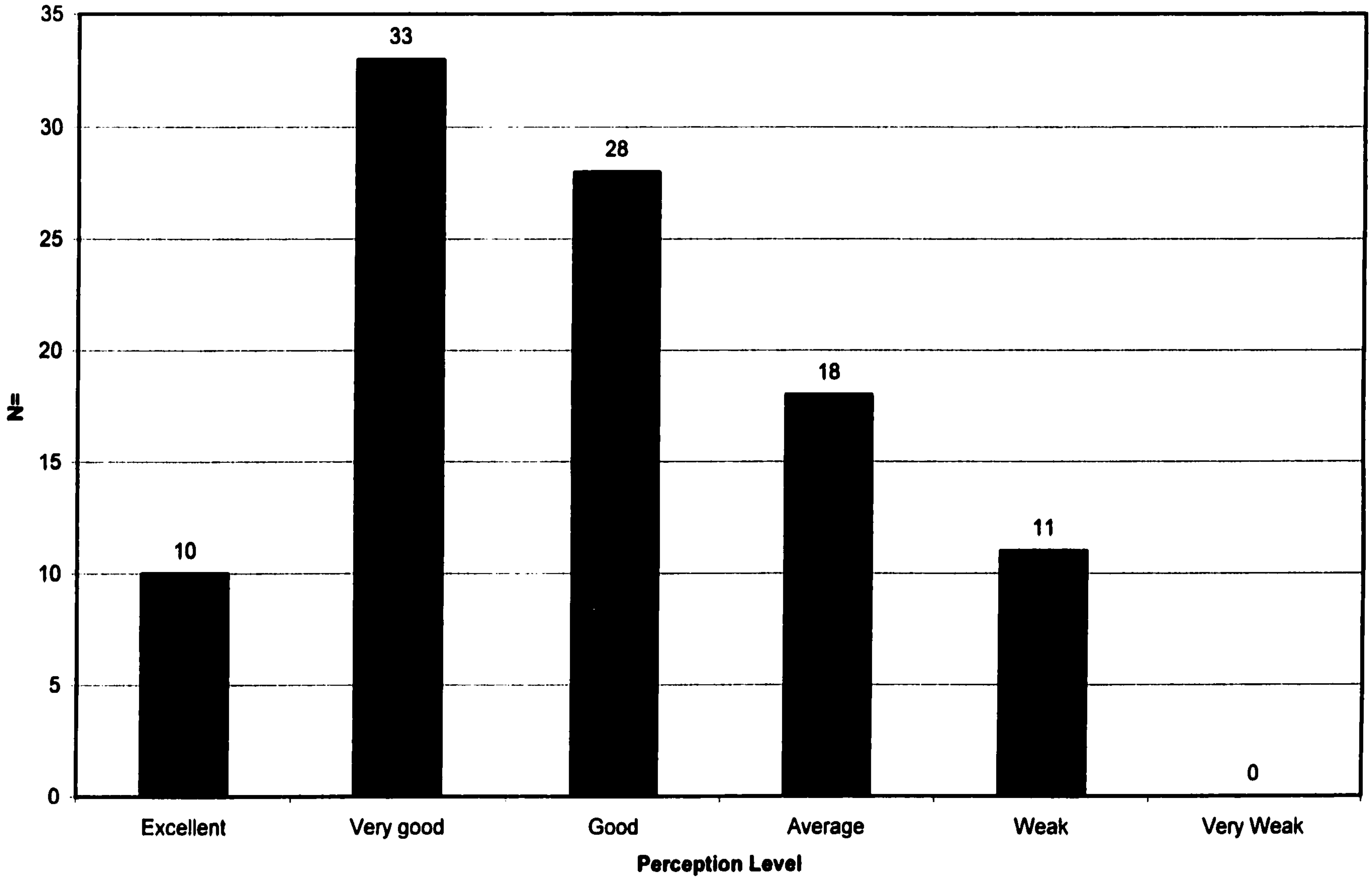
Figure 3.2 Participants' Mother Tongue and Language of Schooling (N=100)



According to the results on the item (N = 100), 44 respondents indicated English as their mother tongue, 15 indicated French, and the remainder indicated another language (N = 41) (see Figure 3.2) . The other languages listed are: Arabic (N = 7), Armenian (N = 2), Bengali (N = 1), Chinese (N = 3), Greek (N = 1), Gujrati (N = 2), Italian (N = 6), Korean (N = 1), Mandarin (N = 4), Persian (N = 1), Polish (N = 2), Spanish (N = 4), Tamil (N = 1) and Urdu (N = 1). Five participants did not specify their mother tongue. On the other hand, approximately half of the respondents (N = 52) reported they did most of their schooling (K-12) in English, 36 respondents indicated French and 12 indicated another language. From these results, we may say that approximately 14 allophones were required to attend French school by law (Bill 101). We may suppose that those students were studying in their third language. Finally, 13 respondents indicated a foreign student status as opposed to 84 Canadian students.

This information is relevant in so far as engineering students are not used to writing texts during their university education. They do not practice their writing skills to the same extent that a student in history or political science would for example. Accordingly, the students were also asked to evaluate their perception of the level of English writing skills. The results on the item are presented in Figure 3.3.

Figure 3.3 Perception of English Writing Skills (N=100)

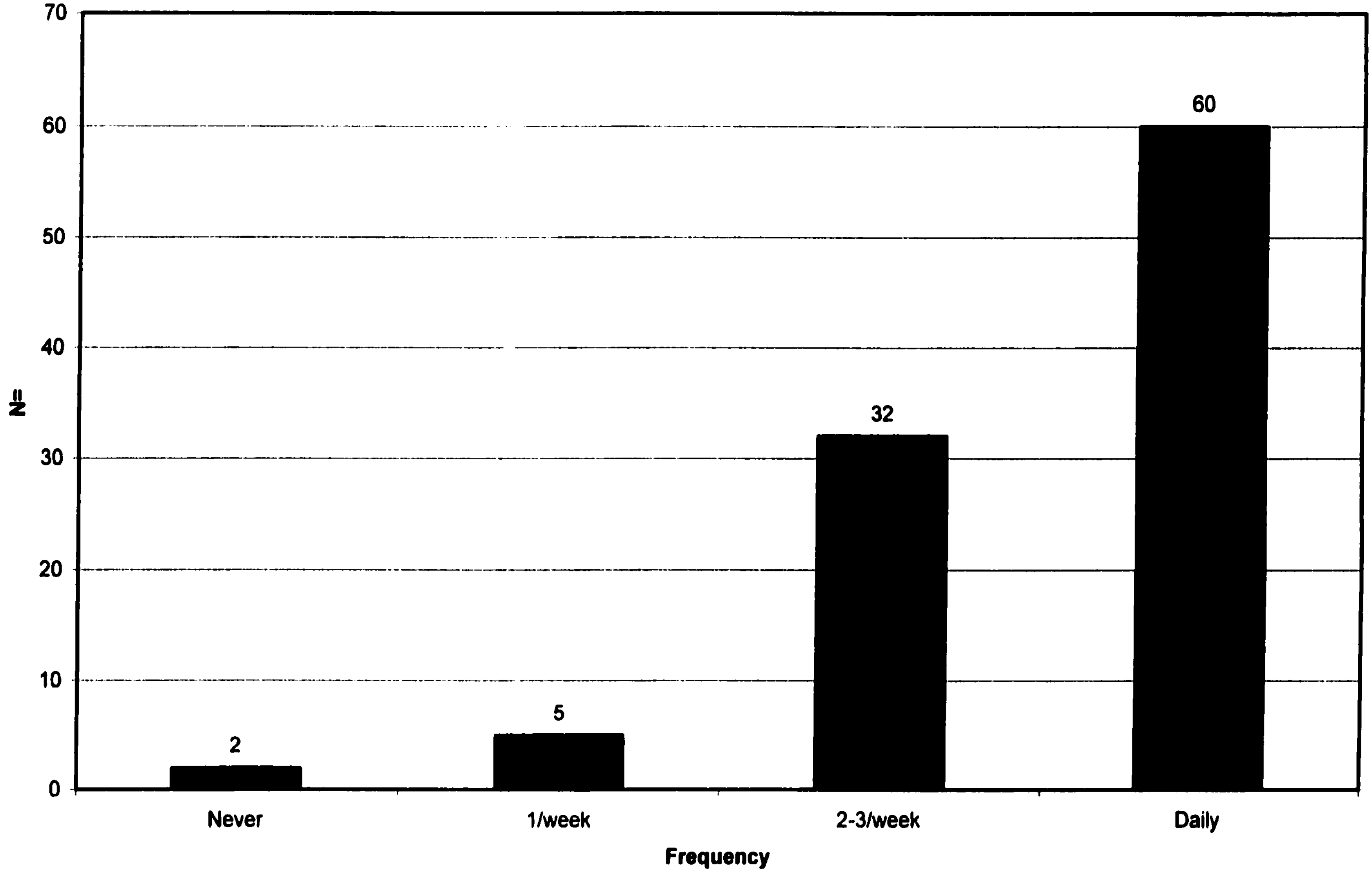


Out of 100 respondents, the majority perceived their English writing skills as "Very Good" (N = 33) or "Good" (N = 28). On the other hand, no student perceived their skills as being "Very Weak".

Respondents were also asked to indicate their current grade point average (GPA). The overall GPA for the sample is 3.01. The mean is slighlty higher for the non-treatment group (N = 29) with 3.13, than in the treatment group (N = 74) with 2.96.

Although computer skills are not the object of the study, it was still important to know to what extent the participants used e-mail tools, since the Tasks were to be posted in WebCT. As we can see on Figure 3.4, the majority of students (N = 60) use e-mail on a daily basis, and 32 do so two or three times per week. Only two never use e-mail.

Figure 3.4 Use of E-Mail (N=99)



Context of the study

The course ENGR 492 is usually taught by part-time faculty. The instructors were approached by the head of the Social Aspects programme about the possibility of conducting a study that would make use of WebCT. Two instructors agreed to participate and one declined. The latter nevertheless agreed to provide access to students' assignments, if there need be. The students did not know prior to registration in a section which professors were participating in the study. Finally, the current project was designed in collaboration with the intsructors and the head of the programme, taking into consideration both practical constraints and research requirements. The details of the project are provided below.

Number and type of assignments.

The team agreed that the WebCT activities would be individual written assignments as opposed to on-line collaborative discussions on given themes. This decision was made in light of the target population. Although engineering students occasionally have to work in teams during their university education, they are not used to exchange ideas and discuss social issues. In short, on-line group discussions would have entailed extensive modelling just to instil the culture of discussion.

In addition, we agreed that there would be five Tasks and the last one would count as the Short Paper. The first four Tasks were perceived by the instructors as a "practice run" for the Short Paper (Task 5). Furthermore, the team agreed that Task 4 should have the same structure as the Short Paper. The specific type and the order of Tasks was the responsibility of the researcher. In other words, the instructors did not know what exactly the students would have to do. The sequence of Tasks and the rationale for them is discussed in the Research Design section of this chapter.

WebCT was used mainly as a mode of communication to deliver the Tasks to the participants. All five tasks were made accessible to student at a specific time on WebCT site. Students could browse or download the instructions for their project (see Appendix IV). Since WebCT is a secure site, no other students had access to the course than those registered in the class.

Topics for the Tasks

The selection of topics for the Tasks was done in collaboration with the instructors. We used the following criteria to guide the choice of topic:

1. Relevance. The topic should be related to the content of the course. Once the topic was selected however, it could not be addressed directly in the classroom. In other words, the professor could not use the selected topic as a theme of discussion or as an example during the semester.
2. Richness and complexity. The proposed topic should be rich enough to allow for developing a theme from several perspectives. Hence, in light of the conceptual framework of the course, the topic could be analyzed from an environmental, social, political, economical, or ethical perspectives. Furthermore, positions on issues pertaining to the selected technology or the technological domain could not simply be right or wrong.

3. Not too technical. This criterion is related to the specificity of the population. Indeed, the team did not want to use topics that would render descriptions of how something works, instead of fostering an analysis of its impacts on society. Furthermore, it was essential to avoid specific topics that would have favored a given segment of the sample. For example, telecommunication technologies would have favored electrical engineering students, bridges, the civil engineerring students and so on.
4. Enough information available. This criterion was used mostly for instructional design purposes. The quantity of information available had to be sufficient enough to sustain four Tasks. The sources of information (articles, books and websites) had to be varied, as well as being accessible to both the researcher and the students. Finally, some of the information should be presented as graphics, tables or charts.
5. Interest. The topic had to be interesting enough to sustain the interest of the students over a four week period.

In addition to these criteria, we deliberately avoided topics that were considered sensitive or that drew on emotions or religious beliefs. Topics such as "human cloning" or "reproductive techniques" for instance, would be better discussed in class. Suitable topics for the project included: "Energy" and the specific types of power (hydro-electricity, clean energies, nuclear power, petrolium etc.); modes of transportation for goods (rail, air, water, land); communication technologies, medical and pharmaceutical and time keeping devices.

Partition of work

Another agreement reached among the team members is that the portion of the course conducted on WebCT was the responsibility of the researcher. In other words, as a strategy to avoid introducing counfounding factors in the study, the instructors agreed not to look at the outputs of the students. The professors avoided talking or prompting about the Tasks in class or giving clarifications to the students. In fact, questions and inquiries were redirected to the researcher via WebCT e-mail. Technical problems (registration or access to WebCT) were also sorted out by the researcher. Students were told so during the introductory session of the project (see Appendix IV). Any inquiries or complaints about the participation were directed to the researcher. The only exception to this was misbehavior on the part of the student.

Consequently, the introduction of the project, the presentation of the WebCT environment, the holding of the drop-in session, the posting and the grading of the Tasks and the debriefing session were the responsibility of the researcher.

Assessment of tasks

To ensure that the students would carry out the Tasks, we decided to allocate a participation grade for completing the Tasks. Hence, the first four Tasks were given a participation grade by the researcher, whereas the fifth Task was graded for content and quality of writing by both the instructor and the researcher (see Table 3.2). As mentioned earlier, since the Tasks were an integral part of the course load, students who declined to participate in the study still had to carry out the Tasks.

Table 3. 2 Assessment of the Tasks

Tasks	Type of assessment	Assessed by
• Tasks 1 to Tasks 4	• Participation worth 10%: • Technical aspects: doing the task on time and according to the norms • Effort	• Researcher
• Task 5 (Short Paper)	• Content assessment worth 20%	• Course instructor and researcher (marked all paper and shared comments. Adjusted the grades by discussing reasons).

The first four Tasks were worth 2.5% each. Two criteria were used to assess participation: a) conducting the task on time and according to the directives outlined in the Tasks Instructions document (see Appendix IV) and b) Effort put in the task as perceived by the researcher. The participation grades were given after the completion of Task 5 as a means not to bias the student during the data collection process. The first four tasks were not graded for content and the students were given no feedback indicating how well they were doing. The short papers on the other hand, received no participation grade and were assessed for their content according to the following criteria: relevance of the sources, clarity of analysis, synthesis, and so on. These criteria have been suggested by the programme director and are usually the grounds for evaluating student performance.

The rationale behind this "participation" strategy is that the students would get a chance to practice their critical thinking and writing skills before writing their Short Paper (Task 5). It was assumed that by not providing feedback, the students would take greater risks, become more creative while practicing. However, the absence of feedback probably had a negative impact on their motivation. Some students were probably wondering about how well they were doing. This in turn may have led the students to question the relevance of the learning experience and therefore spending less time on the tasks.

Debriefing session

At the origins of the research project, we planned the study without any feedback to the students. The participants would have received individual comments after Task 5. Given the pedagogical context, it seemed unfair to the students to carry out the Tasks without any feedback as to how well they were doing. Taking into account the research design, we could not give individual comments to the students, in case that triggered the use of critical thinking skills, thus causing a threat to the study. As a compromise, it was decided to run a debriefing session as a means to provide group feedback (see Appendix VI).

The debriefing session was led by the researcher and was planned between the fourth and the fifth tasks in both treatment groups. The pedagogical goal of the debriefing session was to increase the chances of transfer from the first four tasks to the Short Paper. Content of the debriefing included an overview about the sub-themes selected by the students, the different perspectives taken and the solutions proposed. Questions and feedback from the students were also addressed during the discussion. They could express their opinion about the topic, the type of documents and the purpose of the exercise. However, in order to prevent confounding the study, a specific definition of critical thinking skills was not provided to the students. The outline of the debriefing session is presented in Appendix VI.

Pilot study

A pilot study was run during the Fall semester 2000. The structure of the pilot was essentially the same as the study and served the double purpose of improving the instructional aspects of the project, as well as fine tuning the research procedures. A summary of the most important changes carried out is explained below. The complete list of changes is presented in Appendix VII.

<u>Design of Task 3</u>

One of the most important change conducted as a result of the pilot is the design of Task 3. Originally, Task 3 was an essay. The complexity variable was embedded in the information package provided with the Task instructions. For Task 3, the students were required to improve Task 2. The information package was in a separate document and no specific directives were given as to how to use it. What I expected is that the students would use the information as a way to support their arguments, but very few students made use of the information package. As a result, Task 3 was redesigned as a case study.

Another reason motivated the redesign of Task 3. It seems that moving from a regular essay (Task 3) to writing a scenario (Task 4) represented a gap that was too important to surmount for the perceived value of the exercise. In the debriefing session, the students clearly emphasized that the effort that had to be put in the exercise was greater than the grade value of the Task. Their comments led us to modify the grading scheme of the Tasks.

Grading scheme

As a result of the feedback given during the debriefing, the grading scheme was modified. In the pilot, the participation grade was given for carrying out the Task on time and according to the norms. In the study, we partitioned the grade into "effort" and "doing the Task on time and according to the norms". This strategy was adopted to improve the perceived value of the exercise.

Choice of topics

Another significant change that was made as a result of the pilot concerns the change of topics between Task 4 and Task 5. In the pilot, we used the topic "Impacts of the Motorcycles" in Tasks 2, 3, 4 and "Impacts of Printing" in Task 5. The main problems that arose is that first, the topic of printing was too broad, especially given the time allowed to carry out the task and second, the students did not have prior knowledge about the topic. Consequently, the students remained on the descriptive level instead of conducting an analysis of the impacts of printing. Furthermore, in several instances the students quickly surveyed the history of printing, leaped forward in time to offer clichés or refer to common knowledge about the internet. Clearly, the change of topic between Tasks 4 and 5 hindered the thinking and effort of the majority of the students. Consequently, instead of changing the topic for Task 5, we decided to use a related one.

Instruments

Two instruments were added in the final study: a log of time spent on the tasks (self-reported measure) and a post-task questionnaire (Appendix VIII). Both instruments were created using the "quizz" feature available on WebCT. Once the questionnaire or the time-logs were completed and officially saved, the students could not go back and change their answers.

Drop-in sessions

In addition to the two in-class presentations conducted at the beginning of the semester, two informal drop-in sessions were planned for students who might have difficulty creating their account or logging on to WebCT. The two sessions were held on Friday mornings of the first two weeks of classes. No student droped-in to seek help.

Methodology

Experimental design

The study utilized a quasi-experimental time-series design with the participants in the treatment condition treated as a single-group. The treatment consists of five WebCT tasks to be carried out over a period of five weeks. Outputs of the tasks provided the research data, as shown in figure 3.5.

X1O1 X2O2 X3O3 X4O4 X5O5

O(Non-Trtmt)

Where X = Treatment and O = Output

Figure 3. 5 Experimental Design

O1 and O2 were used to establish a baseline for the current critical thinking skills of the group. Complexity is introduced in O3, O4 and O5. A significantly greater number of occurrences of critical thinking skills should appear between O2 and O3 and between O3 and O4. Evidence of transfer should appear between O4 and O5.

Treatment variables

The manipulated independent variable

The manipulated independent variable is the level of complexity integrated into the design of the tasks. Complexity in this study is characterized by difficulty in the instructional task, which shifts from a typical essay assignment (Tasks 1 and 2) to a case study (Tasks 3, 4, and 5). According to Leenders and Erskine (1989), the degree of difficulty in a case can be understood by using three major dimensions: analytical, conceptual, and presentation. Each dimension is subdivided into three levels of difficulty. When designing a case study for students, any combination of level of difficulty and dimensions may be applied, depending on the desired learning outcome. A summary of their guidelines is presented in Table 3.3.

Table 3. 3 Three Dimensions of Case Difficulty

Level of difficulty	Analytical	Conceptual	Presentation
1	• Problem and solution is presented • Ask student what they think	• Concepts are easily grasped by the majority	• Limited extraneous materials • Data neatly presented and straightforward
2	• Problem is given • Students provide a reasonable solution	• Concepts require further clarification through class discussion, practice or reinforcement	• Average amount of information • Some extraneous data
3	• A situation is given • Students have to find problems and solutions	• Concepts require extensive clarification through class lectures • A minority is expected to understand	• Large amount of extraneous data • Information is less organized • Requires sorting data

In this study, the focus is on the analytical and the presentation dimensions. In Task 3, complexity is introduced as level two of difficulty on both dimensions. Hence, students are given a problem for which they have to find a reasonable solution. Data are presented in a relatively straightforward fashion and include some extraneous information.

Tasks 4 and 5, in contrast, reflect a level three of difficulty on both the analytical and presentation dimensions. A situation is depicted to the students. Their goal is to identify the problem(s) and offer relevant solution(s). A set of data, from different sources, some which are more or less relevant, is presented to them. They might or might not use the data; if they do, they have to select the appropriate facts and figures in order to support their positions. Again, engineering students are trained to find the best possible solution to a well-defined problem. Within this study, the researcher creates an ill-defined problem for which there might be more than one acceptable solution.

Feedback in this study is a controlled variable. The facilitator (i.-e. the researcher) does not provide comments about the content, skills or approaches that might influence the students' performance on the tasks. The only type of individual feedback given concerns presentation: students might be told how to format their messages or how to send attached files properly. If broader advice or instructions need to be delivered, a comment (document or message) would be posted on WebCT for the benefit of the whole group.

Non-manipulated independent variables

A short survey gathered demographics, student characteristics, such as mother tongue, language of schooling, programme of study and other variables which might affect performance (see Appendix III). In addition, the students could give feedback in a post-task questionnaire avalaible on WebCT, about their interest in the topic, the type of tasks they found most engaging or least relevant, and so on. (see Appendix VIII).

Another non-manipulated independent variable is English writing skills (hereafter E.W. skills) to see whether or not there is a relationship between writing skills and critical thinking skills. The students' writing skills were evaluated by an expert in "Teaching English as a Second Language". More details are provided later in this chapter and the results of the E.W. skills assessment are presented in the next chapter.

Dependent variables

The content of the students' work on the five tasks was coded using a revised version of Herrington and Oliver's indicators of critical thinking skills. Frequencies of critical and non-critical thinking skills were recorded. The main indicators of critical thinking skills were: "Judgement and Interpretation", "Multiple Perspectives", "Imposing Meaning" and "Metacognition". The Non-C.T. units categories were "Statements" which includes factual and declarative sentences, opinions and beliefs, and "Varia" which is comprised of any statement which is "Organisational", "Procedural", "Not a Sentence", or "Direct Quotes". The coding book is presented in Appendix XX.

It must be added that although the instructional approaches (essay, case study, simulation) selected to conduct this research project is not conducive to foster metacognitive skills, the indicator "metacognition" was kept from the original model. In rare occasions, some students were being reflective about their own thought processes. Finally, another exploratory dependant variable was added to the coding book which is the "Use of visuals". "Visuals", in this study, include any graphic representation of data such as tables, graphs, figures, pie charts of even pictures. To a certain extent, "Visuals" may be used in the same way as a quote, that is, to illustrate, to support, to challenge a statement.

As mentioned earlier, time spent on task has been recorded. Students were asked, via WebCT, to keep track of their time spent on researching, reflecting and writing. The results are presented in the next chapter.

Other variables

The coders' log, in which the time spent coding a Task and the sequence of coding were recorded, were collected (see Appendix IX). The purpose of the coding logs was to gather information about potential order effects of coding. In addition, the raters were instructed to use the "memo" feature in Nud*ist to document any hesitation between codes and subsequently to outline their reasonning supporting their decision. Finally, the raters kept a journal in which they noted their reflections, ideas and suggestions. The inter-rater reliability checker used the same note keeping techniques—memos and journal--as the coders, but filled out a different coding log (see Appendix X).

Instruments

A complete list of the instruments used in this study is presented in Table 3.4

Table 3. 4 List of Instruments

Instrument	Data generated by	Appendix
Consent form	Participants	II
Short Survey questionnaire eliciting the following information: Personal information, mother tongue and language of schooling, computer skills, job etc.	Participants	III
Log book for students: Time-spent on Task (in WebCT)	Participants	VIII (Item A)
Post treatment questionnaire to students: Effort put in Tasks, Interest in Topic, etc. (in WebCT)	Participants	VIII
Coding Book based on Herrington & Oliver's model of indicators of critical thinking skills	Participants	XX
Coder's log book for time spent and order of coding	Coders	IX
Coder's notebooks	Coders Reliability checker	N/A
Nud*ist memos to document coding hesitations	Coders Reliability checker	N/A
Coders' Post-Task Questionnaire	Coders Reliability checker	XXI
Inter-rater reliability checker's log book	Reliability checker	X
E.W. skills criteria (include this in instruments?)	E.W. skills evaluator	In text

Originally, I planned to collect data about the WebCT frequencies, that is to identify which pages were visited and how often. However, the WebCT tool is not precise enough to follow the specific itinerary of a given subject. At best, it reveals information about login frequencies or last access to WebCT.

The order and type of tasks

The students in the treatment condition had to complete five different assignments over a period of six weeks. There were two essays, one case study and two simulations. Students had one week to complete their task and return it to the researcher via WebCT. The order and type of Tasks is presented in Table 3.5.

Table 3. 5 Order and Type of Tasks

Number of Tasks	**Type of Task**	**Topic**
Task 1	Essay	Technology
Task 2	Essay	Food Production
Task 3	Case Study	Food Production
Task 4	Simulation	Food Production
Task 5	Simulation	Genetically Modified Foods

The Tasks increased in difficulty from relatively simple as represented by the essays (Tasks 1 and 2) to the more complex and time consuming ones (Tasks 3, 4, 5). The rationale was as follows: Engineering students tend to have very little experience in writing essays and other "freer form of assignments". Their learning is usually assessed during exams or quizzes (Hawe, 1991). Their only other writing experience occurs in the technical writing course. So the essay format was selected to provide a warm up to the students from a pedagogical standpoint, and to create a baseline of critical thinking skills.

Procedure

In the next section, the procedure used in each component of the study are described, respecting the logical unfolding of the project. The content revolves around three main axes: design of the Tasks, data collection and coding and reliability procedures. First, an outline of the process of instructional design of the Tasks is presented. The details of the data collection procedure follow. Thirdly, the steps involved in a content analysis are explained. These include data management strategies, testing the instrument, training the coders, coding procedures, inter-rater reliability measures, and reliability checking procedures. The last process described is the assessment of E.W. skills.

Instructional Design procedures

For all five tasks, the students were given few directions, so they knew what to do, but had to decide what to discuss and how to approach the assignment. The design of the assignments offering minimal directives is in line with what Resnick (1997) has suggested about "school culture", to prevent students trying to please the teacher ("what do you want me to do"?). By allocating more freedom, it was believed they would select a motivating approach and choose their own strategies to carry out the task. For example, some students for Tasks 1 and 2, wrote a position paper, others were more reflective, and the majority analysed pros and cons. For Tasks 4 and 5, some students used the opportunity to write a dialogue or a scenario; others wrote a more classic essay. The tasks instructions are presented in Appendix IV.

As explained previously, since the students had only one week to carry out each task, it was decided to provide only one topic for Tasks 2 to 4 (food production), and a related topic for Task 5 (genetically modified foods), so they would spend less time seeking information, facts and data, and focus more on the processing of the information they had. This choice is in keeping with the literature, which suggests that "successful application of critical thinking skills requires, among other things, a knowledge of the subject matter and experience in the area in question [...]" (Norris, 1985: p.44; see also Garrison, (1991); and McMillan, (1987)). However, the danger with keeping the same topic over at least one month is that the students might get bored, and lose interest in the issues, especially if they don't like the topic.

Instructional design of the tasks

Since Tasks 1 and 2 were essays, they required little instructional design and consisted of instructions provided to the students (see Appendices XI and XII). Task 3, on the other hand, was designed from a real situation (Appendix XIII). A food emergency report was found on the website of Food and Agricultural Organisation (FAO)[1], a sub-branch of the United Nations Organisation. For this study, the case of Armenia was selected (October 2000). The name of the country was changed to "Sayahn" in order to prevent the students from looking for supplementary information and the content of the report was simplified to match Lenders and Erskine's (1989) complexity variable level two.

Tasks 4 and 5 were created using the main guidelines for designing educational simulations as proposed by Jones (1985) and Thiagarajan and Stolovitch (1978).

---

[1]http://www.fao.org/WAICENT/faoinfo/economic/giews/english/alertes/2000/srarm100.htm.
Several parts of the Emergency Report were used with little editing, which raised the issue of copyrights. In order not to divulgate the exact address of the document, the general FAO addressed was used in the

Accordingly, the content of the tasks was organized around four elements: the role or the function of the student as represented by the Curriculum Vitae (consultant in communication), the mandate, which was explained in the producer's letter to the consultant, a description of the context, depicted in the "Network Bio", and the facts necessary to conduct the simulation (information file). The complete simulations are provided in Appendices XIV and XV.

Regarding the information file, the selection and the organization of the facts included the document were also done following Lenders and Erskine's model (Complexity variable level 3). Most information, both relevant and extraneous, was gathered on the Web and sorted by type: graphics, full articles, references for print-based materials and web-site addresses.

It should be added here that despite the design guidelines used for Tasks 4 and 5, the term "simulation", in this study, is used rather loosely. In fact, "simulations" or "role plays", are usually enacted as a group. In instructional contexts, they are used to draw primarily on emotion and the learning experience is supposed to be catalyzed in the debriefing period. Since Tasks 4 and 5 were designed as individual tasks, it would be more appropriate to call them "*mise en situation*", that is "complex case studies". However, the term "simulations" is used in this study for reasons of clarity.

The tasks were piloted in the Fall 2000 (see Appendix VII for list of changes made to the tasks). In addition, all modified tasks used for the data collection were formatively evaluated by experts. An instructional expert assessed the clarity of the

---

document. The complete source of information, including the country's name, was given to the students at the debriefing session.

directrives in Tasks 1 and 2. Tasks 3, 4 and 5, on the other hand, were evaluated by at least two individuals: one subject-matter expert in agricultural and environmental issues and one instructional expert. The subject-matter expert was asked whether the information was complete and if not, to identify which issues were missing. The instructional expert validated the format, edited the content, and checked for potential bias (leading the students). In addition, the expert validated the treatment variable, that is, he verified the level of complexity (see Table 3.3) embedded in the assignments. In addition, Task 3 was validated by a second instructional expert who was asked whether one could find more than one solution given the information provided. In each instances, the experts evaluated a printed version of the html document that the students would get (see Appendices XI to XV). The final version of the tasks however, was not tested with individuals of the target population.

Data collection procedures

Sequence of events

On the first week of classes, I visited the two sections of Engr492 which by prior agreement with the instructors were in the treatment condition. The instructor explained the course outline, then let me describe the purpose and the grading of the WebCT tasks and how the students would benefit from doing the exercise (see Appendix V).

The students were given a handout developed by the Instructional and Information Technology Services (IITS) which contained all information about how to use WebCT. Then followed a live demonstration of how to logon. The students were also shown how to use the various WebCT features of the site and how to navigate within the site.

Students were then told that all instructions for assignments and Tasks instructions would be posted using WebCT, including the course outline. Finally, directives for formatting assignments and the procedures for sending the Tasks to the researcher concluded the demonstration session. The Tasks instructions are presented in Appendix V. A question period concluded the demonstration session.

Then, I presented the project and how their work would be used to conduct the research. Students were assured that their participation was entirely voluntary and that their non-participation would not affect their final grade. They were also told that agreeing to participate did not mean extra work or additional meetings other than the ones planned. The students then signed the consent forms and completed the short survey (see Appendix III). The questionnaire was not anonymous but students were assured confidentiality. Finally, the students were told that if they had problem logging on, a drop-in session had been planned on the next two Fridays.

I went back in the groups on the second day of class for a shorter presentation for students who missed the first class. The students who were there the first time could either stay in the classroom or take a hour break. The following week the first Task was posted on WebCT. As mentioned previously, with the exception of the last Task, the students had one week to complete their work. Tasks were posted on Wednesday by midnight (23:55) and they had to be returned by the following Wednesday (23:55), as shown in Table 3.6.

Table 3. 6 Calendar of Events

Week of:	Event
January 1	• Class presentation of WebCT project (section A) • Consent Forms and Short questionnaires
January 8	• Class presentation of WebCT project (section B) • Consent Forms and Short questionnaires • Second class presentations for both sections • Wednesday 10 (23:55): Posting of Task 1 • Friday 12: Drop-in for coaching on logging in WebCT
January 15	• Wednesday 17 (23:55): • Posting of Task 2 • Task 1 due • Friday 19: Drop-in for coaching on logging in WebCT
January 22	• Wednesday 24 (23:55): • Posting of Task 3 • Task 2 due
January 29	• Wednesday 31 (23:55): • Posting of Task 4 • Task 3 due
February 5	• Wednesday 7 (23:55): • Posting of Short Paper Assignment (Task 5) • Task 4 due
February 12	• Debriefing about Tasks 1 to 4 • Sunday 18 (23:55)* • Short Papers (Task 5) due • Completion of the questionnaire post hoc questionnaire about the tasks.
February 19	• Posting of the participation grade (Tasks 1 to 4) on WebCT.
March 8	• Researcher visited the third section (non-treatment group) to obtain consent
March 15	• Collection of the assignments non-treatment group

*Originally, Task 5 was due on Wednesday February 14. Because of the Mid-Term exam period, both instructors gave an extension until Sunday, February 18.

Presentation of the project to the non-treatment group

On March 8, I visited the third section of Engr 492 which was used as a non-treatment group, in order to establish the baseline of C.T. skills of the sample. The goal of the visit was to obtain permission to access student's second assignment in order to conduct a content analysis for critical thinking skills. Thus, I introduced myself, presented the main goal of the study and, as for the treatment group, I briefly explained what were critical thinking skills (see Appendix IV) and how this study would contribute to knowledge. Then the students were given the consent form and the short survey to fill out. Students were told that participation in the study was voluntary and that their non-participation would not affect their grade for the course. They were also told that the information on the short survey, although not anonymous, was confidential. Thirty-two out of a potential 51 students were present on the day of the presentation; 29 accepted to give the researcher access to their assignments. Out of the 29 participants, 25 handed in their assignment on time; therefore only 25 were used in the study.

Coding procedures

Selection of coders

For the study, I created a team of three coders, including the researcher. To select the candidates, I posted a message on the FirstClass server in the Education Department of Concordia University. Then I invited the potential candidates to an information meeting where I presented the research project, described the coders' task and what were my expectations, and outlined the constraints of the projects. Following a question period, the candidates were asked to complete a short questionnaire (see Appendix XVII).

Approximately 12 graduate students of all Education programmes attended the information meeting and eight of them completed the questionnaire. The selection criteria were "desire to gain from a new experience", "understanding of the responsibilities of the coder", "commitment to the project", and "availability".

The selected candidates were both second year Master's students registered in the Educational Technology programme. Rater B has a bachelor's degree in Teaching English as a Second Language. At the time of the training, she was working as an elementary school teacher. She had prior experience in coding, but not in the field of critical thinking. Rater C also has an undergraduate degree in Teaching English as a Second Language, as well as graduate diplomas in Linguistics and French Literature. Rater A (the researcher) has a double major in History and Liberal Arts and a Master's degree in Educational Technology. I also know the student population, for having taught as a part-time faculty in the Social Aspects of Engineering Programme. None of the three raters had prior experience in the field of critical thinking.

The inter-reliability checker was also selected during this process. A Ph.D. student in the Educational Technology Programme, she has a Bachelor's and Master's degrees in Psychology. In contrast to the coders, she had prior experience in conducting her own study on critical thinking skills, but in different conditions (online environments) and with another critical thinking model.

The background of the raters is relevant insofar as the validity of the instrument is concerned. Harwell (1999) contends that similar experience and background of the coders is an important factor to consider but is seldomly documented. In addition, it would have been preferable for the researcher not to participate in the coding to control for severity threat. However, practical constraints –the relatively large number of essays to code in comparison with the small honorarium received and the limited time frame—had to be considered. Dividing the work in two would have created other problems affecting the motivation of the coders (Kneuendorf, 2001) and the quality of work.

<u>Training of coders</u>

In order to reach an acceptable level of agreement between coders, a training session was conducted. Coding for critical thinking skills is a challenge since the content is open to interpretation. The challenge is to establish a common understanding of the codes so the raters will be able to apply the coding book reliably.

The raters participated in a training session designed by the researcher. The training workshop was planned as a series of individual practice and group debriefing sessions. There was no specific duration allocated for the training. The coders were told that when the inter-rater agreement reached a satisfactory level, we would be able to move to Round 1 (coding used to estimate the inter-rater reliablity; for this, 10% of the Tasks were coded by all three raters). Though three practice sessions had been planned originally, we conducted six, which in the end focused on skills that were more troublesome to code or types of Tasks that were technically more difficult to assess, such as a discussion scenario. Although the training session focused mainly on becoming fluent at recognizing the critical thinking skills, other methodological procedures were presented, such as note taking, reflections on the state of mind and using the Nud*ist software to view raw data and record codes.

The coding practices were conducted using work produced by students who participated in the pilot project, thus preserving the research data for analysis. After practicing, the coders met in the debriefing session to discuss the difficulties encountered with the critical thinking definitions. One of the benefits of the practice-debrief approach is that the definitions were enriched by each of the rater's experience, instead of following a unilateral view imposed by the researcher.

The ultimate goal of the training session was that all three coders, as well as the inter-rater reliability checker, understood each other's interpretation of a given definition in order to reach a common agreement over practical examples: "In this situation, one would do this".

It should be added that Raters B and C and the inter-rater reliability checker were not aware of the research hypotheses of the study. Information about the population, the procedures, the rationale for the instructional tasks and the conditions of the data collection was not conveyed until after the coding was completed.

Finally, the training session was evaluated by two instructional experts. The first one conducted an informal formative evaluation. She offered comments on the clarity of the directives, the length, the interest, the format, the sequence of information. The second expert evaluated the replicability of the training session, whether or not a person could take the training manual and be comfortable and able enough to be able to code essays. The Coder's Training Manual is presented in Appendix XVI.

Raw data management

A code number was assigned to each document in order to protect the student's identity and prevent, during coding, any biases related to the nationality of the students.

The new codes included information about the course section (treatment and control), the student's identity number, the rater's identity and the Task number and type.

The raters were randomly assigned to Tasks. To do so, a 3 X 3 matrix (A,B, C; B, C, A; C, A, B) was used to generate the codes. The matrix was converted into a string which was assigned to all Tasks 1. Then the "if-then" formula in Excel was used to generate automatically the rater's code for the subsequent tasks. By using this strategy, the raters coded at least one Task per student, but never more than two. Also, the type of Tasks was evenly distributed among the raters.

Test of the instrument

The original instrument was tested on Tasks drawn from the pilot study. The operational definitions were further defined and the less frequent and more complex categories were identified. Examples of each coding category were extracted from the documents and included in the Coder's Training Manual (see Appendix XVI).

As a result of testing the instrument, I realized that it was necessary to code not only the critical thinking skills but also the units that were not considered to be critical thinking. Consequently, the categories "Varia" and "Statements" were added to the coding book. This strategy was followed to alleviate the cognitive load of the coder. In other words, it was easier to code every text unit, than to first decide whether or not a unit should be coded and second, to decide which code it should bear.

As the instrument was tested, strategies for handling more difficult codes were outlined. Specifically, three strategies were used: 1) identify the key word in the unit; 2) contextualize the unit; and 3) question student's intention. The details of this strategy are outlined in the Coder's Training Manual in Appendix XVI.

Coding procedures

Only the summary of the coding procedures is presented below. The detailed procedures can be found in the Coder's Training Manual in Appendix XIV.

- No multiple coding allowed, that is one code per text unit;
- In the case of indecision, the coder was required to indicate the two codes in a Nud*ist memo and then decide on one. Reason supporting the decision had to be provided;
- Coding on paper was permitted but the codes had to be transferred in Nud*ist without changing th original codes (this allowed the coders to work at home);
- Changing coded work was not allowed unless technical mistakes were detected;
- Coders were asked to vary the type and length of Tasks (not allowed to code all Tasks 1 for example);
- Coders were asked not to consult each other;
- Coders were trained to disregard the quality of expression;
- Coders had to keep a journal of thoughts and a checklist to indicate the order of coding.

Reliability procedures

In order to estimate the inter-rater percentage of agreement, the following steps were taken. First, a sample of 10% (N = 37) of Tasks was drawn from the original pool of data. The selection was random, although the distribution took into account "representativeness" of the data. In short, a proportional number of tasks was selected taking into account the type of task, the rater's code and the group identity. Only one task per student was selected. Although no specific measure was taken to compensate for unusual or less frequent codes, at least three instances of more challenging formats, such as dialogues and letters, were included in the sample.

The coders were then reminded of the coding procedures. In particular, the coders were asked to avoid consulting one another, unless a technical problem arose. We agreed on a 10-day deadline to conduct the rating of the tasks. The coding was done primarily in Nud*ist, although some was carried out on paper and then transferred into Nud*ist.

Once the tasks (N = 37) were coded, I exported the frequencies on the coding categories into an Excel sheet. Then, the raw frequencies for each category were converted into proportions of text units. This was done for each Task and for each coder. For example, the total frequencies of document "101b_t4" for the category "Judgement and Interpretation" were divided by the total number of units in the document, as coded by Rater A. The same steps were carried out for each category and for the other two coders. Once all proportions were calculated, a coefficient of reliablility was established by pairing the coders for each of the coding categories. The overall reliability coefficients were obtained by suming the proportions of each of the C.T. categories ("Judgement and Interpretation", "Multiple Perspectives" and "Imposing Meaning") and the Non-C.T. categories ("Statements" and "Varia"). The reliability coefficients are presented in Table 3.7.

Table 3. 7 Estimate of the Inter-Rater Reliability of the Coders

Critical Thinking Categories		
JUDGEMENT AND INTERPRETATION (1)	PAIRS	Correlation coefficient
	A_B	0.68
	A_C	0.66
	B_C	0.63
MULTIPLE PERSPECTIVES (2)		
	A_B	0.48
	A_C	0.57
	B_C	0.64
IMPOSING MEANING (3)		
	A_B	0.55
	A_C	0.54
	B_C	0.69
Non-Crititical Thinking Categories		
STATEMENTS (5)		
	A_B	0.36*
	A_C	0.51
	B_C	0.48
VARIA (6)		
	A_B	0.96
	A_C	0.95
	B_C	0.97
Overall Coding		
CRITICAL THINKING	A_B	0.58
	A_C	0.63
	B_C	0.76
NON-CRITICAL THINKING	A_B	0.25
	A_C	0.48
	B_C	0.72

*p<.05

It should be added that two categories were excluded for the estimation of the reliability coefficient: "Use of Visuals" and "Metacognition". As mentioned earlier those categories were included on the coding book for exploratory purposes only and consequently are not an object of interest in this study.

Although the reliability coefficients are moderate, all pairs but two are significant at a confidence level of .01. In the "Statements" category, the pair "A_B" is significant at .05, whereas the same pair is not significant in the overall section "Non-critical thinking skills".

Given the relatively low reliablity generated by Rater A, that person should have been removed as a coder. The pre-assigned Tasks of Rater A should have been redistributed to the other two coders for the final round of coding. However, as mentioned before, practical constraints prevented the adoption of the option.

Before we proceeded with the coding of the Tasks, a brief refresher on the unusual codes and the technical procedures was provided orally to the raters. The coding of the Tasks was carried out over six weeks. Rater B worked part-time on the coding, whereas Raters A and C worked full time.

Finally, the sample used in the reliability round was integrated with the rest of the data. Only one Task per coder had to be included though. The documents were introduced in the final sample by using the rater's identity.

Inter-rater reliability checker procedures

It is often said that even triangulation of measures is not enough to ensure reliability in a content analysis study or in observational studies. In fact, Bakeman and Gottman suggest that the use of a reliablity checker has a psychological effet: the raters tend to be more consistent and focused on the job at hand. They suggest that the checker focuses on infrequent codes and rating consistency. However, given the nature of the study, it was not really feasible to ask a person to check for the right or wrong answer.

Consequently, the procedures were modified to fit our situation. The checker had to select a sample of 51 tasks and had to use the following criteria. She could not select a student's work more than once; she had to have an equal number of type of tasks (i.-e., Task 1, Task 2, etc.) and she had to have an equal number of tasks coded by Rater A, Rater B and Rater C. She was not coding the entire text, rather only ten units selected either randomly or within the first third, the second third or the third third of the text. The inter-rater reliability checker procedures are provided in Appendix XVIII.

The last step to estimate the inter-rater reliability would have been to compare the codes generated by the checker with the work of the raters. However, due to the complexity of the procedures and the difficulty of assessing partial sections of text, the measures failed. Specifically, the strategy to select ten units in the first, second of third third of the text created a halo effect (Harwell, 1999). The reliability checker, especially when coding in the last third of the paper, had difficulty making sense of the units. To compensate, she had to go back at the beginning of the text to understand the main arguement of the student. For that reason, the data generated by the inter-rater reliability checker were not used in the study.

Preparation of data

In the treatment group, students's Tasks were sent by e-mail via the WebCT environment. The electronic files were either Word, Word Perfect, HTML, ZIP or text documents. In order to analyze the Tasks in Nud*ist, the documents had to be converted into text files. Also, recognizable text units (sentences) had to be inserted in the text. The steps involved in the preparation of the documents are presented in Appendix XIX.

For the non-treatment group, we received the paper version of the assignments. The documents were first scanned and then processed, following the steps outlined in Appendix XIX.

*Nud*ist set up*

All documents were analyzed and coded using Nud*ist. The three raters were assigned their own set of tasks to code. All three projects used the same index tree, which had been created a priori by the researcher (see Figure 3.6). The index tree of each project was exactly the same for all three raters. The index tree was created to match the coding book (Appendix XX). For example, the code for the skill "Challenges a point" under the category "Multiple Perspectives" corresponded to the node (2 1). It should be added that the index tree of the codes is not representative of any hierarchy or taxonomy of the critical thinking skills. Once the coding was completed, the three Nud*ist projects were merged into one using the merging patch. Finally, the sample of documents that was used to estimate the inter-rater reliability were merged into the main Nud*ist project.

