

## ABSTRACT

Title of Document: A MULTIDIMENSIONAL INVESTIGATION  
OF DEEP AND SURFACE PROCESSING

Daniel L. Dinsmore  
Doctor of Philosophy, 2011

Directed By: Professor Patricia A. Alexander, Department of  
Human Development

Within the psychological literature, there is the general assumption that deeper-level processing should be consistently associated with better academic performance (e.g., Block, 2009). However, studies exploring the relation between depth of processing (e.g., deep and surface processing) and learning outcomes have not produced consistent results. The argument guiding the current investigation was that these inconsistent findings regarding the relation between depth of processing and performance may be attributable to individual and situational factors that moderate this relation.

Specifically, in this study, the potential moderating effects of a situational factor (i.e., text type) and an individual factor (i.e., subject-matter knowledge) on the relation between depth of processing and performance were investigated. Support for these individual and situational factors were derived from a systematic review of the literature that also examined conceptualization and operationalization of deep and surface processing.

The participants for this study were 151 college undergraduates from a wide variety of majors. Participants completed measures of subject-matter

knowledge, read either an expository or persuasive text about the existence of extraterrestrials while thinking aloud, and then completed both a passage recall task and an open-ended task in which they were asked to justify their position on the existence of extraterrestrials. Participants' verbal reports were coded using a scheme based on Pressley and Afflerbach's (1995) *Verbal Protocols of Reading* and the Construction-Integration Model (Kintsch, 2004). The open-ended task was coded using Biggs and Collis's (1982) Structured Observation of the Learning Outcome (SOLO).

Three findings of interest emerged. First, results indicated that the relation between depth of processing and the open-ended tasks were moderated by the type of text (expository or persuasive) participants read. Second, no significant interaction of depth of processing and subject-matter knowledge on either the recall or open-ended learning emerged. Third, significant differences were found in the interaction of depth of processing and type of text between the passage recall measure and open-ended task. Plausible explanations for these findings and implications for future research and instructional practice are discussed.

A MULTIDIMENSIONAL INVESTIGATION OF DEEP AND SURFACE  
PROCESSING

by

Daniel L. Dinsmore

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2011

Advisory Committee:  
Professor Patricia A. Alexander, Chair  
Professor Peter Afflerbach  
Ann A. Battle, Ph.D.  
Professor Gregory Hancock  
Professor Kathryn Wentzel

© Copyright by  
Daniel L. Dinsmore  
2011

## Acknowledgments

I began my graduate school career with what I thought was a pretty clear distinction between who my colleagues, friends, and family were. Colleagues were who you worked with, friends were who you spent your free time with, and family were the ones who were there to guide and protect you. However, as I progressed throughout my graduate school career, I noticed that for me at least, these distinctions had become somewhat fuzzy. I quickly realized that I was becoming a part of another family, an academic family.

I first noticed a blurring of these distinctions when I was assigned a mentor upon entering the lab. Emily Fox is perhaps one of the finest “academic big sisters” one could hope for. Just as my brother Andrew had done for me throughout my childhood, Emily guided me through the inevitable gauntlets that academia threw at me with the patience that only an older sibling could possess. As the lab family grew throughout my time here at Maryland, I was constantly challenged and supported by so many eclectic individuals, particularly during the dissertation process.

It was only through the guidance and support of Patricia Alexander, my advisor, which opened up so many doors in my academic career, just as my parents, Tom and Linda, have been doing for me for the past three decades. The knowledge that my parents and my “academic parent” always have my best interests at heart has provided a place for me to grow and explore the world. It is the ever present love and support of these three individuals that have enabled me to contribute to the pursuit of scientific inquiry through my own research and in collaboration with many others.

Finally, I wish to acknowledge someone who has been a member of both my family and academic family, my wife and sometimes coauthor, Meghan Parkinson. Not only does she listen patiently to all of my sometimes crazy research and teaching ideas, she does the same with my personal opinions as well. Meghan's greatest gift to me is her ability to understand me and translate that for the rest of the world. The successful completion of this dissertation is due in large part to having someone like her that I can rely on implicitly.