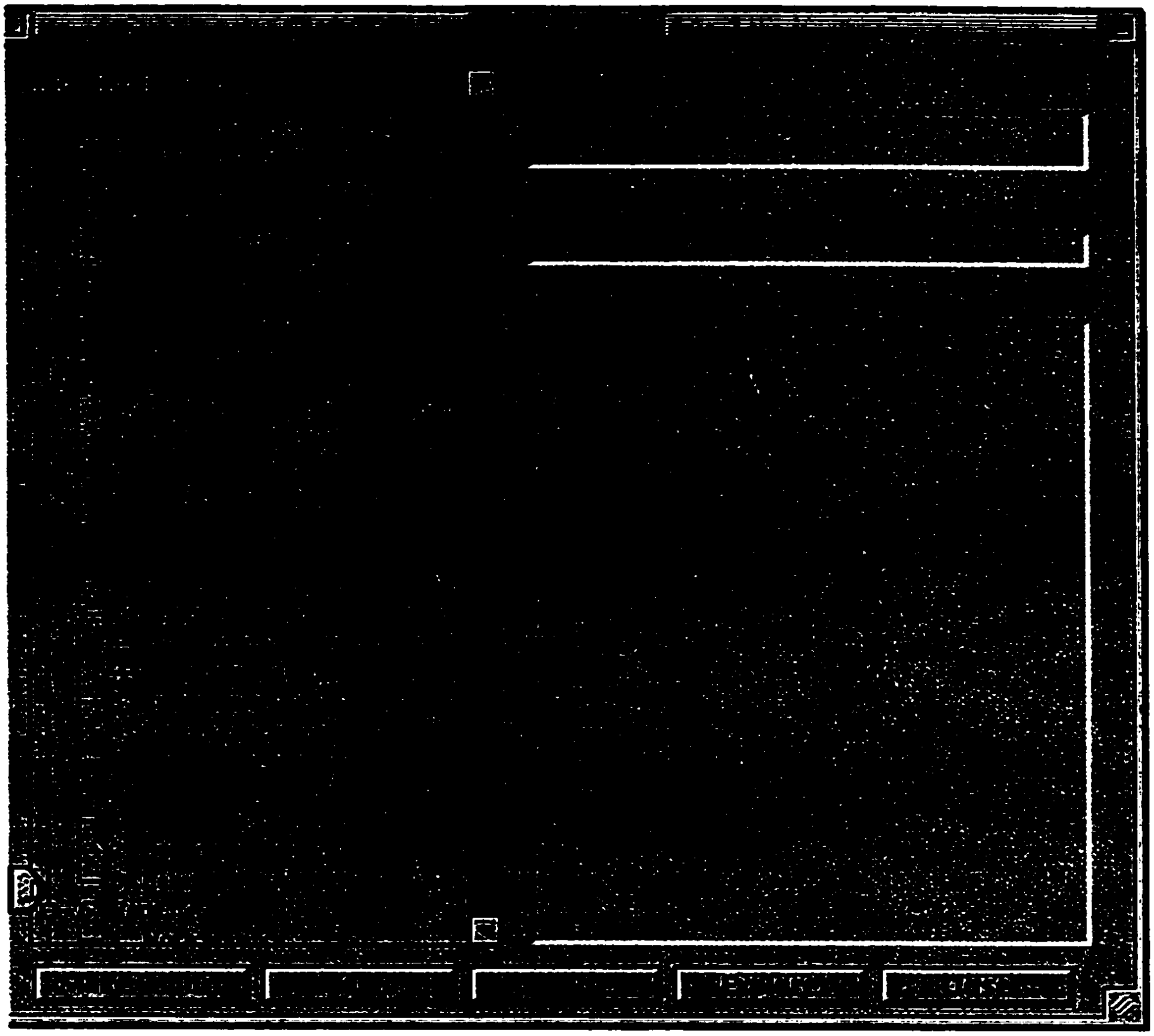


Figure 3.6 Nud*ist Interface

Export of data

The frequencies obtained on each C.T. category were exported as a CSV file, using the "Export Frequency Tables" tool. Data exports had to be conducted for each separate coding node, that is to export one table per skill. In Nud*ist, however, when a category has sub-categories, additional exports had to be conducted. This was the case for the skills:

- "Looks at the Other Side" (2 2 1) / "Suggests Alternative Approach" (2 2 2) in the "Multiple Perspectives" category;
- "Adopts Questioner's Role" (2 3 1) / "Considers Viewer's Perspective" (2 3 2) in the "Multiple Perspectives" category;
- "Predicts or Hypothesizes" (3 2 1) / "Offers Recommendations" (3 2 2) in the "Imposing Meaning" category.

Furthermore, export tables can only be created for nodes that contain at least one level of sub-nodes. In other words, if a category such as "Metacognition" does not have sub-categories of skills, the data are not exportable. To resolve this problem, the node "Metacognition" was moved temporarily to the "Multiple Perspectives" category, in order to be able to export the data.

Another downside of Nud*ist, is that only documents coded at the node appear in the export table. For example, if there is no incidence of the critical thinking skills "Challenges a point (2 1)", the document will not appear on the export list. Consequently, the "missing" document had to be added in manually into the corresponding Excel sheets. To do so, I matched the outputs of each export tables with a master list of the documents.

The missing documents were inserted in Excel and then the value "0" was added in the corresponding column of the skill.

Assessment of English writing skills

The English writing skills of the participants were evaluated at the entry level. The evaluator was a Master's student in Educational Studies. Her undergraduate degree is in Teaching English as a Second Language. She has experience in Teaching English as a Second Language to immigrants and has been using the evaluation criteria outlined hereafter to assess both oral and written skills of the learners.

All assignments of the non-treatment group and the first tasks of the treatment group were used to conduct the assessment. To prevent bias, the identity of the students was stripped from the documents, that is the evaluator did not know the ethnic background of the students. The evaluator allocated a maximum of five points for each of the following criteria: "Grammar and Spelling", "Style", "Vocabulary Richness", and "Development of Ideas". A score on 20 was then obtained and converted into "High" (> 17), "Medium" (13 to 16), and "Low" (< 12). The reliability of the English writing skills assessment has not been established in this study. However, the technique is an established part of Concordia's literacy test for undergraduate students. Reliability testing of trained evaluators is part of the literacy testing process. Results of the English writing skills assessment are presented in the next chapter.

CHAPTER 4: RESULTS

Preparation of data

Attrition

As we saw in the previous chapter, 103 participants out of 108 students either in the treatment or the non-treatment group agreed to participate in the study. This represents a participation rate of 95%. Among those participants, two students (N = 2) in the treatment group abandonned the course and five students (N = 5) in the control condition failed to produce their assignment. The attrition for this study is 7%.

Screened participants

In addition to the students who did not complete the course, some participants were screened out from the original sample. Three criteria lead to the removal of participants from the data: 1) missing Task; 2) late submission of Task and 3) text written in French.

1. Missing Task. Given the structure of the study, participants who failed to carry out at least one Task were removed from the sample. Students who, for instance, failed to submit Task 3 did not receive the first treatment variable (Complexity Level 1). Their outputs for the subsequent Tasks would have been biased. Eleven students did not carry out one, two or three Tasks and were therefore removed from the sample.
2. Late submission of Task. Given the nature of the research design, both the order of the Tasks and the time lapse in between production of the Tasks had to be

respected. Students could not change the sequence of Tasks, produce two Tasks during the same week or submit their Task after the deadline. In the treatment condition, one student carried out all five Tasks but was systematically late. His productions were excluded from the sample.

3. Text written in French. Texts written in French were removed from the sample for procedural reasons. The assignment of Tasks to the raters was conducted using randomization procedures (see chapter 3). Texts in French would have introduced a bias --the ability to read French-- in the coding sample. In the treatment condition, one student wrote the three middle Tasks in French and Tasks 1 and 5 in English. In the non-treatment group, one student submitted the essay in French. The outputs of those students were left out of the final sample.

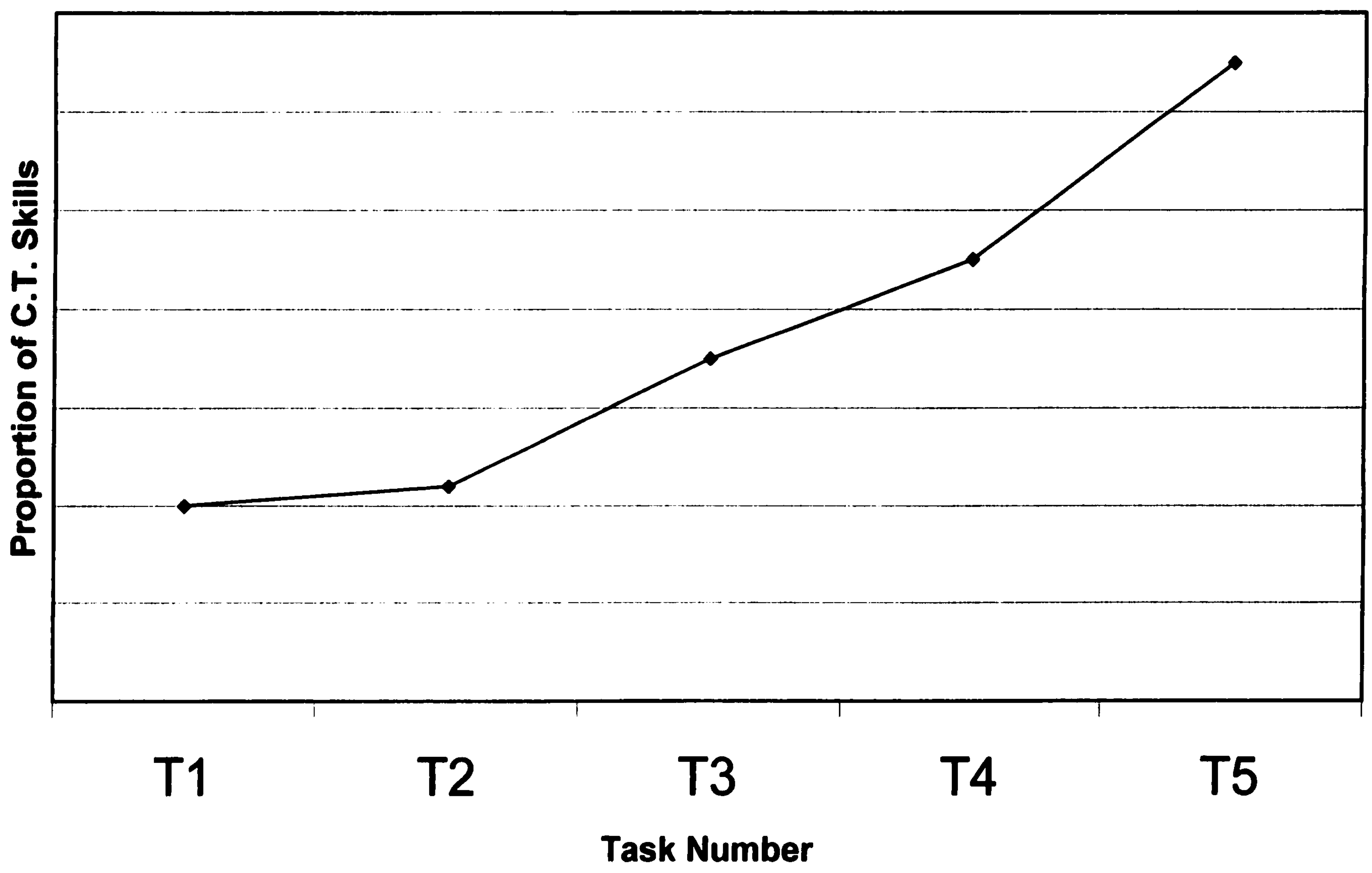
In summary, after applying the screening procedures, the sample of this study consists of 24 participants in the control condition and 65 in the treatment condition. Consequently, the results presented hereafter are based on the analysis of 349 documents produced by 89 participants.

Description of the results

Predicted outcomes

As we have seen in the previous chapter, it was expected that with the introduction of a complexity factor in the instructional Task, the incidence of critical thinking would increase. The tasks rendered more difficult would force the students to put more effort into the assignment, thus demanding a more frequent use of higher order thinking skills. As shown in Figure 4.1, I predicted that use of critical thinking skills would increase as of Task 3.

Figure 4.1 Predicted Outcomes



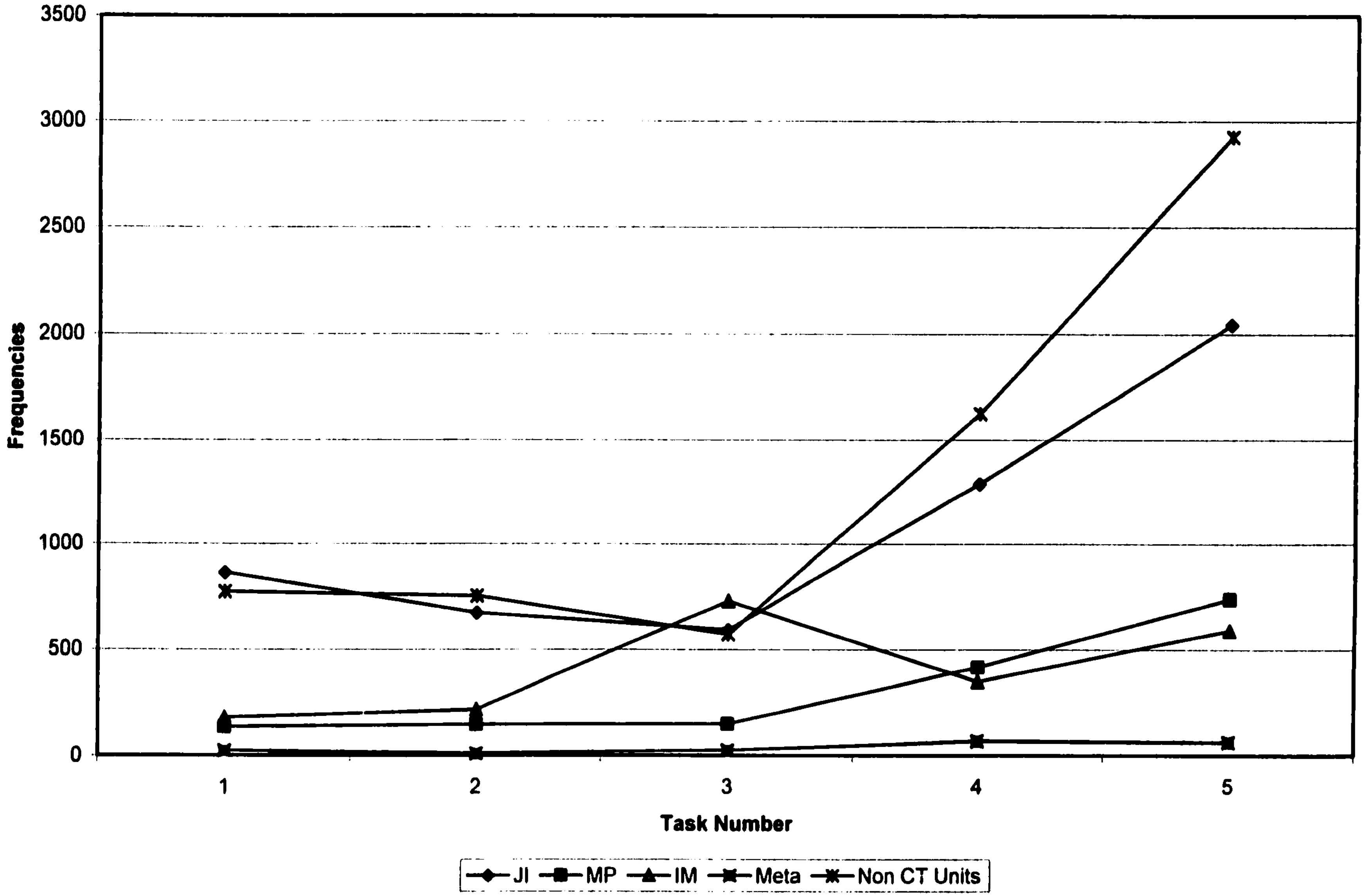
More specifically, Hypothesis 1 predicted that there would be a significant difference between Task 2 and Task 3 (complexity variable Level 1). Hypothesis 2 predicted that there would be a significant difference between Tasks 3 and 4 (complexity variable Level 2). Hypothesis 3 suggested evidence of transfer between Task 4 and Task 5. Finally, since Tasks 1 and 2 were used as the baseline for critical thinking skills, no difference was expected between the means (Hypothesis 4).

The raw frequencies obtained for each coding categories and for each Task are presented in Table 4.1a. One important aspect to point out is the fact that the number of recorded units is not identical for each Task. For example, as we can see in Table 4.1a, there is an important increase of "Judgement and Interpretation" skills from Task 3 (N_{JI} = 593) to Task 4 (N_{JI} = 1 290) and to Task 5 (N_{JI} = 2 045). However, the total number of units in each Task is also increasing from Task 3 ($N_{total\#units}$ = 2 068), to Task 4 ($N_{total\#units}$ = 3 747) and finally, to Task 5 ($N_{total\#units}$ = 6 361). This phenomenon is illustrated in Figure 4.2a.

Table 4.1a Raw Frequencies per Task for All Coding Categories

Task	N	CT skills				NCT units		TOTAL # units
		JI	MP	IM	Meta	Stmts	Varia	
Task 1	65	861	136	180	24	650	121	1 972
Task 2	65	672	147	217	10	636	115	1 797
Task 3	65	593	150	728	26	394	177	2 068
Task 4	65	1290	416	348	69	900	724	3 747
Task 5	65	2045	738	590	63	1757	1168	6 361
TOTAL		5461	1587	2063	192	4337	2305	15 945

Figure 4.2a Trends: Raw Frequencies per Task (N=65)



To solve the problem of dealing with absolute numbers, the raw frequencies were converted into proportions of skills as a mean to create a common denominator to establish a basis for comparison. Consequently, the findings presented in both the results and the discussion chapters are discussed and interpretated as proportions of critical thinking skills. To obtain the proportions of critical thinking skills, the C.T. scores were calculated as follows:

Group C.T. score in Task 1:

$$\text{Group C.T. score in Task 1} = \frac{\text{Total Number of C.T. skills in Task 1}}{\text{Total Number of Units in T1}}$$

To obtain a participant's C.T. score in Task 1:

$$\text{C.T. (p101\_T1)} = \frac{f\,\text{JI (p101)} + f\,\text{MP (p101)} + f\,\text{IM (p101)} + f\,\text{Meta (p101)}}{\text{Total number of units (p101) in Task 1}}$$

Where,

(p) = participant
(f) = frequencies
(JI) = Judgement and Intepretation skills
(MP) = Multiple Perspective skills
(IM) = Imposing Meaning skills
(Meta) = Metacognitive skills

The critical thinking skills "Judgement and Interpretation", "Multiple Perspectives" and "Imposing Meaning" skills are composites of indicators in the coding book (see Appendix XX). To obtain the individual score on a given skill, the following calculations were applied:

Participant's 101 score on JI skills in Task 1:

JI Score (p101_T1) = f(1,1) + f(1,2) + f(1,3) + f(1,4) + f(1,5)

Where,

(f) = frequencies
(1,1) = "Defines"
(1,2) = "Identifies assumptions"
(1,3) = "Makes connections"
(1,4) = "Evaluates"
(1,5) = "Supports"

Participant's 101 score on MP skills in Task :

MP Score (p101_T1) = f(2,1) + f(2,2,1) + f(2,2,2) + f(2,3,1) + f(2,3,2)

Where,

(f) = frequencies
(2,1) = "Challenges"
(2,2,1) = "Looks at the other side"
(2,2,2) = "Suggests alternative"
(2,3,1) = "Adopts questioner's role"
(2,3,2) = "Considers viewer's perspective"

Participant's 101 score on IM skills in Task 1

IM Score (p101_T1) = f(3,1) + f(3,2,1) + f(3,2,2) + f(3,3) + f(3,4) + f(3,5)

Where,

(f) = frequencies
(3,1) = "Recognizes various impacts"
(3,2,1) = "Predicts or hypothesizes"
(3,2,2) = "Makes recommendations"
(3,3) = "Summarizes"
(3,4) = "Concludes"
(3,5) = "Generates new idea"

The non-critical thinking units, "Statements" and "Varia" categories, are not presented here. It should be added here that the category "Use of Visuals" had originally been inserted in the coding book for exploratory purposes. However, because of very low frequencies within a Task and wide variation in frequencies between Tasks, those events have not been included in the analysis. The text units allocated for "Use of Visuals" were not included in the original compilation of data. Their exclusion has no bearing on the ratios presented hereafter.

A summary of the proportion of critical thinking skills per Task is presented in Table 4.1b.

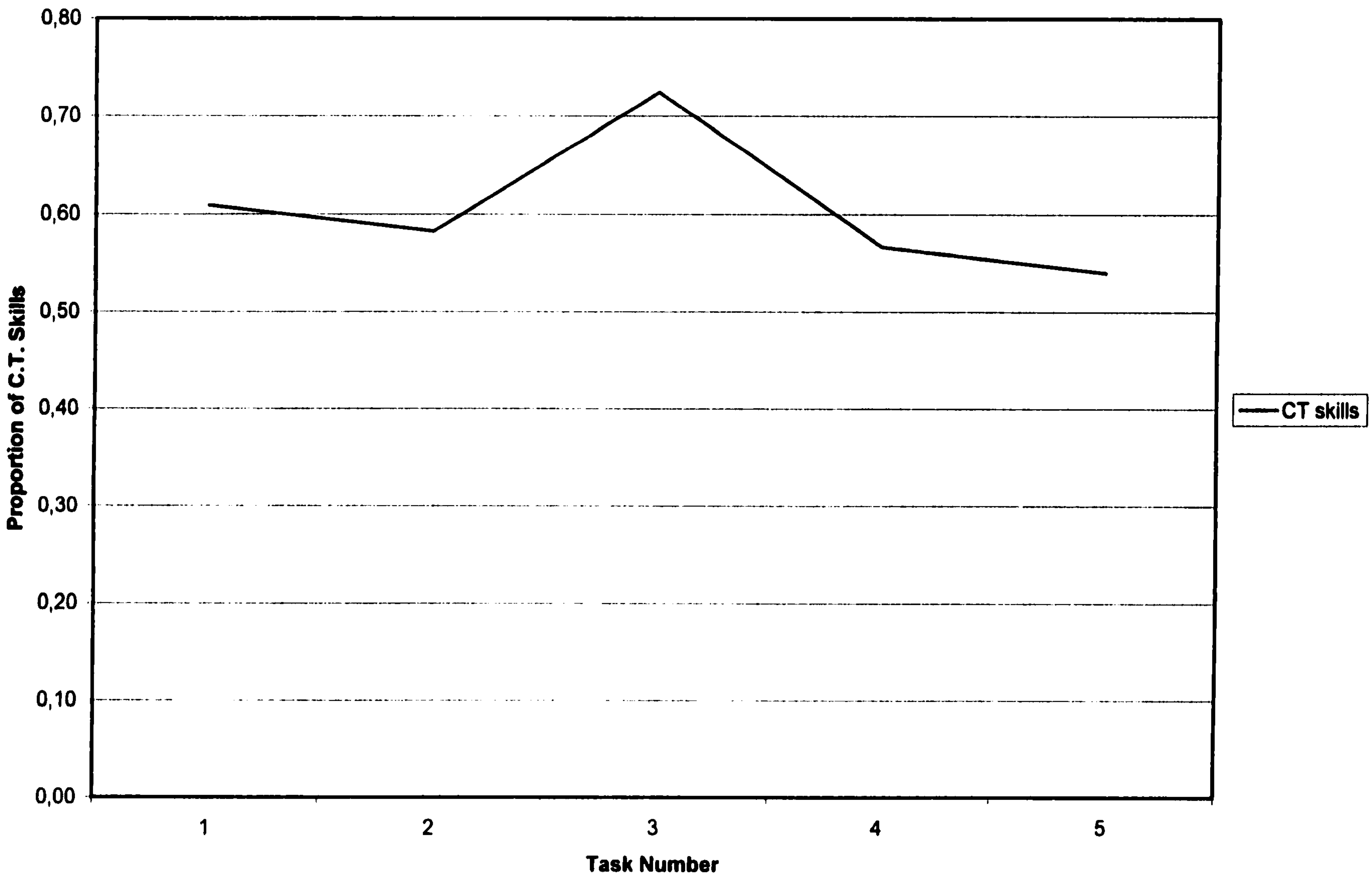
Table 4. 1b Descriptive Statistics: Proportion of C.T. Skills per Task.

Task	CT skills	NCT units	TOTAL # units	AV (CT/65)	Proportion (CT/Total N. of Units)*
Task 1	1201	771	1972	18,48	0,61
Task 2	1046	751	1797	16,09	0,58
Task 3	1497	571	2068	23,03	0,72
Task 4	2123	1624	3747	32,66	0,57
Task 5	3436	2925	6361	52,86	0,54
TOTAL	9303	6642	15945	143.12	0,58

* Results presented in Figure 4.2b

Perhaps the most interesting general finding lies in the contrast between the ratio of critical thinking skills used in Task 3 (72%) and the ratio of critical thinking skills in the other Tasks. Tasks 1 and 2 generated 61% and 58% of critical thinking skills, respectively. The last two Tasks, on the other hand, generated a slightly lower result on the critical thinking skills, 57% in Task 4 and 54% in Task 5. The smallest difference is between Task 3 and and Task 1 with a gap of 11% and the greatest difference is 18% between Tasks 3 and 5. Figure 4.2b shows the actual results from the study. The curve indicates the proportion of critical thinking skills for each Task

Figure 4.2 Trend of C.T. Skills Across Tasks



An analysis of variance for repeated measures (GLM-repeated measures) on the five Tasks was conducted. As shown in Table 4.2, sphericity is assumed $F(4, 256) = 16.426$, $p = .001$. Overall, the treatment generated a significant difference in C.T. scores.

Table 4. 2 Source Table: Trend Analysis for Overall C.T. skills Across Tasks

Source	SS	df	MS	F	Sign.	η^2
CTSCORT						
Sphericity Assumed	15611.860	4	3902.965	16.426	.000	.204
Greenhouse-Geisser	15611.860	3.519	4436.386	16.426	.000	.204
Huyn-Feldt	15611.860	3.749	4164.363	16.426	.000	.204
Lower-Bound	15611.860	1.000	15611.860	16.426	.000	.204
Linear	1142.738	1	1142.738	5.759	.019	.083
Quadratic	4082.051	1	4082.051	14.452	.000	.184
Cubic	23.584	1	23.584	.085	.772	.001
Order 4	10363.487	1	10363.487	53.977	.000	.458
Error (CTSCORT)						
Sphericity Assumed	60827.804	256	237.609			
Greenhouse-Geisser	60827.804	225.219	270.083			
Huyn-Feldt	60827.804	239.931	253.522			
Lower-Bound	60827.804	64.000	950.434			
Linear	12698.594	64	198.416			
Quadratic	18077.048	64	282.454			
Cubic	17764.258	64	277.567			
Order 4	12287.905	64	191.999			
TOTAL	76439.664					

An examination of the polynomial contrasts allows us to describe the shape of the curve presented in Figure 4.2. The Order 4 component of the trend is statistically significant $F(1,64) = 53.977$, $p = .001$. The percentage of the variance explained ($\eta^2 = .458$) is probably attributable to the important rise in critical thinking skills in Task 3. Similarly, the hyperbolic shape as depicted by the quadratic trend, $F(1,64) = 14.452$; p =

.001, is also partly explained by the importance of Task 3 ($\eta^2 = .184$). Finally, the linear component represents 8% of the explained variance ($\eta^2 = .083$), and is statistically significant, $F(1,64) = 5.759$, $p = .001$.

To test hypotheses 1, 2, 3, and 4, pairwise comparisons have been conducted to assess the mean differences between Tasks. Since four pairs of comparisons were planned in the study, out of a possibility 20 pairwise comparisons (if we take into account direction), inflation of familywise alpha is not considered to be a problem. Results are presented in Table 4.3.

Table 4. 3 Pairwise Comparisons: Overall C.T. Means

Measure: MEASURE_1

(I) CTSCORT	(J) CTSCORT	Mean difference (I-J)	Std. Error	Sig.[a]
2	1	-3.089	2.999	.307
3	2	16.265*	2.236	.000
4	3	-18.155*	2.932	.000
5	4	-.705	2.498	.779

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Hypothesis 1 predicted that the introduction of a complexity variable Level 1 would trigger a greater use of critical thinking skills. It is immediately apparent that the pairwise comparison of Task 3 with Task 2 is statistically significant. The mean difference ($M_{diffT3-T2} = 16.265$) is significant ($p = .001$). The null hypothesis 1 is therefore rejected.

The mean difference between Tasks 3 and 4 ($M_{diffT4-T3}$ = -18.155) is also significant (p = .001). Although the null hypothesis 2 is also rejected, it is necessary to point out that the results are not in the direction that was predicted. The students performed better in Task 3 (M_{T3} = 73.0), than in Task 4 (M_{T4} = 54.85).

Hypothesis 3 predicted that there would be evidence of transfer between Task 4 and Task 5. In addition to being negative, the mean difference ($M_{diffT5-T4}$ = -.705) is not statistically significant (p = .779). We fail to reject the null hypothesis. This result is quite puzzling not only because the students performed better in Task 4 (M_{T4} = 59.85) than in Task 5 (M_{T5} = 54.14), but also because the lowest mean across all tasks is obtained in the last assignment.

Finally, the mean difference between Task 1 and Task 2 is not significant ($M_{diffT2-T1}$ = -3.089, p = .307). This finding supports the assumption that the tasks were equal and that they could be used as baseline. To reinforce this finding, a comparison of the outcome of critical skills of the treatment group (M_{trmt} = 56.74) with the outcome of the non-treatment group (M_{ctrl} = 57.77) was performed (see Table 4.4). The difference between the groups is not statistically significant (t (2,87) = -.310; p = .757.

Table 4. 4 Descriptive Statistics: Treatment Group and Non-treatment group on Second Assignment

Condition	N	Mean	Std. Deviation	Std. Error Mean
Treatment	65	56.74	14.72	1.83
Control	24	57.77	11.24	2.29

Originally, the comparison was planned at the entry level, that is to contrast the first assignment of the non-treatment group with Task 1 of the treatment group. However, practical constraints prevented the researcher from obtaining the non-treatment group's first essay. Consequently, the comparison presented here uses assignment 2 of the non-treatment group against Task 2 of the treatment group's.

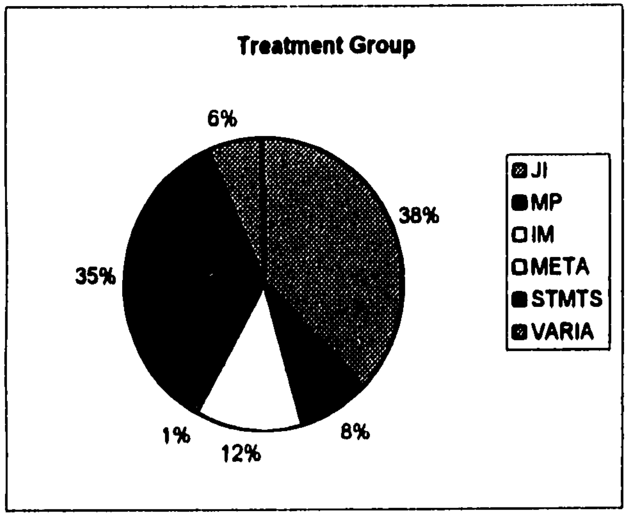
Further, Task 2 and assignment 2 were not the same. In the non-treatment group, the assignments were reflective essays. The student had to select one out of seven questions (See appendix Y) proposed by the instructor. They were required to use the course textbook[2] as their main source of reference. In fact, the students had to comment on the researcher's position on the selected topic. It should be noted, however, that the treatment group also made use of the same textbook.

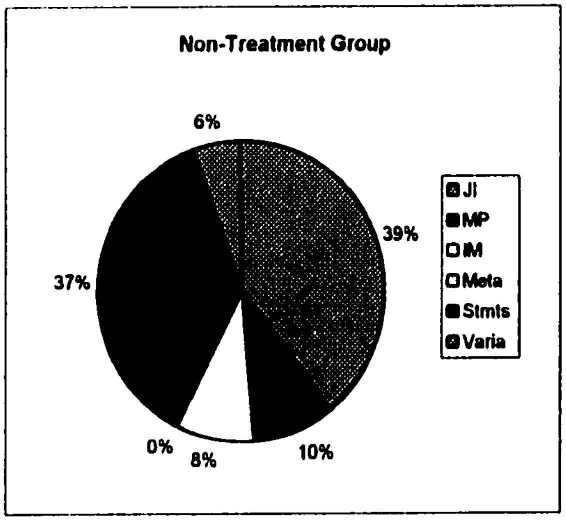
Despite the difference in the type of assignment, the distribution of the skills is relatively similar as can be seen in Figure 4.3.

---

[2] McGinn, Robert, E. (1991). Science, Technology, and Society. Englewood Cliffs, N.J.: Prentice Hall.

Figure 4.3 Distribution of Skills for Task 2





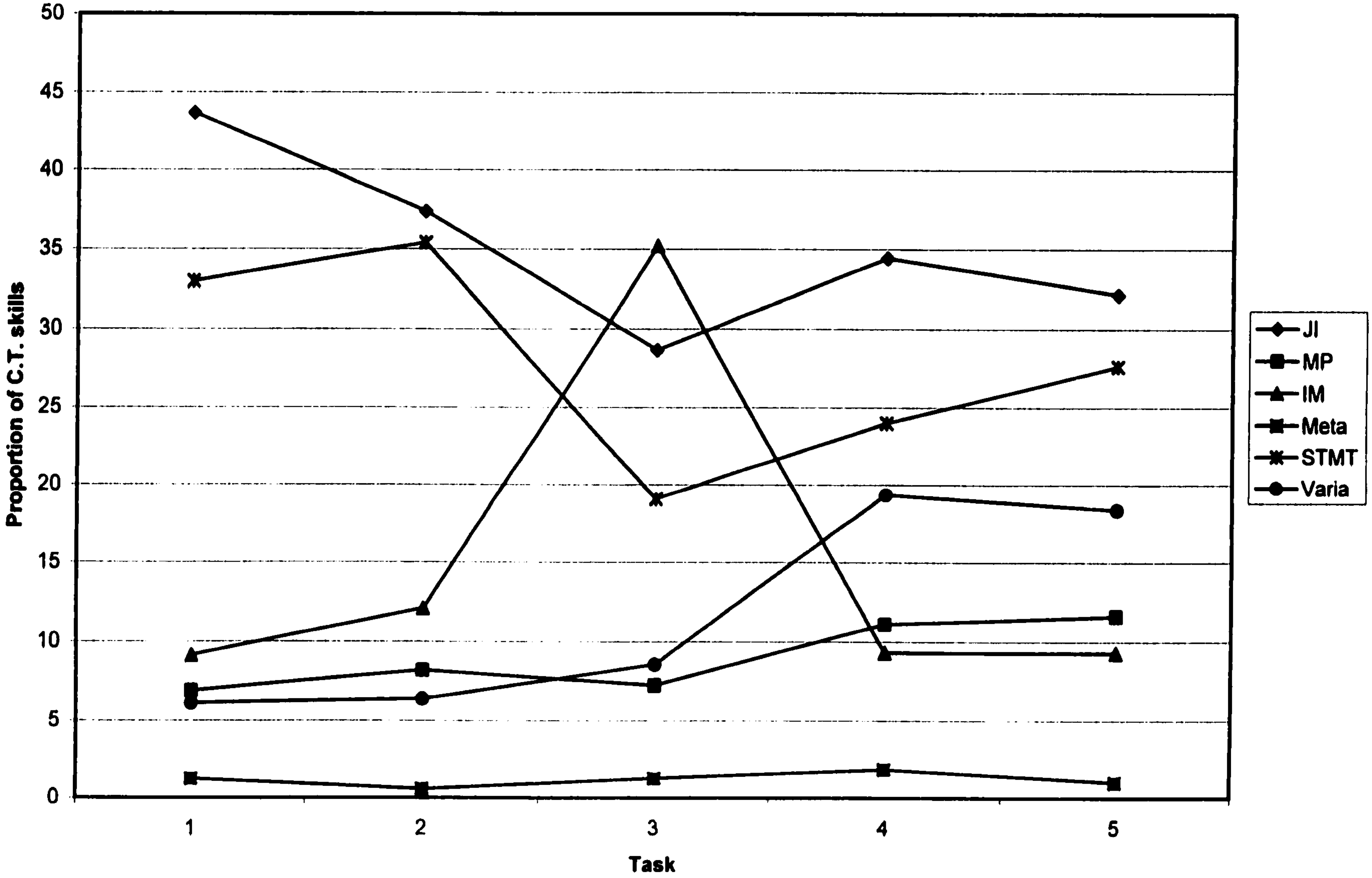
The main difference lies in the "Imposing Meaning" category, with 12% for the treatment group and 8% for the non-treatment group. "Multiple Perspectives" is slightly higher in the control condition (10%) than for the treatment condition (8%). As for the Non-critical thinking skills, there is a slightly greater incidence of "Statements" in the control condition (37%) than in the treatment one (35%).

In summary, I conclude that given the similarity in terms of the ratios of critical thinking skills, the treatment group is representative of the population, despite the differences observed in the context of the study. The distribution of the skills is similar for both conditions, notwithstanding the different course content, the instructional strategies and the time difference between the deadlines of the respective assignments. In the next section, I describe which specific critical thinking skills are contributing to the results described earlier.

Description of the results by type of Skills

Figure 4.4 shows the trends of the six coding categories over the five tasks. For the purpose of this study, only the trends observed for "Judgment and Interpretation", "Multiple Perspectives" and "Imposing Meaning" categories will be described. Because the skill "Metacognition" is not an object of interest in this study, I will not pursue the description of the outcomes for that skill.

Figure 4.4 Trends: All Coding Categories Across Tasks



Judgement and Interpretation skills

An analysis of variance for repeated measures has been conducted to test the effectiveness of the treatment on the skill "Judgement and Interpretation". Results are presented in Table 4.5.

Table 4. 5 Source Table: Trend Analysis for Judgement and Interpretation Skills Across Tasks.

Source	SS	df	MS	F	Sign.	η^2
TASK						
Sphericity Assumed	8636.323	4	2159.081	12.468	.000	.163
Greenhouse-Geisser	8636.323	3.583	2410.440	12.468	.000	.163
Huyn-Feldt	8636.323	3.822	2259.987	12.468	.000	.163
Lower-Bound	8636.323	1.000	8636.323	12.468	.001	.163
Linear	3458.406	1	3458.406	18.382	.000	.223
Quadratic	3475.740	1	3475.740	21.719	.000	.253
Cubic	201.638	1	201.638	1.375	.245	.021
Order 4	1500.538	1	1500.538	7.582	.008	.106
Error (TASK)						
Sphericity Assumed	44330.489	256	173.166			
Greenhouse-Geisser	44330.489	229.304	193.326			
Huyn-Feldt	44330.489	244.580	181.252			
Lower-Bound	44330.489	64.000	692.664			
Linear	12040.762	64	188.137			
Quadratic	10241.930	64	160.030			
Cubic	9382.490	64	146.601			
Order 4	12665.307	64	197.895			
TOTAL	52966.812					

Sphericity is assumed, $F(4, 256) = 12.47$, $p = .001$, $\eta^2 = .163$. The treatment had a significant effect on the incidence of "Judgement and Interpretation" skills. The quadratic component of the trend is statistically significant, $F(1,64) = 21.719$, $p = .001$, $\eta^2 = .253$. The hyperbolic shape is partly explained by the important decrease of "Judgement and Interpretation" skills in Task 3 (see Figure 4.4). The linear trend is significant, $F(1,64) = 18.38$, $p = .001$, $\eta^2 = .223$. Finally, the Order 4 component of the trend is also statistically significant, $F(1,64) = 7.58$, $p = .008$, $\eta^2 = .106$.

The highest score obtained on "Judgement and Interpretation" skills is in Task 1 ($M_{T1} = 42.82$), whereas the lowest score is observed in Task 3 ($M_{T3} = 26.98$) (see Table 4.6).

Table 4. 6 Descriptive Statistics: Judgement and Interpretation Skills.

Task	N	Mean	Std. Deviation
Task 1	65	42.82	16.36
Task 2	65	35.85	12.63
Task 3	65	26.98	12.88
Task 4	65	33.46	14.01
Task 5	65	32.48	11.54

Interestingly, pairwise comparisons reveal that all comparisons with Task 1 and all comparisons with Task 3 are significant (see Table 4.7). The percentage of variance explained by the quadratic and the linear trends are probably explained by those differences. Hence, the most important mean difference is observed between Task 1 and Task 3 ($M_{diffT1T3} = -15.86$, $p = .001$). As well, the second most important mean difference between Task 1 and Task 5 ($M_{diffT1T5} = -10.34$, $p = .001$). The mean difference between

Task 3 and Task 5 ($M_{diffT3T5} = -5.52$, $p = .008$) is also significant. This three point comparison would explain the quadratic trend of the results outlined earlier. However, given the number of comparisons conducted, the chances of finding significant statistical differences (Type I error) between means are increasing. Consequently, the significant differences found for relatively small mean differences (e.g.difference between Task 3 and Task 5, $M_{diff}= 5.5$) have to be considered with caution. The same is true for pairwise comparisons conducted for "Multiple Perspectives" and "Imposing Meaning" skills.

As we can see on Figure 4.4, the linear trend component is explained by the significant differences observed between Task 1 and Task 2 ($M_{diffT1T2} = -6.97$, $p = .003$), between Task 2 and Task 3 ($M_{diffT2-T3} = 8.89$, $p = .001$) and between Task 1 and Task 3 ($M_{diffT1T3} = -15.86$, $p = .001$). It should be added that the significant result obtained between Task 1 and Task 2 is unexpected given the fact that the tasks are considered equivalent. The significant difference is probably due to a high result obtained for the indicator "Supports" in Task 1 (17.2) in contrast with Task 2 (9.5) (See Table 4.8).

Table 4. 7 Pairwise Comparisons: Judgement and Interpretation Skills

Measure: MEASURE_1

(I) TASK	(J) TASK	Mean difference (I-J)	Std. Error	Sig.[a]
1	2	6.969*	2.262	.003
	3	15.859*	2.759	.000
	4	9.355*	2.510	.000
	5	10.341*	2.457	.000
2	1	-6.969*	2.262	.003
	3	8.890*	2.396	.000
	4	2.385	2.065	.252
	5	3.371	1.931	.086
3	1	-15.859*	2.759	.000
	2	-8.890*	2.396	.000
	4	-6.504*	2.417	.009
	5	-5.519*	2.002	.008
4	1	-9.355*	2.510	.000
	2	-2.385	2.065	.252
	3	6.504*	2.417	.009
	5	.986	2.155	.649
5	1	-10.341*	2.457	.000
	2	-3.371	1.931	.086
	3	5.519*	2.002	.008
	4	-.986	2.155	.649

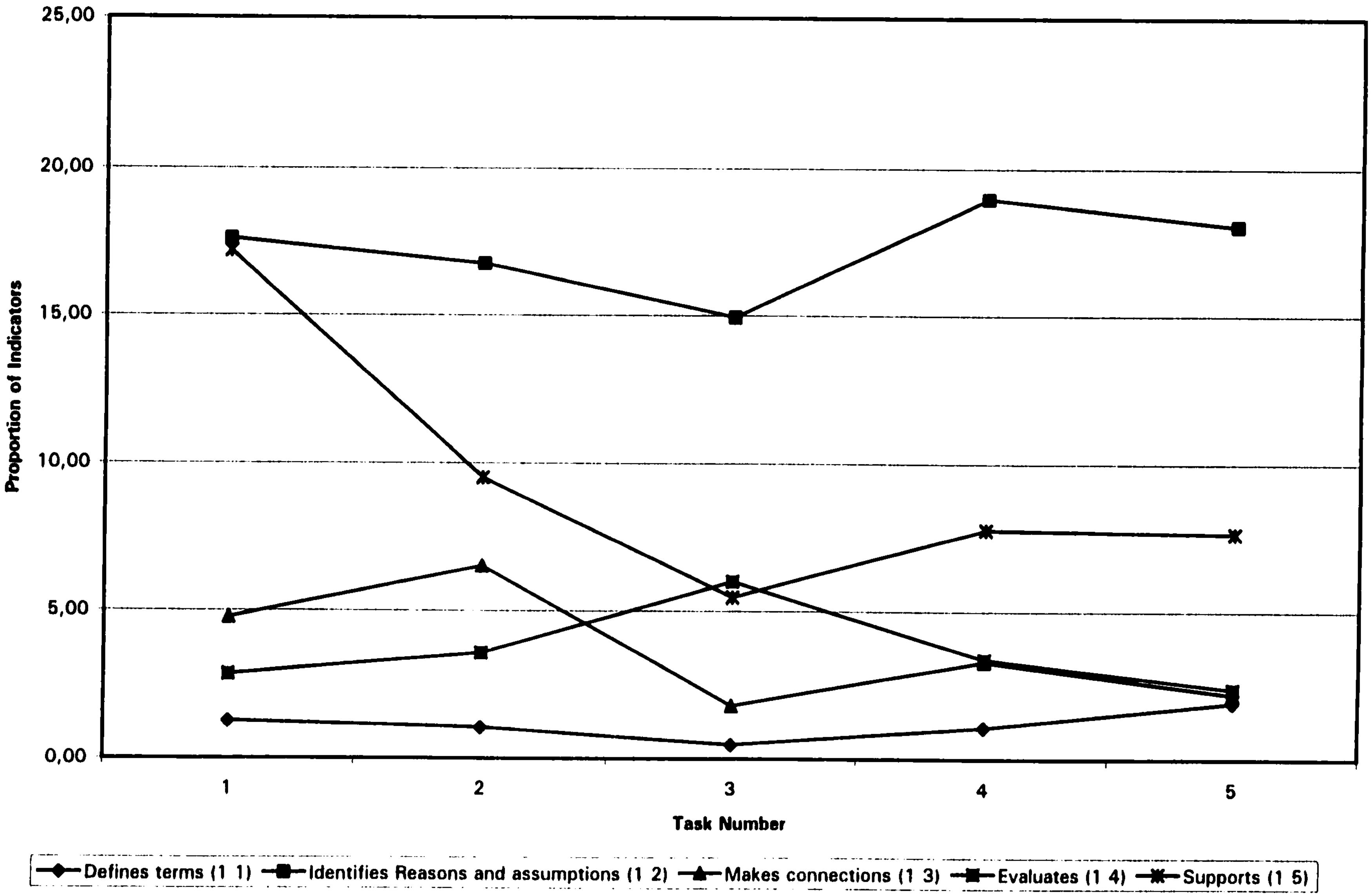
Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Finally, with the exception of Task 3, the "Judgement and Interpretation" skills are the most frequently used across Tasks.

Figure 4.5 Trends: Judgement & Interpretation Indicators Across Tasks



Indicators of Judgement and Interpretation skills

The highest proportions of units in the "Judgement and Interpretation" category is observed for the indicator "Identifies Reasons and Assumptions". The highest score observed is in Task 4 (M_{T4} = 18.95), whereas the lowest score is observed in Task 3 (M_{T3} = 14.94) (see Table 4.8).

Table 4. 8 Proportion* of Judgement and Interpretation Indicators by Tasks

	T1	T2	T3	T4	T5
Defines terms (1 1)	1.27	1.06	0.48	1.07	1.92
Identifies Reasons and assumptions (1 2)	17.60	16.75	14.94	18.95	18.03
Makes connections (1 3)	4.77	6.51	1.79	3.28	2.19
Evaluates (1 4)	2.84	3.56	6.00	3.36	2.37
Supports (1 5)	17.19	9.52	5.46	7.77	7.64

* Proportion = Frequencies of indicator /total number of units in Task, multiplied by 100.

An important decrease for the indicator "Supports" is observed between Task 1 (M_{T1} = 17.19) and Task 2 (M_{T2} = 9.52) and reaches the lowest point in Task 3 (M_{T3} = 5.46). The difference between Task 1 and Task 3 ($M_{diffT1T3}$ = 11.73) is the most important one observed in the category "Judgement and Interpretation". Another important decrease is observed for the indicator "Making Connections", between the highest score in Task 2 (M_{T2} = 6.51) and the lowest score in Task 3 (M_{T3} = 1.79). Finally, a small increase is noticed for the indicator "Evaluates" between Task 2 (M_{T2} = 3.56) and Task 3 (M_{T3} = 6.0). The trend goes down again in Task 4 (M_{T4} = 3.36).

It should be noted that for most indicators within the "Judgement and Interpretation" category, Task 3 scored the lowest scores across Tasks. The only exception to this trend is the indicator "Supports", which obtained the highest score across Tasks. These results support the phenomenon observed in Figure 4.4, where "Judgement and Interpretation" plummeted in Task 3. Similarly, the relatively linear descending trend observed between Task 1 and Task 3 in Figure 4.4, is attributable to important descending slopes on four indicators: "Identify Reasons and Assumptions", "Supports", "Making Connections" and to a lesser extent, "Defines Terms".

Multiple Perspectives skills

The results of the analysis of variance for repeated measures conducted to test the effectiveness of the treatment on the skill "Multiple Perspectives" are presented in Table 4.9.