## Table of Contents

CHAPTER 1: INTRODUCTION .....	1
Statement of the Problem.....	3
What is Deep and Surface Processing?.....	5
How are Levels of Processing Measured?.....	7
Where and When is Deep and Surface Processing Measured?.....	8
What are the Expected Outcomes of the Process.....	9
Purposes of the Study.....	10
Research Questions.....	14
Definitions of Key Terms .....	15
CHAPTER 2: REVIEW OF RELEVANT LITERATURE.....	17
Four Areas for Systematic Investigation.....	19
Conceptualization .....	19
Operationalization.....	20
Situational Considerations .....	22
Model Specification.....	23
Guiding Questions .....	23
Systematic Review Procedure.....	24
Search Criteria .....	24
A Priori Coding Schemes.....	25
Clarity of Definitions .....	25
Measure Type.....	26
Validity Evidence.....	27
Presence of a Task and Academic Domain.....	28
Findings from the Empirical Literature .....	30
The Theoretical Frameworks and Their Impact on the Four Guiding Questions .	30
Intentionality and Approaches to Learning.....	31
An Information Processing Approach .....	33
A Developmental Approach .....	33
Impact on the Four Guiding Questions.....	34
Conceptualization .....	35
Clarity of Definitions .....	35
Definitional Content.....	38
Operationalization.....	43
Types of Measures .....	43
Validity Evidence.....	48
Situational Consideration.....	50
Model Specification.....	52
Concluding Remarks and Indications for Future Research .....	55
Instantiating an Investigation of Deep and Surface Processing with Text .....	60
Theoretical Framing.....	60
The Topographical Perspective.....	61
The Model of Domain Learning (MDL).....	63
The Construction-Integration Model (CIM) .....	64

Measuring Deep and Surface Processing.....	65
CHAPTER 3: METHODOLOGY .....	68
Participants.....	68
Materials .....	69
Measures .....	71
Demographics .....	71
Subject-Matter Knowledge .....	72
Beliefs About ETCs .....	74
Passage Recall Measure.....	75
Measurements .....	75
Think-aloud Data .....	76
Open-ended Response.....	77
Procedures.....	81
Overview of Analyses.....	81
CHAPTER 4: RESULTS AND DISCUSSION.....	87
Preliminary Data Analyses .....	87
Interaction of Depth of Processing and Type of Text on Performance .....	90
Interaction of Depth of Processing and Subject-Matter Knowledge on Performance .....	94
Differences Between the Passage Recall Task and Open-ended Task .....	97
CHAPTER 5: CONCLUSIONS .....	100
The Three Theoretical Frames .....	100
The Construction-Integration Model .....	101
The Model of Domain Learning .....	102
Topographical Perspective on Learning .....	104
The Extant Literature on Deep and Surface Processing .....	106
Conceptualization and Operationalization .....	107
Model Specification and Situational Considerations.....	107
Limitations .....	108
Future Research on Deep and Surface Processing.....	109
Measuring Deep and Surface Processing.....	110
Modeling Deep and Surface Processing .....	111
Implications for Practice .....	114
Appendices.....	117
Appendix A: Summary Table for the Systematic Literature Review .....	117
Appendix B: Expository and Refutational Text Passages .....	162
Appendix C: Demographics Measure.....	174
Appendix D: Subject-Matter Knowledge Measure.....	175
Appendix E: Beliefs About the Existence of ETCs .....	178
Appendix F: Passage Recall Measure.....	179
Appendix G: Instructions for the Verbal Protocol Analysis .....	181
Appendix H: Coding Scheme for the Verbal Protocol Analysis .....	182
References.....	185



## List of Tables

Table 1	Summary Table for Constructs Included in the Study.....	13
Table 2	Definitional Keywords in Explicit Studies by Percentage.....	39
Table 3	Length and Difficulty Data for the Expository and Persuasive Text Passages ....	71
Table 4	SOLO Taxonomy for the Problem-Solving Task Outcome .....	79
Table 5	Standardized Focal and Peripheral Parameter Values of the A Priori Power Analyses.....	84
Table 6	Model Covariance and Mean Matrix for the Measured and Latent Variables .....	88

## List of Figures

Figure 1	Percentages of Each Definition Category Coded from the Tabled Studies.....	36
Figure 2	Studies of Deep and Surface Processing by Year .....	37
Figure 3	Percentages of Each Measure Category Coded from the Tabled Studies .....	44
Figure 4	Percentages of Validity Evidence Categories from the Tabled Studies .....	49
Figure 5	Percentages of Task Type from the Tabled Studies .....	51
Figure 6	Percentages of Outcome Categories from the Tabled Studies .....	53
Figure 7	Percentages of Predictor Categories from the Tabled Studies .....	54
Figure 8	Distribution of Depth of Processing Percentages .....	77
Figure 9	Distribution of Open-ended Response Scores .....	80
Figure 10	Structural Model for Analyses.....	83
Figure 11	Standardized Path Coefficients for the Structural Model.....	90
Figure 12	Interaction Graph for Type of Text on the Relation Between Depth of Processing and Open-ended Response.....	92
Figure 13	Interaction Graph for Subject-Matter Knowledge on the Relation Between Depth of processing and Open-ended Response.....	96

## CHAPTER I

### INTRODUCTION

*While it is not the business of education to prove every statement made, any more than to teach every possible item of information, it is its business to cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions; to lively, sincere, and open-minded preference for conclusions that are properly grounded, and to ingrain into the individual's working habits, methods of inquiry, and reasoning appropriate to the various problems that present themselves.*

(Dewey, 1910/1997, pp. 27-28)

In Dewey's influential book, *How We Think*, he outlined the prerequisites for effective thought and laid out ways in which education can foster these ideals. These constructs Dewey discussed foreshadowed many of the major research areas in the contemporary literature. These areas include the regulation of thinking, epistemic beliefs, interest, and the importance of prior knowledge in the development of effective thought. These areas are important for learning at all ages or stages of academic development.

In recent decades, however, the focus of U.S. education more widely has been on the high-stakes standardized testing that has become so ubiquitous (e.g., No Child Left Behind; H.R. 1, 2001). Although there are likely positives to these high-stakes standardized tests (William, 2010) and while their use appears well intentioned, there are also many questions surrounding the validity, bias, and fairness of these tests (e.g., Sackett, Borneman, & Connelly, 2009). Further, the increased emphasis on high-stakes testing places the focus of education squarely on the outcomes of learning, instead of the

training of effective thought or the *process* of learning that Dewey described more than a century earlier. Dewey specifically warned of the dangers of such a *product* orientation:

In instruction, the external standard manifests itself in the importance attached to the “correct answer.” No one other thing, probably, works so fatally against focusing attention of teachers upon the training of mind as the domination of *their* minds by the idea that the chief thing is to get pupils to recite their lessons correctly. (Dewey, 1910/1997, p. 53)

The effects of this emphasis on testing reach far beyond assessment.

Standardized testing has become the cornerstone of educational reform with seemingly deleterious consequences including “test induced changes in classroom practices and the reproduction of raced-based and class-based inequalities in education” (Au, 2008, p. 639). So, not only has the focus been on whether students get the correct answer, but it has also been on a correct method as to how to get these answers and how to teach these correct methods (i.e., instructional methods). Good, Wiley, and Sabers (2010) contended that standards-based education has the effect of devaluing teachers’ judgments of student success; instead placing the emphasis solely on test scores. In other words, one good source of information about the processes of learning, the teacher, has been largely ignored. One potential outcome of this focus on student proficiency may be the loss of deep learning in schools (Au, 2007; Schraw, 2010).

Additionally, classroom practices aimed exclusively at educational outcomes may have the unintended consequence of turning educational attention almost wholly toward problems for which there are clear-cut correct and incorrect answers to be recalled from memory (e.g., typical multiple-choice tests). Unlike problems typically presented in

educational settings, open-ended problems (i.e., problems whose structure are less well defined with multiple possible goal states) present opportunities in which individuals have to evaluate both their prior knowledge and large sources of external knowledge to solve a problem (Frederickson, 1990; Simon, 1978). Even domains in which many problems are solved algorithmically, such as mathematics, problems can still be considered complex and beyond simple recall, allowing students the opportunity for effective thinking (Kulikowich & DeFranco, 2003). If core academic subjects, such as mathematics and science, are inherently open ended, what advantage would there be to teaching and assessing them using primarily through a recall-based assessment?