Table 4. 9 Source Table: Trend Analysis for Multiple Perspective Skills Across Tasks

Source	SS	df	MS	F	Sign.	η^2
TASK						
Sphericity Assumed	1308.302	4	327.075	6.223	.000	.089
Greenhouse-Geisser	1308.302	3.062	427.329	6.223	.000	.089
Huyn-Feldt	1308.302	3.233	404.694	6.223	.000	.089
Lower-Bound	1308.302	1.000	1308.302	6.223	.015	.089
Linear	814.714	1	814.714	21.454	.000	.251
Quadratic	140.783	1	140.783	2.824	.098	.042
Cubic	.520	1	.520	.014	.907	.000
Order 4	352.286	1	352.286	4.153	.046	.061
Error (TASK)						
Sphericity Assumed	13455.089	256	52.559			
Greenhouse-Geisser	13455.089	195.941	68.669			
Huyn-Feldt	13455.089	206.900	65.032			
Lower-Bound	13455.089	64.000	210.236			
Linear	2430.438	64	37.976			
Quadratic	3190.718	64	49.855			
Cubic	2404.403	64	37.569			
Order 4	5429.530	64	84.836			
TOTAL	14763.391					

Despite the relative flatness of the trend observed in Figure 4.4, the overall treatment is statistically significant for the skill "Multiple Perspectives". Sphericity is assumed, $F(4,256) = 6.22$, $p = .001$, $\eta^2 = .089$. The linear trend component is the most important, $F(1,64) = 21.45$, $p = .001$; $\eta^2 = .251$. A small Order 4 component is also significant with $F(1,64) = 4.15$, $p = .046$; $\eta^2 = .061$.

As shown in Table 4.10, the highest scores for Mutliple Perspectives is observed in Task 5 ($M_{T5} = 11.29$), whereas the lowest scores are observed in Task 3 ($M_{T3} = 6.11$).

Table 4. 10 Descriptive Statistics: Multiple Perspectives Skills

Task	N	Mean	Std. Deviation
Task 1	65	6.86	6.29
Task 2	65	8.11	6.76
Task 3	65	6.11	9.02
Task 4	65	10.46	7.26
Task 5	65	11.29	6.41

Table 4.11 presents the results of pairwise comparisons. The comparisons of Task 1, Task 2, and Task 3 against Task 4, as well as the comparisons of Task 1, Task 2, and Task 3 against Task 5 are all statistically significant. However, there is no significant difference between Task 4 ($M_{T4} = 10.46$) and Task 5 ($M_{T5} = 11.29$), which would have indicated evidence of transfer. These results however, have to be viewed with caution. As we will see in the next section, the significant differences between the means are probably attributable to a confounding factor in the instructional design of the simulations tasks.

Table 4. 11 Pairwise Comparisons: Multiple Perspective Skills

Measure: MEASURE_1

(I) TASK	(J) TASK	Mean difference (I-J)	Std. Error	Sig.[a]
1	2	-1.246	1.034	.233
	3	.754	1.407	.594
	4	-3.598*	1.109	.002
	5	-4.422*	1.030	.000
2	1	1.246	1.034	.233
	3	2.000	1.531	.196
	4	-2.352*	1.124	.040
	5	-3.176*	1.045	.003
3	1	-.754	1.407	.594
	2	-2.000	1.531	.196
	4	-4.352*	1.640	.010
	5	-5.175*	1.536	.001
4	1	3.598*	1.109	.002
	2	2.352*	1.124	.040
	3	4.352*	1.640	.010
	5	-.823	1.042	.432
5	1	4.422*	1.030	.000
	2	3.176*	1.045	.003
	3	5.175*	1.536	.001
	4	.823	1.042	.432

Based on estimated marginal means

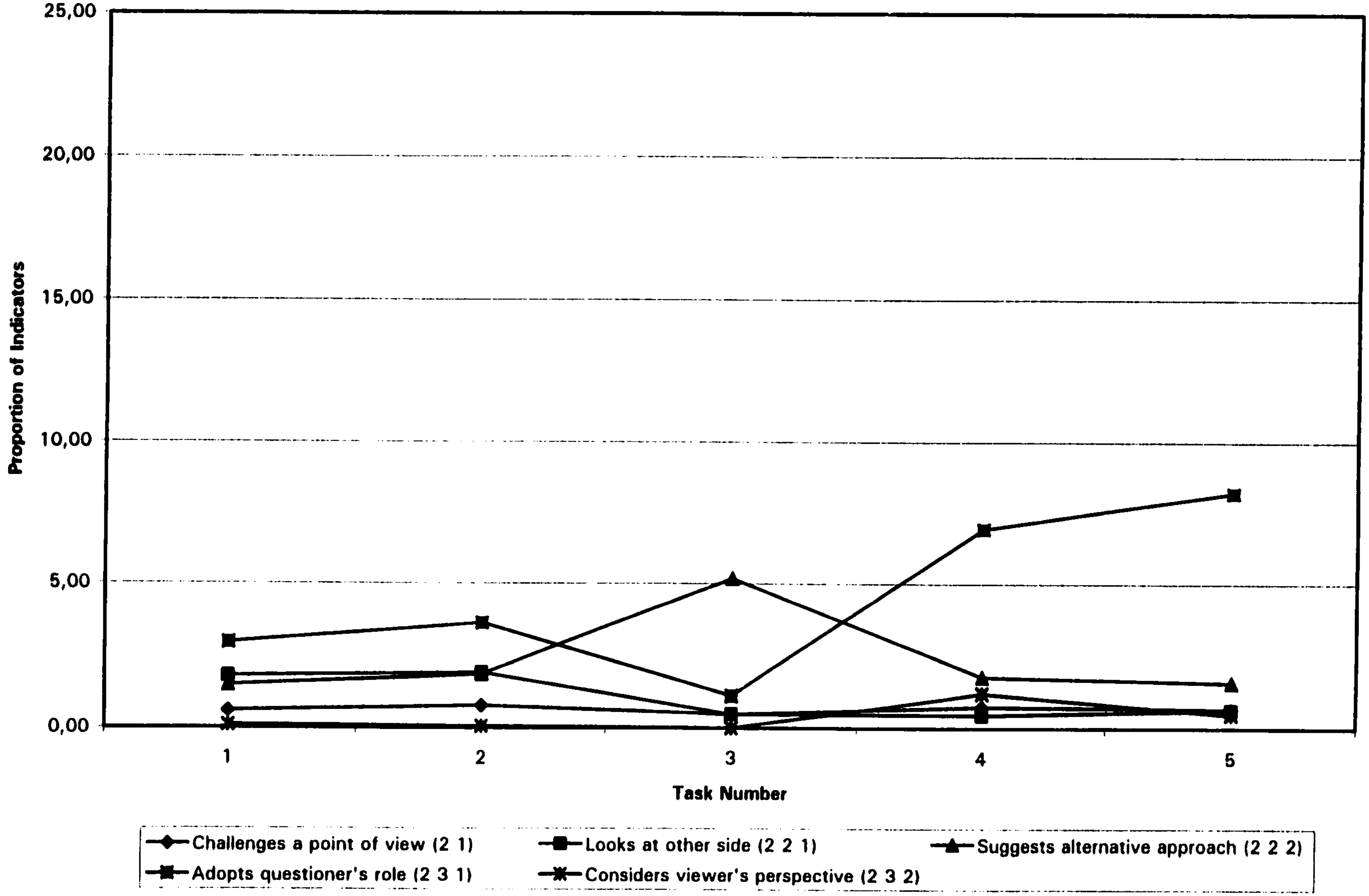
*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Indicators of Multiple Perspective skills

As we can see on Figure 4.6, the significant differences obtained in Task 4 and Task 5 in comparison with the earlier tasks are explained by the dramatic increase from Task 3 ($M_{T3} = 1.11$) to Task 4 ($M_{T4} = 6.94$) of the indicator "Adopts Questioner's Role" (see Table 4.10). This indicator reaches a peak in Task 5 ($M_{T5} = 8.22$). This important difference however, cannot be exclusively attributed to the effect of the treatment variable (Complexity Level 2). Indeed, part of the mandate in Task 4 and 5 was to prepare questions for the chatroom (see Appendices XIV & XV).

Figure 4.6 Trends: Multiple Perspectives Indicators Across Tasks



Despite the confound of having to produce questions for the chatroom (see Appendices XIV and XV), there is one result worth pointing out. The indicator "Suggests an Alternative Approach" peaks in Task 3 ($M_{T3} = 5.17$) (see Table 4.12). The difference with Task 2 is $M_{diffT2T3} = 3.33$, whereas the difference with Task 4 is $M_{diffT3T4} = -3.41$. The shape of the curve is similar to the one observed for the "Judgement and Interpretation" indicator "Evaluates" in Figure 4.4. However, as we have seen this difference is not important enough to create significance in the overall effect of the treatment on Multiple Perspectives in Task 3.

Table 4. 12 Proportion* of Multiple Perspectives Indicators Across Tasks

Indicators	T1	T2	T3	T4	T5
Challenges a point of view (2 1)	0.61	0.78	0.48	0.75	0.68
Looks at other side (2 2 1)	1.77	1.89	0.48	0.45	0.64
Suggests alternative approach (2 2 2)	1.47	1.84	5.17	1.76	1.57
Adopts questioner's role (2 3 1)	2.94	3.62	1.11	6.94	8.22
Considers viewer's perspective (2 3 2)	0.10	0.06	0.00	1.20	0.49

* Proportion = Frequencies of indicator /total number of units in Task, multiplied by 100.

Imposing Meaning skills

The treatment has a statistically significant impact on the incidence of "Imposing Meaning" skills. As shown in Table 4.13, sphericity is assumed $F(4,256) = 77.022$, $p = .001$, $\eta^2 = .546$.

Table 4. 13 Source Table: Trend Analysis for Imposing Meaning Skills Across Tasks

Source	SS	df	MS	F	Sign.	η^2
TASK						
Sphericity Assumed	44691.943	4	11172.986	77.022	.000	.546
Greenhouse-Geisser	44691.943	2.011	22224.404	77.022	.000	.546
Huyn-Feldt	44691.943	2.076	21524.263	77.022	.000	.546
Lower-Bound	44691.943	1.000	44691.943	77.022	.000	.546
Linear	62.264	1	62.264	1.499	.225	.023
Quadratic	18432.470	1	18432.470	109.840	.000	.632
Cubic	269.896	1	269.896	3.030	.087	.045
Order 4	25927.313	1	25927.313	91.998	.000	.590
Error (TASK)						
Sphericity Assumed	37135.766	256	145.062			
Greenhouse-Geisser	37135.766	128.700	288.545			
Huyn-Feldt	37135.766	132.887	279.455			
Lower-Bound	37135.766	64.000	580.246			
Linear	2657.732	64	41.527			
Quadratic	10739.992	64	167.812			
Cubic	5701.255	64	89.082			
Order 4	18036.787	64	281.825			
TOTAL	81827.709					

The trend analysis reveals an important quadratic trend, $F(1,64) = 109.840$, $p = .001$, $\eta^2 = .632$. The Order 4 component is also statistically significant, $F(1,64) = 91.998$, $p = .001$, $\eta^2 = .590$. The percentage of the variance explained by the quadratic and the Order 4 trend components are visible in Figure 4.4. The contribution of Task 3 (M_{T3} =

39.05) in contrast with Task 2 ($M_{diffT2T3} = 26.77$) and in contrast with Task 4 ($M_{diffT3T4} = -29.97$) is undeniable (see Table 4.15).

Table 4. 14 Descriptive Statistics: Imposing Meaning Skills

Task	N	Mean	Std. Deviation
Task 1	65	9.09	8.23
Task 2	65	12.28	8.99
Task 3	65	39.05	20.29
Task 4	65	9.08	7.18
Task 5	65	9.14	6.90

As we can see on Table 4.15, all pairwise comparisons with Task 3 are statistically significant.

Table 4. 15 Pairwise Comparisons: Imposing Meaning Skills

Measure: MEASURE_1

(I) TASK	(J) TASK	Mean difference (I-J)	Std. Error	Sig.[a]
1	2	-3.190	1.764	.075
	3	-29.963*	2.715	.000
	4	6.628E-03	1.185	.996
	5	-5.075E-02	1.307	.969
2	1	3.190	1.764	.075
	3	-26.773*	2.988	.000
	4	3.197*	1.520	.039
	5	3.139*	1.350	.023
3	1	29.963*	2.715	.000
	2	26.773*	2.988	.000
	4	29.970*	2.827	.000
	5	29.913*	2.914	.000
4	1	-6.628E-03	1.185	.996
	2	-3.197*	1.520	.039
	3	-29.970*	2.827	.000
	5	-5.738E-02	1.222	.963
5	1	5.075E-02	1.307	.969
	2	-3.139*	1.350	.023
	3	-29.913*	2.914	.000
	4	5.738E-02	1.222	.963

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

[a]. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

An unexpected result however, is visible in the comparisons of Task 2 (M_{T2} = 12.28) and Task 4 (M_{T4} = 9.08) and similarly between Task 2 and Task 5 (M_{T5} = 9.14) (see Table 4.14). The mean differences are statistically significant ($M_{diffT2T4} = -3.2$ and $M_{diffT2T5} = -3.14$) (see Table 4.15). In short, the students performed better in Task 2 than in Tasks 4 and 5.

Indicators of Imposing Meaning skills

The contribution of each indicators of the "Imposing Meaning" category are presented in Figure 4.7. The scores for each indicator are presented in Table 4.16.

Figure 4.7 Trends: Imposing Meaning Indicators Across Tasks

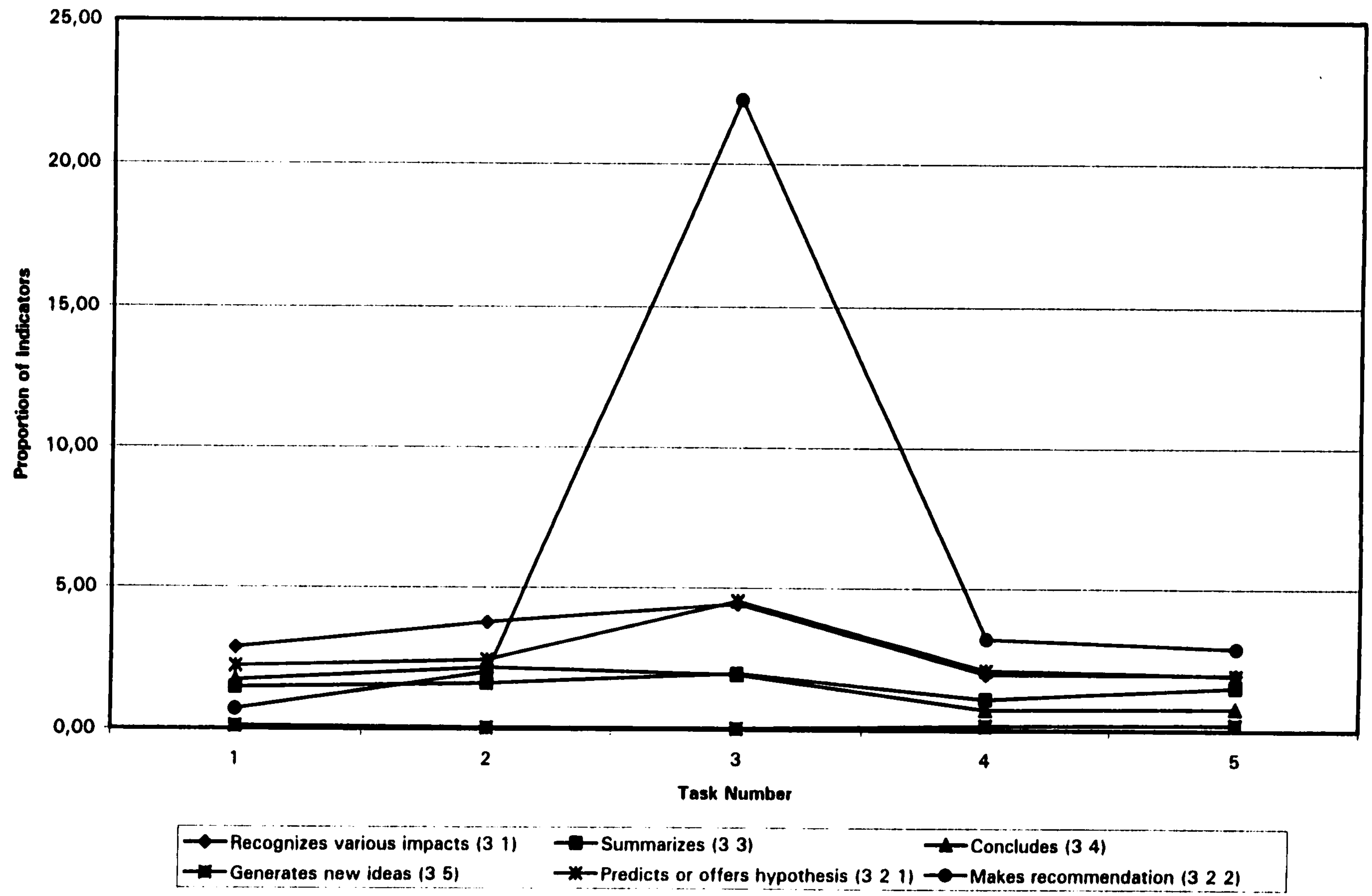


Table 4. 16 Proportion* of Imposing Meaning Indicators Across Tasks

Indicators	T1	T2	T3	T4	T5
Recognizes various impacts (3 1)	2.89	3.78	4.45	1.97	1.97
Summarizes (3 3)	1.47	1.61	1.98	1.09	1.51
Concludes (3 4)	1.72	2.17	1.93	0.72	0.77
Generates new ideas (3 5)	0.10	0.06	0.05	0.16	0.20
Predicts or hypothesizes (3 2 1)	2.23	2.45	4.55	2.11	1.93
Makes recommendation (3 2 2)	0.71	2.00	22.24	3.23	2.89

* Proportion = Frequencies of indicator /total number of units in Task, multiplied by 100.

As shown on Figure 4.7, the indicator "Making Recommendations" contributes most to the overall shape (Quadratic and Order 4 trend components) of the "Imposing Meaning" category. The shape is very similar to the one observed in Figure 4.4. The difference between Task 3 ($M_{T3} = 22.24$) and Task 2 ($M_{T2} = 2.00$) is $M_{diffT2T3} = 20.24$ (see Table 4.16). The difference with Task 4 ($M_{T4} = 3.23$) is $M_{diffT3T4} = -19.01$. This is the most important difference observed for a critical thinking skills indicator.

Although the difference is not as important, we also note a small increase in Task 3 for the indicator "Predicts or Hypothesizes" ($M_{T3} = 4.55$) as opposed to $M_{T2} = 2.45$ in Task 2. Task 3 also scores the highest incidence for the indicator "Recognizes Various Impacts" ($M_{T3} = 4.45$). It should be noted that with the obvious exception of "Making Recommendations", all the indicators scores are lower than 5%.

Finally, the incidence of critical thinking has no relationship with the length of the Task. Results of Pearson correlations are presented in Table 4.17.

Table 4. 17 Correlations: Incidence of C.T. Skills and Length of Tasks

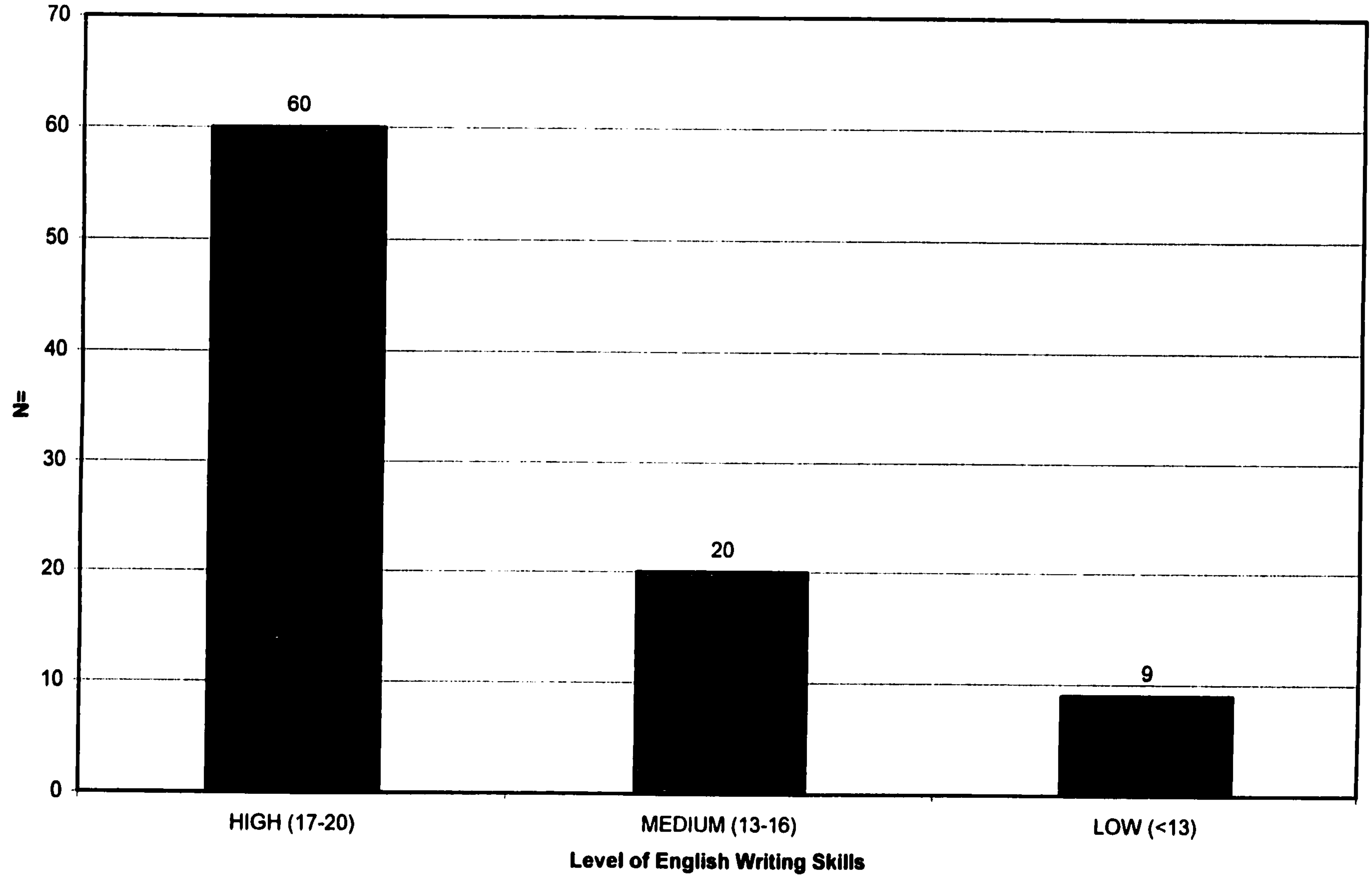
Task	CT-skills Units	Mean	Std. Deviation	Correlations		N =
				Pearson	Sign. (2-tail)*	
Task 1	C.T. Skills	59.83	16.93	.117	.355	65
	N. of Units	30.34	16.78			
Task 2	C.T. Skills	56.74	14.72	.201	.109	65
	N. of Units	27.65	13.97			
Task 3	C.T. Skills	73.00	16.78	-.070	.578	65
	N. of Units	31.82	16.83			
Task 4	C.T. Skills	54.85	16.42	.218	.081	65
	N. of Units	57.65	29.56			
Task 5	C.T. Skills	54.14	13.17	-.022	.861	65
	N. of Units	97.86	43.18			

*p = .05.

English writing skills

Figure 4.8 shows the distribution of students by level of English writing skills. The majority of students (N = 60) scored high (> 17) on the evaluation of their writing skills. Twenty students obtained a medium score (13 to 16) and nine were evaluated as weak writers (< 12). Given the skewed results (M_{TRMT} = 17.12), nonparametric correlations were conducted to explore the relationship between language ability and performance on critical thinking skills. Both Kendall's tau_b (r = .140, p = .119) and Spearman's rho correlation coefficients (r = .191, p = .127) are not statistically significant.

Figure 4.8 Distribution of Students by Level of English Writing Skills (N=89)



Post-task questionnaire

The following section presents the results on the post-task questionnaire (see Appendix VIII). It should be noted that only 36 students filled out the questionnaire. Consequently, the results serve an illustrative purpose only.

The majority of students found the topic "Genetically Modified Foods" either "interesting" (N = 16) or "very interesting" (N = 10). On the other hand, 14 students found the topic "Food Production" "interesting" (N = 14) or "somewhat interesting" (N = 12).

Figure 4.9 Interest in Topic

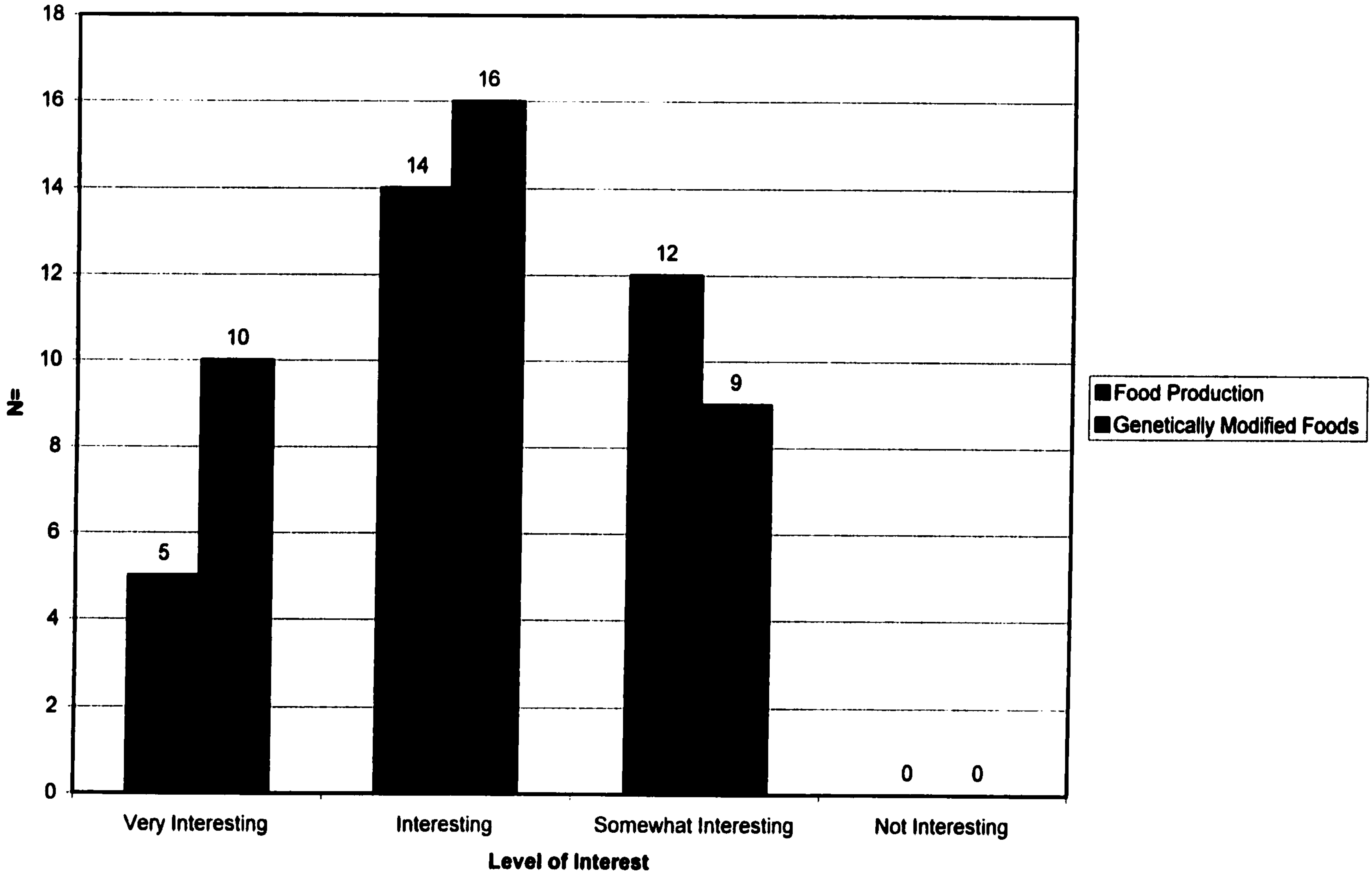
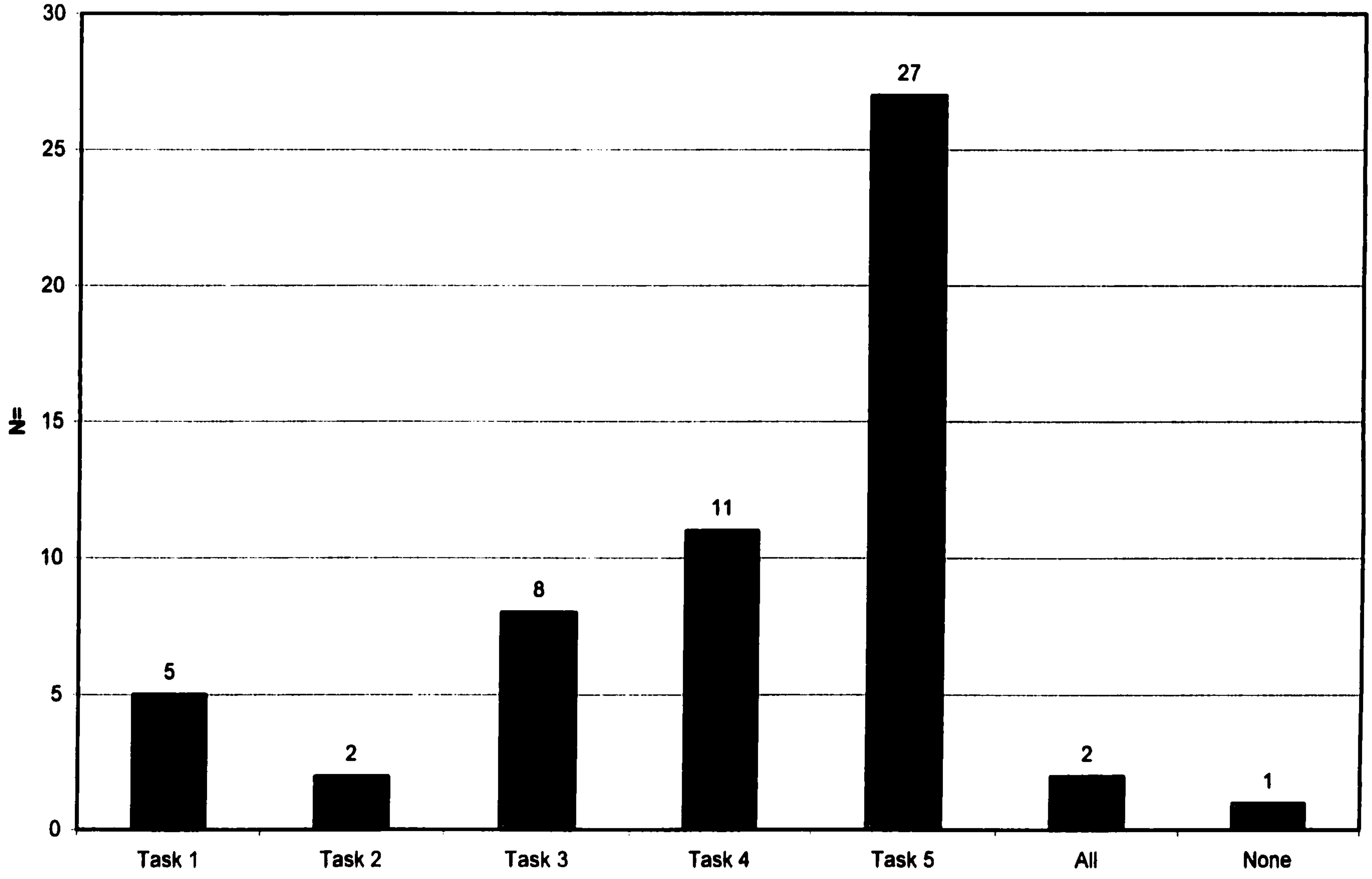


Figure 4.10 shows the results on the item "Which Task (s) did you find most engaging". Multiple answers were allowed. The majority of students (N = 27) found that Task 5 was the most engaging task in contrast with Task 4 (N = 11). Task 2, according to the respondents is the least engaging Task (N = 2).

Figure 4.10 Most Engaging Task



While the most engaging Tasks were Task 4 and 5, the simulation is the format the students least enjoyed (N = 20) (see Figure 4.11). The essays and the case study generated the same level of enjoyment (N = 4). Eight students enjoyed all formats. Finally, the majority of students invested more effort in Task 4 (N = 25) than in any other Tasks (see Figure 4.12).

Figure 4.11 Format Least Enjoyed

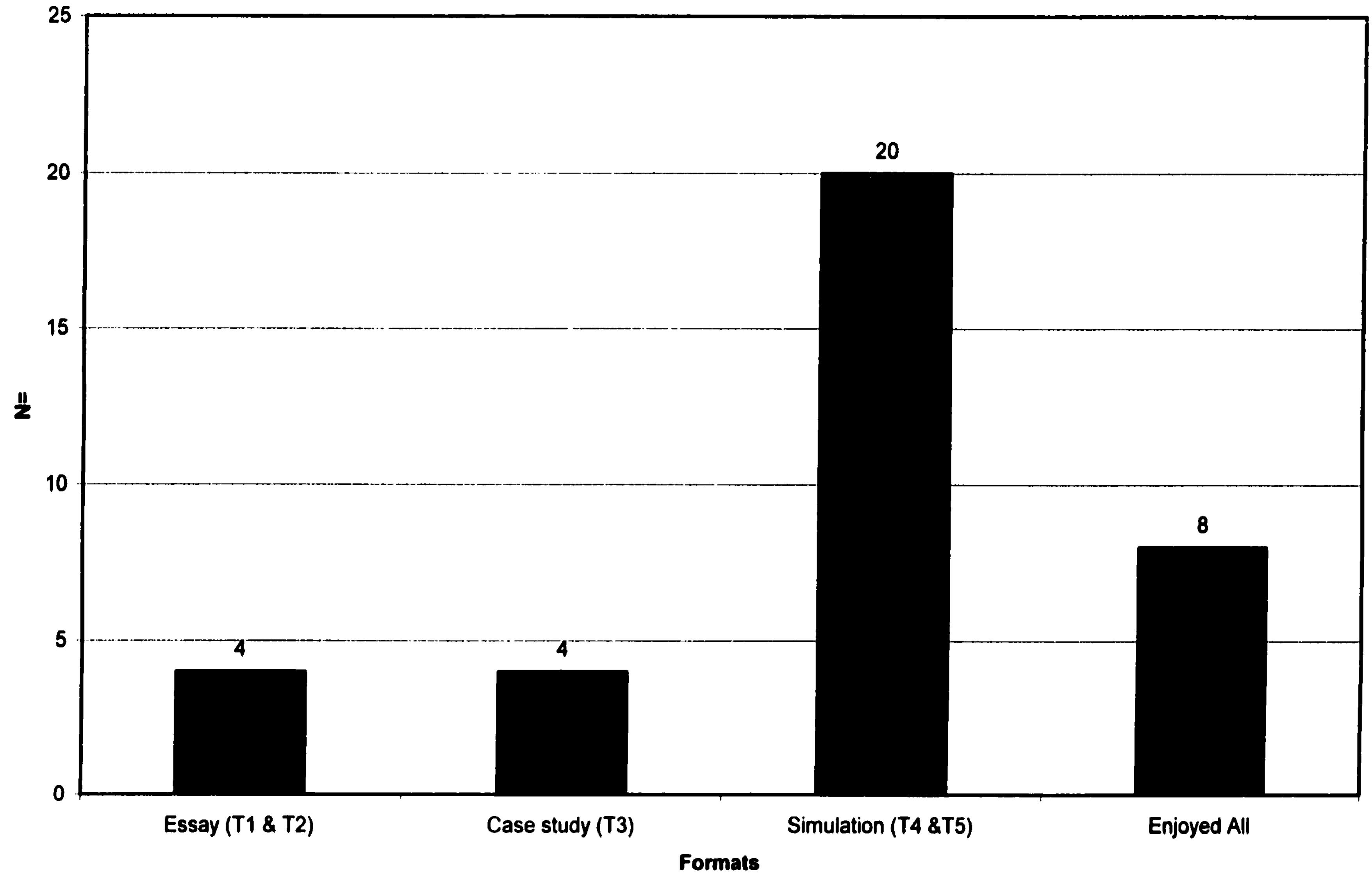
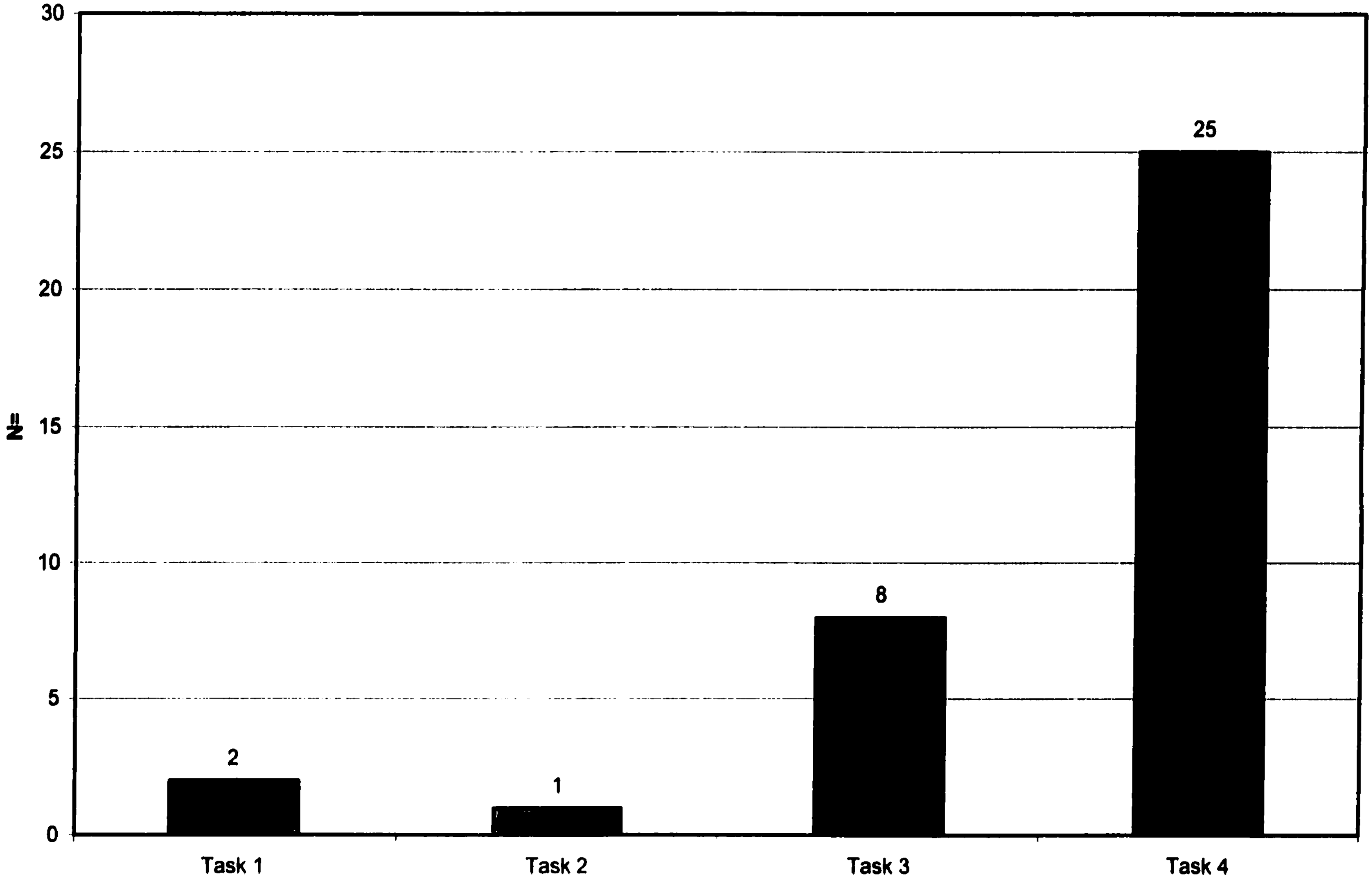


Figure 4.12 Most Demanding Task



CHAPTER 5 : DISCUSSSION

This chapter is organized in two sections. In the first section, the main results on Tasks 3, 4 and 5 are analyzed. The second section reflects on the process of conducting research on critical thinking skills.

Summary of Results

There are three main observations to draw from the preceding chapter. First, the substantial and significant change obtained in transition from Task 2 to Task 3 indicates that complexity level 1 triggered a higher incidence of critical thinking skills. In addition, the set of skills used is different than the ones used in other Tasks.

Secondly, although there is a significant difference in the use of critical thinking skills in the transition between Task 3 and Task 4, the direction of the results is not the one that was predicted. In addition, the absence of evidence of transfer between Task 4 and Task 5 suggests that complexity level 2 had detrimental effects. The fact that Task 5 generated the lowest score among all Tasks suggests there are other factors at work here. Since some students used the option in Task 5 of creating scenarios or dialogues, while others wrote typical essays, I have analysed below the distribution of skills for these different types of outputs in order to examine competing hypotheses.

Thirdly, the absence of a significant difference between Task 1 and Task 2 in the treatment group, plus the absence of difference on Task 2 between treatment and non-treatment groups suggest, on one hand, that essays tend to generate a certain type of skills; and on the other hand, that these scores might legitimally be used as a baseline against which we can test differences. We observed, however, that for the category "Judgement and Interpretation", there is a difference between Task 1 and all other tasks, attributable to scores on the indicator "Support". Since this outcome is not of interest in this study, I will not pursue the issue further.

Analysis of results

Task 3

As we have seen previously, Task 3 obtained the highest ratio of critical thinking skills (73%) across Tasks. There are striking differences in the types of skills used in this case study. For many skills, Task 3 generated the greatest number of extremely high scores, and interestingly, the highest number of extremely low scores as well. Table 5.1 presents the highest and lowest scores for each skill for all Tasks..

Table 5. 1 Highest and Lowest Scores Across Tasks

CT	Sub-Categories	Highest Scores		Lowest Scores	
		Task	Score	Task	Score
JI	Defines	T5	1.92	T3	0.48**
	Identifies assumptions	T4	18.95*	T3	14.94
	Makes connections	T2	6.51	T3	1.79
	Evaluates	T3	6.00	T5	2.38
	Supports	T1	17.91	T3	5.46
MP	Challenges	T2	0.78	T3	0.48
	Looks at the other side	T2	1.89	T4	0.45**
	Suggests alternative	T3	5.17	T1	1.47
	Assumes questioner's role	T5	8.22*	T3	1.11
	Considers viewer's perspective	T4	1.2	T3	0***
IM	Recognizes	T3	4.45	T5	1.96
	Predicts or hypothesizes	T3	4.54	T5	1.93
	Recommends	**T3**	**22.24***	T1	0.71
	Summarizes	T3	1.98	T4	1.09
	Concludes	T5	0.20	**T3**	**0.05****
	Generates new ideas	T2	2.17	T4	0.72

NOTES: Scores in bold indicate the highest and lowest scores across all C.T. categories.
* Highest score within the C.T. category
** Lowest score within the C.T. category
*** Null score

Task 3 generated the highest scores for six indicators : "Evaluates", "Suggests alternative approach", "Recognizes", "Predicts and Hypothesizes", "Makes Recommendations" and "Summarizes". Four of these fall in the "Imposing Meaning" category.

Conversely, Task 3 generated the lowest scores for eight indicators. In the "Judgement and Interpretation" category these are : "Defines", "Identitifies Reasons and Assumptions", "Makes Connections" and "Supports"; in the "Multiple Perspectives" category, "Challenges", "Assumes Questioner's Role" and "Adopt Viewer's Perspective"; and in the "Imposing Meaning", there is only "Concludes".

In Task 3, the increase in "Imposing Meaning" from Task 2 is without doubt, the most important finding. The ratio rises approximately 28%, to 39%, (see Table 4.14). With the exception of "Judgement and Interpretation" skills in Task 1 (M_{T1} = 42.82, see Table 4.6), this score is the highest obtained in any C.T. category across all tasks. If we look at the distribution of the skills within the "Imposing Meaning" category, we clearly see a surge in "Makes Recommendations" (over 22%). This does have to be intepreted with caution. Since the students were required to identify the problem and find reasonable solutions, the nature of the task might explain the propensity to "Make Recommendations".

Nevertheless, the magnitude of the result obtained for the "Makes Recommendations" skill is interesting for another reason. If we compare the score for "Makes Recommendations" against the next highest ratio within the "Imposing Meaning" category, we notice that both "Recognizes Various Impacts" and "Predicts or Hypothesizes" scored 4.5% (see Table 4.16). In other words, all the other scores are below 4.5%, which suggests that the ratios within the "Imposing Meaning" category normally tend to be small.

Another phenomenon observed in Task 3 is that "Judgement and Interpretation" accounted for 29% of all the skills used in the case study, a lower proportion than in any other task. We note also that the incidence of statements went down as well in Task 3 (see Figure 4.4). Here, I suspect that there might be a relationship between the use of "Statements" and "Judgement and Interpretation" skills. To that effect, Allegretti and Frederick (1995) suggests that a claim is an integral part of building an arguement. Logically, when analyzing facts, one needs a basis upon which to establish a judgement. Task 3 results suggest that the students probably bypassed the "Judgement and Interpretation" requirement and jumped into "Imposing Meaning" right away. In other words, given the clarity of the instructions and the goal of the case, the majority of students might have assumed that the reader knew what he or she was talking about and proceeded immediately with the interpretation of the situation without establishing premises for discussion. This interpretaion demands further analysis, beyond the scope of this study.

To summarize Task 3, we might say that the complexity variable level 1 as embedded in the case study approach triggers not only the specific skill of "Making Recommendations", but also most skills under the "Imposing Meaning" category, even if the ratios tend to be small. In addition, we note that "Statements" and "Judgement and Interpretation" skills tend to go down in case study. This evidence suggests that case study produced not only quantitative differences in the incidence of critical thinking skills, but also a qualtitative differences: case study triggered "Imposing Meaning" skills but not "Judgement and Interpretation" nor "Multiple Perspectives" skills.

Task 4 and Task 5

We have seen in the preceding chapter that the results of the simulations run contrary to expectations. The students did not increase their use of critical thinking skills from Task 3 to Task 4 and worse, the mean score in Task 5 is the lowest of all Tasks.

There are three possible reasons for the results. First, the complexity Level 2 variable is too complex to trigger critical thinking skills, at least under the conditions of the study (individual work). Secondly, the nature of the task might trigger more statements and varia units than critical thinking skills.

The students were relatively free to organize the content of the Task and to select the desired format. As a result, some students wrote regular essays, dialogues, or even a "letter" to the "Television Show Producer". Such a diversity of outputs was unexpected. In addition, as we will see subsequently, some students kept the same strategy for Task 4 and Task 5, and others changed approaches for the two assignments. Given the diversity of outputs on the Tasks, we may wonder whether there is a difference in the ratios of critical thinking skills, according to the type of outputs. For instance, do "Essays" tend to trigger larger ratios of critical thinking skills than the "Scenarios"?

I suspect that while the dialogue format as a whole is a more creative endeavour, the unit of analysis used in this study is not likely to reveal the incidence of critical thinking skills. A comparison of the three types of outputs is offered as a means to support this suggestion.

Finally, Task 5 was carried out in different conditions, which might have had an impact on the treatment. The factors will be explored subsequently.

Complexity Level 2

The decrease in use of critical thinking skills in the transition from Task 3 to Task 4 might be related to different types of difficulty in the tasks. While complexity Level 1 was well handled in Task 3, complexity Level 2, embedded in Task 4, might have been too demanding under individual learning conditions. The components of the simulations (role, mandate, context), as well as the quantity of information to sift through (see Appendices XIV & XV), might have played a role. To create the simulations, a situation as proposed by Leenders and Erskine (1989), was created using guidelines from Jones (1985) and Thiagarajan and Stolovitch (1978).

The main feature of the simulation is that the components extracted from reality (roles or functions, context and mandate) are distributed among the participants. In other words, the complexity of the situation is shared between the players. The complete picture is pulled together during the interaction. In this study, the components of the simulation were handled individually, thus affecting the level of engagement of the student in the Task: The role of consultant for example, was easy to disregard without affecting the outcome on the Task. In sum, it is suggested that the level of complexity embedded in the simulation tasks might be efficient only in group interactions, not in individual learning situations.

Furthermore, the negative outcome with regards to Tasks 4 and 5 might be related to decisions made about the research design of the study: Since the main object of research was to investigate the effect of task design on the incidence of critical thinking skills, I chose to control variables such as the instructor's interventions and individual feedback that may have threatened the internal validity of the study. For that reason, I excluded the conceptual dimension of Leenders and Erskine's model of complexity (see Table 3.3). As we can see on level three of the case difficulty, the conceptual dimension requires "extensive clarification through class lectures" (Leenders and Erskine, 1989; p. 117). In other words, the contradictory findings of this study tend to support the fact that the intervention of the instructor is necessary to trigger the use of critical thinking skills in complex learning tasks such as the simulations.