### **Statement of the Problem**

These aforementioned issues can be informed by a wealth of research in educational psychology and closely related fields. A promising research area that has examined the role of cognitive processing and outcome is the levels of processing approach (LOP; e.g., deep processing versus surface processing). One view of LOP that has been forwarded in the literature contends that deep processing (i.e., intentional learning of the material; Marton & Saljo, 1976) should lead to more desirable outcomes, whereas surface or shallow processing (i.e., cognitive processes aimed at rote memorization; Marton & Saljo, 1976a; 1976b) should lead to less desirable outcomes (e.g., Chamorro-Premuzic & Furnham, 2008; Phan, 2008). However, these relations have often been weakly correlated in empirical studies (e.g., Cano, 2007).

At issue here is disentangling the relation between the processes and products of learning, specifically the complex relation between depth of processing and performance. In addition to the mixed findings in the literature, there are implicit assumptions by

Marton and Saljo (1976a) as well as in related theories (e.g., *Approaches to Learning*; Biggs, 1978) that levels of processing (i.e., deep versus surface) are stable across situations.

These assumptions that deeper levels of processing (i.e., the use of more deep processes in relation to surface processes) should lead to better academic outcomes and that an individual's processing is stable across tasks had led to a further assumption that individuals can be categorized as either "deep processors" (i.e., those who use primarily deep-level processing) or "surface processors" (i.e., those who use primarily surface-level processing) and an expectation that the deep processors should have better outcomes than the surface processors. For example Tian (2007) posited that "A good assessment method should be able to distinguish between deep learners and surface learners in a way so the former are rewarded while the latter are punished" (p. 387). The suggestion here is that if individuals are encouraged to use deeper-level processing they will perform better on outcome measures.

While the *Approaches to Learning* (Biggs, 1978) view has been the predominate view in the literature on deep and surface processing, alternate frameworks such as developmental models like Alexander's Model of Domain Learning (MDL; Alexander, 1997, 2004) have also been espoused. Models such as the MDL contend that the level of processing shifts as one moves from acclimation (i.e., the beginning stages of domain expertise) to proficiency (i.e., the end stages of domain expertise; Alexander, 1997). In fact, the MDL predicted that surface-level strategies are invoked often in acclimation and diminish as one moves through competence and towards proficiency, while deep-processing strategies do the opposite. However, both surface-level and deep-processing

strategies are evident at all stages of domain expertise and their interplay depends in part of the specific nature of the task and the purpose for which that task is engaged. In other words, the most advantageous level of processing for two individuals may not be the same depending on their stage of development (i.e., acclimation, competence, or expertise) which is described in more detail in Chapter 2. This view directly countered the notion that deep processing should *always* be the goal and that surface processing strategies should be discouraged or punished.

Despite the promise of levels of processing to explore the nature of learning, there are four issues obscuring useful directions for research and practice from the literature. These issues include (a) varied and somewhat conflicting conceptions of deep and surface processing that populate the literature, (b) an overreliance on self-report questionnaires and lack of validity evidence for processing measures, (c) lack of data pertaining to situational considerations during task performance, and (d) a restriction on the types of task performance or outcomes that are assessed. Each of these issues has the potential to at least partly explained the inconsistent findings between depth of processing and performance.

### **What is Deep and Surface Processing?**

Conceptual or construct clarity within the educational research literature has long been a problem (e.g., Alexander, Schallert, & Hare, 1991; Murphy & Alexander, 2000; Van Houtte, 2005). This concern over conceptual clarity has also been associated with the strategy literature, especially related to metacognition, self-regulation, and self-regulated learning, where explicit or well-stated definitions have often been lacking for core constructs of cognitive processing (Dinsmore, Alexander, & Loughlin, 2008). A

review of the empirical literature identified issues including the lack of explicit definitions for deep and surface processing as well as the di- or trichotomous nature (i.e., categorized into two or three groups) of many of these definitions associated with the approaches to learning view.

For example, Phan (2009) references the Approaches to Learning view by defining deep processing as “an intention to understand the authors’ meaning and linking it to their prior knowledge and personal experience” and surface processing as an “approach where the main emphasis is on studying merely for the intention of reproducing information without any further analysis” (p. 159). This stands in contrast to Murphy and Alexander’s (2002) definition of deep processing that consisted of, “such procedures as relating the text to prior knowledge or building a mental image, entail a personalization or transformation of the printed message” and “surface-level processing as strategies that, “refer to techniques, such as rereading or omitting unfamiliar words, that facilitate the initial apprehension or deciphering of a written text” (p. 199).