Rival hypotheses : Task 5

In this section we explore reasons possibly underlying the low level of critical thinking skills in Task 5. Specifically, I suggest that two factors had a reverse effect on the outcome on critical thinking skills: the perceived value of Task 5 and the time-spent on task. Contrary to expectations, the value of the assignment and the time-spent on task were probably deterring factors, rather than extrinsically motivating.

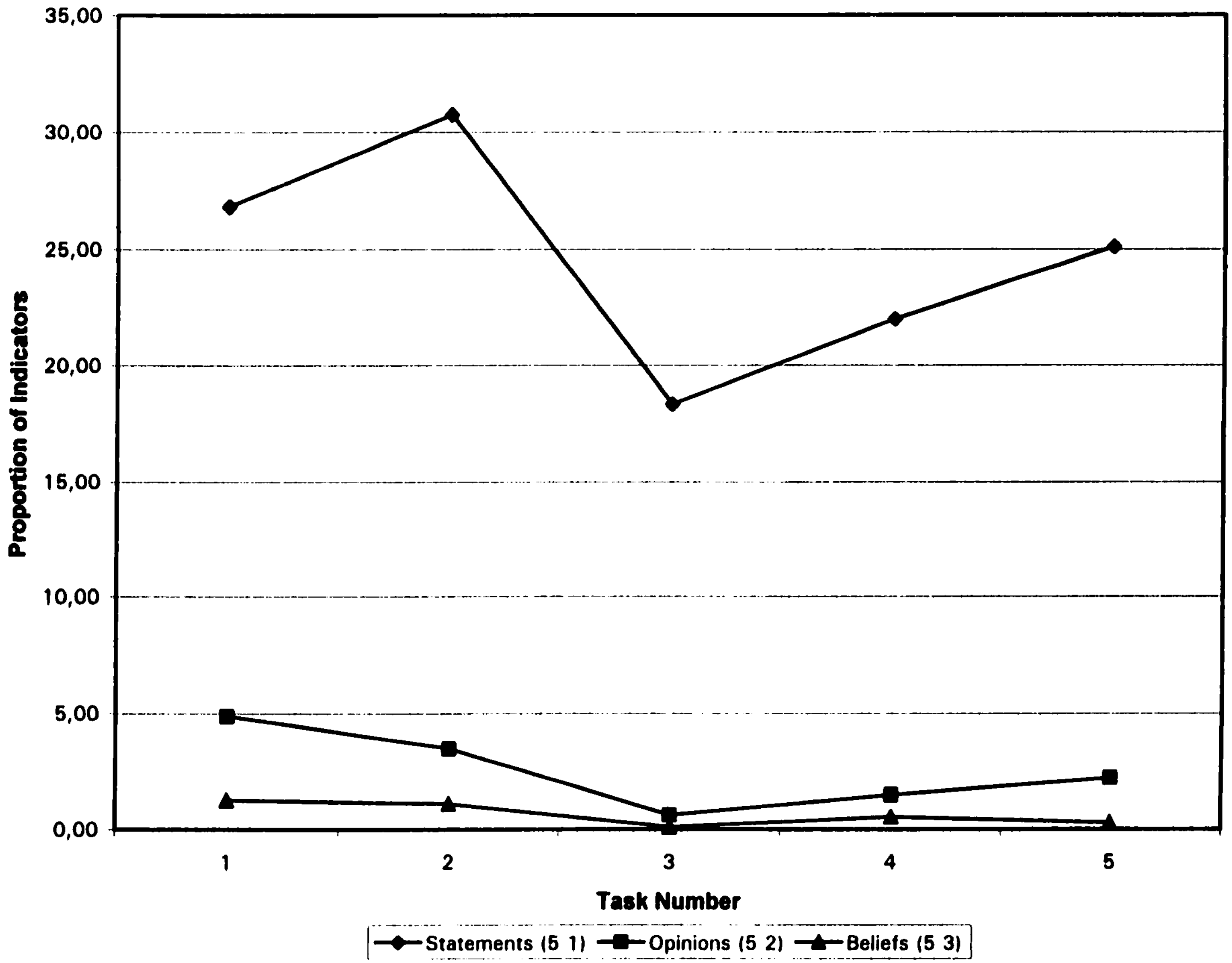
Tasks 1 to 4 were graded for participation and effort, for a total of 10%, whereas Task 5 was worth 20% of the final grade (see Table 3.6: Calendar of events). The lower outcome on the critical thinking skills is counter-intuitive since we generally assume that students would put more effort into an assignment which is worth more in terms of grades.

Tasks 1 to 4 were read by the researcher only. The participation grades were posted in WebCT after the Short Papers (Task 5) had been handed in. This procedure had been adopted to prevent sending any message or stimuli that could influence the level of effort put into a given Task. In other words, I tried to control individual feedback.

Task 5, on the other hand, was marked by both the researcher and the professor, graded for content. I suggest, tentatively, that the higher value of the assignment encouraged the students to put more effort into Non-critical thinking skills, such as the organisation of ideas and the formatting of the paper, than in the first four less formal Tasks. I speculate that the students spent more time formatting the paper for the benefit of the markers than they did in the less formal assignments, to obtain a better grade. In other words, the value of the paper deters the use of critical thinking skills in favour of better organized content. This claim seems to be supported by the increase in the number of "Procedural" statements in the simulation assignments.

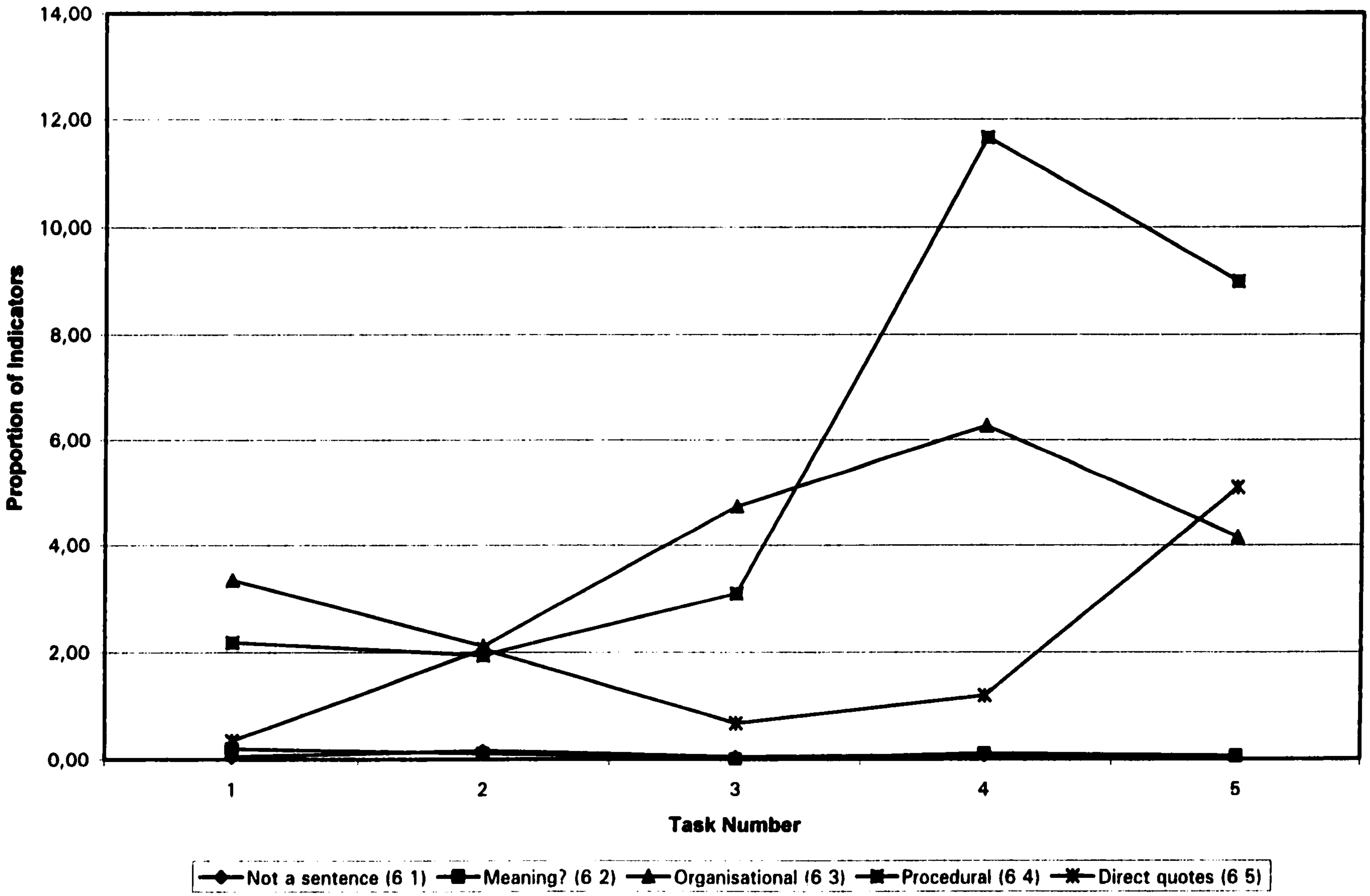
If we compare Figure 4.5 (J.I. trends), Figure 4. 6 (M.P. trends), and Figure 4.7 (I.M. trends), the difference between Task 4 and Task 5 on the indicators is minimal. What changes is the level of the Non-C.T. categories. In Figure 5.1, we note a slight increase for the indicator "Statements".

Figure 5.1 Trends: Statements Indicators



In addition, the proportion of procedural units went up from 3% in Task 3 to 11% in Task 4 (see Figure 5.2). Although the incidence of procedural statements decreased to 9% in Task 5, the simulation format tends to trigger more procedural statements relative to the other types of tasks.

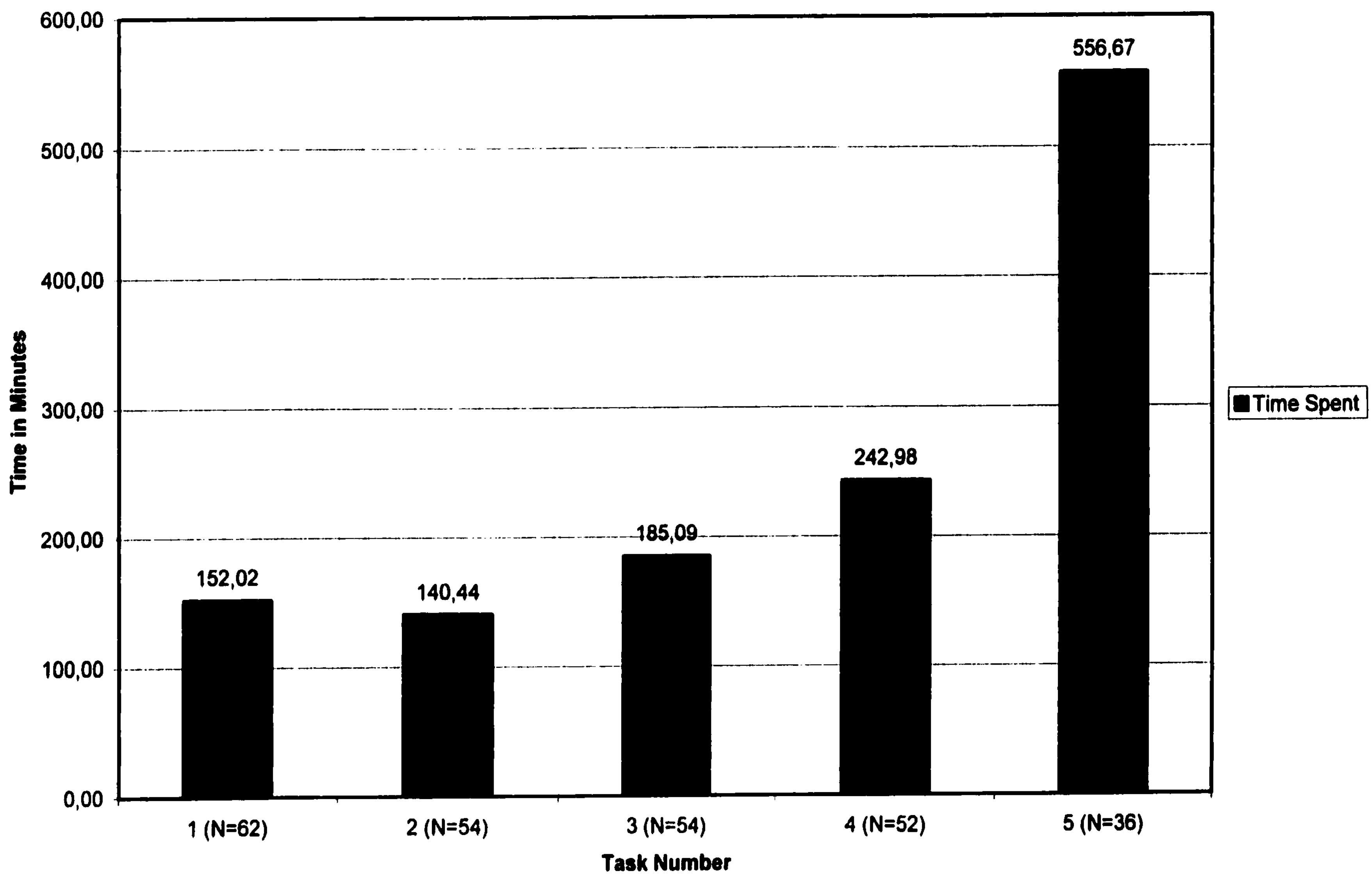
Figure 5.2 Trends: Varia Indicators



But perhaps a clearer indicator supporting the "formal organization" process is the increased use in Task 5 of "Direct Quotes", which constituted 5.1% of the text units as opposed to 1.2% in Task 4. The use of quotes might be related to the way students perceive how to write a paper that will be graded for content. The difference on quotes between Tasks 4 and 5 provides a hint of a rival hypothesis.

Another contrary result is that the students spent more time on Task 5 than in any other Tasks. Normally, we might relate time-spent to effortful thinking, which is a disposition towards critical thinking skills. In fact, the average time spent on Task 5 is more than 2.5 times greater than that spent on Task 4 as shown in Figure 5.3.

Figure 5.3 Time Spent on Task



However, this result has to be evaluated with caution. First, not all students completed the questionnaire, with only 36 reporting their time spent on Task 5. Secondly, this is a self-reported measure, so there was no way to verify whether the indicated times represent the actual time spent. Thirdly and most importantly, the validity of the measure is weak. Although a definition indicated what to include (searching for information, reading, brainstorming, writing and editing; see first item of the Post-Task Questionaire in Appendix VIII) we cannot be sure what a student's definition of "time spent on Task" comprises.

The rival hypothesis has one important implication for educational practice. The impact of grading, the effort that the students will put into formalizing a paper have to be considered when conceiving insruction that aims at fostering critical thinking skills. In this regard, Wade (1995) provides only feedback on written assignments, but does not mark the assignments.

Type of outputs: Tasks 4 and 5

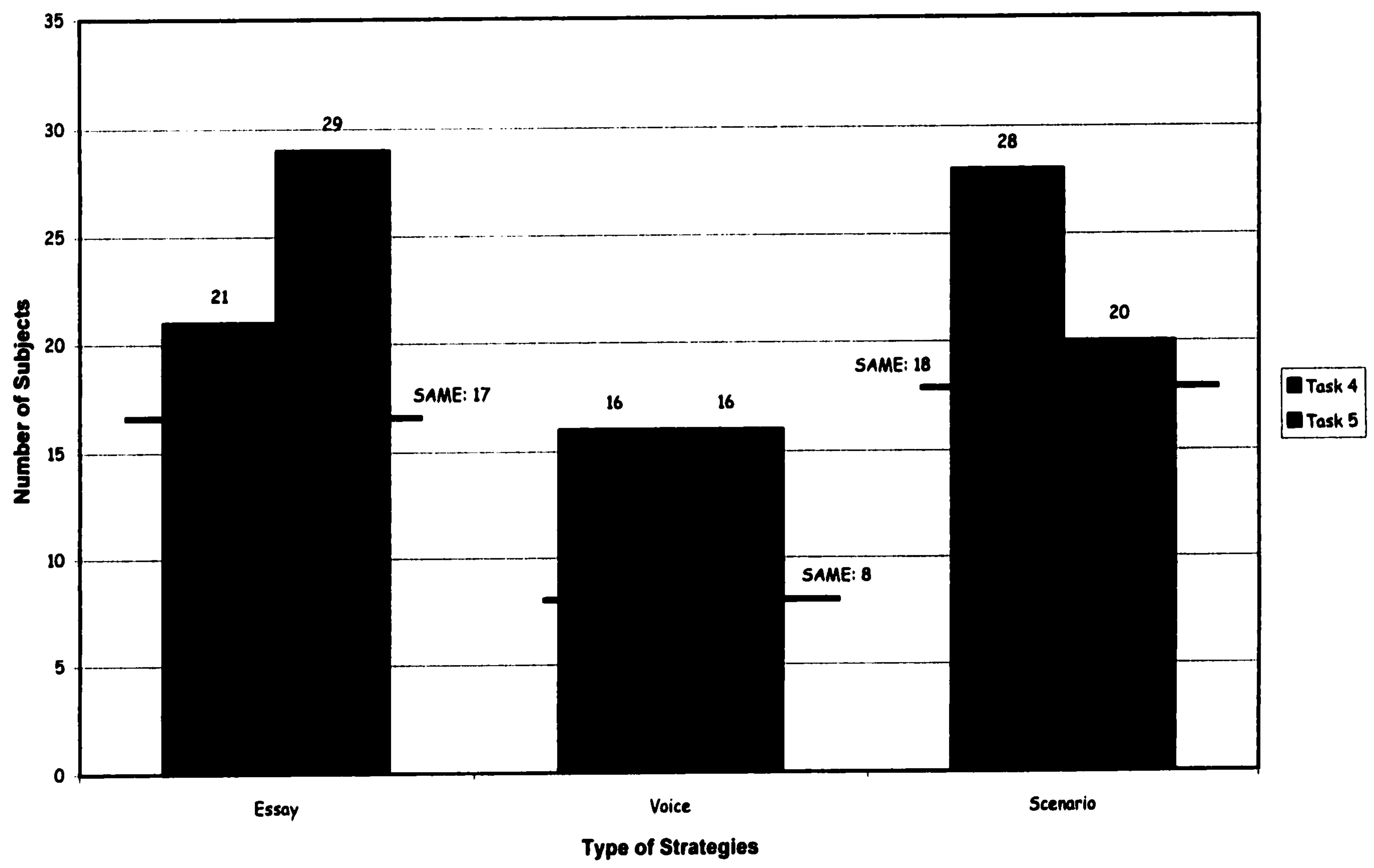
The diversity of outputs for Tasks 4 and 5 might have an impact on the use of critical thinking skills. In this section, I compare the differences obtained under three different conditions.

Three categories were used to sort the student's outputs: "Essay", "Voice", and "Scenario". Output were classified as "Essay" if there was absolutely no mention of the role of consultant or writer of a TV episode. Both content and format defined essays on the topic of "Food Production" or "Genetically Modified Foods".

The "Scenario" condition, on the other hand, was defined as either a dialogue or a synopsis. In a dialogue, at least two characters were depicted talking together. In a synopsis, the students outlined their intention by providing a plan that contained specifications for the course of the TV episode, the characters portrayed, the content covered, sometimes time allocations for each segment, or even recommendations for visuals that should appear on the screen.

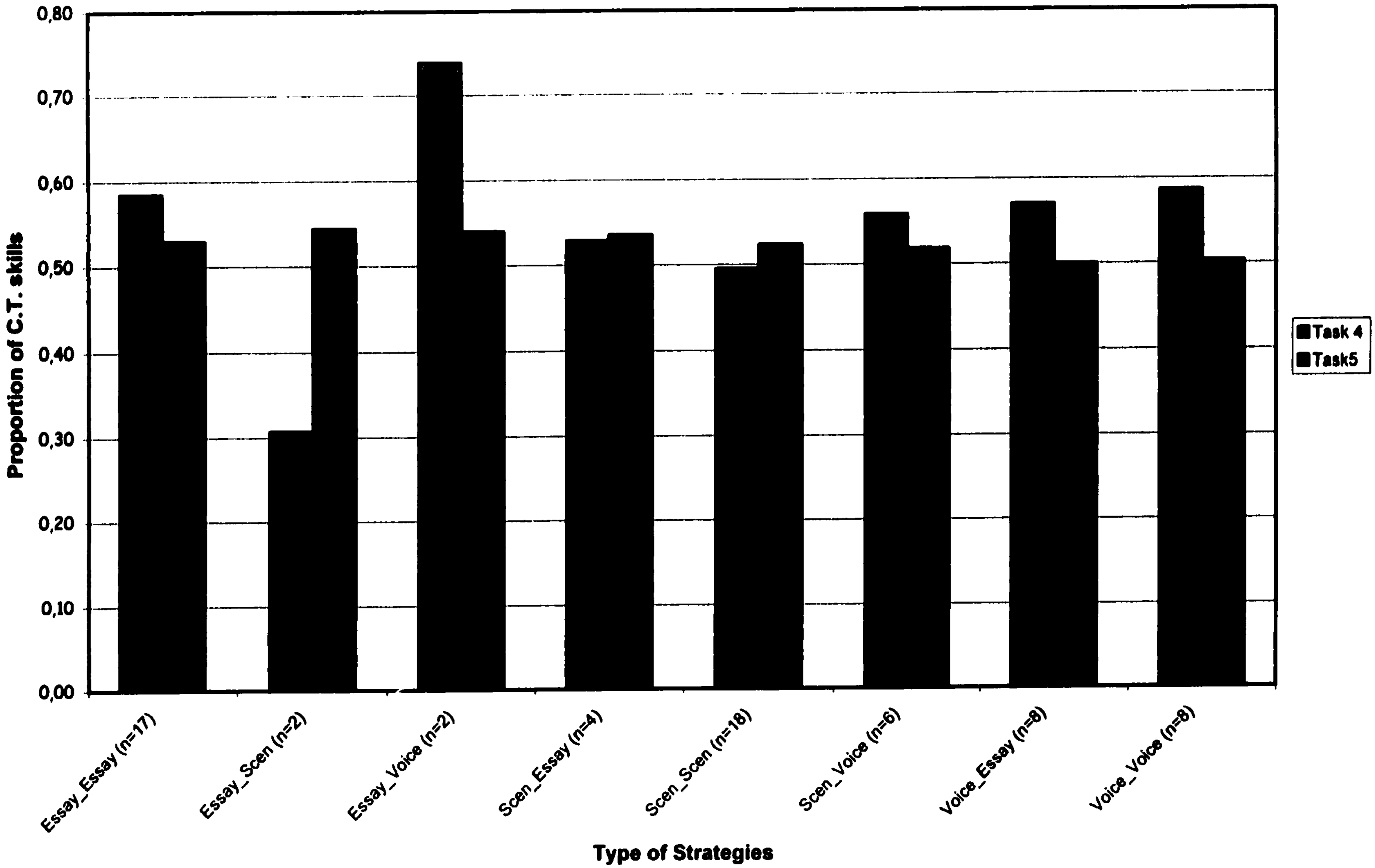
The "Voice" category included outputs which depicted some kind of acknowledgement of the consulting role or mandate. Typically, the student would mention, in the introduction or in the rationale (Part 2 of the Task, see Appendix XV), the topic of the show or refer to the previous mandate (i.e. Task 4) but would proceed with a regular essay for the remainder of the text. One unexpected type of output that also falls under the "Voice" category is a letter to the "show producer" in lieu of holding a more formal discourse, as required in part 2 of the Task. With one exception, the students who wrote "Scenarios" did not write letters to the "show producer". The distribution of the number of students per selected strategy is presented in Figure 5.4.

Figure 5.4 Strategies Selected for Tasks 4 & 5 (N=65)



One interesting feature of this classification is that a greater number of students (N = 29) wrote essays for Task 5 as opposed to Task 4 (N = 21). Conversely, 28 students created a "Scenario" for Task 4, but only 20 chose that strategy for Task 5. Sixteen students chose the "Voice" strategy for both Tasks. On Figure 5.4 we see that the majority of students chose the same strategy for Tasks 4 and 5 (N = 43), whereas some made a different choice (N = 22). Among those who chose the same strategy, 17 students wrote essays, eight adopted the "Voice", and the largest group created "Scenarios" (N = 18).

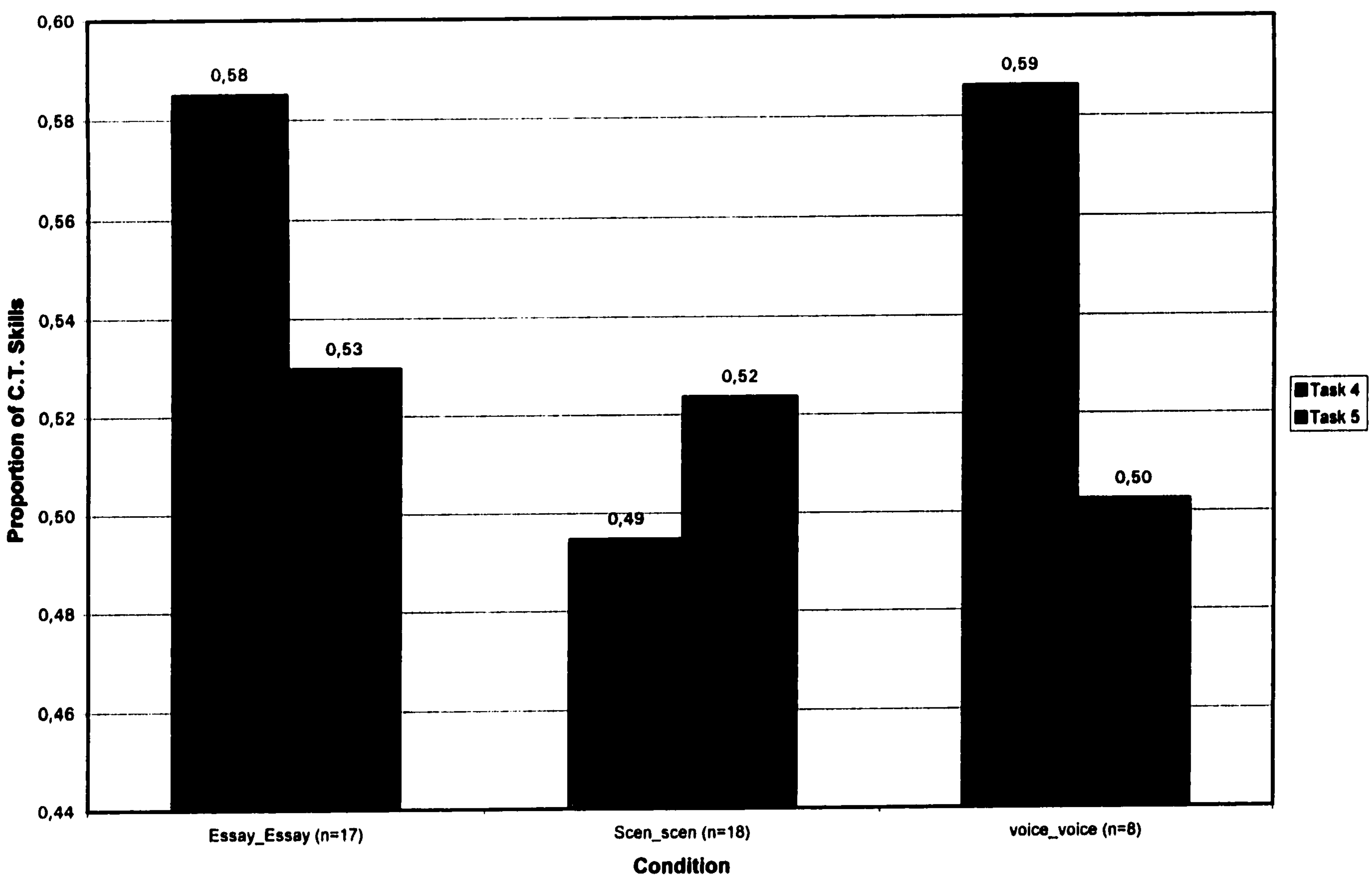
Figure 5.5 Distribution of C.T. Skills by Type of Strategies for Tasks 4 & 5 (N=65)



In contrast, the smallest groups are "Essay-Scenario" (N = 2) and "Essay-Voice" (N = 2). No student chose the strategy "Voice-Scenario". The combination does not appear on the graph. Given the wide variety of outputs, only the outcomes of those who kept the same strategy will be analyzed (see Figure 5.5).

In Figure 5.6, we note that the students in the "Essay-Essay" condition produced more critical thinking skills in Task 4 (58%) than in Task 5 (53%). The same is true for students in the "Voice-Voice" condition, with 59% in Task 4 and 50% in Task 5. Only the "Scenario-Scenario" group improved in this respect, by a modest 3%.

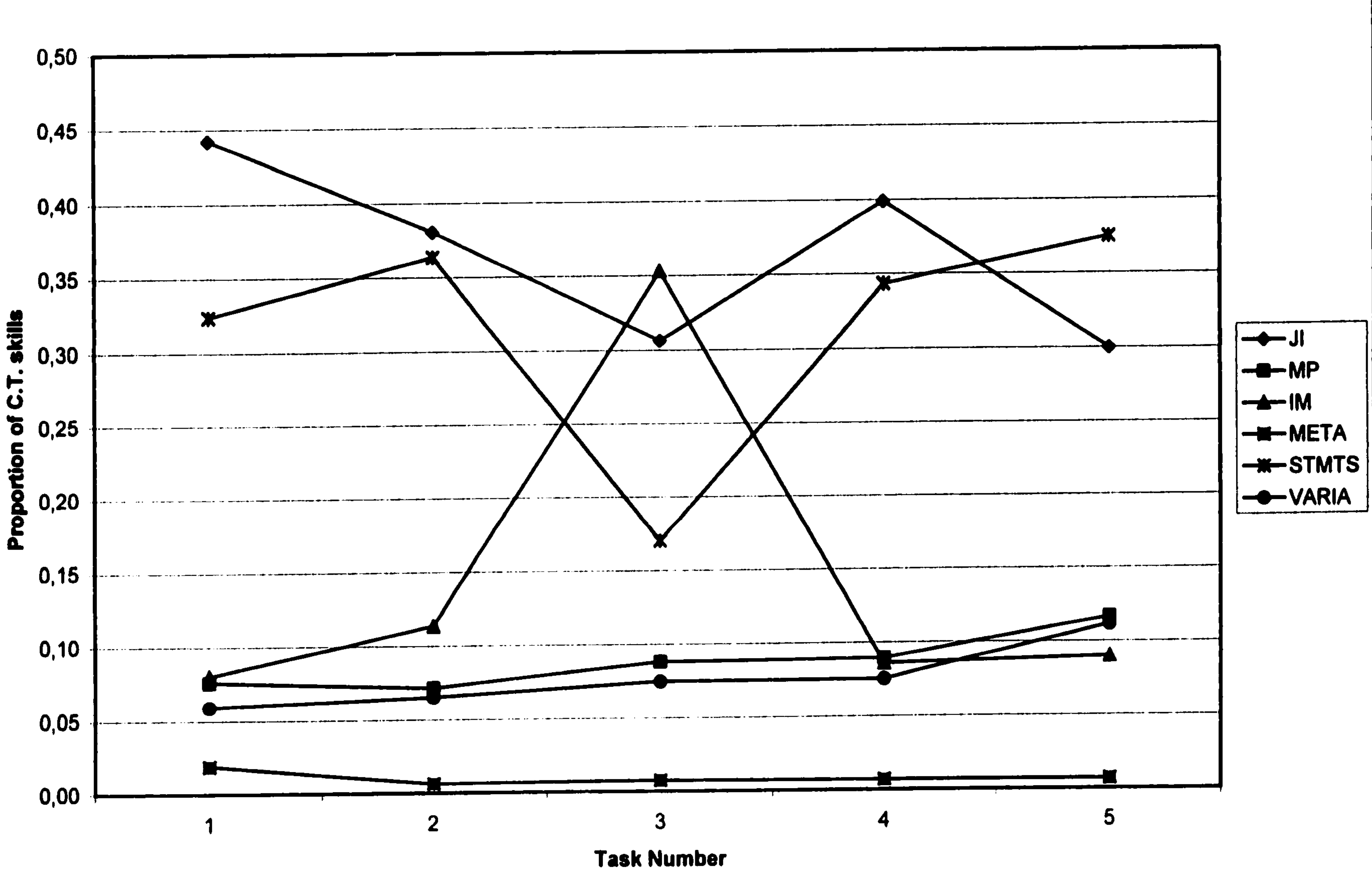
Figure 5.6 Proportion of C.T. Skills by Condition for Tasks 4 & 5 (N=43)



The slight variations in total critical thinking skills between Tasks 4 and 5 in same-strategy groups, and the fact that the critical thinking score is more or less the same in Task 5 for the three conditions tend to support the rival hypothesis suggested earlier, that given the nature of the task, the unit of analysis might not be appropriate to reflect actual use of critical thinking skills.

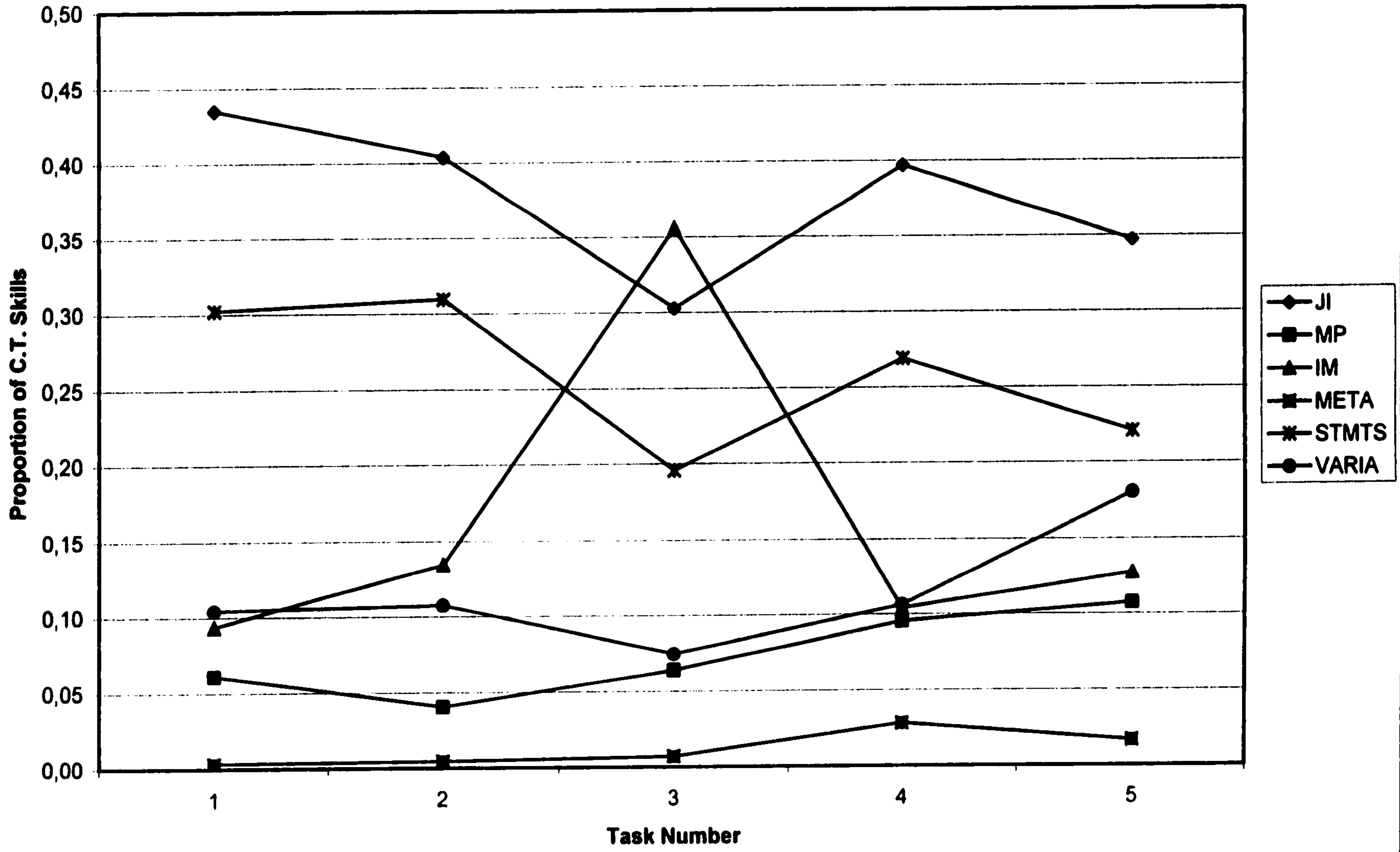
On the other hand, there might be a difference in the type of skills used by type of output. Figure 5.7 presents the trends for all coding categories for the "Essay-Essay" condition (N = 17).

Figure 5.7 Trends: All Coding Categories
ESSAY-ESSAY Condition (N=17)



If we compare the trend with Figure 4.4, the most important observation is that both the shape of the slopes and the magnitude of the scores are strikingly similar for all the critical thinking skills categories. The only variations occur in Non-C.T. categories. There is an increase of approximately 22% of "Statements" between Task 3 and Task 4. The incidence of statements peaks in Task 5 at 35%. "Varia" units on the other hand are relatively stable in the Essay-Essay condition. This suggests that the increase in the "Varia" units observed in Figure 4.4 is not attributable to the students in the "Essay-Essay" condition.

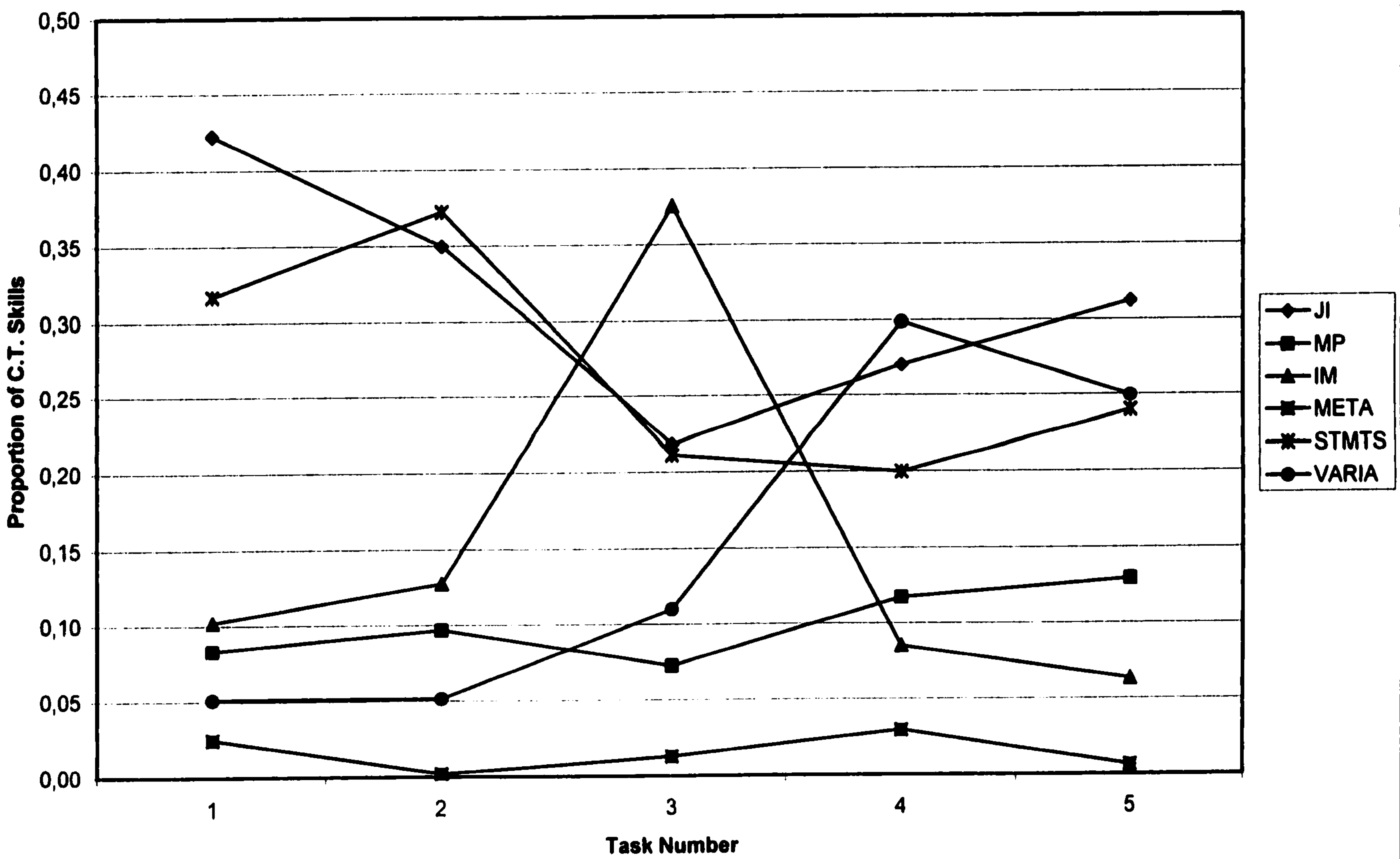
Figure 5.8 Trends: All Coding Categories
VOICE-VOICE Condition (N=8)



In Figure 5.8, the trends for the "Voice-Voice" condition (N = 8) are presented. We note the almost perfect similarity of the "Judgement and Interpretation" trend and the "Statements" trend. This finding suggests that events in these categories are related and occur together. If we compare Figure 5.8 with the general trends (Figure 4.4), we note that, as in the "Essay-Essay" condition, the main differences occur in the Non-C.T. categories. The students in the "Voice-Voice" condition produced more statements in Task 4 (27%) than the overall sample (24%). On the other hand, they produced less "Varia" units in Task 4 (10%), than the overall sample (19%). We note however, that the students in the "Voice-Voice" condition used more "Varia" units in Tasks 1 and 2 (11%) than the overall sample (6%). Similarly, the incidence of "Judgement and Interpretation" skills tends to be higher in the "Voice-Voice" condition across tasks than the whole sample. For instance, a 3% difference is observed in Task 2 and Task 5, and a 5% difference in Task 4. If we compare the "Voice-Voice" condition with the "Essay-Essay" condition, we note that aside from the differences observed for the Non-C.T. units, the trends are relatively similar.

The greatest variations on the trends on the coding categories are observed in the "Scenario-Scenario" condition (N = 18) (see figure 5.9). Moreover, the variations occur not only in the simulation tasks as expected, but in Tasks 1, 2, 3 as well.

Figure 5.9 Trends: All Coding Categories
SCENARIO-SCENARIO Condition (N=18)



Once again, the greatest variations occur in the Non-C.T. units categories. In the Scenario-Scenario condition, fewer statements were produced in Task 4 (20%) and in Task 5 (24%), than in the overall sample (T4 = 24%, T5 = 27%) (See Figure 4.4). Curiously however, the proportion of statement units in Task 2 (37%) is marginally greater than the proportion of "Judgement and Interpretation" skills (35%). In other words, the students in the "Scenario-Scenario" condition produced as much or more statements than critical thinking skills. This is not true for the other conditions (see figures 5.7 and 5.8), where there is always a greater incidence of "Judgement and Interpretation" skills than "Statements".

A similar phenomenon is observed in Task 4 (Figure 5.9). There is always a greater proportion of "Varia" units used in Task 4 (30%) than critical thinking skills. This represents a difference of 11% between the scenario students and the group (see figure 4.4). The proportion of "Varia" units in Task 5 is still higher higher in the "Scenario-Scenario" condition (25%) than in the overall sample (18%).

We have seen so far that the outcome of the Non-C.T. categories in the simulation Tasks vary from one condition to another. Specifically, the students in the "Scenario-Scenario" condition produced more "Varia" units than the students in the other two conditions. This however, can be related to the nature of the output. To create a dialogue for example, requires a different organisation of the text. In addition, the sentence as a unit of analysis to code the content of students' assignments creates a more conservative outcome. In other words, the overall creative aspect of the dialogue is not measured by the current choice of text units. In the following example, the facts have been skillfully integrated in a dialogue. The entire excerpt suggests evidence of critical thinking.

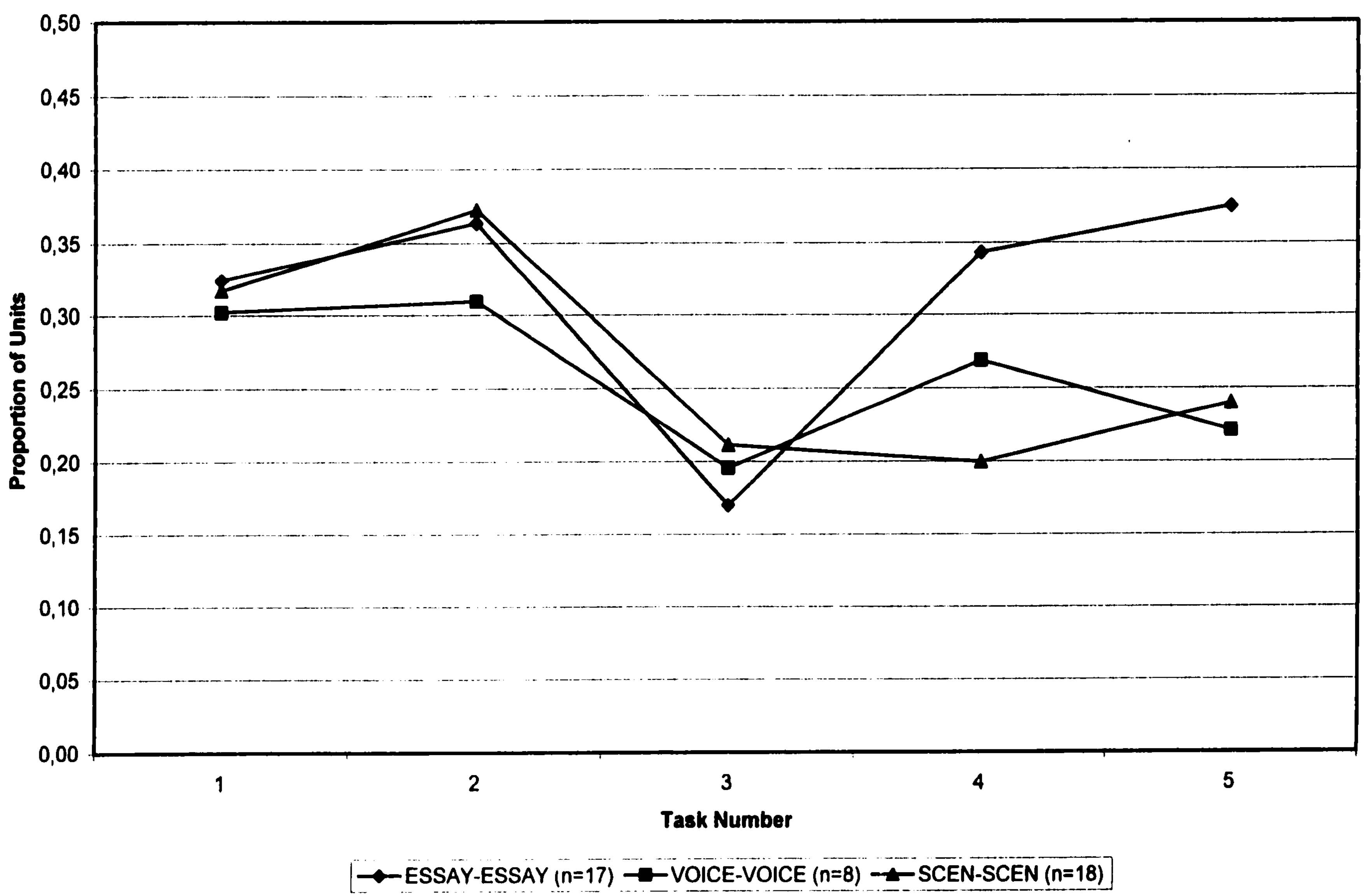
However, from a technical point of view, all the sentences but the questions asked by « Impacts » are Non-C.T. units.

> We have invited an expert from the FAO (Food and Agriculture Organization) from the UN to talk about the subject.
>
> "Impacts": Exactly how many people in the world are affected by hunger?
>
> Expert: About 800 million people in the world go hungry every day.
> The worst part about these statistics is that 200 million are children.
> In the next 30 years, the population is expected to rise by three billion people.
>
> "Impacts": These are alarming statistics!
> Will there be enough food produced to keep up with this rapid growth?
>
> Expert: According to a researcher, as of today there is actually a surplus of food being produced.
>
> **He proclaims that if the global food supply were converted to calories and divided by the world's population, there would be enough food for roughly 12 percent more than the actual population.**
>
> "Impacts": How is this possible?
>
> It seems like the world population is growing too fast to keep up.
>
> Expert: Actually, although world population has rapidly grown in recent years, food production has risen even faster.(FROM 117_T4).

It is clear that the counter-intuitive results observed in Tasks 4 and 5 are attributable to the nature of the Tasks as well as to extraneous variables. If we look at Figure 5.10, we note a decrease in statements for both "Essay-Essay" (N = 17) and

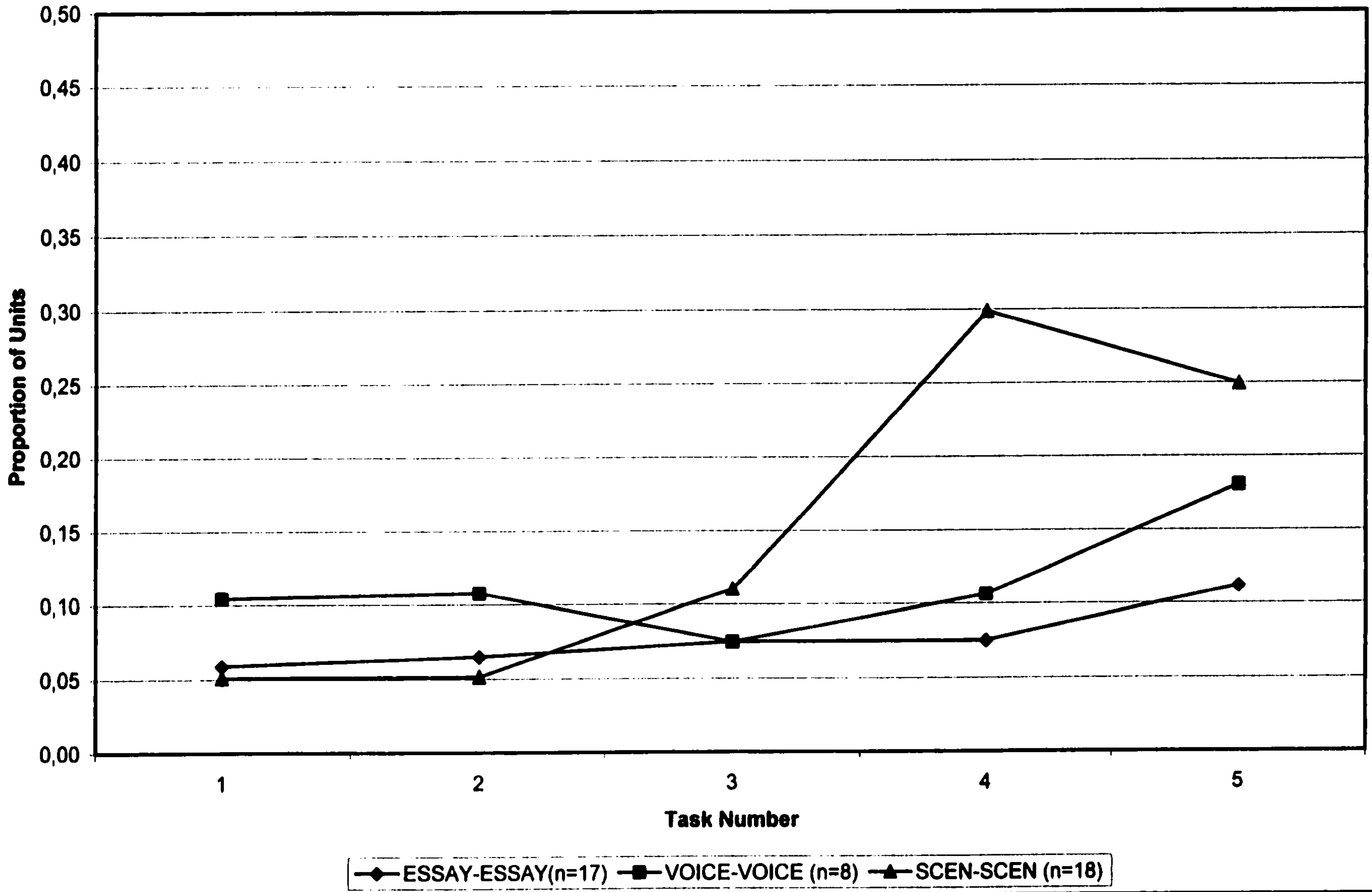
“Voice-Voice” (N = 8) and “Scenario-Scenario” condition (N = 18) in Task 3; only the “Essay-Essay” rebounds from this decline in Tasks 4 and 5. Improvement on critical thinking skills is not evident here.

Figure 5.10 Statements: Proportion of Units per Task by Condition (N=43)



In Figure 5.11 we note an increase over Tasks 4 and 5 of the "Varia" units for all three conditions, most prominently in the "Scenario-Scenario" case. Simulations evidently trigger more use of "Varia" units than other type of tasks. This result, combined with the high "Varia" score of the "Scenario-Scenario" condition, supports the rival hypothesis that since Task 5 was evaluated for content and was worth 20%, the students were more attentive to formal aspects of their assignments. Practical implications will be discussed later in this chapter.

Figure 5.11 Varia: Proportion of Units per Task by Condition (N=43)



In short, I tentatively suggest that the low level of critical thinking skills in Task 5, is mainly due to the fact that the paper was worth more than the other Tasks. Furthermore, grading for content causes the students to spend more time on formatting as well as searching and integrating quotes in the text. This hypothesis partly explains the rise on Non-critical thinking skills in Task 5.

In summary, we have shown that there are some qualitative differences between type of outputs in Tasks 4 and 5. While the distribution of critical thinking skills tend to be similar in all three conditions, differences are observed in the frequency of Non-C.T. units.

Reflections on the research process

In the literature review, we have seen that critical thinking is not a common event (Ennis, 1993) and that it is believed to happen in novel situations (Halpern, 1998; Garrison, 1991). This implies that the frequencies expected in sampling should generally be small. When conducting research we are looking for significant numerical differences in the quantitative paradigm, or salient characteristics of a phenomenon in the qualitative paradigm. Unfortunately, in the field of critical thinking, the salient skills are less frequent and often overshadowed by the ones that are more frequent. In addition, the more discrete skills are often ignored in research (see for example Anderson et al., 2001; Norris, 1989). In the classroom, instructors focus on specific set of skills (Wade, 1995; King, 1995; Brovey, 1988; Hanley, 1995; Allegretti & Frederick, 1995), despite the vast aray of skills to cover (see Ennis' taxonomy; 1987). Consequently, the full repertoire of critical thinking skills is seldom completely covered.

In the next section, I address the issue of the least frequent skills. Given the structure of the study, I will be able to show which skills tend to be least frequent across tasks and which skills are not triggered in any mode of instruction. This analysis should provide clues as to how to design instruction to address those skills.

Least frequent skills

Table 5.2 presents a synthesis of the skills that accounted for less than 5% in each type of Task. The scores over 5% as well as the results on Non-C.T. units have been included for illustrative purposes.

Table 5. 2 Synthesis of the Skills That Accounted for Less Than 5% of the Units per Type of Task

C.T. Sub-Categories	TYPE OF TASK		
	Essays (T1-T2)	C.S.(T3)	Simulations (T4-T5)
JUDGEMENT & INTERPRETATION			
Defines	•	•	•
Identifies assumptions	18%	14.9%	19%
Makes connections	6.5%	•	•
Evaluates	•	6%	•
Supports	17.2%	5.5%	7.8%
MULTIPLE PERSPECTIVES			
Challenges	•	•	•
Looks at the other side	•	•	•
Suggests alternative	•	•	•
Assumes questioner's role	•	•	8.2%
Considers viewer's perspective	•	•	•
IMPOSING MEANING			
Recognizes	•	•	•
Predicts or hypothesizes	•	•	•
Recommends	•	22.2%	•
Summarizes	•	•	•
Concludes	•	•	•
Generates new ideas	•	•	•
NON-CRITICAL THINKING SKILLS			
STATEMENTS			
Statements	30.8%	18.3%	25.1%
Opinions	•	•	•
Beliefs	•	•	•
VARIA			
Organisational	•	•	6.3%
Procedural	•	•	11.7%
Direct quotes	•	•	5.1%

• Indicates less than 5% of the units demonstrated the skill.

As expected, critical thinking skills are not used frequently. In fact, the majority of the skills accounted for less than 5% across types of tasks. The category "Judgement and

Interpretation" has the most frequent scores above 5% which would support the findings of Herrinton and Oliver (1999). In other words, the macroability "Judgement and Interpretation" could be the most frequent category of skills generated during instructional activities. Two skills appear to be common to all types of tasks: "Identifies Reasons and Assumptions" and "Supports". For the former, the numbers are stable across types of task, always relatively high. We may conclude that the skill "Identifies Reasons and Assumptions" is necessary for building arguments and that it occurs in any type of Tasks. The model used by Allegretti & Frederick (1995) is a good example of such a claim.

More surprising however, is the fact that the categories "Multiple Perspectives" and "Imposing Meaning" almost never account for more than 5% of skills across type of Tasks. The only exceptions are "Assumes Questioner's Role" in the simulations with a ratio of 8.2%, but as explained before that result might be an artefact of the nature of the Task. The other exception is "Makes Recommendations" in the category "Imposing Meaning", which accounts for more than 22% of the total units used in Task 3.

Table 5.3 presents a synthesis of the skills that accounted for less than 1% of the units. The purpose of this synthesis is to identify which skills are not triggered by any type of task design. This will lead us to identify instructional design strategies that might generate the least frequent skills.

Table 5. 3 Synthesis: Least Frequent Skills per Type of Task (Less than 1% Accounted for).

C.T. Sub-Categories	TYPE OF TASK		
	Essays (T1-T2)	C.S.(T3)	Simulations (T4-T5)
JUDGEMENT & INTERPRETATION			
Defines		•	
Identifies assumptions			
Makes connections		•	
Evaluates			
Supports			
ADOPTING MULTIPLE PERSPECTIVES			
Challenges	•	•	•
Looks at the other side		•	•
Suggests alternative			
Assumes questioner's role		•	
Considers viewer's perspective	•	•	
IMPOSING MEANING			
Recognizes			•
Predicts or hypothesizes			
Recommends			
Summarizes			
Concludes			
Generates new ideas	•	•	•
NON-CRITICAL THINKING SKILLS			
STATEMENTS			
Statements			
Opinions		•	
Beliefs		•	•
VARIA			
Organisational			
Procedural			
Direct quotes		•	

• Indicates that less than 1% of the units is accounted for.

We note that the case study most frequently generated scores lower than 1%. As suggested earlier, the case study triggers a set of critical thinking skills to the exclusion of others. Case study as an individual instructional approach encourages the students to use skills that are more complex ("Imposing Meaning"), but at the expense of skills that are more analytical ("Judgement and Interpretation"). The only skills that are not least frequent are those that tend to be more evaluative such as "Evaluates" and "Suggests Alternative Approaches".

A second observation is that the category "Multiple Perspectives" most frequently generated the lowest scores, over all types of Task; the only skill within the category that generated more than 1% but lower than 5% across all Tasks is "Suggests an Alternative Approach". Once again, individual assignments might not trigger the use of "Multiple Perspectives" skills. As mentioned earlier, it was feasible for the students to disregard some information (e.g. context, network "bio" etc.) presented in Tasks 4 and 5 and still carry out the tasks successfully. Alternately, conducting the simulations as a group would have entailed the participants to gather and to share relavant pieces of information by asking questions or considering another point of view, in order to complete the picture and eventually solve the given problem. In short, it seems that instructional activities that engage the zone of proximal development, such as group simulations, might encourage different points of view, thus triggering the adoption of "Multiple Perspectives". In essence, collaboration and classroom interventions might be necessary to create cognitive dissonance in the students (King, 1994; Baron, 1987).

A third observation is that two categories accounted for less than 1% of skills across all types of Task. They are "Challenges" and "Generates New Ideas". By definition, these skills are complex and should not happen very often. This finding suggests that individual instruction and the absence of feedback is not sufficient to trigger these skills. Once again, collaborative work might be required. An instructor or facilitator might have to set up the conditions favorable for the emergence of these skills. The instructor might be able to manipulate the context to catalyze the use of those skills by adopting, for example, a modelling strategy. However, the incidence of "Generating New Ideas" will always remain low unless it is directly fostered.

Finally, a fourth observation regarding skills that are less frequent than 1% concerns the "Statements" category. We noticed that the frequency of "Beliefs" units went below 1% for both the case study and the simulations. In other words, " Beliefs" were stated less often as time went on. This could be construed as a positive outcome since most of the time beliefs are not generally supported by reason or logical argument in academic papers. In others words, the students improved over time and used "Beliefs" less frequently.

The study of critical thinking skills

These four observations have important implications for studying and promoting the use of critical thinking skills in students. As we have seen in this study, "Judgement and Interpretation" skills occur more frequently than "Multiple Perspectives" or "Imposing Meaning" skills. In a research context, high scores such as the ones observed for Judgement and Interpretation skills would normally overshadow outcomes on other critical thinking categories. In this study, we saw that Judgement and Interpretation skills accounted for over 15% of units across all tasks (see Figure 4.5). On the other hand, Imposing Meaning skills always accounted for less than 5% of the units (see Figure 4.7) with the exception of the indicator "Making Recommendations". Such differences in the outcome might bring the researcher to focus on more important findings, overlooking more subtle results. In this study, we have found that small mean differences were significant. As an example, all the comparisons with Task 4 and 5 were significant for the skill "Multiple Perspectives". Some mean differences are as small as 3% (see Table 4.11).

"Judgement and Interpretation" skills will tend to happen more often, at least in written assignments. In the example below, the student presents an alternative approach ("Multiple Perspectives" in italics) but the other sentences are all "Judgement and Interpretation" skills (in bold).

> *Such contradictory facts about labeling only one part of products leaves us to desire the real intentions behind this industry that makes astronomical profits for the sake of the human kind.*
>
> **Not only will it bring profit to food sales on the market, but also it will create new investments in genetic research, which will benefit a large part of the biotechnology sector.**
>
> **It is due to this fact that the majority of people are skeptical about such technology, as its humanitarian principles (such as to improve nutrition) are driven by profit.**
>
> **In fact, by looking at the following figure, we can see that only 65% of people 5 believe that genetic engineering will make life better.**

"Judgement and Interpretation" skills are not only used more frequently, they are also easier to identify. As we can attest by looking at Table 3.7, there is less inter-coder variation for the "Judgement and Interpretation" category than there are for "Multiple Perspectives" and "Imposing Meaning" skills. This suggests that identifying "Judgement and Interpretation" skills entails less interpretation than the other two coding categories. In other words, "Multiple Perspectives" and "Imposing Meaning" categories might be more connotative in nature than the "Judgement and Interpretation" skills. As Ahuvia (2001) suggests, deciphering latent meaning in focal texts is more difficult than identifying denotative units. Connotative meaning is a challenge when attempting to achieve inter-rater agreement. Training of coders becomes important especially if we consider the fact that those codes identify unusual skills.

Unit of analysis

Another important consideration when conducting a content analysis of critical thinking skills is the selection of a unit of analysis. For this study, a sentence was chosen as the basic unit of analysis. As we have seen however, the unit of analysis is not appropriate in the case of scenarios since a lot of "noise" is created by the "Varia" units. But perhaps more interestingly I suggest that a sentence as the sole unit of analysis is not sufficient to assess the more complex skills such as "Multiple Perspectives" and "Imposing Meaning". In fact, the sentence is probably valid as a unit of analysis for "Judgement and Interpretation" skills only, since they tend to be more simple.

Alternative units of analysis might be considered. Paragraphs might be a viable alternative, although not for the population in this study. Paragraphs are often ill-constructed, that is, there is more than one main idea in the paragraph. Worse, some students do not use paragraphs to organize their texts, or they create single sentence paragraphs. In addition, paragraphs often contains more than one critical thinking skills. Multiple coding could be allowed for either sentence of pragraph units but that would increase the difficulty of interpreting data.

Another alternative would be to use « units of meaning » (Henri, 1992) as it is often the case in computer supported discusion groups (Newman, Webb, & Clive,1995; Herrington & Oliver, 1999). The coder would organize the text following the perceived meaning. The problem with this is that is places a greater cognitive load on the coder, who not only has to identify the proper skill, but also has to "create" a unit of meaning. In addition, the conceptual organization around the meaning of units by the coder, at least in written assignment, might alter the original intention of the student. In other words, the content is somewhat re-interpreted by the coder.

Another alternative might be to code sentences first, and then look at the « peripheral » context of the skills. In other words, when a complex critical thinking skills occurs in a sentence, the strategy would be to look at the proximate sentences. A cluster would thus be created. This strategy would have the advantage of providing the context in which a skill is occuring. The main results would take the form of frequencies as in this study, but occasional clusters would be created to establish a relationship between an unusual code and the condition in which it happened. The challenge with this strategy is once again the role of the coder. Interpretation of what is contained in a unit would threaten the level of agreement between coders.

Reliability and internal validity in content analysis

Both issues of complex codes and the choice of unit of analysis speak to the challenge of establishing a satisfactory percentage of rater agreement. As we have seen in this study, despite the training of the coders, the negotiation of meaning for unusual codes and embedded strategies to address the internal validity aspects, the percentage of agreement achieved is still moderate. More complex skills such as "Imposing Meaning" and "Multiple Perspectives" lend themselves to connotative interpretation and open up the issue of test-restest reliablity of the instrument as well.

One solution might to be to triangulate measures. Although a content analysis of students' assignments revealed interesting information about the effect of task design on critical thinking skills, other means of evaluation could be used in order to understand the phenomenon better. Anderson et al. (2001), for instance, used classroom observations and related the results to written assignments. Another strategy might be to include a self-reported measure (see Tsui, 1999). A carefully designed questionnaire using a Likert scale in which the C.T. items would correspond to specific instructional events might inform the educational community about the context and the instructional factors contributing to the development of critical thinking (McMillan, 1987; Coffman, 1987). This in turn would provide an indication as to how and what to improve (Norris, 1985) in order to become, as Ennis (1987), suggests not only a critical thinker but a good critical thinker.

CHAPTER 6 : CONCLUSION

The objective of the study was to explore the effect of various instructional task designs on the incidence of critical thinking skills. Specifically, given that critical thinking happens in novel situations (Halpern, 1998), the goal was to introduce a complexity factor in the instructional tasks in order to trigger the use of critical thinking skills. In order to understand better the sole impact of task design, important aspects of teaching and learning, such as the provision of individual feedback and the teaching of critical thinking were deliberately left out of the study as a strategy to control for threats to validity. In other words, the underlying assumption was to find how to design tasks that would trigger the use of critical thinking skills, as opposed to focus on teaching methods to foster critical thinking.

The participants in this study carried out five tasks. The first two tasks were essays and their content was used to establish the baseline of critical thinking skills. The first level of complexity was introduced in the third Task, which was presented as a fairly well structured case study. The second level of complexity was introduced in Tasks 4 and 5. These latter tasks took the form of a simulation wherein a situation was provided and the students had to identify the problems and solutions (Leenders and Erskine, 1989). All the tasks were carried out individually. No instruction about the definition of critical thinking, nor feedback, was given to the participants during the exercises.

The introduction of a complexity variable in the case study made a significant difference on the incidence of the critical thinking skills. As opposed to the other instructional methods, the case study triggered less “Judgement and Interpretation” skills, which tend to be relatively common across type of tasks, and more "Imposing Meaning" skills, such as "Making Recommendations". In addition, the case study tended to trigger a specific set of skills. Incidence of the indicators "Evaluates", "Suggests Alternative Approaches", "Summarizes", "Concludes", "Predicts and Hypothesizes" and "Recognizes Various Impacts" all increased. On the other hand, the case study triggered very few incidences of "Multiple Perspective" skills.

The introduction of level 2 of the complexity variable in Tasks 4 and 5 did not have a positive impact on the outcome of critical thinking skills. I have suggested that the amount of information presented in the simulation combined with the lack of clear directives and goals prevented the students from using critical thinking skills. In short, the simulations were too complex to trigger critical thinking, at least under the conditions of the study. In fact, Leenders and Erskine (1989) suggest that such a complex situation requires the intervention of the instructor to catalyse learning.

The lowest level of critical thinking skills occured in Task 5. Two rival hypotheses may partly explain this negative result. First, Task 5 was carried out in different conditions, that is, the student output was graded for content by the instructor and was worth 20% of the final grade, whereas the first four Tasks were given a participation grade and read by the researcher only. For that reason, I suggest that the students took fewer risks and became more conservative in organizing their assignment. The second rival hypothesis concerns the type of outputs produced by the students. As we have seen, several students wrote dialogues. I have speculated that the sentence as a unit of analysis is inapropriate to reveal the true incidence of critical thinking skills.

The findings in this study have implications for both instructional and research practice. From an instructional point of view, the research design could be transposed with some modifications, to the classroom or an on-line environment. Using an experiential approach (Kolb, 1984), a practitioner could, for example, use the first three Tasks as means to provide experience to the learners. The outcome on the third Task could serve as the basis of discussion about critical thinking skills. The learning from this discussion could then be applied in a more complex learning situation, either individually or collaboratively. The outcome of the experience would then be brought back into the classroom and reinvested in the next assignment or exercise. In short, the first three tasks can be carried out independently, but some form of scaffolding is required to foster critical thinking in complex learning situations. Critical thinking in complex situations is not likely to happen without the intervention of the instructor or a scaffolding strategy embedded in instruction.

The unexpected outcomes of Task 4 and Task 5 show that the content and the directives need to be simplified to trigger critical thinking skills. This finding has some implications for the design of online environments. In resource-based learning for example, where only general guidelines are provided, the students have to find their own information to carry out their assignments (MacDonald & Mason, 1998). The assumption behind the approach is that students are self-directed. In fact, this type of open learning environment might be suitable to a limited range of learners such as graduate students (Laurillard, 1993).

On the other hand, the results of this study indicate that the design of content alone is not enough to trigger critical thinking or even learning. More instructional guidance is needed to address the issue of complexity, especially with regards to the use of simulations.

It should be added that while the essays were used to establish the baseline for the incidence of critical thinking skills, this does not negate the importance of this instructional approach as a means to trigger critical thinking skills. On one hand, one of the challenge in this study was to establish a basis of comparison in order to gage the effect of task design on the incidence of critical thinking skills. On the other hand, to my knowledge, there is no study that suggests a normal "target" for the incidence of critical thinking skills; we don't know what should be the expected proportion of critical thinking skills versus non-critical units in written assignments such as essays for instance. Should the proportion be 30%-70%, 40%-60%, 50%-50%? Although an interesting issue, the purpose of the study was not to propose such a target for given task designs, but rather use what can be considered as the most common type of written assignment, that is the essay. However, the results of the study show that the proportion of critical thinking skills in essays was quite high (T1 = 60% and Task 2 = 58%). Given the conditions in which the tasks were carried out, the use of the essay, especially as an individual mode of learning, should not be disregarded as an instructional approach to trigger the use of critical thinking skills.

The implications for research stem from the limitations of the study. The results presented here are applicable to individual learning situations. Important variables such as individualized feedback, instruction about critical thinking and the intervention of the instructors were controlled in order to isolate the instructional methods. Further research is needed to understand the role of those variables on the process of becoming a critical thinker.

Another limitation is the decision to look at the incidence of critical thinking skills only. Restricting the content analysis to counting frequencies obviates evaluation of the quality of the thinking. In short, we observed the incidence of critical thinking as measured by the indicators but we do not know whether the use of the skill is appropriate or effective. From a research point of view, evaluating reliably the quality of all the reported instances of critical thinking is quite a challenge. Perhaps the most substantial difficulty would be the cognitive load imposed on the coders. Coding critical skills is already intellectually demanding, especially when the texts are long and the sample is large. Adding a qualitative component to the coding might distract the coders from the main task. One solution might be to use two coders : one to identify the skills and the other to evaluate the quality. The coders could change tasks during the coding process to ensure inter-coder reliability. Another solution might be to assess the overall quality of the thinking in an assignment, using a Likert scale for example, but that too brings up the issue of inter-rater reliability.

Because of the way the tasks were designed in this study, using written assignments as the focal texts for the content analysis implied neglecting skills such as "Dealing with Uncertainty", "Deciding of a Path of Action" and to a lesser extent "Metacognition". The tasks might have triggered those skills, but the instructional mode is not appropriate for the observation of mental processes. Other techniques such as verbal protocol analysis or classroom observations might reveal the presence of those skills. One could integrate in the instructional strategies the provision of a journal, documenting their learning for example; or specifically ask the students to include a reflective component in their assignment. But the disadvantage is that prompting reflection introduces a confounding variable.

In a similar vein, we have seen that the incidence of "Adopting the Questioner's Role" increased in Tasks 4 and 5 because it was part of the mandate of the student to prepare questions for the chatroom. Obviously, the incidence of "Making Recommendations" also increased in Task 3, because of the case study format. Other aspects of the learning context also have an impact. The fact that Task 5 was graded for content is a good example of the challenges that have to be taken into consideration by the researcher. In short, designing instruction for the purpose of conducting research is never completely free of biases.

The choice of a unit of analysis is also important when conducting a content analysis in the field of critical thinking. The type of focal texts used (student assignments, on-line discussions, videotaped classroom interactions) (Ahuvia, 2001), the type of text (essays, case studies, dialogues, outlines) and to a certain extent the background of the participants (undergraduate and graduate students in science or liberal arts domain) have to be taken into consideration when selecting a unit of analysis. Furthemore, the sentence as a text unit, as explained in the discussion chapter, seems to be appropriate for "Judgement and Interpretation" skills, but not for more complex skills such as "Multiple Perspectives" or "Imposing Meaning" skills. A clustering strategy, that is, looking at the peripheral context of the "incidence" of the skill, might reveal information about the nature and the circumstances of use of the given skill. Such a strategy could inform instructional designers about the type of context in which more complex critical thinking skills are likely to happen and about the components of such complex critical thinking skills.

Finally, another interesting avenue of research on critical thinking would be the exploration of the role of knowledge. Several authors contend that prior knowledge is necessary to trigger the use of critical thinking skills (Halonen, 1995; Halpern, 1998). While addressing the issue of prior knowledge in an experimental study might represent an unsurmountable challenge (Underbakke, Borg, & Peterson, 1993), it would still be possible to observe how students use factual knowledge in their assignments. In this study, I created a knowledge basis by using the same topics across tasks and yet, the topics selected were quite diverse. A follow-up study might look at the students who chose the topic of "Genetically Modified Foods" in Tasks 2 or 4 and see how they did in Task 5. Another study might compare the way knowledge is handled by social students in comparison with students in science.

In conclusion, perhaps the most important consideration for future research on the development of critical thinking skills would be the use of different measures. Although conducting research in educational settings entails striking a balance betwen practical constraints and research requirements, triangulation of measures would address the issues of reliability and validity which are pivotal for the advancement of the field.

REFERENCES

Adelsköld, G., Aleklett, K., Axelsson, R., & Blomgren, J. (1999). Problem-based distance learning of energy issues via computer network. Distance Education, 20(1), 129-143.

Ahuvia, A. (2001). Traditional, interpretive, and reception based content analyses: Improving the ability of content analysis to address the issues of pragmatic and theoretical concerns. Social Indicators Research, 54(2), 139-172.

Allegretti, C. L., & Frederick, J. N. (1995). A model for thinking critically about ethical issues. Teaching of psychology, 22(1), 46-48.

Anderson, T., Howe, C., Soden, R., Halliday, J., & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. Instructional Science, 29, 33-44.

Andrews, J. K. J., McBride, G. T. J., & Sloan, E. D. J. (1993). Humanities for undergraduate engineers: a rich paradox. Journal of Engineering Education, 82(3), 181-184.

Andrews, J. N. (1990). General thinking skills: Are there such things? Journal of philosophy of education, 24(1), 71-79.

Bakeman, R., & Gottman, J. M. (1997). Observing interaction: an introduction to sequential analysis (Second ed.). New York: Cambridge university press.

Baron, J. B. (1987). Evaluating thinking skills in the classroom. In J. B. Baron & R. J. Sternberg (Eds.), Teaching thinking skills: Theory and practice (pp. 221-247). New

York: W.H. Freeman and Company.

Baron, J. B., & Sternberg, R. J. (Eds.). (1987). Teaching thinking skills: Theory and practice. New York: W.H. Freeman and Company.

Bensley, A. D., & Haynes, C. (1995). The acquisition of general purpose strategic knowledge for argumentation. Teaching of psychology, 22(1), 41-45.

Bereiter, C., & Scardamalia, M.Beyond Bloom's taxonomy: Rethinking knowledge for the knowledge age. Developing higher-level approaches to knowledge., [(Web)]. Available: http://csile.oise.utoronto.ca/abstracts [2000, June].

Bernstein, D. A. (1995). A negotiation model for teaching critical thinking. Teaching of psychology, 22(1), 22-24.

Bloom, B. S. (1956). Taxonomy of educational objectives, handbook 1: cognitive domain.

Bodgan, R. C., & Biklen, S. K. (1992). Qualitative research for education: An introduction to its theory and methods. Boston: Allyn and Bacon.

Borg, W. R., & Gall, M. D. (1989). Educational research: an introduction. New York: Longman.

Bransford, J., Sherwood, R., Vye, N., & Rieser, J. (1986). Teaching thinking and problem solving: Research foundations. American Psychologist, 41(10), 1078-1089.

Braten, I. (1993). Cognitive strategies: A multi-componential conception of strategy use and strategy instruction. Scandinavian Journal of Educational Research, 37(3), 217-242.

Brovey, D. J. (1988). Critical thinking, literary criticism, and concept mapping: Tools for teaching STS courses. Bull. Sci.Tech.Soc., 8, 524-527.

Brown, A. L. (1994). The Advancement of learning. Educational Researcher(November), 4-12.

Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. American Psychologist, 52(4), 399-413.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher(January-February), 32-42.

Bullen, M. (1998). Participation and critical thinking in online university distance education. Journal of Distance Education/Revue de l'éducation à distance, 13(2), 1-32.

Calkins, M. S. J., & Armstrong, D. (2000). A CD/Web-based team project for undergraduate business ethics students. Paper presented at the Ed-Média, Montréal, Canada.

Campbell, D. T., & Stanley, J., C. (1966). Experimental and quasi-experimental designs for research. Boston: Houghton Mifflin Company.

Carey, J. W., Morgan, M., & Oxtoby, M. J. Intercoder agreement in analysis of responses to open-ended interview questions: Examples from tuberculosis research. Cultural Anthropology Methods, 8(3), 1-5.

Carlson, E. R. (1995). Evaluating the cerdibility of sources: A missing link in the teaching of critical thinking. Teaching of psychology, 22(1), 39-41.

Case, R., & Bereiter, C. (1984). From Behaviourism to cognitive behaviourism to cognitive development: Steps in the evolution of instructional design. Instructional Science, 13, 141-158.

Christen, R., Angermeyer, J., Davison, M. L., & Anderson, K. (Spring 1994). Performance assessments of critical thinking: Reflective judgement approach, [Web].

Center for Applied Research and Educational Improvement. University of Minnesota. Available: http://www.education.umn.edu/carei [2002.

Coffman, W. E. (1987). Measurement of thinking skills: An historical perspective. Iowa Testing Programs occasional papers Number 34. Paper presented at the Meeting to discuss future directions for the California Assessment Program, Bekerley, CA.

Cognition and technology group at Vanderbilt. (1990). Anchored instruction and its relationship to situated cognition.

Cognition and technology group at Vanderbilt. (1991). Some thoughts about constructivism and instructional design. Educational Technology, September, 16-18.

Collison, G., Elbaum, B., Harind, S., & TInker, R. (2000). Critical-thinking strategies, Facilitating on-line learning: Effective strategies for moderators (pp. 127-164). Madison, WI: Atwood Publishing.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design & analysis issues for field settings. Boston: Houghton Mifflin Company.

Dick, W. (1987). A history of instructional design and its impact on educational psychology. In J. A. Glover & R. R. Ronning (Eds.), Historical foundations of educational psychology (pp. 183-202). New York: Plnum Press.

Doolittle, J. H. (1995). Using riddles and interactive computer games to teach problem-solving skills. Teaching of psychology, 22(1), 33-36.

Ellington, H. I., Addinall, E., & Percival, F. (1981). Games and simulation is science education. London: Kogan Page.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), Teaching thinking skills: Theory and practice (pp. 9-

26). New York: W.H. Freeman and Company.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. Educational Researcher, 18(3), 4-10.

Ennis, R. H. (1990). The extent to which critical thinking is subject-specific: further clarification. Educational Researcher, 19(4), 13-16.

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

Ennis, R. H. (2002). Issues pertaining to the assessment of critical thinking skills.

Eraut, M. (1985). Knowledge creation and knowledge in professional contexts. Studies in Higher Education, 10(2), 117-133.

Facione, P. A.Critical thinking: What is it and why it counts, [Web]. California Academic Press. Available: http://www.calpress.com/critical.html [1999.

Fogarty, R., & McTighe, J. (1993). Educating teachers for higher order thinking: The three-story intellect. Theory into Practice, 32(3), 161-169.

Garrison, D. R. (1991). Critical thinking and adult education: A conceptual model for developing critical thinking in adult learners. International Journal of Lifelong Education, 10(4), 287-303.

Gibbs, L. E. (1985). Teaching critical thinking skills at the university level: A review of some empirical evidence. Informal Logic, VII(2&3), 137-149.

Halonen, J. S. (1995). Demystifying critical thinking. Teaching of psychology, 22(1), 75-81.

Halpern, D. F. (1997). Critical thinking across the curriculum: A brief edition of thought and knowledge. Mahwah, NJ: Lawrence Erlbaum Associates.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. Dispositions, skills, structure training, and metacogniive monitoring. American Psychologist, 53(4), 449-455.

Hanley, G. L. (1995). Teaching critical thinking: focusing on metacognitive skills and problem solving. Teaching of psychology, 22(1), 68-71.

Harwell, M. (1999). Evaluating the validity of educational rating data. Educational and Psychological Measurement, 59(1), 26-37.

Hawe, L. Q. (1991). Added value using application exercises in engineering degree courses. Journal of Further and higher Education, 15(2), 54-64.

Henderson, B. B. (1995). Critical-thinking exercises for the history of psychology course. Teaching of pyschology, 22(1), 60-63.

Henri, F. (1989). La téléconférence assistée par ordinateur dans une activité de formation à distance. Unpublished doctoral thesis, Concordia University, Montréal.

Henri, F. (1992). Formation à distance et téléconférence assistée par ordinateur: Interactivité, quasi-interactivité, ou monologue? Journal of Distance Education/Revue de l'éducation à distance, VII(1), 5-24.

Herrington, J., & Oliver, R. (1999). Using situated learning and multimedia to investigate higher-order thinking. Journal of Interactive Learning Research, 10(1), 3-24.

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. Educational and Psychological Measurement, 60(4), 523-531.

Jakoubek, J. (1995). Developing critical thinking skills in psychology content course. Teaching of Psychology, 22(1), 57-59.

Jones, E. A., & Ratcliff, G. (1993). Critical thinking for college students. University Park: National Center on Postsecondary Teaching, Learning & Assessment. The Pennsylvania State University.

Jones, K. (1985). Designing your own simulations. London: Methuen.

Jones, P. C., Merritt, J. Q., & Palmer, C. (1999). Critical thinking and interdisciplinary in environmental higher education: the case of epistemological and values awareness. Journal of Geography in Higher Education, 23(3), 349-357.

Keeley, S. M., & Haas, P. F. (Coping with faculty resistance to teach critical thinking). 1997. College Teaching, 46(2), 63-67.

Keppel, G. (1982). Design and analysis: A researcher's handbook (Second ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kindfield, A. C. H., & Hickey, D. T. (1999). Tools for scaffolding inquiry in the domain of introductory genetics. Paper presented at the Annual Meeting of the American Educational Research Association, Montréal.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. American Educational Research Journal, 31(2), 338-368.

King, A. (1995). Inquiring minds really do want to know: Using questioning to teach critical thinking. Teaching of psychology, 22(1), 13-17.

Kolb, D. A. (1984). Experiential Learning: Experience as the source of Learning and development. Englewood Cliffs, NJ: Prentice-Hall.

Krippendorf, K. (1980). Content analysis: An introduction to its methodology. Beverly Hills: Sage.

Kuhn, D. (1999). A developmental model of critical thinking. Educational Researcher, 28(2), 16-26, 46.

Landry, R. (1993). L'analyse de contenu. In B. Gauthier (Ed.), Recherche sociale (Deuxième ed., pp. 337-359). Montréal: Presses de l'Université du Québec.

Laurillard, D. (1992). Learning through collaborative computer simulations. British Journal of Educational Technology, 23(3), 164-171.

Laurillard, D. (1993). Rethinking university teaching framework for the effective use of educational technology. London: Routledge.

Layton, D. (Ed.). (1990). Innovations in science and technology education (Vol. III). Paris: Unesco.

Leenders, M. R., & Erskine, J., A. (1989). Case research: The case writing process (third ed.). London, Ontario: School of Business Administration, University of Western Ontario.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. Theory into Practice, 32(2), 131-137.

Macdonald, J., & Mason, R. (1998). Information handling skills and resource-based learning in an Open University course. Open Learning(February), 38-42.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. Harvard Educational Review, 62(3), 279-300.

McBurney, D. H. (1995). The problem method of teaching research methods. Teaching of psychology, 22(1).

McLoughlin, C. (1997). Investigating conditions for higher-order thinking in telematics environments. Unpublished Doctoral dissertation, Edith Cowan, Perth.

McLoughlin, C., Baird, J. R., Pigdon, K., & Woolley, M. (2000). Fostering teacher inquiry and reflective learning processes through technology enhanced scaffolding in a multimedia environment. Paper presented at the Edmedia, Montreal, Canada.

McLoughlin, C., & Oliver, R. (1998). Planning a telelearning environment to foster higher order thinking. Distance Education, 19(2), 242-264.

McMillan, J. H. (1987). Enhancing college students' critical thinking: A review of studies. Research in Higher Education, 26(1), 3-27.

McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. Educational Researcher, 19(4), 10-12.

Merrill, M. D. (1991). Constructivism and instructional design. Educational Technology, May, 45-52.

Neueundorf, K. A. (2002). The content analysis guidebook. Thousand Oaks, CA: Sage publications.

Newman, D. R., Webb, B., & Clive, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. Interpersonal computing and technology: An Electronic journal for the 21st century, 3(2), 56-77.

Norris, S. P. (1985). Synthesis of research on critical thinking. Educational Leadership, May, 40-45.

Norris, S. P. (1989). Can we test validly for critical thinking? Educational Researcher, 18(9), 21-26.

Oliver, R. (1999). Exploring strategies for online teaching and learning. Distance Education, 20(2), 240-254.

Osana, H. P., Bennett, T., & Tucker, B. J. (in press). Decision making in problems of equity: Implications for reform in social studies. Contemporary Educational Psychology.

Ostheimer, M. W., Mylrea, K. C., & Lonsdale, E. M. (1994). An integrated course in fundamental engineering and english composition using interactive and process learning methodologies. IEEE Transactions on education, 37(2), 189-193.

Paul, R. (1987). Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions. In J. B. Baron & R. J. Sternberg (Eds.), Teaching thinking skills: Theory and Practice (pp. 127-148). New York: W. H. Freeman and Company.

Perkins, D. N., & Solomon, G. (1989). Are cognitive skills context-bound? Educational Researcher(January-February), 16-25.

Portelli, J. P. (1994). The challenge of teaching critical thinking. McGill Journal of Education, 29(2), 137-151.

Quellmalz, E. S. (1987). Developing reasoning skills. In J. B. Baron & R. J. Sternberg (Eds.), Teaching thinking skills: Theory and practice (pp. 86-105). New York: W.H. Freeman and Company.

Resnick, L. B. (1987). The 1987 presidential address: Learning In school and out. Educational Researcher, December, 13-20.

Simich-Dudgeon, C. (1998). Developing a college web-based course: lessons learned. Distance Education, 19(2), 337-357.

Singer, J., Marx, R. W., Krajcik, J., & Chambers, J. C. (2000). Constructing extended inquiry projects: Curriculum materials for science education reform.

Educational Psychologist, 35(3), 165-178.

Smith, D. G. (1977). College classroom interactions and critical thinking. Journal of Educational Psychology, 69(2), 180-190.

Thiagarajan, S., & Stolovitch, H. D. (1978). Instructional Simulation Games (Vol. 12). Englewood Cliffs, NJ: Educational Technology Publications.

Tishman, S., & Andrade, A.Thinking dispositions: A review of current theories, practices, and issues, [Web]. Available: http://learnweb.harvard.edu/alps/thinking/docs/Dispositions.htm [2001.

Tishman, S., Jay, E., & Perkins, D. N. (1993). Teaching thinking dispositions: From transmission to enculturation. Theory into Practice, 32(Summer), 147-153.

Tsui, L. (1999). Courses and instruction affecting critical thinking. Research in Higher Education, 40(2), 185-200.

Underbakke, M., Borg, J. M., & Perterson, D. (1993). Researching and developing the knowledge base for teaching higher order thinking. Theory into Practice, 32(3), 138-146.

Underwood, M. K., & Wald, R. L. (1995). Conference-style learning: A method for fostering critical thinking with heart. Teaching of psychology, 22(1), 17-21.

Van der Maren, J.-M. (1993). Méthodes de recherche pour l'éducation. Montréal, Qc: Université de Montréal.

Voss, J., F., & Wiley, J. (1997). Developing understanding while writing essays. International Journal of Educational Research, 27, 255-265.

Wade, C. (1995). Using writing to develop and assess critical thinking skills. Teaching of psychology, 22(1), 24-28.

Winn, W. (1991). The assumptions of constructivism for instructional design. Educational Technology, September, 38-40.

Winnips, K., Collis, B., & Moonen, J. (2000). Implementing a "scaffolding by design" model in a WWW-based course considering cost and benefits. Paper presented at the Ed-Media, Montréal, Canada.

Wolfe, C. R. (1995). Homespun hypertext: Student constructed hypertext as a tool for teaching critical thinking. Teaching of psychology, 22(1), 29-33.

Young, A. C. (1997). Higher-order learning and thinking: What is it and how is it taught? Educational Technology, July-August, 38-41.

Young, M. F. (19?). Instructional design for situated learning. ETR&D, 41(1), 43-58.

Zech, L., Gause-Vega, C. L., Bray, M. H., Secules, T., & Goldman, S. R. (2000). Content-based collaborative inquiry: A professional development model for sustaining educational reform. Educational Psychologist, 35(3), 207-217.

APPENDICES

APPENDIX I OVERVIEW OF STRATEGIES TO TEACH CRITICAL THINKING

Author	Type of CT skills	Teaching Strategies	Domain
Allegretti C. L. & Frederick, J.N.(1995)	Argumentation Assessment : Validity of sources	Discussion	Ethical Issues Psychology
Anderson et al. (2001)	Justification of arguments	Classroom discussion Written assignments	Vocational education
Bernstein, D.A (1995)		Negotiation Role plays and simulations	Critical thinking Psychology
Brovey, D.J. (1988)	Assessment of information	Concept mapping	STS course
Bullen, M. (1998)	Argumentation	Discussion Collaboration	On-line environemnt (Graduate students)
Calkins & Armstrong (2000)	Assessment	Case-based	Business Ethics
Carlson, E.R. (1995)	Evaluation: Credibility of sources	NS	Philosophy Psychology
Collison et al. (2000)	Sharpening strategies to focus Developing In depth dialogue	On-line discussion Facilitation Dialogue	On-line environments
Doolittle, J.H. (1995)	Problem solving	Computer games	(College students)
Grigg, L. (2001)	Explores Models of CT Reflection and discussion	N.S.	Course on C.T.
Hanley, G.L. (1995)	Metacognition	Self-reflection-evaluation Problem solving	Course on C.T.(undergraduate)
Henderson, Bruce B. (1995)	Analysis Comparisons	Exercises of reading and questioning	History of Psychology
Herrington, J. & Oliver, R. (1999)	Uncertainty Deciding of a path of action Judgement and Interpretation Multiple Perspectives Imposing Meaning Metacognition	Online discussions Collaborative Situated Learning	Educational Technology (Graduate students)
Jazoubek J.(1995)	Argumentation Assessment : Validity of sources	N.S.	Psychology (Sophomore)
Jones, P. C., Merrit, J.Q., & Palmer (1999)	Epistemological and values awareness (basis for CT)	N.S.	Environmental Education
King, A. (1995)	Inquiry Questioning	Reciprocal peer questioning Classroom discussion	College students

Author	Type of CT skills	Teaching Strategies	Domain
McBurney, D.H. (1995)	Ill-defined Problem Solving	Case study/ problem method	Research Methods
McLoughlin, C.	General Model	Scaffolding	Teacher education
McLoughlin, C., Baird, J., Pigdon, K., & Woolley, M. (2000)	Inquiry Reflective Learning	Scaffolding Multimedia environment	Teacher education
Osana, H.P., Bennett, T., & Tucker, B.J. (2001)	Decision-Making	NS (research)	Social Studies; adolescents
Tishman, S., Jay, E., & Perkins, D.N. (1993)	C.T. Dispositions	Enculturation	General
Underbakke, M., Borg, J.M., & Peterson, D. (1993)	Hypothesis Testing Assessing arguments Solving interpersonal problems Probalisitic thinking	Model	Teacher Training
Wade, C. (1995)	Asking questions Examination of evidence Problem definition Analysis of assumptions	Written assignments	General psychology Human sexuality courses
Winnips, K., Collis, B., & Moonen, J. (2000)	Argumentation	Scaffolding Web-Based	Educational Technology (Graduate Students)
Wolfe C.R. (1995)	Informal reasoning Ill-structured problem solving	Student-constructed hypertexts to create conceptual links	Course on cognition and computers

NOTE: Author specified « Critical thinking » in title or abstract of the article.

APPENDIX II: CONSENT FORM

WebCT tasks and critical thinking skills
Geneviève Légaré is exploring how instructional strategies can contribute to the development of critical thinking skills while using WebCT. From a theoretical standpoint, this study will contribute to the understanding of key characteristics of the development of critical thinking skills. From a practical point of view, this study will help individuals use higher cognitive skills effectively. In addition, practitioners will gain an understanding as to how to design instruction to enhance higher cognitive skills.

Your instructor is participating in this study. I am asking you to voluntary take part. No extra work will be requested of you other than regular course assignments. If you agree to participate, rest assured that all information collected for the purposes of the research project will be kept confidential.

Your participation is voluntary; you are under no obligation to take part. Also, you may choose to discontinue your participation at any time. If you choose to discontinue your participation, no data you have provided will be used in this study. Furthermore, your decision will not affect your participation grade in the WebCT portion of this course.

If you agree to participate, please read and sign this consent form. The study will be conducted over a period of five weeks. I will be using three sources of information for the research. First, I will look at the four WebCT tasks. Second, I will look at the short paper which is basically the same as task 4, but with a different topic. Finally, after the short paper has been handed in, 5 students will be selected at random to participate in an interview with me for approximately 30 minutes. Again, you participation is entirely voluntary. You may choose to opt out of the study without impact on your grades.

Once the information is compiled, a brief summary of results will be made available to you and a more complete explanation of the study will be provided if you so desire. Any questions or concern you have with respect to this research should be addressed to Geneviève Légaré via e-mail at legare@alcor.concordia.ca.

Thanking you for your time,
Geneviève Légaré

<u>Student consent to participate in research</u>

This is to state that I agree to participate in data collection for research conducted by Geneviève Légaré for her doctoral dissertation in Educational Technology. I have read the above description and understand the agreement. I freely consent and agree to participate in the collection of data for this research project.