Differences in these definitions were due primarily to the developmental framework (i.e., the MDL) used by Murphy and Alexander in their investigation. In other words, a major difference between these views rested in the differences between one’s *intention* before a task (e.g., intent to process the task in a surface manner) versus the *actual* processes one evokes during a task (e.g., use of surface-processing strategies while performing a task).

Further, in the former view (i.e., approaches to learning), there has often been a dichotomization made between deep and surface approaches where they seem to be rather incompatible. For instance, Arteche, Chamorro-Premuzic, Ackerman, and Furnham (2009) defined these approaches as:



...the individual's preferred style for learning and can be classified into three categories: deep (intrinsic motivation, engagement with the subject matter, and desire to know everything about the topic); surface (aim at learning the minimal amount necessary to pass); and achievement (goal-orientated study strategies). (p. 359)

The conceptualization of levels of processing as categorical is problematic. As Thorndike and Gates (1929) pointed out, there is likely a distribution of people that fall along a continuum of any variable (in this case their level of processing) rather than a true dichotomy. While it may be the case that dichotomization of the construct makes discussing it and statistical analysis easier, these are not legitimate reasons to conceptualize the construct in a way that impedes the clarity of the findings of an investigation (e.g., Pedhazur, 1997). Pedhazur (1997) conjectured that, "Some of the conflicting results in the research literature of a given area may be due to the practice of categorization of continuous variables" (p. 575).

### **How are Levels of Processing Measured?**

The predominance of dichotomous conceptualizations associated with the approaches to learning view may hamper our ability to measure deep and surface processing as well. This issue in the literature has been particularly problematic because it may be depressing the true amount of variability in the samples used for these investigations; thereby making it difficult to determine the actual correlations between process and product. Additionally, the focus of measurement becomes a search for those who engage in more deep-level processing, versus those who engage in surface-level

processing because of the implicit assumption that surface-processing strategies equate to poor learning.

Beyond problems with the dichotomous measurement approach, there has been an overwhelming reliance on self-report instruments of strategy use or in one's approach to learning. As with the conceptual clarity issue, the overuse of self-report instruments has plagued numerous research literatures in education (e.g., Dinsmore, Alexander, & Loughlin, 2008; Fischer, 2007; Gress, Fior, Hadwin, & Winne, 2010). The ability to list what strategies one has used while performing a task retrospectively measures the extraneous variable of remembering (or realizing) what strategies one used while performing a task. This issue is even more problematic in the approaches to learning framework because these self-reported approaches were applied to situations not specified in these instruments.

### **Where and When are Levels of Processing Measured?**

This latter point raises yet another shortcoming within the extant literature on level of processing; that is, many studies have not taken into account the situation (or context) in which the learner has performed a given task (if there was a performance task at all). Recently, Alexander, Schallert, and Reynolds (2009) described a learning topography in which they outlined four dimensions of learning: the *what*, *where*, *who*, and *when*. A consideration of the *where* dimension is what has been lacking in the extant literature. The *where* referred to the physical, social, or cultural aspects of the environment, both concrete and abstract. Without due consideration of the context or situation, a rather narrow view of what individuals do (e.g., use of deep- or surface-processing strategies) may be presented. For instance, the type of text (e.g., expository

versus persuasive) an individual reads may have a large influence on their level of processing. There is also evidence of these effects with respect to metacognitive monitoring (Dinsmore, Loughlin, Parkinson, & Alexander, 2009) as well as their knowledge, interests, and beliefs (Buehl, Alexander, Murphy, & Sperl, 2001; Murphy, Long, Holleran, & Esterly, 2003).

In addition, there has been little consideration of the *when* (i.e., changes associated with maturation and experience; Alexander et al., 2009). A typical example of an investigation of deep and surface processing has been to correlate a self-report measure (typically approaches to learning) to one's performance in a particular class (e.g., Phan, 2009; Thomas & Gadbois, 2007). In effect, this has restricted the range of personal characteristics, or as Alexander et al. (2009) called it, the *who* (e.g., knowledge, interest, or beliefs) that may be present. Given that individuals have been taking the same class, it can be assumed that they were somewhat more homogenous than even a random university sample on many variables. The lack of many longitudinal or cross-sectional investigations of deep and surface processing (with a few exceptions) has limited our knowledge of the developmental aspects of the construct.