____I agree to participate. ___I do not agree to participate.
Name (please print)______________________ Student ID:_____________
I would like a copy of the study findings when they are available. ___Yes ___No

Signature___________________________ Date_____________

APPENDIX III SOCIAL ASPECTS OF ENGINEERING:

SHORT SURVEY

Last name:__________________ First name:_________________________

Home tel: (___) _____________ E-mail: _________________

Concentration: Civil___ Computer___ Mechanical___ Electrical___
Building:___ Industrial___

Current GPA :_______

Status: Canadian student ___ Foreign student:___

Please answer all the following questions. All information provided will remain confidential.

1. What is your mother tongue? English ___French: ___Other (specify):____

2. You did <u>most</u> of you schooling (up to grade 12) in: English ___French: ___Other (specify):_____

3. How would you consider your English <u>writing</u> skills (tick one only):
a) Excellent: ___ b) Very Good: ___ c) Good: ___
d) Average: ___ e) Weak: ___ f) Very weak:___

4. When was the last time you wrote and essay?_______________

5a) Do you have a part-time job? Yes__No__
5b) If you work part-time, is your job related to the engineering field? Yes__ No__

6. Do you have a computer with modem at home? Yes___No___

7. Do you surf the Web? a) Never__ b) Once a week___ c) 2-3 times per week____ d) Daily____

8. Do you know how to use:
a) Chat utilities: Yes: ___ No:___ b) E-mail: Yes: __ No: ___
c) Bulletin boards and newsgroups: Yes: ___ No:___ d) Groupware: Yes: __ No: ___

9. How often do you use e-mail? Never__ Once a week___ 2-3 times per week____ Daily____

10. What do you expect to get out of this course? (Please use the reverse if not enough space).

11. How do you think this course will help you or contribute to your career as a professional engineer? (Please use the reverse if not enough space).

APPENDIX IV TASK INSTRUCTIONS

ENGR 492 WebCT Tasks

Instructions to students

MENU

General presentation

The WebCT Tasks are designed to develop your thinking skills. In fact, the WebCT exercises are occasions to practice applying your skills before you dive into the other course assignments (short paper and major report).

Why thinking skills

The completion of the WebCT Tasks will be beneficial both in the long term and the short term. As a future engineer, you will be required to analyze facts in order to make a decision, you will need to conduct a synthesis of the information and gather data in order to make a recommendation, or evaluate expert opinions in order to take a position on a given issue.

The short term perspective has to do with this course. The ENGR492 instructors and I designed the Tasks so they can be applied directly to your assignments. In other words, part of the grade for the short paper and the major report will be allocated to the skills that you have practised on WebCT.

What to expect

- The Tasks are in fact "arguments" that you have to develop or problems that may call for one or many solutions in order to be solved.

- You will not be told exactly what to do or what is "good" to write. Also, it is up to you to decide how much time should be spent on the Task and how much you should write about the issue. Some Tasks might puzzle you, you might have to sort out information or have to make decisions about what to include to support your argument.
- In fact, you are your own judge: You decide which angle you're going to take; you decide when the quantity of text is sufficient and finally, you decide when your work corresponds to your own criteria and is ready to send in. It is up to you to decide when your response is good enough to be sent.
- The Task will be posted every Wednesday at midnight. We believe that a week should leave you plenty of time to look at the Task, estimate the amount of time required to carry it out and then return it by midnight on the following Wednesday.

Assessment of Tasks

- The four WebCT Tasks are worth 15% of the final grade for the course. Each individual task is worth 3.75%.
- Assessment of the the Tasks consists of two elements: a) doing the task on time and according to the directives and; b) effort put in the Task.
    - If you fail to submit your Task on time you will loose a portion of the 15% (1.88%). No argumentation allowed. No excuses accepted
    - You may loose points if directives are not followed properly (see instructions for writing document and sending mails).
    - Effort put in the Task will be assessed. I will give a "doing fine" or "you're in trouble" comment (if you prefer, a "pass-fail" grade) for each individual Task.
- I will read every input and assess the effort put into them. Feedback will be provided during a debriefing session, after Task 4 has been completed but before the short paper is due.

Late comers:

Depending on individual registration date, hand in the two Tasks on the same day.

How to access WebCT

- To get the complete Web site address, please refer to the IITS WebCT information sheet distributed in class.
- Computer access: You may use a computer station in your respective lab.
- For those who need help to log on for the first time on WebCT, a drop in session will be held on Friday January 12 from 9:00 to 12:00 and from 13:00 to 16:00 in room ER-103-3, 2155 Guy St. (In Royal Bank building).

What if there are technical problems

- Server failure: University servers might go down. This possibility may cause trouble if it happens on a Wednesday. In that case, Tasks can be e-mailed via WebCT on Thursday and the next Task will be posted on that Thursday as well.

- Should any other unanticipated problem occur, we will deal with them as they come along and as efficiently as possible.
- If you have technical problems or any other trouble that might prevent you from doing your Task on time, please do not wait at the last minute to tell me. Send me an e-mail (legare@alcor.concordia.ca), outlining what is the issue (be clear and precise) and we'll take it from there.
- For any other technical problems (login, access to WebCT etc.) contact Technical Support: **848-7613 Monday to Friday, 9:00 to 5:00.**

How to write up the Tasks

Note: This advice is also good for the short term paper and the major report.

- Be as clear as possible: write short and simple sentences.
- Always have the reader in mind: readers should not have to guess what you mean.
- Your opinions should not just be stated: they should be supported or documented.
- Reminder: There are always two sides to a medal: What's the reverse argument.
- Structure of ideas: create links between paragraphs
- One idea per paragraph
- Try to have an introduction and a conclusion. Even when writing a short piece, have an opening and a closing statement.
- Run the spell checker before submitting the Task.
- Run the spell checker after running the spell checker.

How to format your document

- Always keep a copy of your work until the course is over.
- Use a style that is common and easy to read such as Arial, New York Times, or Comic.
- **Your name, student ID number and course section should appear on the top left corner of your document.**
- **Put the number of the Task in the top left corner of the document.**
- If your document is longer than one (1) page, please number the pages at the bottom right corner.
- Word processors: Word 7.0 is preferred. Always save your file as "Word 7.0" document. Your extension should always be ".doc". You can use other word processors such as WordPerfect or ClarisWorks. If you do so, save your document in "Rich Text Format". Do not forget to check the extension which should be ".rtf". If you are working on a Mac, you have to put in the extension manually.
- Compressed files or documents are also accepted.

Important note: You may loose up to 1.88% if your document does not comply with these instructions. In some cases, I may return the document and request a resubmission.

How to send the Task by e-mail via WebCT

- The Task should be an attached file. Do not send your task as a message in the composer box.
- Save your document with the file name being composed of your last name, the first initial of your given name, and the number of the Task (here Task 1). All extensions should be ".doc" for Word 7.0 documents or ".rtf" for other documents.

  For example Mark Jones:

  JonesM1.doc

  Or

  JonesM1.rtf
- Then follow these steps:
  "X" stands for "click":
  1. From Main page, X on Mail icon.
  2. X "Compose"
  3. Send to: X "Browse"
  4. Scroll down and select "Geneviève Légaré"
  5. Subject of the e-mail: Specify the subject as follows:

     Task #1: Your last name and your first name

  6. In your e-mail message, write the following message:

     {Your last name} + {your first name}+{student ID number } + {Course section ENGR 492 section D or BB} sending you Task 1.

  7. To attach a file, X "Browse" in the lower right corner. Select your file.
  8. X "Attach". The name of your file should appear in blue. X the box to attach the file. If you wish, you may verify the document by double clicking on the file name. In that case X "close" to go back in the mail composer.
  9. In the mail composer, you can X "Preview" to make sure that your message is okay.
  10. X "Send".
- After sending the task, do not forget to indicate the time it took you to carry out the assignment (back to homepage, time log icon).

IMPORTANT NOTE: It is your responsibility to apply the required style. I may return incomplete messages or documents. With the exception of Task 1, I will not give a full grade for Tasks that do not respect the directives.

Reception Notice

I will send a reception notice after receiving each Task. There are three types of messages:

- "Your message has been received successfully".
- "Your message has been received successfully, but some directives were not followed properly. Please, next time refer to the "Task Instructions document" for adequate formatting."
    - Note: With the exception of Task 1, you will loose a portion of the grade if instructions are not applied properly.
- "Your message has not been received successfully. The problem is so and so. Please resubmit your document."

If you have not received my notice by Saturday, I may not have received your Task or I may have accidentally forgotten to send you a notice. It is your responsibility to ensure that I have all your Tasks in hand.

Other things you may want to do

- Scan all documents with an anti-virus including the ones you receive.
- Check the file size allowed by your service provider (some servers are limited to 1.5 meg).
- If your document is larger than 1.5 meg it would be a good idea to compress it.
- If you include pictures or graphs in your document, you should definitely compress it.
- I you are sending a message other than the Task, be as clear and concise as possible: Present yourself, explain your problem and outline the history of our correspondance, if any.
- Check new postings on WebCT (messages and documents) on a regular basis. I may decide to send a message of general interest (e.g.: Delay in sending reception notice due to server failure).

Allowed/ not allowed:

<u>You are allowed to</u>:

- Quote sources and make explicit references.
- Use other sources (books, magazines, the Web etc.). The direct references should be integrated in your text.
- Refer to information in previous Tasks
- Refer to what you said in previous Tasks
- Write as much as you need.
- Do a good job!
- Work off-line and then post the text.
- Write drafts or edit your text before sending.
- Send in your Task early.
- Use other features in WebCT.

<u>You are not allowed to:</u>

- Have someone else write the text for you. This is ghost writing and the university has strict rules for that.

- University also has strict rules about plagiarism (see calendar).
- Collaborate. The Tasks are individual.
- Recycle and/or reuse Tasks and short papers from previous semesters (either your work or someone else's).

APPENDIX V ESSAY QUESTIONS : NON-TREATMENT GROUP

Discuss one of the following quotations from chapters 1-6 and 9 of McGinn's Science, Technology, and Society.

1. « Much social conflict, not reducible to economic or racial differences, has been occasioned by developments in technology and science ». Chap.1, pg.5.

2. « ...students attracted to or repelled by the study of mathematics, natural sciences, and engineering tend to have distinctive personality traits. Conversely, involvement in these specialties may foster or reinforce development of certain personality traits and behavioural patterns, such as introversion, diligence, and a problem solving mentality. » Chap.4, pg.60.

3. « ...accounts of technological development and of technological change as a « cause » of social change must take on board the fact that the activities of sciene and technology exist and unfold in rich social contexts. These contexts contribute powerfully to the development of these twin forces and interact with them and their products in an ongoing process of ociotechnical transformation ». Chap.5, pg.101.

4. « The unprecedetend opportunities for individual development afforded by the Industrial Revolution have in many cases been turned to good account. To a large extent however, this cultural potential has been dissipated as the institution of leisure

has been co-opted and made to serve the imperative and interests of large scale techno-economic systems, capitalist and socialist alike. » Chap.6, pg.116.

5. « Besides challenges to its intellectual tenability, contemporary science and technology are calling into question the utility of traditional absolutist ethical theories—that is, their ability to serve as intelligent guides to action in a wolrd of rapid and profound technical and social change. » chap.9, pg.163.

Another topic discussed by McGinn not mentioned above.

APPENDIX VI CONTENT OF THE DEBRIEFING SESSION

Debriefing: Tasks

Engr492 Impacts of technology on society

Announcements

- Deadline: Sunday Feb. 18, midnight
- Send to GL WebCT
- Time spent + short survey
- Further comments by WebCT mail
- Weekend newspapers

Introduction

- Your feedback
- Rationale: Tasks
- Overview of what happened
- Transfer to Short Paper

Your feedback

- WebCT environment
- Tasks
- Type
- Enjoyment, interest
- Topic: Food production
- Grading scheme and motivation

Instructional Design Principles

- Develop methological skills (trial-&-error)
- Provide freedom
- choice of issues
- approach
- Challenge: Balance direction and responsibility
- Encourage criticism about own work
- Create « uncertainty »

Tasks 1& 2: Rationale

- Task 1:
  - Warm up: Topic and activity
  - « Academic »
- Task 2:
  - « Academic »
  - Warm up about the topic
  - Why food production?

Task 3: Rationale

- Case study: Self-contained
- Main goal: Identify reasonable solutions
- « Voice » expert/consultant : Preparing for T4
- Use of data: How would you use data?
- Country?

Rationale Task 4: Simulation

- Form: TV Show?
- Audience:
- Pedag.: Able to explain in layperson 's terms
- Real life: Always consider an audience (letters, reports, bids, proposals)
- Client: Network bio (mandate--framework)
- Creativity and motivation
- Practice for the SP

Rationale Task 4: Food file

- Amount of Data:
- Sort out and select data
- Challenge the data
- Question sources :
  - Clues : Issues and concepts
  - What 's there? What 's not there?
  - Do you need to do extra research?

Rationale Task 4: Part 2

- Part 2: Rationale
- Become objective: Assess your arguments
- How well do you know/understand the topic:
  - Subjects left out
  - Questions for chat
  - Understanding of the audience:
    - Sequencing
    - Strategy

Task 2: Main Topics

Task 2: Strategies

Task 3: Solutions

Task 3: Strategies

Task 4: Main Topics

Task 4: Strategies

Transfer: Short Paper

- Your task: Identify problems and solutions
- Keep the voice (audience)
- Part 2: Rationale
- Use facts and data: 2 options
- Use what's there
- Develop a novel perspective
- Referencing

Goal: **Problems** and solutions

- Problems: Identify and analyze
- Select your problems
- Beyond description; demonstrate your understanding
  - What are the basic concepts?
  - What are the issues?
  - Who are the stakeholders?
  - Are you sure about that? (question your sources)

Goal: Problems and **solutions**

- In light of the problems; If /then...
- What are the potential solutions?
- What is the relationship between solutions and problems? (Strength, clarity…)
- How feasable are your solutions? (Eg. Govrnt Task 3)
- What are the impacts of your solutions?
- Use caution

Last word

- Short survey: WebCT about tasks
- Suggestions for next Fall: WebCT mail to GL
- Topics
- Grading scheme
- Types of Tasks

Results:

- Fall (by e-mail)

APPENDIX VII CHANGES MADE FROM THE PILOT PROJECT

Item	Remarks	Suggestions for Improvement
Survey questionnaires	• Missing items on the questionnaire	• Added items such as student status (foreign students) and GPA.
Log-on WebCT	• Up until the third week some students had problems logging on	• Added a "drop-in" session during the first two weeks with author
Choice to send tasks via WebCT or e-mail	• Some students did not use WebCT at all. Received the instructions from a friend	• Used WebCT only • Used e-mail for problems with accessing
Posting & receiving tasks	• Originally, posting was done in the morning and receiving in the same evening…students saw the content of the following task	• Posted and received the tasks by midnight
Processing the tasks	• The reception of attached files, sorting, opening the files, returning confirmation message and entering the grade was very time consuming	• Used only WebCT to receive tasks • Created template for the OK ones
Tracking of students	• Tracking exceptions was often a problem since the students did not provide adequate or complete information about their situation; Some students produced their work late, got sick, etc.	• Created a better filing system and process cases right away • Added instructions in the WebCT document regarding etiquette and information needed (facilitator and/or technical support staff needed to understand the problem at hand.

Item	Remarks	Suggestions for Improvement
Tasks schedule and short paper	• The tasks represent quite some work for the students but the majority of them completed their tasks on time. However, as one student said, they put as much time as was available • The short paper however comes in conflict with mid-term exams; for the pilot, and extension was given to both sections	• See grading scheme item to improve effort • One instructor also suggested to drop one task. Not possible (research design) • Keep an opening for late papers (announce an official deadline but give an extension to all).
Tasks design	• Task instructions document: additions • T1, T2 and SP OK • T3: Two documents (Improvement of T2 and stats on topic in separate document) • T4: Simulation too long for the timing in term; students fail to endorse the role. Part 2: Rationale for episode • Great but reformulate questions • Call them "Tasks"? Last 2 are case studies/simulations; "Simulation" appears on the course outline	• Minor changes in instructions • First 2 tasks: Minor changes • T3: Assumption for separate document was to avoid biasing the students, let them search on their own and select relevant information. None of that worked out. Not enough directives or structure. Use data for a problem to solve (multiple possible solutions). Add part 2: rationale? Could even give the role right from the beginning but cognitive overload? • T4: Transfer some parts of the simulation into task 3 (student is a consultant—make a recommendation to a board). Keep the writing of the episode but reinforce the "voice"
Task format	• Longer tasks are difficult to read on screen	• Use web page design tool for tasks; easier to read on screen (create menus and anchors) but keep "light" otherwise too long to download.
Task topics	• T1: General • T2-4: Same Topic (motorcycle); • SP: Same structure; different topic • T2-4: Motorcycle: Problem of interest but data OK • SP: Printing: perfect	• T1: Keep same topic • T2-4: Need topic with data from different "domains"; suitable for problem solving; positions are not clear cut; sustain interest • SP: Need topic than can be researched easily and somehow relates to the previous tasks.
Tasks grading scheme	• 15% allocated for doing the task on time and according to rules; represents 3,75%/Task. Students complained (no feedback and relevance of norms...) • Posting of grades delayed	• Kept same % and same norms but allocate points for "perceived effort". Pass or fail.

Item	Remarks	Suggestions for Improvement
Short papers: grading	• Grading: 2 markers: OK • Problem with description vs analysis • Time available for topic: would students spend more time if made possible? • Follow-up on resubmits • Plagiarism (cut and paste) • 20% or 25%?	• Worked well (use same scale!) • Descriptive papers: a)Rule of thumb: 1:3 or 1:4 (1 for description) b) Post 2 positive examples from Fall term; one annotated. • Time: Need to discuss this issue • Resubmits: Who does what? • Provide better instruction about "cut&paste =plagiarism) • Increase to 25%?
Student evaluation forms	• Evaluation criteria related to CT skills • "Weight" of each item	• Both major report and short paper criteria need reinforcement (cohesion) with CT skills- To be reviewed • Weight needs review
Debriefing session	• Lasts 45 minutes • Note taking	• Make more concise and spend less time on separate tasks (no time to look at task 4) • Frequencies (topics and approaches used) are a good strategy; use to facilitate transfer toward Short Paper • Keep notes of students comments
Feedback on the tasks	• No feedback received except during the debriefing sessions	• Created a post-task questionnaire available via WebCT
Time spent on task	• No time keeping.	• Created a log book so student could indicate how much time they spent on their tasks.

APPENDIX VIII POST-TASK QUESTIONNAIRE

(Questionnaire posted on WebCT).

Please, answer all eight questions. Your answers will have no bearing on your grades. This information will be used to improve the design of instructional activities.

A) Indicate in minutes how much time you spent on Task 5. Include activities such as reading and searching, brainstorming, writing and editing.

B) How interesting did you find the topic of « food production »?
1. very interesting
2. interesting
3. not interesting
4. not interesting at all

C) How interesting did you find the topic of « genetically modified foods »?
1. very interesting
2. interesting
3. not interesting
4. not interesting at all

D) In terms of course content, which Task was most engaging? You may select more than one.
1. Task 1
2. Task 2
3. Task 3
4. Task 4
5. Short Paper
6. All Tasks
7. None

E) Which Task format did you least enjoy?
1. Essay (Tasks 1 & 2)
2. Case study (Task 3)
3. Simulation (Task 4 & SP)
4. Enjoyed all formats

F) In which Task did you put the most effort?
1. Task 1
2. Task 2
3. Task 3
4. Task 4

G) How would you rate the quality of your Short Paper? Note: Your answer has no positive or negative impact your grade.
1. Outstanding
2. Excellent
3. Very Good
4. Good
5. Poor
6. Fail
7. I don't know

APPENDIX IX CODER'S CHECKLIST AND LOGBOOK

Checklist instructions

The checklist has been created for two reasons:
First, it should facilitate the management of your work. With the checklist, you will be able to:

- track the tasks that have been coded
- decide which task you want to do next
- evaluate how many tasks you can code on a good/not so good day, which in turn should help in planning your schedule
- indicate a reminder about a task (need to be reviewed, file is damaged etc.)

Secondly, the checklist will be used to "collect data". I may use the dates and the order of the coding in order to explain some of the results.

Description

Numbers: The number in a cell indicates a task to be coded. According to this example, Rater A will be coding Task 3 for subject 203, 205, 207 and Task 4 for subject 202 and 204.

Shaded cells: The shaded cells indicate that the task has been selected for the reliability test (Round 1). In this example, all three raters will code Task 3 of subject 204 and Task 4 of subject 205.

The reason why some shaded cells don't have a "subject" number is because that specific Task (here: 204_T3) was originally assigned to another rater for usual coding.

Note: In Nud*ist the file name will appear as "204b_T3.txt" or "204c_t3.txt".

Rater A

[illegible]	[illegible]	[illegible]
		202
203		
		204
205		205
207		

"Housekeeping" procedures (dating, length of sessions etc.)

In the Excel document, you will find a sheet called "Journal". The purpose of the Journal is to keep track of the information about one working session, as well as to keep track of the tasks coded in one session. A working session can be of any length and you can have more than one session on the same day. In other words, when you take a break, the session is over.

For each session:

- Indicate the date (day and month)
- The length of the session: in decimals, rounded up to 15 minutes. E.g.: You worked 1h45. Enter 1.75.
- Indicate the total number of tasks for that session.
- Indicate the filename of all the tasks you coded during that session. Refer to the filename in Nud*ist or to the hard copy.
- You may indicate comments in the last column.

What if...

There are so many little steps to consider when organizing the logistics that I may have done some mistakes. In any case, you should contact me by e-mail and briefly explain the problem. Here are some of the things that could happen:

- A number is on your list but you cannot find the file in Nud*ist, but you have it on paper: I probably forgot to import it into Nud*ist.
- A number is on your list but you cannot find the file in Nud*ist nor on paper: It is quite possible that the student did not do that specific assignment.
- You find a task in Nud*ist or on paper that is not assigned to you. I made a sorting mistake.
- A file is corrupted and you cannot read it in Nud*ist.

For any of the above reasons, send me an e-mail to outline the problem. Indicate the full file name.

Final remarks

- You may use either the paper version of the checklist or work directly in Excel, as long as you always use the same format.
- The checklists will be collected at the end of the project.

JOURNAL

RATER ID:

(Eg) 1	*07-07-2001*	*0.45*	*1*	*204a_t4*	*First Dialogue I coded*

APPENDIX X RELIABILITY CHECKER'S LOGBOOK

Reliability Checker's Logbook

First round

Date:

[redacted]					
Eg.	*2nd*	*112c_t3*	*80*	*26-36*	
1	Random				
2	2nd				
3	Random				
4	1st				
5	3rd				
6	2nd				
7	Random				
8	1st				
9	3rd				
10	Random				
11	Random				
12	1st				
13	3rd				
14	Random				
15	2nd				
16	1st				
17	3rd				

2nd round

Reliability Checker's Logbook

Date:

18	Random				
19	2nd				
20	Random				
21	3rd				
22	Random				
23	2nd				
24	Random				
25	1st				
26	Random				
27	2nd				
28	Random				
29	1st				
30	3rd				
31	2nd				
32	Random				
33	1st				
34	3rd				

3rd round

Reliability Checker's Logbook

Date:

35	Random				
36	Random				
37	1st				
38	3rd				
39	Random				
40	2nd				
41	1st				
42	3rd				
43	Random				
44	2nd				
45	Random				
46	3rd				
47	Random				
48	2nd				
49	Random				
50	1st				
51	Random				

Appendix XI Task 1

TASK 1

A reminder:

- Task 1 is due Wednesday January 17, midnight.
- Send the task via WebCT to Geneviève Légaré.
- Please refer to the "Tasks Instructions" document for formatting instructions.

You are allowed to:

- Do a good job!
- Write as much as you need.
- Quote sources and make explicit references.
- Use other sources (books, magazines, journals, Web sites etc.).
- Refer to class discussions.

Your task for this week:

Write an essay, as long or short as you like, explaining your views about the impact of Technology on society.

Appendix XII Task 2

TASK 2

A reminder:

- Task 2 is due Wednesday January 24, midnight.
- Send Task 2 via WebCT to Geneviève Légaré.
- Please refer to the "Tasks Instructions" document for formatting instructions.
- Marks will be deducted for documents and messages that do not comply with the instructions.

You are allowed to:

- Refer to information in previous Task
- Refer to what you said in previous Task

You are still allowed to:

- Do a good job!
- Write as much as you need.
- Quote sources and make explicit references.
- Use other sources (books, magazines, journals, Web sites etc.).
- Refer to class discussions.

Your task for this week:

Write an essay, as long or short as you like, explaining your views about the impacts of food production technologies.

Appendix XIII Task 3

TASK 3

A reminder:

- Task 3 is due Wednesday January 31, **midnight**.
- Send task to Geneviève Légaré via **WebCT**.
- Please refer to the "Tasks Instructions" document for formatting instructions.
- Marks will be deducted for late Tasks.

You are still allowed to:

- Do a good job!
- Write as much as you need.
- Quote sources and make explicit references.
- Use other sources (books, magazines, journals, Web sites etc.).
- Refer to class discussions.

Your task for this week:

Assess the food crisis situation of Sayahn and provide a reasonable and feasible solution (or solutions) to solve the problem. The challenge consists of ensuring a stable and adequate amount of food each year. To help you out in this mandate, facts and data are presented in the Emergency Report included below.

Food Crisis in Sayahn

Contents:

Problem statement

The country of Sayahn is currently in a food crisis situation. In the past ten years, food production of the country has been barely sufficient to insure the minimum level of subsistence for its population. In addition, a severe drought affected the 2000 agricultural season of the region. As a result, experts are calling for "a timely intervention to prevent widespread starvation". The problem is to ensure a stable and adequate amount of suitable food each year in Sayahn.

Your mandate

As a consultant who has extensive expertise on issues pertaining to the domain of food production, you have been mandated by an international agency to help in the assessment of the situation. Your specific task consists of offering a reasonable and feasible solution (or solutions) that would solve Sayahn's food crisis problem. An Emergency Report has been produced by field observers and is included below. Facts and data should constitute the basis of your analysis of the problem. Given the urgency of the situation, the board is expecting your report in one week.

Emergency Report: Food Crisis In Sayahn*

15 October 2000

* Source: UN-Food and Agriculture Organisation Reports.

1. General description of the country

The resident population of Sayahn is officially estimated at 2.5 million people. After a decade of high levels of unemployment, low wages, and the steady erosion of safeguards against poverty, living conditions are precarious for most of the population, particularly in rural areas. With the added effects of the current drought, a large number of the rural population is expected to face serious food shortages in the current year (2000/2001).

Agriculture and the economy

Agriculture contributes about one-third to GDP and accounts for 42 percent of employment. Crop production accounts for roughly 60 percent of agricultural output and livestock production about 40 percent.

Factors affecting agricultural production

1.1 Physical constraints

Arable land is limited. The country is mountainous, with only 28 percent of land below 1500 metres altitude. Large number of farmers cultivate land which is on steep slopes and highly eroded and may have been suited for animal grazing and/or forestry. Topography and climatic conditions, soil fertility and the access to irrigation water vary greatly and affect yields. With inadequate use of fertilizer over a decade, lack of adequate drainage in the valleys (30 000 hectares are saline), soil erosion (many trees were felled in 1993-1994 for fuel to cope with energy shortages), degradation of pastures, and pollution, yield potential is undermined.

The country's irrigation systems are largely surface-water based. Ground water is also used in some areas. Snow cover is the main source of water feeding into the country's river systems. In addition, the extensive use of ground water for irrigation purposes has significantly lowered the water table in most areas. The state of disrepair of irrigation infrastructure and consequent loss of limited water has also exacerbated the problem.

1.2 Financial constraints

Farmers, like the rest of the population, have been impoverished by hyperinflation in the early nineties and limited earning opportunities. Access to formal credit for many small farmers remains difficult. As a result, access to quality inputs and machinery is difficult.

1.3 Institutional constraints

Lack of basic farming knowledge and management skills is a major bottleneck. The use of fertilizers, improved seeds and pesticides is limited and are usually acquired by barter. Agricultural research and the limited extension services have not been functioning properly for many years. Only four varieties of wheat were produced in the last 20 years. Lack of major plant breeding and seed production in the country is a serious drawback.

1.4 Marketing constraints

Marketing constraints are possibly the chief obstacle to increasing farm income and the most difficult for farmers to overcome. Roads are poor and transport is expensive. As a result, there are seasonal surpluses of perishable produce (e.g. of potatoes) while shortages persist elsewhere. The bulk of the population has very limited purchasing power and auto-consumption and barter probably account for up to two thirds of consumption of all domestic produce.

Effects of the drought

Both winter and spring rainfed cereal crops, harvested from July, were drastically reduced. The potato crop, harvested from September, was also severely affected. The problem of low water levels in rivers and canals due to the drought was compounded by the poor condition of the irrigation systems.

2. Food production

Official data is currently based on the reports of farming organizations and a survey of 7 000 small farmers but actual yields are not measured. Table 1 depicts the trends in agricultural production since 1991.

Table 1. Trends in agricultural production (area '000 hectares--production '000 tonnes)

	1991	1996	1997	1998	1999	2000/2001
Total area sown	417	347	335	324	328	371
Fodder crops area	213	105	80	69	94	105
TOTAL GRAINS						
Area	153	184	199	199	176	184
Yield (kg/ha)	1 954	1 734	1 296	1 638	1 710	1 174
Production	299	319	258	326	301	216
WHEAT						
Area	80	93	113	124	112	104
Yield (kg/ha)	2 138	2 108	1 584	1 927	1 911	1 452
Production	171	196	179	239	214	151
BARLEY						
Area	65	77	74	64	51	78
Yield (kg/ha)	1 785	1 325	797	1 109	1 275	692
Production	116	102	59	71	65	56
POTATOES						
Area	23	33	33	33	32	35
Yield (kg/ha)	11 957	12 818	10 909	13 333	12 938	7 143
Production	275	423	360	440	414	250
VEGETABLES						
Area	22	21	20	19	21	20
Yield (kg/ha)	20 136	21 190	18 450	20 789	21 381	19 000
Production	443	445	369	395	449	380
FRUITS						
Area	46	34	25	21	22	21
Yield (kg/ha	4 717	4 770	4 378	6 050	4 000	6 190
Production	217	158	109	126	88	130
GRAPES						
Area	27	21.8	17.8	15.7	15.7	15.2
Yield (kg/ha)	7 815	7 310	6 097	6 752	7 325	7 480
Production	211	158	108	106	115	113
LIVESTOCK PRODUCTION						
Meat (liveweight '000 tonnes)	133	83	84	86	86	160

Milk ('000 litres)	412	431	435	457	466	240
Eggs (millions)	482	192	191	219	326	350

Cereal production

Wheat production is estimated at about 151 000 tonnes. The drastic fall in feed is expected to lead to widespread under-nutrition in livestock, which in turn may result in a significant increase in mortality in the coming winter months.

Potato and vegetable production

Potato is one of the major staple crops in Sayahn. Potatoes account for about 10 percent of the total annual sown area, whereas vegetables represent approximately 5 percent of total sown area. This year, field visits have indicated that the potato crop is in a rather poor state.

Fruit production

Apple, apricot and peaches are some of the main fruits grown. Fruit production was spared from this year's severe drought, mainly due to its heavy reliance on irrigation. Current prices of these fruits are so low that the farmers are not even picking them.

Livestock production

In 1999, the livestock population consisted of nearly 479 000 head of cattle (of which 262 000 were cows), 549 000 sheep and goats, 71 000 pigs and some 2.9 million poultry. Virtually every rural family has at least one cow. In the single cow mixed income families the livestock products often represent about 40 percent of the household income.

Table 2. Comparison of yield between 1998-1999 (kg/ha)

	Cereals	Wheat		Barley	Potatoes	Grape
		Winter	Spring			
1998 Official	1 020	1 130	720	780	12 680	2 880
1999 Official	800	790	1 690	800	14 360	7 470

3. Food supply situation

The cereal supply/demand balance for 2000/01 (Table 3) is based on the following assumptions and parameters:

- As of December 2000, the resident population is 2.5 million;
- per capita wheat consumption (wheat flour in grain equivalent) of 170 kg/person/annum, 10 percent higher than last year, to offset drought-induced shortages in other food;
- wheat grain equivalent of the deficit in the staple potato crop of 33 000 tonnes, (based on a per

capita potato consumption of 73 kg/person/annum6) will add an extra 13 kg/person/annum;

- barley consumption, mainly used for feed, is expected to increase from an average of about 5 000 tonnes in the last few years to 10 000 tonnes due to food shortages;
- increased feed-use of barley to 60 000 tonnes from an estimated 40 000 tonnes last year. Feed use of wheat is estimated at 104 000 tonnes, nearly 30 percent below last year's estimate;
- seed requirement in 2000/2001 assumes similar area planted as last year: barley 78 000 hectares and potatoes 35 000 hectares. However, area under wheat is anticipated to be reduced by about 25 percent to 80 000 hectares (seeding rates: 300 kg/hectare for wheat, 250 kg/hectare for barley and 3 500 kg/hectare for potatoes);
- post-harvest losses and other uses are assumed to be about 15 percent of harvest;
- commercial import of wheat in 2000/01 is estimated at the average level of the previous three years of about 350 000 tonnes of wheat and wheat flour (in grain equivalent), 5 000 tonnes of barley and 1 000 tonnes of potatoes.

Table 3. Staple food supply/Demand balance for 2000/2001 ('000 tonnes)

	Wheat	Barley	Potatoes
Domestic availability	151	54	250
Stock draw-down	--	--	--
Domestic production	151	54	250
Total utilization	256	98	394
Food used feed seed	42 510 424	106 020	18 350 123
Other uses/losses	23	8	38
Stock buildup	--	--	--
Import requirements	425	44	144
(Cereal equivalent)	425	44	33
Anticipated commercial	350	7	1
Food aid pledges	70	--	--
Uncovered deficit	38	37	--

4. Food assistance requirements

The overall socio-economic and agricultural situation is alarming with the large shortfall in cereal and in potato production, the fragile nutritional status of many people and the limited purchasing power and capacity to cope for most, particularly female-headed households. Due to economic hardship, the diet of the vulnerable population has been mainly bread, potatoes and cabbages. The nutritional status of the vulnerable population could be further exacerbated by Sayahn's usually harsh winter.

Table 4. Number of people critically in need of food assistance and the amount of wheat required as food assistance from Nov.2000 to June 2001.

Regions most seriously affected	Rural population			People in critical need of assistance in rural areas	% of people in critical need of assistance in rural areas	Total food commodities allocated
	Female	Male	Total			
L.	65 630	55 906	121 536	68 125	56.1%	4 651
S.	53 789	45 820	99 609	65 659	65.9%	4 477
T.	56 354	48 004	104 358	51 648	49.5%	3 522
G.	104 707	89 194	193 901	52 132	26.9%	3 555
A.	28 981	24 686	53 667	25 259	47.1%	1 722
K.	44 580	37 976	82 556	34 089	41.3%	2 324
TOTAL	354 041	30 586	655 627	297 000	45.3%	20 252

Appendix XIV Task 4

TASK 4

A reminder:

- Task 4 is due Wednesday February 7, **midnight**.
- Send task to Geneviève Légaré via **WebCT**.
- Please refer to the "Tasks Instructions" document for formatting instructions.
- Marks will be deducted for late Tasks.

You are still allowed to:

- Do a good job!
- Write as much as you need.
- Quote sources and make explicit references.
- Use other sources (books, magazines, journals, Web sites etc.).
- Refer to class discussions.

Your task for this week:

Your task for this week is twofold. First, you have to write an episode of the television series *Impacts!* Second, you have to provide a rationale for the episode. The topic for the show is "Food Production". You are expected to outline some of the specific problems and offer potential solutions.

TV Episode

Contents:

Your credentials

A few years ago, you graduated from Concordia University with a Bachelor's degree in engineering. Since then, you have worked in large companies, but after a while you decided that it was time to start your own business. You felt that you needed to use your expertise in a different way. As a result, you created a consulting firm called *«TechTalks»*.

The main mission of *TechTalks* is to popularize, to explain in layman's terms, the relative significance of tools, techniques and technologies to various clienteles. In other words, you became a professional communicator whose personal goal is to get people to understand the impacts of technology. *TechTalks* is doing quite well and your accomplishments are indeed impressive. Your professional portfolio includes:

Several articles for magazines. To name but a few, you have contributed to :*

- Technology Review for which your wrote "Writing Devices: A Revolution?";
- Québec Science : "Les Secrets de la fabrication du papier journal");
- and for the prestigious Smithsonian : "The Role of Ink in the Development of the Printing Press" and also, "Modern Paint Pigments: How Chemistry Affected the Way Artists View Colours".

* The magazines really do exist but the articles are fictitious...

- You developped a communication strategy for the *International Motorist Association*. The main goal of the strategic plan was to explain the technical and social issues of the electric car. The target audience was the corporate and individual members of IMA.
- As a regular guest on radio shows, you are invited to debate ethical issues about medical and bio-medical developments. Topics surveyed include "feeding antibiotics to cattle", "human cloning", "genetic engineering of crops" and even "doping strategies for athletes".
- For the count of a science museum, you participated as a content expert in the development of a museum exhibition intitled "0". The main goal of the exhibit was to explain the concept of "zero" as well as to present several calculating devices.
- When you have spare time, you make a point of visiting schools to explain to young children the role(s) and responsabilities of professionnal engineers. Of course, this activitiy is volunteer work.

Although your accomplishments prove to be quite fulfilling both on the personal and professional sides, you have not yet achieved financial freedom. But this is about to change...

Recently, *Seen-Us? TV*, a national television network called for proposals to get content writers involved in a project. The broadcasting company is launching a new television series called *Impacts!* When you received the invitation to submit a bid, you could not believe it: The topic was right up your alley! You know that television networks pay very well and that if you succeed with the first project, you might get new contracts. Your financial security would thus be insured for years to come.

The first step of the contest was to present an "essay" on the topic of your choice. Apparently, the topic did not really matter, as long as it was somehow related to Technology. You have submitted a piece about some issues concerning food production technologies. This morning, you received a letter from the producer of *Impacts!*.

Letter of invitation

January 31, 200x

<u>To</u>:
Mr., Mrs. So-and-So
President,
TechTalks

<u>From</u>:
Mr. Art Bukowski
Project Manager and
Producer of the television series: *Impacts!*
SeenUs? TV Network

Dear Mr., Ms. So-and-So,

Thank you for having taken the time to submit a bid in our call for proposals. As you may have figured out, we use the strategy of essay writing as a means to select a successful candidate. Your piece about food production technologies was quite interesting. I enjoyed very much reading your essay and I think that as a content expert, you have great potential, especially because of the original perspectives you have to offer.

This is why I, as the main producer of *Impacts!*, am delighted to welcome you as a new and precious collaborator on our team. Your task for now consists of writing the content of one episode of *Impacts!* The topic? Well, given your expertise on the subject of food production, as well as the novel perspectives you have demonstrated, the team decided to keep it. I know! I know! You have already written several drafts on the subject for the bidding phase. But if I may say, to produce something of great quality in the broadcasting industry or, as far as I am concerned, in any other fields, requires several iterations. This is what we call "the cost of excellence". Anyway, your mandate is twofold. I have outlined the specifications below. I took the liberty of feeding you a few tips of the trade in order to avoid usual pitfalls.

Impacts!: <u>Content Writer's Mandate</u>

Food Production Episode

PART 1: Write the content for one episode of *Impacts!* about food production.

Tips of the trade:

- Although one episode lasts between 45 and 55 minutes, you can write as much as you want;
- You can choose any of the following organisational strategy:
  - Dialogue (for example, a discussion between an interviewer and interviewees)
  - Narrative (story)
  - The classic strategy is of course the well known "Introduction-Body-Conclusion";
  - There is also the "Thesis-Antithesis-Synthesis" approach;
  - You can proceed in a scientific way by presenting a set of facts, formulating an hypothesis and then proceed to reject or accept the proposed hypothesis;
  - You could even be more adventurous and creative by selecting what I call the "string of pearls" approach. The pearls are a series of facts or arguments that appear more or less related, but once you put them together, your artistic treatment subtly reveals the "string" to viewers. It is what we call the second degree of a text ; or even better,
  - You can use your own organisational strategy.
- About the content: You have *carte blanche*. You treat any issues you wish to cover. However, before outlining the content, I would advise you to consider the *SeenUs? TV* "Introductory Documentation". Since it is your first collaboration with us, I have included the documents for your benefit. The documentation covers important points such as the mission statement, the goals, the profile of our target clientele and includes a description of the series *Impacts!*.
- I also included a "food file" which contains some information about the topic.
- Suggest, whenever relevant, a recommandation for visuals (archives, pictures, charts and graphics) to be used in the show. Of course, you can include the visuals themselves if they are available. In any case, you need to indicate the source of the references so my research assistant can clear the copyrights.
- Do not forget to think of a great title!

PART 2: The second part of you mandate is the rationale for the episode.

Your job consists in describing the strategic decisions you have taken while writing the content for the show. You may want to use the following questions as a guide to develop the rationale:

- Which issues about food production did you leave out and why?
- Why did you choose to treat the selected issues?
- Considering our target population, explain why the viewers will watch the *Impacts!* episode about food production?
- Why did you sequence the information the way you did?
- Finally, if you were an on-line facilitator of our *Impacts!* chat groups, what questions would you ask the participants to enhance their understanding about food production?

The reason I am asking to produce this rationale is that I have a meeting with the Vice-President Programming for official approval of the food production episode. As a promoter of excellence, he always challenges the producers' and the writers' plans. He will ask us these questions. If our answers prove to be satisfactory, we will get the green light for the project.

I am meeting the Vice-President for breakfast Thursday morning February 8. So please, send the episode via WebCT to Ms. Légaré, my executive assistant, no later than **midnight on February 7**. She will diligently forward your attached file to me.

I hope that my information will have been useful. I am looking forward to see you in person. I am sure, Mr., Mrs. So-and-So, that this is the dawn of a fruitful collaboration.

Sincerely,

Art Bukowski
Project Manager and
Producer of the television series: *Impacts!*
Seen-Us? TV Network

Network Bio

SeenUs? TV Network

Introductory Documentation

Mission statement

Educate--rather than entertain-- the general public through the comprehensive treatment of issues pertaining to a wide variety of topics.

Goals

- Challenge the basic assumptions of viewers;
- Get our viewers to pursue their interest in topics we propose;
- Focus on content rather than form in order to promote comprehensive understanding of issues and topics;
- Avoid entertainment frills to grab viewers' attention;
- Get the people to talk about us: We want to trigger discussions about the topics we present;
- Use a variety of current technologies to enhance learning potential;
- Succeed by making money with documentaries and educational programmes.

Financial profile

SeenUs? TV is a public broadcasting company. We conduct annual fundraising campains. Basically, our viewers subsidize what they want to see.

Target audience

The *Seen-Us? TV* target audience is somewhat difficult to circumscribe. Our viewers come from all kinds of background: They are from various cultural and linguistic backgrounds, and they have different educational levels and economic means.

However, according to a survey we conducted during our fundraising campain, we do know that our viewers are somewhat tired of sentionnalism. They are turning away from entertainment in favor of in-depth treatment. They want to think, they want facts as well as opinions. In essence our viewers are demanding and we are more than happy to take on the challenge of satisfying them.

Concept of the television series: " *Impacts!*"

The main concept of *Impacts!* is that we want the viewers to rediscover Technology from different points of view. We want them to consider facts, aspects, and impacts that they would have never thought of before. Essentially, through our treatment of the subject, we will bring them to go beyond their basic assumptions, beyond primary thinking. In addition, we expect that after watching the complete series of *Impacts!*, the viewers will have developped a comprehensive understanding of the Impacts of Technology on Society.

Impacts! has a second component and *SeenUs?TV* is quite proud of it. After each episode, the viewers are invited to chat online with our team of experts. They discuss the angles of the show and share their thoughts with us. Our hidden agenda is to get them to learn even more about the issues presented. At the end of the first programme year, we will collapse all commentaries and do a second part for every topic we broadcasted. It will be a real catharsis! But the second component is still in the making. We still have to finalize the logistics of such a great adventure!

Food file

Excerpts	PDF files (access to)	Web Sites	On-line articles	Graphs & Tables

Excerpts:

From "Trends Point to Gains in Human Development, While Many Negative Human Impacts on Vital Ecosystems are Increasing". Sustainable Development Information Service--Global Trends. http://www.igc.org/wri/trends/index.html *(last access: 01-2001)*.

- Projections of future population between 8 and 12 billion by 2050
- "Global food production is generally adequate to meet human nutritional needs, but problems with distribution mean that some 800 million people remain undernourished. World food production is still rising, but everal trends will make it more challenging to feed an additional 3 billion people over the next 30 years."
- "Consumption of natural resources by modern industrial economies remains very high--in the range of 45 to 85 metric tons per person annually".
- "Global energy use, which has increased nearly 70% since 1971, is projected to increase at more than 2% annually for the next 15 years."
- "The world has lost half its forets over the past 8,000 years through converesion to farms, pastures, and human settlements or commercials sites."
- "...water availability is likely to become one of the most pressing and contentious resource issues of the 21st Century."

From: Rome Declaration on World Food Security. Rome Declaration on World Food Security and World Food Summit Plan of Action (1996). From: http://www.fao.org/wfs/homepage.htm *(last access: 10-01-2001)*.

- "We pledge our political will and our common and national commitment to achieving food security for all and to an ongoing effort to eradicate hunger in all countries, with an immediate view to reducing the number of undernourished people to half their present level no later than 2015".
- "Poverty is a major cause of food insecurity".

- "Increased food production, including staple food, must be undertaken. This should happen within the framwork of sustainable management of natural resources, elimination of unsustainable patterns of consumption and production, particularly in industrialized countires, and early stabilization of the world population".
- "We reaffirm the importance of international cooperation and solidarity..."
- "We are determined to make efforts to mobilize, and optimize the allocation and utilization of technical and financial resources from all sources, including external debt relief for developing countries, to reinforce national actions to implement sustainable food security policies".

From: The state of food insecurity in the world, 1999.

- Jacques Diouf, Director-General, FAO
- Goals: Reduce the number of hungry people to around 400 million by 2015; reduce by half or more the number of 34 million hungry in developed countries.
- "If all the world's undernourished people were gathered together, the population of the continent of the hungry would dwarf that of every other continent except Asia."
- Two distinct measures:
    - Undernourishment: data about number of people and the amount of food avaiable to them.
    - Undernutrition: from data about people's weight, height and age. Indicate the outcomes of inadequate food intake, poor health and sanitation conditions.

From: FAO of the United Nations, "World Food Summit: Five years later. International conference to review progress in fight againts hunger". Press Release 00/64.
http://www.fao.org/WAICENT/OIS/PRESS_NE/PRESSENG/2000/pren0064.htm *(last access: 21/01/2001).*

- In a statement to the Council, Dr.Diouf said that the present state of food insecurity in the world "requires action beyond business as usual", adding that several factors have created "an environment that is making it difficult to achieve the goals of the World Food Summit".
- "...the number of hungry people is being reduced only by 8 million per year instead of the 20 million necessary to achieve the goals of the World Food Summit".

From: "OBESITY, AN OVERVIEW", by S. J. McNulty, * G. Williams**

* Clinical Research Fellow in Diabetes and **Professor of Medicine, Diabetes and Endocrinology, Research Group, Department of Medicine, University Hospital Aintree, Liverpool. From:
http://www.rcpe.ac.uk/public/volume29_3a.html

EPIDEMIOLOGY OF OBESITY

The average BMI in the United Kingdom has been increasing steadily since the turn of this century.This is by no means a unique phenomenon, as over half the adult population (aged 35-65 years) of Europe are now overweight and 10-20% of men and 15-25% of women are obese.6 In the United States, the prevalence of clinical obesity is over 50% in certain subgroups of the population (Hispanics and Blacks).5 This steady trend towards ever-increasing weight is illustrated by the aeroplane manufacturers Boeing, who have had to increase the assumed weight of each passenger by over 20 pounds since they first began to produce planes;5 in addition, the impression of those who visit the USA at intervals is that the diameter of the standard American restaurant plate has steadily expanded. Figures 1 and 2 show the rising prevalence of obesity in both the United Kingdom and the United States.

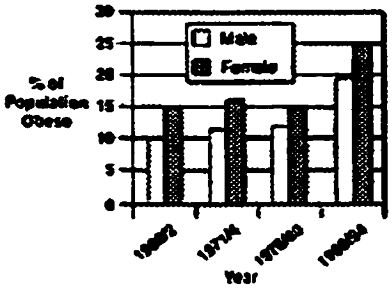


FIGURE 1
Prevalence of obesity (BMI ≥30 kg/m²) in USA

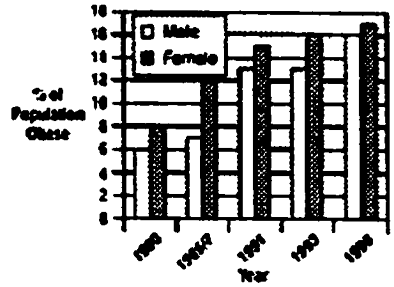


FIGURE 2
Prevalence of obesity (BMI >30 kg/m²) in UK

From: *Canadian food guide*: http://www.hc-sc.gc.ca/hppb/nutrition/pube/foodguid/eguide3.html

Depending on age, gender and activity level, the *Canadian Food Guide* recommands:

- Grains: 5-12 servings per day
- Fruits and vegetables: 5-10 servings per day
- Milk (adults): 2-4 servings per day
- Meat and alternatives: 2-3 servings per day

Dietary Guidelines for Americans: http://vm.cfsan.fda.gov/~dms/nutguide.html

---

Food file

Excerpts	PDF files (access to)	Web Sites	On-line articles	Graphs & Tables

PDF Documents (Access to)

Some effects of global warming on agriculture: http://www.fao.org/News/Factfile/FF9721-E.HTM

Proportion of undersnourished people: http://www.fao.org/NEWS/FACTFILE/FF9809-E.HTM

Main causes of dryland soil degradation by region: http://www.fao.org/News/Factfile/FF9710-E.HTM

Trends in Cassava production: http://www.fao.org/News/Factfile/FF0001-e.htm

People living in urban area, 2005: http://www.fao.org/News/Factfile/FCTLIB-E.HTM

Number of hungry people rising: http://www.fao.org/News/Factfile/FF9808-e.htm

Food Aid recipients: http://www.fao.org/News/Factfile/FF9606-e.htm

Number of people undernourished: http://www.fao.org/News/Factfile/FF9609-e.htm

Undernutrition can be reduced: http://www.fao.org/NEWS/FACTFILE/FF9702-E.HTM

Food file

Excerpts	PDF files (access to)	Web Sites	On-line articles	Graphs & Tables

Web Sites

International Dairy Federation: http://www.fil-idf.org

International Federation of Agricultural Producers (IFAP): http://www.ifap.org

International Fertilizer Industry Association (IFA): http://www.fertilizer.org (check server)

International Meat Secretariat (IMS): http://www.meat-ims.org

Statistics Canada: http://www.statcan.ca/english/Pgdb/

Government Publications Canada: http://publications.pwgsc.gc.ca/publishing/info/govpub-e.html

OECD (biotechnology): http://www.oecd.org/ehs/icgb/

WRI : Water demand trends: http://www.igc.org/wri/trends/water.html

WRI : Food crops and biodiversity : http://www.igc.org/wri/biodiv/foodcrop.html

WRI:Global trends (Sustainable development services): http://www.igc.org/wri/trends/index.html

US Global Change Research Information office: http://www.gcrio.org/

Food file

Excerpts	PDF files (access to)	Web Sites	On-line articles	Graphs & Tables

On-line articles

Definition of Biological Diversity: http://www.igc.org/wri/biodiv/cwb-i.html

Article by Anthony C. Janetos: "Do We Still Need Nature? The Importance of Biological Diversity". *Consequences: The nature & implications of environmental changes*, 3(1). http://www.gcrio.org/CONSEQUENCES/vol3no1/biodiversity.html

"The Fattening of America. More" by Neil Sherman: http://body.subportal.com/health/Nutrition_Fitness_Looks/Weight_Control/Obesity/106551.html

"A hunger for Fatty foods, Europe is catching up US": http://body.subportal.com/health/Nutrition_Fitness_Looks/Weight_Control/Obesity/100438.html

"Foreign Students Balloon on U.S. Diet": http://body.subportal.com/health/Nutrition_Fitness_Looks/Weight_Control/Obesity/104489.html

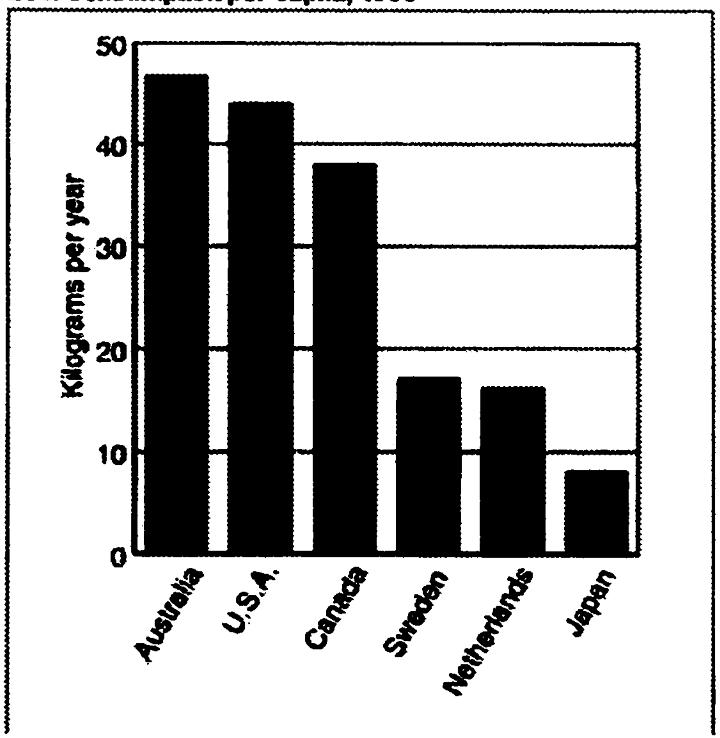
Food Outlooks: http://www.fao.org/WAICENT/faoinfo/economic/giews/english/fo/fotoc.htm

Food file

Excerpts	PDF files (access to)	Web Sites	On-line articles	Graphs & Tables

Graphs and Tables

Figure 3
Beef consumption per capita, 1990

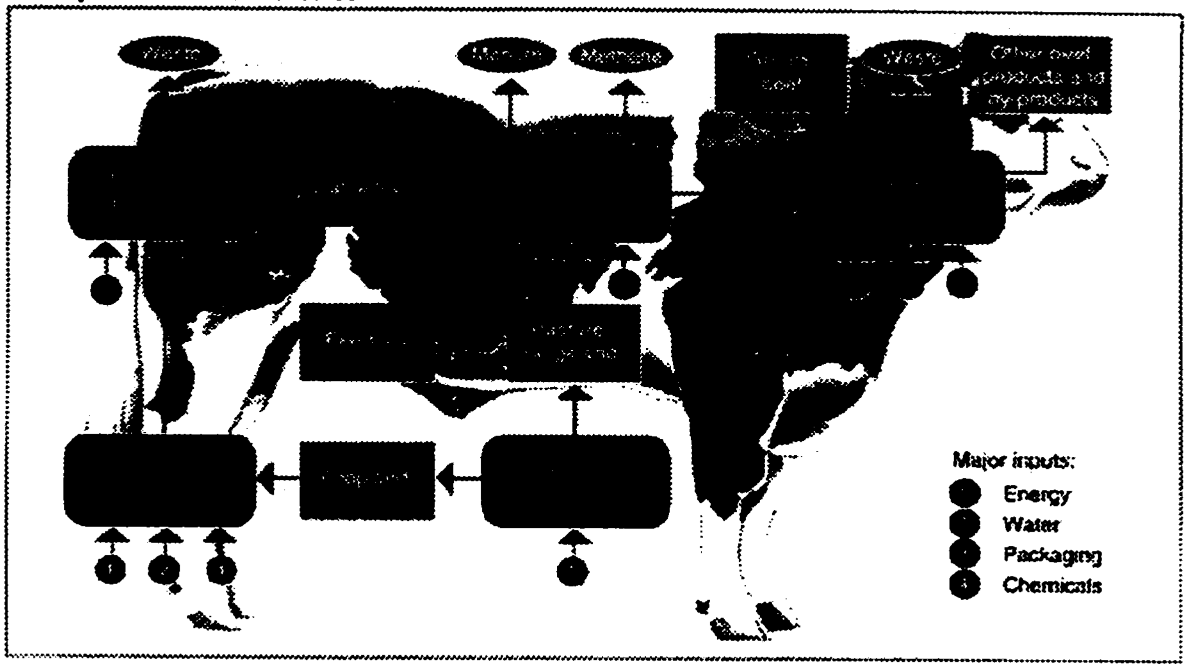


Source : Food and Agriculture Organization, Agrostat database

Original article : Minister of Public Works and Government Services Canada (1995). "Connections: Canadian lifestyle choices and the environment". *SOE Fact Sheet* No.95-1, Environment Canada, p.4. (*beef consum.jpg*).

About cattle in Canada: "About 52% of Canadian farms have cattle, totalling roughly 12 million animals. Around 30% of the country's farmland, or roughly 4 million hectares, is used as pasture".

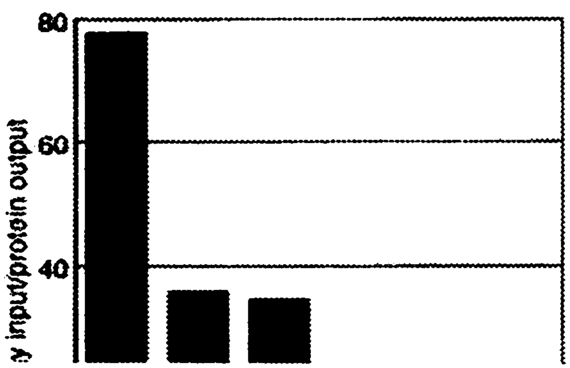
Figure 4
Beef production : an overview

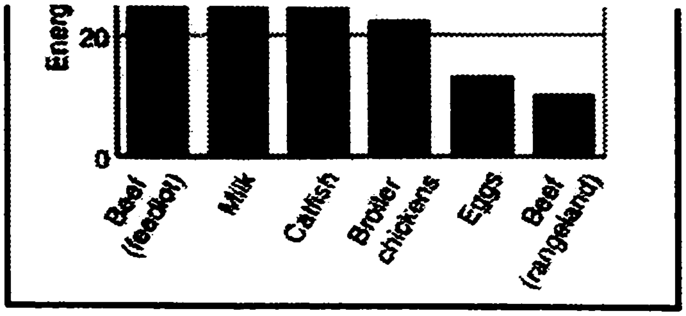


Original article : Minister of Public Works and Government Services Canada (1995). "Connections: Canadian lifestyle choices and the environment". *SOE Fact Sheet* No.95-1, Environment Canada, p.5. (*beefprod.jpg*).

About energy requirement: "There are significant differences in how efficiently various animals and production methods use energy to produce protein."

Figure 5
Energy requirements : production of animal protein

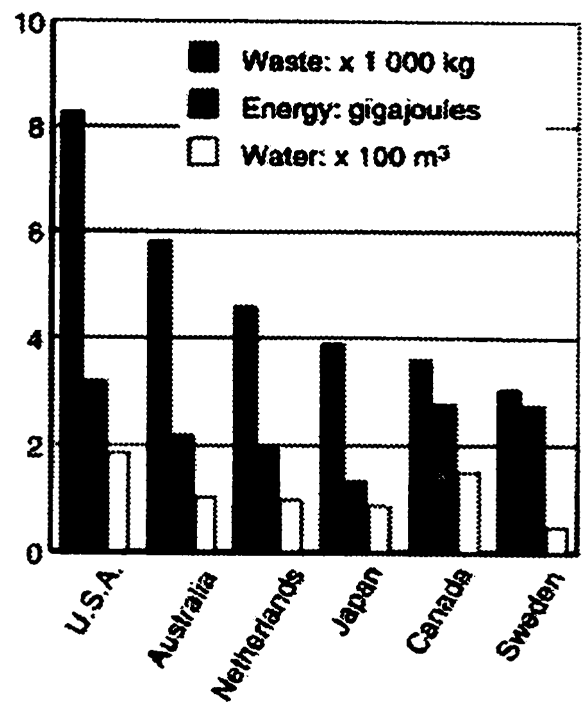




Original article : Minister of Public Works and Government Services Canada (1995). "Connections: Canadian lifestyle choices and the environment". *SOE Fact Sheet* No.95-1, Environment Canada, p.6. (*energyrequire.jpg*).

About urban solid waste production: "Industrialized countries tend to be among the largest producers of waste and among the largest consumers of energy water in the world."

Figure 8
Urban solid waste production, energy use, and water use per capita, 1991



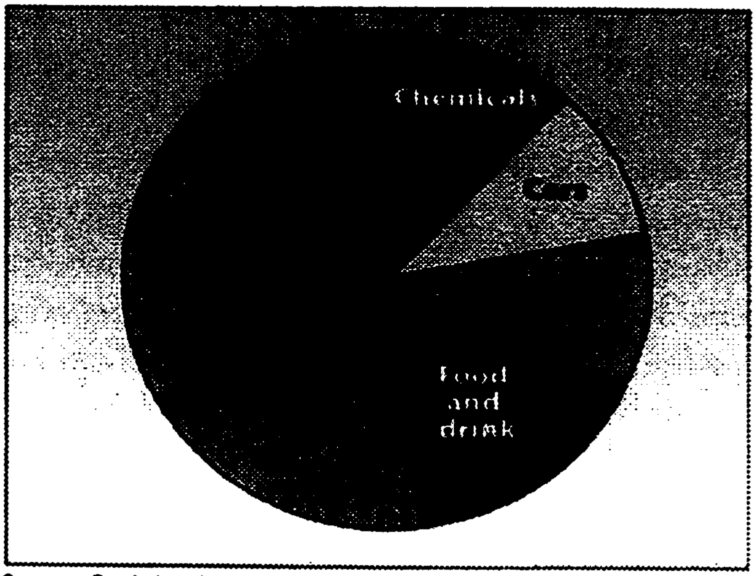
Note: Waste figure for Canada refers to residential waste.

Source: Statistics Canada (1994c).

Original article : Minister of Public Works and Government Services Canada (1995). "Connections: Canadian lifestyle choices and the environment". *SOE Fact Sheet* No.95-1, Environment Canada, p.9. (*urbanwaste.jpg*).

About manufacturing sectors: "The food industry is one of the world's largest industrial sectors" .

Figure 1
European Union manufacturing sectors by production value, 1991



Source : Confederation of the Food and Drink industries of the EEC

Processed food output 1990: 20 largest countries (US$ billion)

1	USA	104	11	Australia	8
2	Japan	75	12	Belgium	7
3	Germany	34	13	South Korea	6
4	France	26	14	Austria	5
5	UK	18	15	India	4
6	China	17	16	Thailand	4
7	Italy	15	17	Argentina	4
8	Brazil	14	18	Indonesia	4
9	Mexico	10	19	Philippines	4
10	Netherlands	9	20	Denmark	4

Source : Confederation of the Food and Drink industries of the EEC

Original article: United Nations Environment Programme (January-March 1995). "Facts and figures: Food processing and the environment". *UNEP Industry and environment*, 18(1), pp.4-7. *(ProcessedFood.jpg)*

Table 1: Food imports and exports by world region, 1992. (p.4)

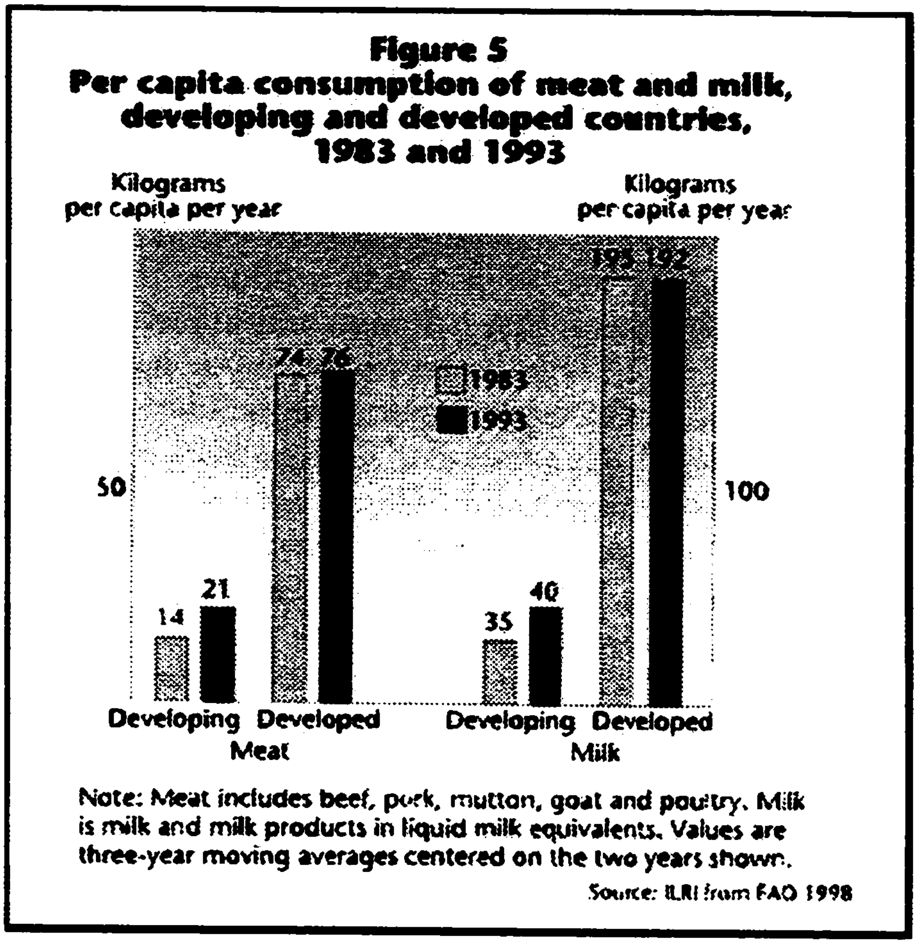
	US$ billion	
	Imports	Exports
Developed Countries	195.1	180.8
North America	22.9	40.8
Europe	135.3	126.4
Oceania	1.6	10.5
CIS	12.6	0.4
Other developed	22.6	2.8
Developing countries	66.8	57.1
Africa	9.8	4.6
Latin America	14.0	21.6
Near East	16.8	5.8
Far East	25.5	24.7
Other developing	0.8	0.4
Source: UN Food and Agriculture Organization		

Original article: United Nations Environment Programme (January-March 1995). "Facts and figures: Food processing and the environment". *UNEP Industry and environment*, 18(1), pp.4-7.

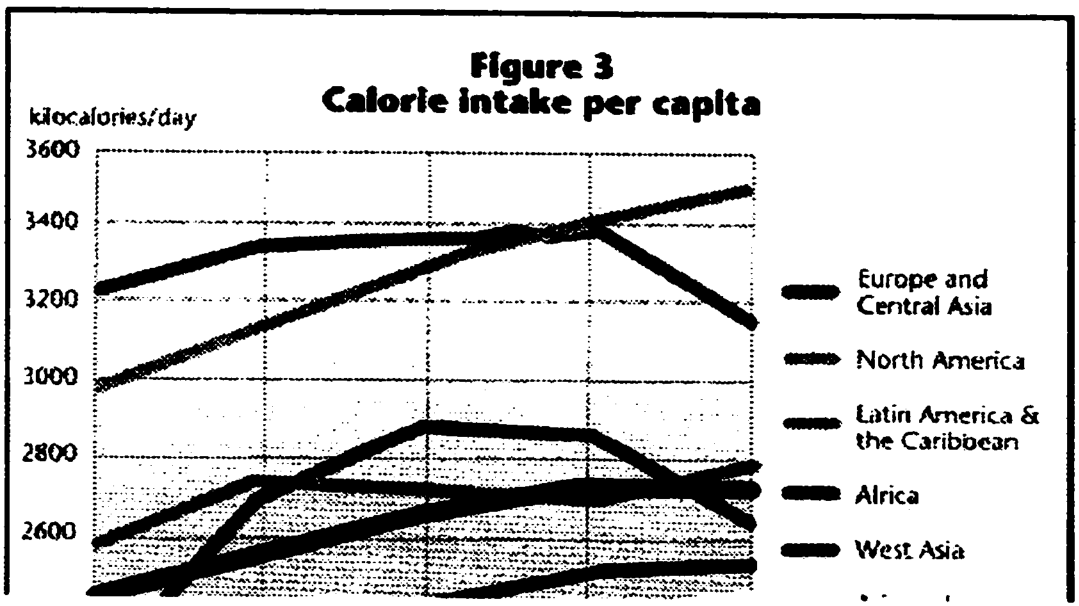
Table 3: World's 12 largest food companies, 1992.

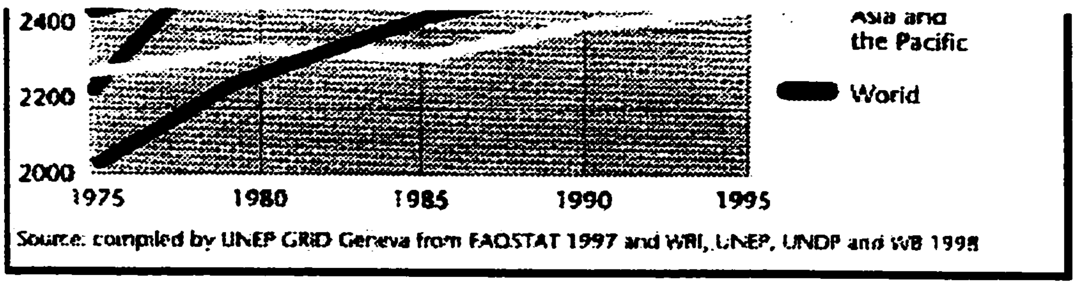
	food sales, US$ billion
Nestlé	37.6
Philip Morris	33.0
Unilever	22.5
PepsiCo	13.7
BSN	13.1
Grand Metropolitan	12.4
RJR Nabisco	9.7
Sara Lee	6.7
CPC INternational	6.6
Heinz	6.6
Campbell	6.3
Source: *Food Business*	

Original article: United Nations Environment Programme (January-March 1995). "Facts and figures: Food processing and the environment". *UNEP Industry and environment*, 18(1), pp.4-7.



Original article: United Nations Environment Programme (April-September, 1999). "Agri-food production: Facts and figures" . *UNEP Industry and environment*, 18(1), pp.4-7. *(consumption.jpg)*





Original article: United Nations Environment Programme (April-September, 1999). "Agri-food production: Facts and figures" . *UNEP Industry and environment*, 18(1), pp.4-7. *(calorie.jpg)*

Table 1 Annual per capita grain use, consumption of livestock products, and life expectancy at birth in selected countries, 1995.

Country	Grain use (1)	Per Capita consumption				
		Beef	Pork	Poultry	Milk (2)288	Life expectancy (years)
United States	800	45	31	46	288	77
Italy	400	26	33	19	197	78
China	300	4	30	6	5	72
India	200	1	0.4 (1)	1	34	60
(1) Rounded to nearest 100 Kg. (2) Estimates based on FAO production figures for 1994.						
Source: FAO 1995; USDA 1996; Population Research Bureau 1999.						

Original article: United Nations Environment Programme (April-September, 1999). "Agri-food production: Facts and figures" . *UNEP Industry and environment*, 18(1), pp.4-7.

FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.1, February 2000. 40p.*(p.26, fertil_FO26b.gif)*

Average Fertilizer Spot Prices (bulk, f.o.b.)

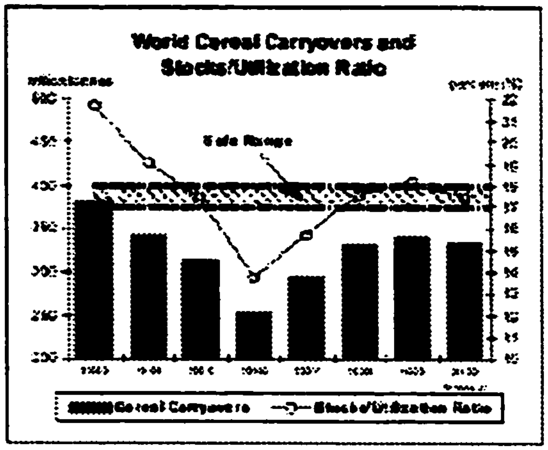
	2000	1999		Change from
	January	January	December	last year 1/
	(..............	US$/tonne	..............)	(percentage)
Urea				
eastern Europe	79-81	63-65	68-69	+ 25.0
Near East	109-113	79-82	95-97	+ 37.9
Ammonium Sulphate				
eastern Europe	42-43	31-34	40-42	+ 30.8
Far East	55-56	50-51	55-56	+ 9.9
U.S. Gulf	42-44	45-55	38-41	- 14.0
western Europe	55-60	41-46	55-60	+ 32.2
Diammonium Phosphate				
Jordan	164-169	205-209	165-170	- 19.6
North Africa	159-164	201-207	162-166	- 20.8
U.S. Gulf	147-151	201-203	145-148	- 28.2
Triple Superphosphate				
North Africa	131-135	158-162	132-135	- 16.9
U.S. Gulf	136-140	163-170	136-140	- 17.1
Muriate of Potash				
eastern Europe	95-100	95-106	95-100	+ 0.5
Vancouver	117-131	115-129	117-131	+ 1.6

western Europe	129-137	129-137	129-137	-

Source: Compiled from Fertilizer Week and Fertilizer Market Bulletin

1/ From mid-point of given ranges.

FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.1, February 2000. 40p.*(p.16, carryover_FO16.gif)*



FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.1, February 2000. 40p. *(p.18, exportprices_FO18.gif)*

Cereal Export Prices *

	2000 Jan.	1999 Dec.	1999 Jan.
	(.....	US$/tonne	.....)
United States			
Wheat 1/	111	105	126
Maize	93	89	98
Sorghum	91	85	96
Argentina 2/			
Wheat	94	81	112
Maize	93	90	93
Thailand 2/			
Rice white 3/	244	240	307
Rice, broken 4/	161	153	230

Source: FAO, see Appendix Table A.9

* Prices refer to the monthly average.
1/ No. 2 Hard Winter (Ordinary Protein).
2/ Indicative traded prices.
3/ 100% second grade, f.o.b. Bangkok.
4/ A1 super, f.o.b. Bangkok.

FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.1, February 2000. 40p.*(p.20, meatexp_FO20.gif)*.

World Meat Exports 1/

	1998	1999	2000 estim.
	(...	thousand tonnes	...)
WORLD	**15 097**	**15 835**	**15 846**
Poultry meat	6 163	6 395	6 483
Pig meat	2 882	3 080	2 983
Bovine meat	5 126	5 417	5 423
Sheep meat and			

Goat meat	678	690	691
Other meat	251	252	266

Source: FAO

Note: Total computed from unrounded data.

1/ Includes meat (fresh, chilled, frozen prepared and canned) in carcass weight equivalent; excludes live animals, offals and EC intra-trade.

FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.1, February 2000. 40p. *(p.38, freight_FO38.gif).*

Table A 8 - **OCEAN FREIGHT RATES FOR WHEAT**

	From U.S. Gulf ports to:					From North Pacific ports to:	
	Rotterdam 1/	CIS Black Sea 1/ 2/	Egypt (Alexandria) 1/	Bangladesh 1/	East Africa Sudan 1/	China 1/	Japan 1/
	(			US$/tonne			)
July/June							
1994/95	15.25	30.46	18.74	23.75	39.65	22.29	32.48
1995/96	12.95	30.00	16.83	21.87	41.65	25.94	35.00
1996/97	11.00	16.65	12.77	20.00	-	27.00	28.29
1997/98	9.50	18.10	11.70	20.17	-	27.50	26.00
1998/99	9.40	25.45	9.25	18.75	-	27.00	29.17
1999 - January	7.50	22.00	8.50	18.50	-	27.00	30.00
June	14.75	22.00	11.00	18.50	-	27.00	30.00
July	14.75	n.q.	11.00	18.50	-	27.00	30.00
August	14.75	40.97	12.00	18.50	-	27.00	31.00
September	12.50	40.97	11.00	18.50	-	27.00	31.50
October	12.00	40.97	13.00	18.50	-	27.00	32.00
November	12.00	40.97	15.00	18.50	-	27.00	32.25
December	11.75	40.97	13.00	18.50	-	27.00	32.25
2000 - January	13.00	40.97	15.05	18.50	-	27.00	32.50

SOURCE: International Grain Council

Note: Estimated mid-month rates based on current chartering practices for vessels ready to load three to four weeks ahead

1/ Size of vessels: Rotterdam over 40 000 tons; CIS 20-40 000 tons; Egypt over 30 000 tons; Bangladesh over 40 000 tons; East Africa 15-25 000 tons; China 20-35 000 tons; Japan 15-24 999 tons.

2/ Excludes CIS and United States flag vessels.

FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.5, November 2000. 40p. *(p.39, commodities_FO39.gif).*

Table A.13 - **SELECTED INTERNATIONAL COMMODITY PRICES**

	Currency and Unit	Effective Date	Latest Quotation	1 month ago	1 year ago	Average 1989-91
Sugar (I.S.A. daily price)	US cents per lb	27.01.00	5.4	6.0	6.3	11.4
Coffee (I.C.O. daily price)	US cents per lb	27.01.00	80.7	86.6	93.5	76.7
Cocoa (I.C.C.O. daily price)	US cents per lb	27.01.00	41.6	41.3	64.3	56.3
Tea (all tea, London, weekly)	US$ per kg.	27.01.00	1.8	1.9	1.7	1.5
Bananas (Central America, f.o.r. Hamburg)	DM per tonne	27.01.00	1 078 1/ 1 369 2/	1 474 1/ 1 202 2/	2 026 1/ 1 436 2/	1 107
Rubber (RSS 1, spot London)	Pence per kg.	27.01.00	46.5	40.5	62.8	54.5
Cotton (COTLOOK, index "A" 1-3/32")	US cents per lb	27.01.00	47.2	44.0	55.7	78.5
Wool (64's, London)	Pence per kg	27.01.00	270	270	315	460

SOURCE: FAO

1/ EC duty paid, estimated. 2/ Estimated price for EFTA markets

FAO, Global Information and Early Warning System on Food an Agriculture. *Food Outlook*, No.1, February 2000. 45p. *(p.20,shipments_FO20.gif).*

Non Cereals - Food Aid Shipments by Destination

	1995	1996	1997	1998	1999 estim.
	(.....	.....	thousand tonnes	.....	.....)
WORLD	**1 281**	**887**	**896**	**916**	**1 627**
LIFDC	670	615	580	627	635
Africa	466	392	313	302	296
Asia	234	255	275	341	313
Latin America and the Caribbean	152	146	170	232	152
Others	420	193	138	39	868

Source: WFP

Note: Totals computed from unrounded data.

Appendix XV Task 5

Short Paper

<u>A reminder</u>:

- The Short Paper is due Wednesday February 14, midnight.
- Send assignment to Geneviève Légaré via WebCT.
- Please refer to the "Tasks Instructions" document for formatting instructions.

<u>You are allowed to</u>:

- Do a good job!
- Write as much as you need.
- Quote sources and make explicit references.
- Use other sources (books, magazines, journals, Web sites etc.).
- Refer to class discussions.
- Refer to previous Tasks content.

<u>Your task for this week</u>:

The Short Paper has the same structure as Task 4. Your task is still twofold. First, you have to write an episode of the television series *Impacts!* Second, you have to provide a rationale for the episode. The topic for the show is "labelling of genetically modified foods". You are expected to outline some of the specific problems and offer potential solutions.

Your Mandate

Contents:

Introduction

As the president of TechTalks, you have recently completed a mandate for the television network Seen-Us? TV. You wrote the content of one episode of *Impacts!* about food production. The producer of the show, Mr. Art Bukowski has written another letter, which you have received this morning.

Art Bukowski's Letter

February 7, 200x

To: Mr., Mrs. So-and-So

President,

TechTalks

From: Mr. Art Bukowski

Project Manager and

Producer of the television series: Impacts!

SeenUs? TV Network

Dear Mr., Ms. So-and-So,

We have received your fabulous episode of *Impacts!* about food production. Our script writing team is now adding the final touches to your text. The next step is the production of the episode. There is no need to tell you that the we are looking forward to see the results on the screen.

As I have mentioned in my previous letter the goal of our television series *Impacts!* is that the viewers develop a comprehensive understanding of the Impacts of Technology on Society. Since your participation in the episode on the food production was such a success, we have decided to grant you another contract. As a group, we have decided that the next topic for *Impacts!* should be "the labelling of genetically modified foods". Again many issues and solutions could be analyzed. We will leave the choice up to you. We are convinced that you will find an interesting and challenging way to address this fascinating topic.

The terms of reference for the contract are exactly the same as the ones outlined for the episode on food production, that is, there is two parts to your mandate. The second part is still very important since as you know, I will have a meeting with the Vice-President Programming. He will still ask questions that will challenge the rationale for the episode. I have another breakfast meeting with the Vice-President on Thursday morning February 15. So please, send the GM episode via WebCT to Ms. Légaré, my executive assistant, no later than midnight on February 14. She will diligently forward your attached file to me.

If you have any questions, please be assured Mr., Mrs. So-and-So, that I will be more that delighted to be of any help.

Sincerely,

Art Bukowski

Project Manager and

Producer of the television series: *Impacts!*

Seen-Us? TV Network

GM Food File

Excerpts	Online Articles	Web Sites	Graphs & Tables

Excerpts

"The fuss over genetically modified food " Leanne Hachey. http://cbc.ca/news/indepth/foodfight/hachey.html (last access : 02-02-2001)

- Rift between Europe and America, also called the " Atlantic Divide " : " Whatever the case, the Atlantic divide can be boiled down to two things. The first is all about experience : the side of the Atlantic hasn't seen a scare comparable to Mad Cow disease. The second is all about dollars : North Americans expect their food to be cheap".

  Agriculture and Agri-Food Canada , the agri-food industry (farmers, suppliers, grocers and restaurant workers) : $91 billion in domestic and retail sales in 1997, accounting for 8.5% of GDP.

" Our Food : Food Labelling " BIOTECanada. From : http://www.biotech.ca/EN/agri_labelling.html (last access: 02-02-2001)

- " In the next generation of foods containing genetically modified ingredients, consumers will see benefits affecting them directly, such as enhanced nutrition, reduced allergens, improved taste and extended freshness. If these changes include health and safety concerns, or significant changes in composition or nutrition, they will be labelled as such. "

  Emerging benefits of biotechnology (partial list) : Biosensors to warn of dangerous bacteria in spoiled foods; Rice with enhanced vitamin A and iron to reduce child blindness in developing countries; Foods used to deliver vaccines; Crops used to inexpensively produce health care products " etc.

" Labelling of Genetically Engineered Foods in Canada ". Canadian Food Inspection Agency : http://www.cfia-acia.agr.ca/english/toc/labetie.shtml (last access : 02-02-2001)

- " Health Canada is responsible for setting food labelling policies with respect to health and safety matters (i.e. nutritional content, allergens, special dietary needs etc.). This applies to all foods, including foods that have been derived through biotechnology. "

  " The governmentís examination of its policy on the labelling of foods from biotechno,logy is under development. " (...) Based on three consultations (1993) a " set of guidelines were developed. They reflect a general consensus to :

  - require mandatory labelling if there is a health or safety concern, i.e. from allergens or significant nutrient or compositional change (these decisions will be made by Health Canada)
  - ensure labelling is understandable, truthful and not misleading
  - permit voluntary positive labelling on the condition that the claim is not misleading or deceptive and the claim itself is factual.

- "Canada's major trading partners, including the United States, support labelling on a case-by-case basis only in instances of health, safety and compositional change. "

" Voluntary Labelling of Foods Derived From Biotechnology " . Canadian Food Inspection Agency : http://www.cfia-acia.agr.ca/english/toc/labetie.shtml (last access : 02-02-2001)

" The main legislation dealing with food labelling in Canada is in the Food and Drugs Act. This legislation prescribes what can be said on food labels so that consumers get essential information. The legislation allows voluntary labelling of all foods, including foods derived from biotechnology. "

" 31 Critical Questions in Agricultural Biotechnology " (March 2000). Answers compiled and edited by Barun Mitra, Andrew Appel and Gregory Conko, AgBioWorld. From : http://www.agbioworld.org/articles/critical.html (last access : 02-02-2001)

" Until the Green Revolution spread to South America and then to Asia, beginning about 40 years ago, the only way for developing world farmers to keep up with population growth was to convert forests, jungles and deserts into farmland. "

- " In the developed Western countries, advances such as hybridization, agricultural chemicals, and farm machinery have boosted production per acre of farmland to the point where it appears that the amount of food per acre has reached the limit of the ability of crop plants to convert sunlight to energy ".

" Genetic engineering is the latest in a number of strategies that have collectively been termed the " Green Revolution (...) "

" In many cases, biotechnology can certainly help farmers get higher yields from their land. If biotechnology is used to provide low cost solutions to improve village farms, then it can help to address world poverty ".

" Somewhat related to the question of the quantity of food is its quality; that is, whether it delivers the vitamins and minerals required to maintain human health. Here, too, genetic engineering can help. Recently, rice has been developed with added beta carotene (which is converted into vitamin A in the human body) and increased iron level ".

" It's important to note, however, that given current estimates, the farmers around the world only produce enough to feed the current population at such levels. "

" A complex approach to the question requires us to look at increasing food production without changing distribution channels. Producing more grains worldwide will not in itself increase the availability of grains to underdeveloped nations ".

" Summary of Public Opinion Surveys Related to Labeling of Genetically Engineered Foods " Consumer Policy Institute/Consumers Union. Consumer Policy Institute: http://www.consumersunion.org/aboutcu/offices/CPI.htm (last access : 02-03-2001)

US Surveys - Genetically Engineered Food

- CNN Poll - May 20, 1999

CNN conducted an on-line, interactive quickvote on May 20, 1999, following a report in the journal

Nature that genetically engineered corn posed a serious threat to the survival of the monarch butterfly. The poll question posed by CNN was "A study found bio-engineered corn can harm butterflies; should such crops be put on hold pending more study?" As of May 24, 1999, of 26,179 responses, 75% or 19,543 people agreed there should be a hold on GE crops until further study had been conducted, 25% or 6,636 people disagreed.

Time Magazine- January 11, 1999

In "Brave New Farm," by James Walsh, reporting on opposition to GEFs, a box titled "What People Think" reported responses to two questions: "Should genetically engineered food be labeled as such? 82% yes, 14% no. If food were labeled as genetically engineered, would you buy if for yourself or your family? 28% yes, 58% no." Sample size was not indicated.

Novartis, Inc. - February 24, 1997

Novartis, Inc., one of the five large international companies that is developing genetically engineered seed, conducted a survey of American consumers which was released on February 24, 1997 at the first International Conference on Regulation of Crop Protection and its Implication for the Food Supply, sponsored by Tufts University and the Foundation for Nutritional Advancement.

According to a press release on the results of the survey, "Most Americans want foods that are genetically altered to be clearly identified with labels. 93% of Americans who responded to a recent survey by the world's largest agribusiness company agree that bioengineered food should be labeled as such, including 73% who strongly agree with the position." Sample size was not indicated.

Canada Survey - Genetically Engineered Food

Toronto Star Poll, "Public Prefers Genetically Modified Food To Be Clearly Labeled" June 2, 1998.

In a poll conducted by the Toronto Star the issue posed was, "In Canada, genetically engineered

food must be labeled only when it changes the nutritional value or could pose a health risk to some people. Should all genetically engineered foods be labeled?" Of those polled, 98% said yes, while 2% said no.

Australia Survey - Genetically Engineered Food

Australian Financial Review, "All Transgenic Food To Be Labeled", by C. Bolt, December 18, 1998.

In a story about how the Australia New Zealand Food Standards Council (Health Ministers) voted 6 to 4 to require mandatory labeling of all genetically engineered foods, the articles states "Surveys showed more than 80% of consumers want full labeling."

•

Johnson, G.L. and Thompson, P.B. (1991). " Ethics and values associated with agricultural biotechnology ". chapter 9. In Agricultural biotechnology : Issues and choices. Baumgardt, B.R. and Martin, M.A. (Eds.).West Lafayette, Indiana : Purdue University Agricultural Experiment Station, pp.121-137.

p. 124 : " Biotechnology as a form of playing God. (...)Biotechnology does not create life from dead matter; instead, it increases our ability to control the forms of already living things ".

God

•

Humanity Nature

CBC Food Fight Magazine. Food for thought " Fighting the fungus "
http://cbc.ca/news/indepth/foodfight/mag_thought4.html (last access :02-02-2001)

" Genetic engineers say that this method -- adding one gene -- is more precise than traditional plant breeding ñwhere great gobs of genes are moved indiscriminately. "

- Ann Clark, GE critic : " You as a genetic engineer, don't have any control if it (the gene) lands here or it lands here or it lands here. Order matters because genes donít act alone. They don't act in isolation. They interact with each other. One of the big problem is the point of insertion is random. We don't know on which chromosome it is going to land . "

"Guidelines for the safety assessment of novel foods, Vol. I : Preamble and guidance scheme for notification". Food Directorate, Health Protection Branch, Health Canada, September 1994.

Novel foods may include (p.4):

- Products and processes that have previously not been used before as food or to process food in Canada
- Food containing microorganisms that have not previously been used as food or to process food
- Foods that result from genetic modification and exhibit new or modified characteristics that have previously not been identified in those foods, or that result from production by organisms exhibiting such new or modified characteristics
- Food that is modified from the traditional product or is produced by a process that has been modified from the traditional process.

Substantial equivalence (p.7) : " The idea that existing organisms used as food or as a source of food can be used as the basis of comparison when assessing the safety of the human consumption of a food or food component that has been modified or is new. If one considers a modified traditional food about which there is extensive knowledge on the range of possible toxicants, critical nutrients or other relevant characteristics, the new product can be compared with the old in simple ways ".

"Guidelines for the safety assessment of novel foods, Vol. II : Genetically modified Microorganisms and plants". Food Directorate, Health Protection Branch, Health Canada, September 1994.

" The safety assessment of whole foods derived from genetically modified microorganisms, plants and animals is more complex than evaluation of single chemical food constituents or defined chemical mixtures. In assessing the safety of whole foods, knowledge of the previous use as food, the level of complexity of the whole food, and the breadth of the modification will be determining factors in establishing information requirements for the evaluation. Where appropriate, the basis of these safety assessments will be comparison of the molecular, compositional, toxicological and nutritional data for the modified organism to those of its traditional counterpart. "

-

Krebs John (2000). GM Food safety : Facts, uncertainties , and assessment. Chairmanís report. The OECD Edimburgh Conference on the Scientific and Health Aspects of Genetically Modified Foods, 28 February-1 March, 2000. http://www.oecd.org/ (last access : 02-02-2001)

Principal Conclusions :

- #8 Worldwide, many people are eating GM foods (especially in North America and China) with no adverse affects on human health having been reported in the peer-reviewed scientific literature.

  #9 There could, in theory, be long term effects on human health that have not yet been detected because GM foods have been available for less than ten years.

  #12 Consumers should be allowed to choose. Labelling of GM foods is important, although there was no agreement on how far this should extend. It is important also to note that the labelling applies to the process by which organisms are created and not to the food product, which in many cases is identical to its conventional counterpart.

  #13 The assessment of the safety of any novel food, including GM food, involves a variety of kinds of evidence. Our commonly used tool is the concept of " substantial equivalence " . The essence of this idea is that a comparison between the novel food and one already in the diet provides the basis for asking questions about the safety of the novel product. Substantial equivalence is not a quantitative criterion or a hurdle, but a framework for thinking. It is continually modified and updated, but it is timely now, after six years of using the tool, to undertake a more detailed review.

  #15 (...) The precautionary principle is now beginning to be discused internationally in relation to food safety, but it has not yet been translated into an agreed operational form.

  #16 The majority of speakers from developing countries stressed the crucial importance of GM technology as part of the armoury for feeding their population in the future (...). However, the view was also expressed that the future application of GM technology in developing countries should be more explicitly tuned to the needs of local people rather than of multinational corporations.

  #19 The first generation of GM crops and foods are perceived as having brought little direct benefit to consumers in developed countries, but this may well change as new products appear with direct quality, health or price benefits.

Goldburg, R. and Greenlee, W.F. (1993). Technical risk assessment and regulations, workshop report. Agricultural biotehnology : A public conversation about risk, NABC Report 5. Ithaca, NY : National Agricultural Biotechnology Council, pp.21-23.

" Products of agricultural biotechnology, such as field tests of genetically engineered crops or foods derived from genetically engineered crops, may pose risks to ecosystems or human health. However, the traditional risk assessment paradigm, developed to assess the carcinogenicity of chemicals, is not easily applied to products of agricultural biotechnology products. "

" Tools (appropriate test systems) should be developed to evaluate the potential hazards of three classes of organisms : Animals, plants and microorganisms. Scientific guidelines need to be developed to ensure that any ecological or human and animal health risks or agricultural biotechnology products are adequately addressed :

Animals :

- containment/ecological releases
- human safety or expressed products
- unforeseen metabolic effects

Plants :

- containment/ecological releases : Aletered disease/insect susceptibility, weediness, and outcrossing
- human safety or expressed products
- unforeseen metabolic effects

Microorganisms:

- containment/ecological releases : colonization, pathogenicity/toxicity to nontarget organisms, and frequency and impact of gene transfer to other microbial species.
- unforeseen metabolic effects

Barton, K.A . (1996). Biotechnology : Catalyst for change in agriculture. *Agricultural biotechnology : Novel Products and new partnerships*, NABC Report 8. Ithaca, NY : National Agricultural Biotechnology Council, pp.47-58.

> " The agricultural industry has historically been comprised of distinct segments including service providers, growers, processors, distributors, and consumer outlets. We have generally relied on the passage of commodity materials from the farm, through processors, and on to consumers in distinct steps that capture increasing incremental value as the refined agricultural materials approach the end-user. A modern trend toward larger farming operations and vertical integration, combining processing and distribution, has been driven to date by economies of scale and other efficiencies provided by channeled flow of materials through a controlled development pathway. However, as biotechnology adds increasing value directly to germplasm, it will catalyze a series of changes in industry value-capture paradigms, which will lead to a more comprehensive transition toward identity preservation of crops from seed to consumer. " (p.53)

- " Biotechnology will contribute to the streamlining of product development cycles, or " cycle time ", in an industry where advancement has been generally slow and methodical. Plant and animal traits, which in the past were objectives of classical breeding programs, are already being generated in considerably shorter time frames through the use of gene mapping and tracking. " (p.54)

> " Through classical breeding, the average strength of cotton fiber across the industry has been increasing at a relatively constant rate of 1.5 percent per year, with a cumulative strength increase of 16 percent achieved from 1980 to 1991. However, we are now able to use genetic engineering to dramatically increase the strength of cotton fiber with the addition of genes from other organisms. In our recent research project, the strength of the major upland cotton variety was increased by more than 60 percent with a single transferred gene. This represent a strength enhancement equivalent to 30 years of classical plant breeding in a fraction of that time, and the fiber strength achieved now exceeds the current premium system for fiber strength. " (p.55)

> " Biotechnology represents a powerful new tool for plant and animal breeding, and the application of this technology will lead to new products and new uses for agriculture. This technology enables very rapid product development. The increased speed of accomplishing genetic improvements also condenses product development expense in much shorter time frames. These issues will necessitate changes in industry practice to ensure that appropriate value can be captured for the increased contribution of genetics as a component of final product value. " (p.57).

Goldburg, R. (1996). Novel crops and other transgenics : How green are they? *Agricultural biotechnology : Novel Products and new partnerships*, NABC Report 8. Ithaca, NY : National Agricultural Biotechnology Council, pp.59-66.

" Biotechnology products frequently are touted as environmentally beneficial. Even in the face of such claims, however, it is unclear to many environmentalists whether biotechnology products will be good for the environment. In part, this is simply because of the increasingly varied nature of biotechnology products, which range from industrial chemicals to transgenic animals. It is unreasonable to expect that as a class, biotechnology products will be " good " or " bad " for the environment. " (p.60)

- " Within the predominant group of engineered organismsóplantsóthe range of species being field-tested until recently has been relatively narrow. As of March 1993, about 85 percent of field-tests of genetically engineered crops were of six species : corn, cotton, tomatoes, potatoes, soybeans, and tobacco. However, the diversity of crops being genetically engineered has increased substantially : Transgenic varieties of more than 40 different crop plants have now been field-tested in the United States. " (p.61)

  " The predominant ecological risk associated with crop plants is that they will transfer, via pollination, their acquired genes to related wild or weedy plants, or to other cultivated non-transgenic varieties of the same crop. How will these gene transfer risks apply to the wider range of crop varieties under development? " (p.62)

Kneen, B. (1996). Constructing food for shareholder value. *Agricultural biotechnology : Novel Products and new partnerships*, NABC Report 8. Ithaca, NY : National Agricultural Biotechnology Council, pp.127-136.