### **What are the Expected Outcomes of the Process?**

To further complexify matters, Alexander et al. (2009) contended that the *who*, *where*, and *when* are likely to interact with their fourth proposed dimension of learning, the *what* (i.e., the target of learning). The links between the process (i.e., levels of processing) and product (i.e., the outcome) have also been relatively constrained. In most cases the outcome of learning has been constrained to recall tasks such as a word recognition (e.g., Ally, Gold, & Budson, 2009; Boldini, Russo, & Avons, 2004), face

recognition task (e.g., Bentley, Driver, & Dolan, 2009; Block, 2009), or passage recall tests (e.g., Murphy & Alexander, 2002; Schommer & Surber, 1986). Without due consideration of other types of outcomes (e.g., open-ended responses), there has been a difficulty in making links between the process and any other product other than recall. Further, these different types of outcomes (recognition and open-ended) have not been tested simultaneously.

### **Purposes of the Study**

The four aforementioned issues informed the design of the current study. The purpose of the current study was to understand the complex interplay of three of these different dimensions of learning (i.e., the *who*, *where*, and *what*) to shed light on the relation between the characteristics of the learner (i.e., subject-matter knowledge), levels of processing (i.e., deep and surface processing), the situation (i.e., type of text), and outcomes of the learning task (i.e., a recall measure and an open-ended response).

The first purpose of this study was to explore the existing literature on levels of processing to help inform the methodology for the empirical study. Four major questions guided this literature review:

How are deep and surface processing conceptualized?

How have deep and surface processing been operationalized in the literature and what validity evidence have been provided for these measures?

What is the nature of the tasks used in these investigations?

What predictors of deep and surface processing have been examined?

This review was used to inform these four questions and helped conceptualize deep and surface processing, create valid measures of deep and surface processing, design

appropriate tasks, and identify relevant variables in the *who*, *what*, *where*, and *when* dimensions as proposed by Alexander et al. (2009) for this investigation.

The empirical study examined three of these dimensions (the *who*, *where*, and *what*) framed by the Topographical Perspective of Learning. Levels of processing was investigated during text processing of one of two types of text (the *where*) and operationalized using a think-aloud protocol, which is a concurrent measure of individuals' monitoring and reading behavior (Pressley & Afflerbach, 1995). This concurrent measure assessed the nature of individuals' processing while doing a task, rather than assessing their intentions for a task prior to completing a task. While using the think-aloud protocol to operationalize levels of processing (the proportion of deep-processing strategies to surface-level strategies) is aimed at *strategic processes* (those that are intentional and purposeful), it is certainly possible that participants may also report skillful performance (automated processes). Since there is no way to know whether these processes reported in the think aloud are strategic or skillful, a more liberal scheme that includes both under the label *strategies* was implemented. This scheme was implemented with the awareness that some of the explicated cognitive processes may well be more precisely classified as explicated skills. However, the terms deep-level strategies and surface-level strategies will be used to be consistent with the CIM, MDL, and Pressley and Afflerbach's (1995) description of reading behaviors. The Topographical Perspective of Learning, MDL (Alexander, 1997) and the CIM (Kintsch, 2004) will be described in further detail in Chapter 2.

Participants read one of two text passages (i.e. either an expository passage meant to inform or a two-sided refutational passage meant to persuade), both about possible

existence of intelligent extraterrestrial life in the universe. The data collected through the think-aloud methodology were coded using a pre-existing coding scheme (e.g., Dinsmore, Fox, & Parkinson, 2010; Fox, Dinsmore, & Alexander, in press; Fox, Dinsmore, Maggioni, & Alexander, 2008) and CIM to classify utterances as a deep-level strategy, a surface-level strategy, or other.

A prior measure of subject-matter knowledge (i.e., prior knowledge about astronomy and extraterrestrial life) was collected to investigate how the characteristics of the individual (the *who*) affected both the level of processing (deep and surface) and outcomes. The outcome task for this investigation (the *what*) consisted of two tasks, a passage recall measure and an open-ended response. The recall task was a multiple-choice test assessing individuals' recall of the passage. The open-ended response was one question designed to assess both their comprehension of the text passage and their ability to draw on prior knowledge to justify their beliefs about the existence of extraterrestrials. This response was assessed using the Structured Observation of the Learning Outcome (SOLO) taxonomy (Biggs & Collis, 1982). This taxonomy assessed responses based on an individual's: capacity to answer to the question (i.e., use of cues and relevant data), relating operations used (i.e., their ability to generalize), consistency of their response, and their level of closure. The SOLO taxonomy was useful in this case because it allowed an examination of the justification used by individuals rather than their response as correct or incorrect. Table 1 presents the constructs along with how they were measured and scored.