One way to frame the discussion : " Food industry " is a cultural expression, one way of seeing the world. " This is the way the wealthy industrialized societies of the North view the world out of their historically and culturally distinct experience. But what about other perspectives and experiences? " (p.127)

- Example of the V-chip and TV sex and violence; the issue is " reduced to a matter of individual parental option (...) it seems to me that the labeling of genetically engineered foods is being approached in much the same way (...) There appear to be no questions raised either about the production of violence on TV or about the production of genetically engineered food. In the case of food, the biotechnology industry promotes the position that it would be absurd, unworkable, and costly to label genetically engineered food. In a 1993 interview, an FDA spokesperson said : " It would not be merely a matter of putting a sticker on a tomato or a banana. Producers would have to segregate the genetically engineered foods from other varieties. Does the label have to follow the food processing chain? It would increase the cost of these foods to consumers and would disrupt our complex food distribution system ." "(p.128)

GM Food File

Excerpts	Online Articles	Web Sites	Graphs & Tables

Online Articles

" Fuelling the Fire : How the Net as Driven the GM Food Debate " just-food.com editorial team, 01-01-2001. From : http://just-food.com/index.asp?c=1 (last access : 10-01-2001)

New Scientist, Genetically modified world. " Unpalathable truths. Demanding proof that genetically modified foods are

safe is all very well, but without a rational system for tetsing conventional foods, we may never get it ". http://www.newscientist.com/nsplus/insight/gmworld/gmfood/gmfood.html (last access : 02-02-2001).

" Focus on National Food Standards Authorities -- Part III : The Canadian Food Inspection Agency ". just-food.com editorial team, 13-12-2000. From : http://just-food.com/index.asp?c=1 (last access : 10-01-2001)

" Percy versus Monsanto ". Martin OíMalley, CBC News Online. From : http://cbc.ca/news/indepth/onthefarm/onthefarm_omalley_index.html (last access : 02-02-2001)

More on Monsanto: http://www.atomz.com/search/?sp-s=1&sp-a=000611d2-sp00000000&sp-w-control=0&sp-date-range=-1&sp-p=any&sp-w=soundalike&sp-k=News&sp-q=Monsanto

Health Canada (1998). *Nutrition Labelling : List of core nutrients.* http://www.hc-sc.gc.ca/food-aliment/english/subjects/food_labelling_and_claims/list_of_nutrients_for_nutritio.html (last access: 02-02-2001)

Greenpeace. " Genetically engineered food : Still Unlabeled and untested. " Genetic engineering campain, Background information. http://www.greenpeaceusa.org/ge/ (PDF Document, last access : 02-02-2001)

GM Food File

Excerpts	Online Articles	Web Sites	Graphs & Tables

Web sites

Nutrition labelling in Canada, NIN and Health Canada : http://www.hc-sc.gc.ca/food-aliment/english/subjects/food_labelling_and_claims/nutrition_labelling_and_nutrie.html

National Institute of Nutrition, " National Forum provides information on GM food labelling issues (press release): http://www.nin.ca/en/index-nin.html

CBC Marketplace : " Labelling genetically modified foods " : http://cbc.ca/consumers/market/files/food/gmfood/cbc.html

Health Canada : Novel Food Information -- Food biotechnology : http://www.hc-sc.gc.ca/food-aliment/

Regulation on biotechnology in Canada : http://www.cfia-acia.agr.ca/english/ppc/biotech/bioteche.shtml

The GMO Crop (mis)Information Page : http://www.cropsoil.uga.edu/~parrottlab/forum.htm

Union of concerned scientists : http://www.ucsusa.org/agriculture/index.html

The alliance for Better foods :http://www.betterfoods.org/

Mothers for Natural Law : http://www.safe-food.org/

The Food Marketing Institute : http://www.fmi.org/

International council of food professionals : http://maxpages.com/iafsp

Natural Foods merchandizer: http://www.nfm-online.com/nfm_backs/Dec_98/monsanto.html

American Dietetic Association : http://www.eatright.org/

Institute for Scientific Information: http://www.isinet.com

Prince of Wales on GM Foods : http://www.princeofwales.gov.uk/speeches/agriculture_01061999.html

Le monde diplomatique -- Food Industry -- Mad Scientists :
http://www.transnationale.org/anglais/sources/alimentation/processed_scientists.html

Carthagena Protocol (Biosafety convention) : http://www.ukabc.org/cartagena.htm http://www.biodiv.org/

Miami group http://www.google.com/search?q=Miami+group+labelling&hl=en&lr=&safe=off

---

GM Food File

Excerpts	Online Articles	Web Sites	Graphs & Tables

Graphs and Tables

Monsanto Company. Sustainability -- Seeds for thought. Learning is discover, Issue number 2.
From :http://www.monsanto.com/monsanto/biotechnology/teaching_science/default.htm (PDF Document, last access : 02-02-2001)

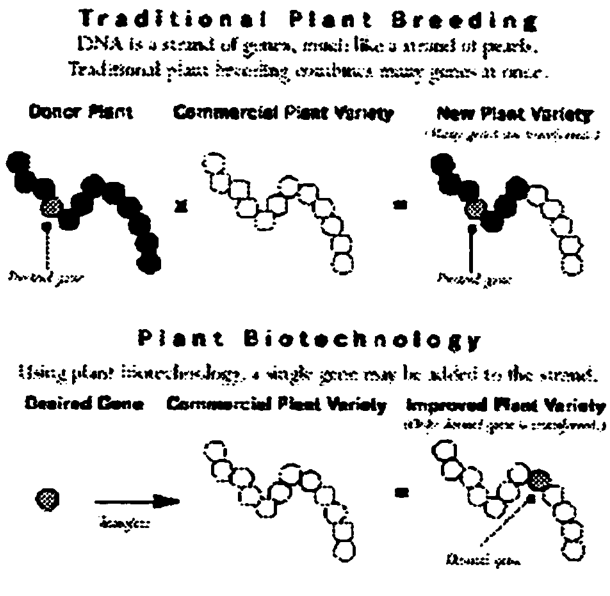


Figure :, Comparison of Traditional Plant Crossing with Biotechnology

Goklany, I.M. (August 2000). " Applying the precautionary principle to genetically modified crops ". Policy Study Number 157, Center for the Study of American Business, St.Louis : Washington Univesrity in St.Louis. http://www.agbioworld.org/ (PDF Document, last access : 28-01-2001)

Table 1 . Current extent of Public Health Problems partly or wholly caused by insufficient food or poor nutrition, and the likelihood that they could be alleviated using GM rather than conventional crops.

Problem	Current Extent (Year)	Likelihood that GM crops would reduce problem
Undernourishment	825 million people (1994-96)	Very high
Malnutrition	6.6 million deaths per year in children under 5 years old (1995)	Very high
Stunting	200 million people (1995)	High
Iron-deficiency anemia	2,000 million people (1995)	High
Vitamin A deficiency	250 million people (1995)	High
Ischemic and cerebrovascular diseases	2.8 million deaths per year in high-income countries (1998) 9.7 million deaths per year in low/mid-income countries (1998) (includes those due to smoking)	Moderate
Cancers	2.0 million deaths per year in high-income countries (1998) 5.2 million deaths per year in low/mid-income countries (1998) (includes those due to smoking)	Moderate

Sources : World Health Organization, The World Health Report 1999 (Geneva : World Health Organization, 1999); Food and Agriculture Organization, FAO Databases, at apps.fao.org; id., " The State of Food Insecurity in the World, " at www.fao.org/FOCUS/E/SOF1/home-e.hm, 1999.

National Institute of Nutrition. *Nutrition labelling : Perceptions and preferences of Canadians.* National Institute of Nutrition, June 1999. From : http://www.hc-sc.gc.ca/food-aliment/english/subjects/food_labelling_and_claims/consumer_research_final_report/consumer_research_final_report.html

Figure 2 – Importance of Nutrition Information on Food Packages

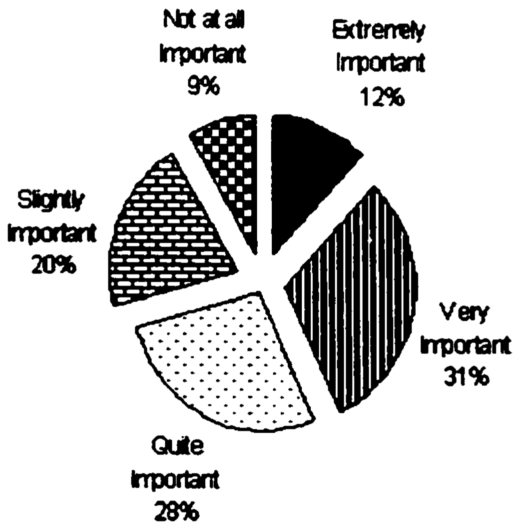


Table 3 . Disliked nutrition information on food labels

	Percent of Canadians (N=378)
Nothing disliked	38
Net : difficult to understand	17
- Difficult to understand	8
- Some words are hard to understand	7
- Scientific words are hard to understand	2
Net : missing/not enough information	11
- Not enough information	8
- Some products are not labelled	2
Net : Deceptive/misleading	8
Net : Difficult to read	9
- Print is too small	7
- Difficult to read	3
Net : Serving sizes are different or not uniform	3

Figure 3 -- Frequency of Use of the Nutrition Information Panel

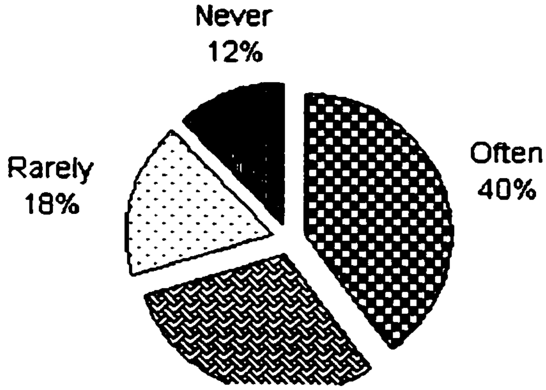
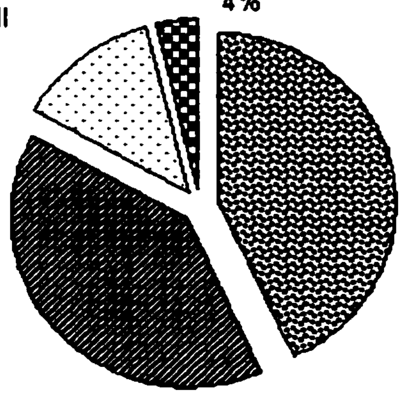


Table 4. Uses of the nutrition information panel

	Percent of Candians who use the panel **Often** or **Sometimes** (n=786)	
	Often	Often or Sometimes
To see how high or low a food is in nutrients like fat or sodium	58	87
To see how high or low a food is in nutrients like fibre, vitamins or minerals	46	83
To get a general idea of the calorie content of a food	45	78
To compare similar types of foods with each other	40	76
To compare different types of foods with each other	39	74
To see if something in the adverstising/on the package is true	33	65
To figure out how much of a food product you should eat	23	54

Figure 4 -- Level of Understanding of the Nutrition Information Panel



Harlander, S.K., BeMiller, J.N., and Steenson, L. (1991). " Impact of biotechnology on food and nonfood uses of agricultural products ", chapter 3. In *Agricultural biotechnology : Issues and choices*. Baumgardt, B.R. and Martin, M.A. (Eds.).West Lafayette, Indiana : Purdue University Agricultural Experiment Station, pp.41-52.

Table 3. Genetic improvement of microorganisms

Type of fermentation	Nature of Improvement	Benefit
Dairy	Bacteriophage (virus) resistance	Elimininate economic losses due to destruction by virus infection
	Accelerated ripening of cheese	Decreased storage cost
	Higher levels of the enzyme beta-gaactosidase	More digestible for lactose-intolerant individuals
Meat	Bacteriocin production (natural preservative)	Inhibition of pathogenic or spoliage organisms
	Addition of cholesterole-reducing enzymes	Reduction of an undesirable dietary component
	Addition of fat-modifying enzymes	Alteration of the saturated to unsaturated fat ratio
Beer	Alpha-amylase production	Production of " lite " or low-calorie beer

---

Doyle, M.P. and Marth, E.H. (1991). " Food safety issues in biotechnology ". chapter 4. In *Agricultural biotechnology : Issues and choices*. Baumgardt, B.R. and Martin, M.A. (Eds.).West Lafayette, Indiana : Purdue University Agricultural Experiment Station, pp.55-65.

> " The most common reasons for lack of (food) safety are : a) foodborne illnesses or microbial origin, b) environmental contaminants in food, c) toxicants naturally present in food, d) additives, and e) specific nutritional and dietary problems of certain consumers for whom some foods are never safe ". (p. 55).

Table 5. Environmental contaminants and natural toxicants that can cause unsafe food.(p.60)

Environmental cotaminants	Examples of foods involved
Heavy metals (Hg, Pb, Cd, As)	Numerous plant products
Polychlorinated bephenyls	Fish
Polybrominated bephenyls	Rare
Pesticides	Numerous plant products
Herbicides (dioxin)	Numerous plant products
Fungicides	Numerous plant products
Radioactivity	Plant products, milk
Shellfish poisons	Shellfish
Natural toxicants	

Oxalates	Rhubarb
Glycoalkaloids	Potatoes and other plant products
Cyanide compounds	Lima beans (some varieties)
Mushroom toxins	Mushrooms
Vasoactive amines	Cheese, wine, certain fish
Psychoactive substances	Certain spices
Allergens	Milk, wheat, fish
Mycotoxins	Numerous plant and animal products

Lacy, W.B., Busch, L. and Lacy, L.R. (1991). " Public perceptions of agricultural biotechnology ". Chapetr 10. In *Agricultural biotechnology : Issues and choices*. Baumgardt, B.R. and Martin, M.A. (Eds.).West Lafayette, Indiana : Purdue University Agricultural Experiment Station, pp.139-161.

Figure 2. Opinions regarding whether particular scientific developments will make life better (p.149).

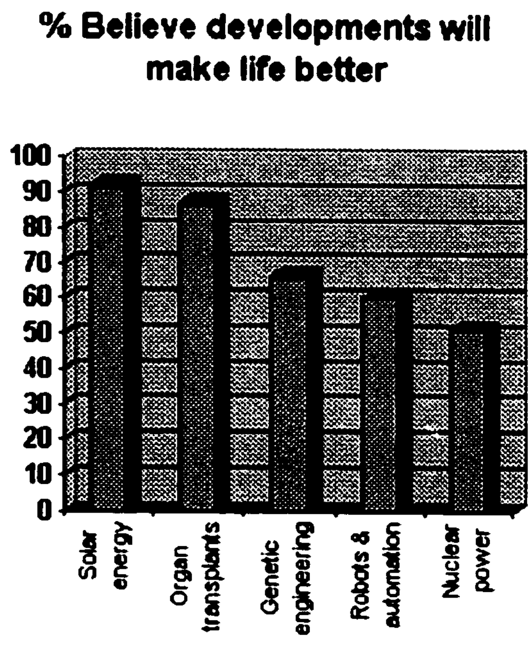
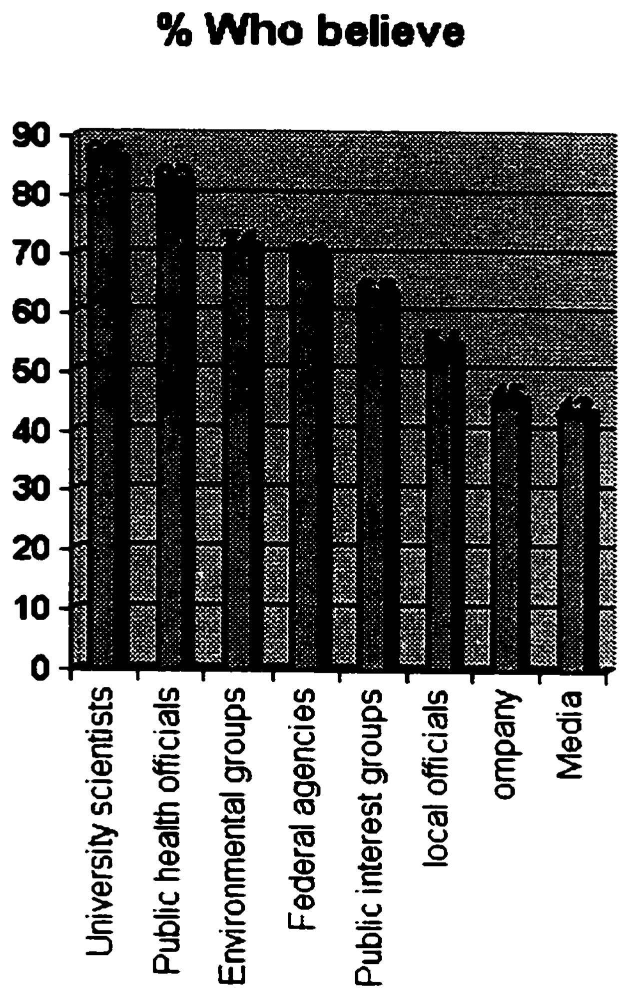


Figure 6. Credibility of sources of information on risks of genetically engineeried products (p.157).


% Who believe
90
80
70
60
50
40
30
20
10
0
University scientists
Public health officials
Environmental groups
Federal agencies
Public interest groups
local officials
ompany
Media

APPENDIX XVI CODER'S TRAINING MANUAL

Coder's Training Manual

An investigation of the effect of instructional strategy on the development of critical thinking skills with engineering students.

Doctoral research project

In partial fulfillment for the Ph.D. in Educational Technology

Geneviève Légaré

Concordia University

May 2001

TABLE OF CONTENT

Training 4: Final definitions

(See section "Content: Coder's Training Manual" for more details)

Appendices

- Original theoritical framework
- List of essay questions (300 series)
- Task instructions (5)

INTRODUCTION

The object of the research project is to study the effect of instructional strategy on critical thinking skills. The research design entails a content analysis of student essays, using a modified version of Herrington and Oliver's model of critical thinking skills. Each unit of analysis (one sentence) will be coded using their model.

One of the most important challenge of this type of investigation is the issue or reliability. The data used in content analyses are subject to interpretation. In this study, three people, including the researcher, will code the data. The three coders will take part in a training session. The goal of the training is to reach an inter-rater reliability rate of 80% and above.

Getting familiar with the C.T. skills coding book and becoming comfortable with the data takes time and practice. This is why the training has been designed as a series of debriefing meetings and three individual practice sessions. The training session should last a total of three non-consecutive days. However, the length of the training session might be adjusted depending on the team's "performance" on the coding task.

The three practice sessions are designed to become familiar not only with the tool and the type of data, but also to get acquainted with the researcher's own understanding of the C.T. skills model. While practicing, the coders will apply the coding procedures described in this manual and test their working methods. In addition, the coders will take notes during the training process. Their reactions and suggestions will be discussed at the next debriefing meeting.

Coder's reactions, questions and suggestions will be discussed during the debriefing meetings. Although the main purpose of the training is to calibrate the definitions of C.T. skills, the coders are expected to bring in their input and to participate in the fine-tuning of the definitions (see "Reviewed definitions" tool). In the final meeting, it is expected that the fine-tuning of the codes will be completed and that the team will have reached an agreement as to how to deal with the challenges of interpretation. The reviewed list of definitions will be used to complement the coding book. Technical, administrative and logistic details related to the task will also be dicussed at debriefing meetings.

Finally, a reliability checker will conduct occasional observations in order to palliate for reliability decay. The person will also take part in the training session.

CONTENT: CODER'S TRAINING MANUAL

To facilitate the practice sessions and the understanding of the process, the training manual is organized into thematic sections. The sections are described below.

Tool section

The "tool section" contains the material that one needs while coding data. Coders should always have the tools with them when working on the data:

- C.T. skills Coding book (in plastic cover)
- Notebook
- List of examples

Guidelines and procedures section

This section is the core information of the project. It contains a general description of the coding task, all the procedures (coding and note taking), tips on working methods, explanation of the *Nud*ist* environment and so on.

Training sections

All three training sessions are organized more or less in the same way. The meeting agenda is presented first, followed by the steps for the practice session. The third component is a sheet called "Reviewed definitions" which will be used during meetings as a tool to calibrate the definitions. Whenever necessary, supplementary materials will be placed as the fourth item. Fifth, you will find the set of data to be used in the practice session. The coders may write directly on the materials. Finally, in Training 1 and 2, there is a coded version of the same set of data. The purpose of the annotated essays is for the trainee to compare their coding with the researcher's. Discrepencies will be discussed at the next debriefing session.

Training 1: Introduction and examples

Content:

- Meeting's agenda
- Practice steps
- "Reviewed definitions" sheet
- Set of data (N = 5)
- Set of annotated data

Training 2: Initial reactions and challenges

Content:

- Meeting's agenda
- Practice steps
- "Reviewed definitions" sheet
- List of challenges of interpretation
- Set of data (N = 5)
- Set of annotated data

***Training 3: Fine tuning and use of* Nud*ist**

Content:

- Meeting's agenda
- Practice steps
- "Reviewed definitions" sheet
- Set of data: "Target" (N = 5)

Training 4: Final definitions

Content:

- Meeting's agenda
- "Reviewed definitions" sheet

Appendices

Additional information is presented in this section. The documents will help in the understanding of the broader context of the research project. There are three different documents included:

- the theoritical framework by Herrington and Oliver. The coders might refer to it to see how the coding book has evolved from the original version;
- a list of essay questions distributed by the instructor. This document will be useful to code the essays belonging to the 300 series;
- a description of the student's tasks (Task 1 to 5). During the coding task, you may need to refer occasionnally to the data provided in those documents, especially when "visuals" are used.

DESCRIPTION OF THE CODING PROJECT

- Main task: Code approximately 170 essays using the critical thinking coding book;
- The task will be done using the computer program *Nud*ist* (Non-numerical Unstructured Data Indexing, Searching and Theorizing;
- All units (one sentence) must be coded, including the non-critical thinking incidences;
- Achieve 80% inter-rater reliability rate or above: The first round (N = 40) will constitute the pool to estimate the percentage of inter-rater agreement;
- Goal: All tasks coded by June 20th.

STEPS OF THE CODING PROJECT

There are two general steps of the coding project: training and the coding itself. The goal of the training is to calibrate all raters in order to achieve 80% inter-rater reliability rate.

<u>Training</u>

Training 1: Introduction and examples

- Introductory meeting (1.5 h)
- Practice with examples (Goal: 50-60% inter-rater reliability)

(Gen : Rating compilation of practice)

Training 2: Initial reactions and challenges

- Meeting: Initial reactions (1.5h)
- Practice challenges (Goal: 70% inter-rater reliability rate)

(Gen: Rating compilation Practice on challenges)

*Training 3: Fine tuning and use of Nud*ist*

- Meeting : Fine tuning definitions (1.5h)
- Practice: Use *Nud*ist* to reach target (Goal: Above 80% Inter-rater reliability rate)

(Gen: Rating compilation Practice on target)

Training 4: Final definitions

Coding Task

- First round (N = 40)
- Computing of inter-rater reliability
    - If the rate is above 80%, then the raters proceed with their own set.
    - If the rate is below 80%, there are two options, depending on the result:
        - trouble shooting and re-calibration of C.T. skills
        - re-training period and move to round 2 (2nd round N = 20) to re-evaluate inter-rater reliability.

CODING PROCEDURES

1. All units of analysis must be coded
2. A unit is one sentence
3. You do not have to assess the "degree" of the C.T. skill (high, medium, low). You have to assess its presence only.
    - There are two parts of coding: critical thinking and the use of visuals (see coding book)
    - Complex sentences where semi-colons (;) and colons (:) are used, constitute one unit of analysis
    - *There is one exception to the above point*: The student is not using the punctuation properly, that is two independent sentences should have been created.

    About the use of parentheses and brackets

    1) When inserted in a sentence, do not consider the content of parentheses as a separate unit of analysis.
    2) When a complete sentence is in parentheses consider it a unit in itself. Use the critical thinking categories and codes provided

    About quotes:

    1) When the student uses a direct quote (quotation marks in paragraph or indented) indicate it as such (code: "quote")
    2) When the student paraphrases or reformulates an author's work (referenced sentence), code it as a C.T. skill.
4. When facing a challenge about the C.T. codes, use one of the three reasoning strategies: 1) Question student's intention; 2) Contextualization, and; 2) Identify key words in unit.
5. When the student uses "visuals" such as graphs, tables etc., indicate the type of use (see code book).
6. A "visual" is any graphic representation of data. "Visuals" include tables, figures, charts, and histograms. Pictures and images do not have to be coded.
7. When a visual is used, two codes have to be identified (C.T. Skill and use of visual).
8. Suspend judgement: Do not try to identify patterns in the data.

NOTE TAKING PROCEDURES

- There are two types of notes:
  - "Memo" function in Nud*ist: Justification of decisions regarding codes (research purpose)
  - Note book: record emerging patterns, metacognitive events, reflections, ideas for your own projects, questions, refinement of definitions etc. (Metacognitive and research purposes)

5. Date all entries for both types of notes
6. The notebooks will be collected at the end of the coding tasks
9. Indicate exceptions and peculiar events in your notebook and/or in Nud*ist, use the "memo" feature

Communicate to researcher any problems encountered (technical and/or conceptual).

CODING SESSION

Before you start a coding session

- Review the coding book before starting
- Read at least one paragraph of an essay before coding
- Establish your goal for the session (number of tasks to be coded)
- Create a "clean slate" (mind set)
- Repeat your "mantra": Let the data speak!

During a coding session

- Suspend judgement
- Jot down emerging ideas in your notebook
- When you have to make an interpretation "call", document your coding decisions using Nud*ist "memo" function
- Keep the coding book at hand
- Review the codes regularly, especially the ones you don't encounter too often
- Do not review or change units that have already been coded
- Refer to reasoning strategies when facing a challenge
- (Save your work on a regular basis)

Closing a coding session

- Apply saving and storing procedures

- Communicate to researcher problems or issues as soon as they arise
- For the next coding session, identify one essay to be read first; select an essay that is short and relatively easy to code.
- Assess your day!
- Write a "To do list" for the next session, if necessary.

WORKING METHODS

- Work directly on the screen. You may, however, code the hard copies and then transfer the results in Nud*ist (Time consuming).

Motivation strategies

- Establish a number of essays to code for the day
- Observe the pile going down!
- Vary the type of essays: alternate essays (T-1, T-2), scenarios (T-4, T-5), case studies (T-3), questions (T-1 300 series).
- Vary the length of essays: Scenarios tend to be longer than essays
- Take a break every hour or so
- Establish a routine that is convenient for you
- Include extrinsic motivational agent in your routine

State of mind

- The main point is to let the data speak
- Suspend judgement. Do not try to identify patterns in the data
- Avoid "marking" papers and "make abstraction" of the writing styles
- Focus on the intention of the writer, not on what you would have done or how you would have done it.
- Clean slate: If you "catch" yourself thinking about patterns, record it in your notebook.
- Clean slate: If ideas are emerging, record them in your notebook
- Do not "self-correct" yourself (E.g.: "I don't see enough of this skill...")
- Trust your judgement
- When the decision is difficult to make, use a reasoning strategy: Intention of the writer, contextualization, identification of key words (see "challenges" section).
- Keep the same level of rigor, toughness, severity/leniency.
- Remember that weak writers may show some critical thinking.

NUD*IST PROCEDURES

- Saving procedures
- Storing procedures
- Categorizing
- Note taking procedures
- Other (creation of categories--varia)

ACCESS TO COMPUTER LAB

- Name on list--security (need student ID card)
- Unlimited access, including evenings and weekends
- Close and lock door when you go out

RELIABILITY CHECKER'S ROLE

- The reliability checker will participate in the training session
- The reliability checks will happen sporadically: she decides when to conduct a verification
- The role of the reliability checker is to watch for decay in the assessment of the raters. Decay might be defined as follows:
  a) Raters are not consistent with their assessment (fatigue):
     <u>Inconsistency is defined as either</u>:
     - Irregular leniency (irregular severity) or
     - Irregular application of a same code: the same code is sometimes applied, sometimes not
  b) Less frequent C.T. codes tend <u>not to be</u> identified when applicable; there is a tendency to apply more common codes.
- If the checker notices that there is evidence of decay in one or more rater, there are two possible options. Depending on the severity of the case, the checker might decide that the rater(s) have to either:
  a) go through a "refresher" session and review codes and reasoning processes
  b) re-train on the coding of the C.T. skills
- NOTE: The researcher and the reliability checker are independent of one another.

TRAINING 1: Introduction and examples

Introductory meeting (1.5 h)

- presentation of the training kit
- explanation of student assignments
- presentation and explanation of the critical thinking indicators
- explanation of the coding procedures
- explanation of note taking procedures
- explanation of reliability checker's role
- explanation of "reviewed definitions" sheet
- explanation Practice #1
- time sheets and payment
- set date and time next meeting
- varia

TRAINING 1: Introduction and examples

Practice: presentation of examples (Goal: 50-60% inter-rater reliability)

Purposes:

- get acquainted with the different levels and styles of writing
- get acquainted with different types of text organization or structure
- practice suspending judgement: writing skills vs. C.T. skills

Steps:

- among the three data sets provided, select at least one sample of each type of essays (T1, T2, T3, T4, T5)
- quickly read your sample without analyzing the texts. The purpose is to get acquainted with the levels and styles of writing
- review the coding book
- review the C.T. skills examples
- using Set #1, apply codes on hard copies
- compare your coding with the annotated version; take note on discrepancies and agreements
- Exercise for the debriefing meeting: With the annotated version of "DA_T1_...", analyze the reasoning behind the decisions (Set #1). Identify agreement and disagreement. Challenge the interpretation and support your arguments.
- (drop Set #1 in mail box)

Documenting the process

- take notes about the process. The questions will be addressed in the next training session.
- questions?
- need for clarification?
- disagreement?
- metacognition: take notes on your reflection process
- reflections on the content, writing styles, essay format
- difficulties encountered?
- may identify examples of problems encountered for the purpose of discussion
- while reading did you:
  - get annoyed with mistakes, writing style?
  - try to find out what you would do to improve the essay?
  - try to identify weaknesses in arguments?
  - try to assess accuracy of content?

G: Rating compilation Practice #1

TRAINING 2: Initial reactions and challenges

Meeting : Initial reactions (1.5 h)

- debriefing: initial reactions, impressions, questions, disagreements etc.
- calibration: exchange on definitions and categories
- specific example "DA..." : Discussion and clarification of researcher's understanding; agreement on definitions
- completion of "reviewed definitions" sheet
- adjustments on procedures
- overview of challenges section
- adjustment on number of tasks for practice #2
- discussion on issue of time/review of calendar
- discussion about motivation, concentration and other working strategies
- review of working and note taking procedures
- set date for next meeting
- varia

TRAINING 2: Initial reactions and challenges

Practice : Challenges of interpretation (Goal: 70% inter-rater reliability rate)

Purposes

- practice with essays which are more challenging and of various length than preceding set
- practice dealing with uncertainty (challenges of interpretation)
- use reasoning strategies (student's intention, contextualization, key words) to make a decision
- get more familiar with subtleties of the codes

Steps

- review coding procedures and working strategies
- read and code Set #2 using the coding book
- compare your coding with annotated version
- keep in mind: avoid looking for patterns; let the data speak!
- specific exercise: With the annotated version of "MC_T1_...", analyze the reasoning behind the decisions (Set #2). Identify agreement and disagreement. Challenge the interpretation and support your arguments.
- take notes on definitions to be "fine tuned" ("reviewed definitions" sheet)
- (Drop Set #2 in my mail box)

Documenting the process

- Take notes:
- identify similarities and differences between your assessment and the annotated version
- indicate disagreement and elements that you feel need to be discussed
- metacognition: focus on challenges of interpretation encountered
- assess your own understanding of the task
- see previous list for note taking

G: Rating compilation Practice #2

TRAINING 3: Fine tuning and use of *Nud*ist*

Meeting: Fine tuning (1.5 h)

- debriefing: challenges encountered
- review of the reasoning strategies
- specific exercise: Collaborative review of "MC_T1..." and discussion of the interpretation challenges.
- update of the "Reviewed definitions" sheet
- calibration: fine tuning the understanding of the indicators
- discussion on motivation and interest
- number of essays for practice #3 to be reviewed
- introduction to Nud*ist interface + procedures for saving documents
- presentation of note taking procedures in Nud*ist

TRAINING 3: Fine tuning and target

Practice: Use *Nud.ist* to reach target (Goal: Above 80% inter-rater reliability rate)

Purposes:
- get familiar with the *Nud.ist* environment
- focus on mind state (clean slate)
- "mentally appropriate" the updated version of the "reviewed definitions" sheet

Steps:
- review updated definition sheet
- review coding and working procedures
- code set #3 (Set #3 is considered to be a "Target" set. There are no annotated essays)
- keep in mind the challenges of interpretation

Documenting the process
- take note on problems with Nud*ist set up
- same as with preceding steps

G: Rating compilation Practice #3

TRAINING 4: Final definitions

Final Meeting: Definitions

- debriefing on Nud*ist environment
- review and improve working methods
- final calibration on C.T. skills: agreement on "reviewed definitions" sheet
- review calendar
- recall reliability checker's role
- presentation of data kits
- assignment of rater's code (ABC)
- explanation of the difference between data used in training and original data (see below)
- set date for first round results (overlap, N=40)

<u>About the codes of each essay</u>

To preserve confidentiality, the subjects' identity has been stripped from the essays. An identity code appears on each essay, in top left corner of the first page. For the coding task, you need only to pay attention to: a) the letter following the three first digit and b) the essay number.

<u>Example</u>: 141C_T3

a) In this example, this essay will be coded by rater "C".

b) "T3" is the type of essay. The information is useful when you want to determine which type of essay you are going to code (refer to section on working strategies: motivation).

All student information has been taken out of the text by the researcher. However, if you come across a student's name or identification number, please delete the information and advise the researcher.

<u>About the original data</u>

There are some differences between the data used for the training and the data used for the study. Here is what to expect:

- About T3: In the training kit, "T3" are essays, whereas in the original data they are case studies
- In the training kit, there are no examples of " 300 series --T1".
- About the topics: In the training kits, the topic for T2, T3 and T4 is "motorcycles", and "printing" for T5. In the original data, the topic is "food production" or "genetically modified foods" (Task 5).

Note: For ethical reasons, the data cannot be used for purposes other than this research project. Data cannot be reproduced without participants' permission. All copies of data, including the sources used in the training session, will be collected once the coding task is completed.

APPENDIX XVII : CODER'S INTERVIEW QUESTIONNAIRE

Name :____________________ Phone : ____________________

Academic background :____

Please answer all of the following questions. You may use point forms. Please use the other side of the sheet if needed.

1. What prior experience do you feel is relevant to this task?

2. What is <u>your</u> interest in the project? How is this project meaningful to you?

3. What do you think you will learn from this experience?

4. What aspect of the task do you anticipate as being the most challenging?

APPENDIX XVIII : INTER-RATER RELIABLITY CHECKER PROCEDURES

- Your function is to code a sample of tasks (N=51).
- Within that sample, you will code a randomized range of units (see Excel sheet : Reliability checker's *Journal*).
- The range of units is always 10.
- A reliability coefficient will be established between your coding and each coder's coding.

Sampling
- All three coders should be checked.
- Equal number of tasks for each coder (N=17)
- At least one type of task for each coder -- one of each T1 to T5 (see sampling grid)
- Do not select the same subject number twice (eg. 114)

Timing
- The sample of 51 should be spread over time (17 per round).
- The deadline for coding is currently July 31st. I have a one-week « buffer ».
- Three checks during the overall coding period: around July 20th, July 25th and August 1st.
- (I'll be working in my office, so nothing in the lab in my file)

Task selection
- To select a task, click "glegare shared group"
- Access the coder's journal
- Verify which tasks have been coded
- Make your selection following the sampling criteria.
- Indicate the filename in your journal
- Proceed with your coding
- Select tasks that have been "freshly" coded. For example, if it's your second round of checks, do not select a task that has been coded prior to your first round

Range of units : Random assignment
- It is important that you follow the randomization strategy.
- In the second column of your journal, there are four possibilities : 1st , 2nd, 3rd and random.
- « 1st », « 2nd » and « 3rd » refers to the portion of the task. In fact, you divide the total number of units within each task in 3.

  <u>For example</u> :
  - If « 2nd » is indicated, it means that you have to select your range of units within the « 2nd-third » of the paper.
  - If the paper has a total of 80 units, the three possible ranges will be approximately (1-25), (26-53), (54-80).
  - In our case, you will have to select 10 units within the 26-53 range.
  - On the example, you have selected units 26-36.
- When « Random » is indicated, you are free to choose the range of units.
- Always select your range of units based on the total number of units. Do not look at the text first.

NOTE : Some texts are shorter than others. In fact, many of them have less than 30 units, which would allow you to divide in « thirds ». A text then will not have a « third » section. In that case, start from the end of the text and code 10 units. Indicate the « glitch » in the « comments » column. Do not alter the random order.

Getting ready to code

- Since you are selecting the files randomly, all the documents have been imported in your Nud*ist project.
- Once you have selected the documents to code, you have to change the filename.
- When the pop-up window appears, add « h » at the end of the filename. For example the file « 114b_t3 » will now be called « 114b_t3h ». (This step is important for merging procedures.)

Coding

- Before you proceed with the coding, indicate the total number of units of the file.
- Select your range of units.
- This should be done « blindly ». Do not read the text a priori. Indicate the range of units first, then jump to those units while in Nud*ist.
- As the other coders, you may double code and use the memo functions whenever necessary (see coding procedures).

Final remarks

Before you start your rounds of checks, you should code at least one full task. The coding will not be included in your « results ».

Checklist

Use the following Checklist to track your sample. Indicate the file number in the cell.

	Check 1	Check 2	Check 3
Coder A	T1:	T1:	T1:
	T2:	T2:	T2:
	T3:	T3:	T3:
	T4:	T4:	T4:
	T5:	T5:	T5:
Coder B	T1:	T1:	T1:
	T2:	T2:	T2:
	T3:	T3:	T3:
	T4:	T4:	T4:
	T5:	T5:	T5:
Coder C	T1:	T1:	T1:
	T2:	T2:	T2:
	T3:	T3:	T3:
	T4:	T4:	T4:
	T5:	T5:	T5:

APPENDIX XIX : STEPS TO PREPARE DATA FOR NUD*IST

1. Added header in document containing the new filename (E.g. 108b_T1)
2. Added periods after titles and subtitles to create a single text units
3. Removed all graphics from the text and inserted appropriate message in header
4. Removed reference lists, footnotes and endnotes from text (useless text units)
5. Scanned the documens for student's names
6. Removed blank spaces which would have created an empty text units (created a macro in Word)
7. Corrected the punctuation to avoid creating meaningless text units. For example, replaced "U.S. Government" with "US Government" (created a macro in Word)
8. Verified the use of quotations marks to avoid creating meaningless text units. Replaced "___." with "___". (created a macro in Word)
9. Ran QSR Macro "Convert sentence into text units"
10. Ran QSR Macro "Format for Nud*ist"
11. Saved the document as a text file using the coding file name to replace the student's name.

APPENDIX XX CODING BOOK

Part 1: Critical Thinking Skills

C.T. Categories based on Herrington and Oliver's synthesis of higher order thinking.

CT Categories	**Any statement which ...**	**Key words**	**Nud*ist codes**
Judgement and interpretation (JI)	**Defines** terms in a way appropriate for the context	May use own terms or paraphrase a dictionary definition. Intention is to provide a common understanding of terms.	(1 1)
	Identifies **reasons** and **assumptions**	Establishes the rationale of the context. Explains the reasons why a statement is made. Establishes a common basis for understanding	(1 2)
	Makes connection between ideas and/or facts	Bridges two worlds or two ideas. Reaches "widely". Beyond a "toddler's" understanding.	(1 3)
	Evaluates, assesses ideas, facts, or statements	Student is weighing or pondering.	(1 4)
	Seeks **to support (to defend)** a position taken on an issue	Purpose may be either explanatory or illustrative. Uses examples or illustrations. In-depth.	(1 5)
Multiple perspectives **(MP)**	**Challenges** a conclusion or a previously made point	Goes beyond "idées reçues". Challenges conventional ideas. Position (or angle) is based on reason.	(2 1)
	Suggests an **alternative approach**; looks at **the other side** of an issue	Looks at the other side: Statement that there is another side.	(2 2 1)
		Suggests alternative approach: An elaboration about the other side. Discussion beyond mere statement. Looking for solution.	(2 2 2)
	Assumes a **questioner's** role and/ or considers the viewer's* perspective	Questioner's role: Uses a questioning strategy in the essay. Distancing from the role of "writer".	(2 3 1)
		Viewer's perspective: Considers the viewer's standpoint. Empathy and understanding.	(2 3 2)

CT Categories	Any statement which ...	Key words	Nud*ist codes
Imposing meaning (IM)	Recognizes that alternative approaches have different impacts (**consequences**)	Considers consequences or causes and effect. Somewhat evaluative but more complex than JI4.	(3 1)
	Offers a **prediction**, an **hypothesis** or a **recommendation**	Prediction or hypothesis: A projection in time	(3 2 1)
		Offers a recommendation or solution to a problem	(3 2 2)
	Summarizes point of views	May be in conclusion or at the end of a paragraph or a section. A shorter reformulation of facts and arguments.	(3 3)
	States a **conclusion**	Ends a paragraph a section or the essay.	(3 4)
	Generates **new** ideas and/or **novel** understanding (original thinking)	Demonstration of a novel understanding (original thinking). "Creates" a new idea. Light!	(3 5)
Metacognition	Expression of an awareness of thinking processes or understanding		(4)
Statements			
	Declarative or factual sentences	Def.: "To make something clear". E.g. Motorcycles is a means of transportation A list of facts. May use bullets. Might be "historical info." TIP: Student is probably using a source of information.	(5 1)
	Opinion (position taken on an issue)	Position taken on an issue. Based on reason.	(5 2)
	Belief (conviction)	Expression of a conviction. Statement not based on reason.	(5 3)
Varia	Not a sentence	Unit does not constitute a proper sentence (syntax or punctuation issues)	(6 1)
	Meaning: The sentence does not make sense	Impossible to extrapolate the sense.	(6 2)
	Organisational	All titles and subtitles	(6 3)
	Procedural sentences	Information given to the reader about the organisation of the text. E.g.: "In the next section...")	(6 4)
	Direct quotes	Citation in quotation marks.	(6 5)

*Please refer to task instructions.

Part 2: Use of "visuals"

A visual is any graphic representation of data: tables, graphs, figures, pie charts etc. When a student uses a visual, insert a new text unit in Nud*ist. Apply the following codes:

Level of analysis: The student...	**Nud*ist Codes**
Referred to data in the text but does not provide the numbers *(E.g.: As we can see in the "food file")*	(7 1)
Cut and pasted a "visual" in the text, but did not refer to it. *(Expl.: The visual is "just there").*	(7 2)
Used visuals in lieu of an explanation *(Expl.: Visuals are used to replace an argument)*	(7 3)
Used visuals to support an argument or to illustrate a point *(Expl.:Visuals is a proof of the argument; e.g.: "...., as shown in Table 1.")*	(7 4)
Analyzed the data *(Expl.: Discuss the data in the text; might focus on a partial set)*	(7 5)
Challenged/questioned the data *(Expl.: Effort in seeing what is not said; restructures the data; towards a novel understanding etc.)*	(7 6)

Appendix XXI Post-Coding Questionnaire

The goal of the questionnaire is to collect your input about the coding process while your reflections are still fresh in your mind. Your comments and suggestions will certainly be useful for the discussion section of the dissertation.
However, I may or may not refer directly to your comments in the dissertation. In any case, your identity will be preserved (I'll be using « Rater A, Rater B, Rater C »).

On another note, I would really like you to be as candid as possible. Please, do not try to spare my feelings or worse, try to « match » what I would like to hear. Your input is really valuable to me. If there are issues or points that I did not think off, feel free to add them at the end of the questionnaire.

Send the questionnaire by e-mail to my FirstClass address. Use your coder's name as a filename. For example, XXX would indicate « raterB_post_quest.doc ».

Again, I would like to thank you for participating in the coding of the data. Without you, this project would not have been possible. I really appreciate your dedication to the task and the fact that I could count on you for getting the job <u>well</u> done. I am convinced that one day, someone will return the favor to you!

Geneviève

I suggest that you read all the questions before completing them. Please provide a short answer for each of the questions.

If some reflections have been indicated in your notebook, there is no need to duplicate your comments. Simply indicate that you have already addressed the issue in your notes.

The questions are organized into four categories : the « coding process », « metacognition », « critical thinking skills » and « your input ». You may answer the questions in the order that is most convenient for you.

However, I came up with more questions that I had originally planned. Since I am a little bit last minute with the questionnaire –I apologize for this--and consequently imposing yet another tight deadline, I selected the questions that I feel should be answered before the debriefing session. In other words, the questions that are marked with an asterisk (*) indicates that you should answer them before the closing session.

CODING PROCESS

The purpose of this section is to gather information about the coding process in order to document what went on in practice.

Which type of task (T1, T2, etc.) did you like best? Did you like least? Why?

*To what extent did the length of the task affect your coding?

What is the best state of mind for coding?

*To what extent did the quality of the language or the writing style affect your coding?

What was your coding approach? For example, did you read a few paragraphs before coding? Did you start coding right away? Etc.

*What was the difference between coding on paper and coding directly in Nud*ist?

*Did the structure of the paper affect your coding? For example, long paragraphs, ill-structured paragraphs etc.

*What is the difference between coding a long task and a short task?

What did you do when you suspected plagiarism? Plagiarism : The student cut-and-pasted from a source, but did not put quotation marks. How did you know?

Is there any technical errors you tended to do as you were coding or entering the codes in Nud*ist? If so, which ones?

Anything I forgot to ask?

METACOGNITION

The purpose of this section is to « tap into » your thought processes as you were coding. Please, feel free to make connections or to add any additional thoughts that you feel I forgot.

How did you « keep yourself in check »? How did you « police » yourself? What were you telling yourself when you were facing a problem of interpretation?

*How did your coding evolve overtime?

*Did you feel inconsistent with some of the CT codes? With what codes?

*Were you, sometimes, thinking about what another coder would say in a given situation? Explain.

When did you become confident in yourself? What was the trigger?

Did you have a « breakthrough » or an « illumination » about the codes? If so, when did it happen?

*Did your attitude about a paper or a student's ideas change during the coding the paper?

Did your state of mind about coding change from the training period to round 1? And from Round 1 to the final round? How?

What were your motivational strategies?

Anything I forgot to ask?

THE C.T. CODES
For this section, the questions are quite general. Nevertheless, I would like you to be as precise as possible with your answers. Examples would be of great help.

*Is there any Critical Thinking skills you feel are missing from the instrument? If so, which ones?

*Is there any codes with which you still feel confused? Uncertain? If so, which ones?

*Is there codes that you feel are interchangeable? If so, which ones?

*Did you take some « liberty » with the interpretation of the codes? In other words, did your interpretation of a code change from the common definition we had established during the training session?

What improvements or changes would you suggest about the instrument?

*Which codes (or category of codes) did you find the easiest to identify? The most challenging to identify? Why?

Anything I forgot to ask?

YOUR INPUT : PROJECT IN GENERAL

What is your opinion about the unit of analysis?

If you were to conduct a similar project, what would do differently?

From a project management point of view, what skills should I improve? This is for my personal benefit!

*Now that you have completed the coding task, what type of results do you think I will get?

*Can you visualize emerging patterns in the data?

What are the difficulties you think I will face with the analysis of the results?

If you were to conduct a research project about critical thinking skills, how would you now approach or design it?

To what extent did you use the strategies outlined in the training?

What were the most useful aspects of the training session?

What would you do differently for the training session?

And finally, what do you retain from the experience?

Anything I forgot to ask?

Thank you for completing the questionnaire.
Geneviève