Table 1

*Summary Table for Constructs Included in the Study*

Construct	Measure/Measurement	Scores
Depth of processing	Verbal protocol and coding scheme (Appendix J)	Percentage of deep-level strategies with a possible range from 0-100% ( $M = .61$ , $SD = .28$ )
Type of text	Manipulated condition: expository or persuasive	N/A
Subject-matter knowledge	16 graduated response multiple choice items	Factor scores ( $M = .00$ , $SD = 1.00$ )
Passage recall	10 graduated response multiple choice items	Latent factor
Justification	Open-ended response coded using the Structured Observation of the Learning Outcome (SOLO) taxonomy	Code categories range from 0 (prestructural) to 4 (extended abstract; $M = 1.57$ , $SD = 1.06$ )

Since this investigation sought to study the interactions of the variables described previously, it fits into the general design category of Attribute-Treatment Interactions (ATI; e.g., Cronbach & Snow, 1977). This purpose of this empirical investigation was to examine the interaction of individuals' aptitude (i.e., their subject-matter knowledge and levels of processing) in a manipulated situation (i.e., expository or two-sided refutational text) on two types of learning outcomes (i.e., a recall task and an open-ended task).

## Research Questions

1. Does the type of text (i.e., expository versus two-sided refutation) moderate the relation between levels of processing (deep and surface) and reading outcomes (i.e., recall and open-ended responses) during the reading of science-related texts?

It was hypothesized that participants reading the two-sided refutational text would engage in more deep-level strategies than those reading the expository text and that a higher proportion of deep-level strategies would result in higher scores on the open-ended task. These differences were expected due to previous interaction effects in the extant literature (Kamalski, Sanders, & Lent, 2008), different amounts of working memory requirements (Kellogg, 2001), and differences on multiple-choice outcome scores (Carrell & Connor, 1991). Specifically, two-sided refutational text has been found to be more effective at changing knowledge and beliefs (Buehl, Alexander, Murphy, & Sperl, 2001; Murphy, Long, Holleran, & Esterly, 2003), as well as changes in metacognitive processing (Dinsmore et al., 2009).

2. Do individual characteristics (i.e., subject-matter knowledge) moderate the relation between levels of processing (deep and surface) and reading outcomes (i.e., recall and open-ended responses) during the reading of science-related texts?

It was hypothesized that increased levels of subject-matter knowledge would moderate the relation between levels of processing and reading outcomes.

Specifically, the relation between the strategies reported (i.e., deep or surface) and reading outcomes was predicted to be greater for those with higher levels of subject-matter knowledge than those participants with lower levels of subject-



matter knowledge. The relation hypotheses were supported by previous investigations of the MDL, all be they through the use strategy checklists (e.g., Alexander, Sperl, Buehl, Fives, & Chiu, 2004; Murphy & Alexander, 2002).

3. Does the relation between depth of processing and learning outcomes differ between a passage-recall task and an open-ended task?

It was hypothesized that participants demonstrating a greater use of deep processing (i.e., increased use of deep-processing strategies) would have higher scores on passage recall and open-ended tasks, but that this difference in performance would be greater for the open-ended task. The larger disparity for the open-ended task was hypothesized since differences between forced choice and open-ended task performance have been found in regards to search strategy, where open-ended problems require complex, expansive, and multidisciplinary knowledge of a field (Laxman, 2010). Additionally it has been found that students engage in more peer learning strategies and critical thinking during open-ended tasks (Lodewyk, Winne, & Jamieson-Noel, 2009).

### **Definitions of Key Terms**

*Deep processing* is the use of strategic and monitoring behavior that involve a more extensive manipulation or interpretation of a task or text (Alexander, Murphy, Buehl, Fives, & Chiu, 2004; Nolen & Haladyna, 1990).

*Domain knowledge* is defined as the breadth and scope of subject-matter knowledge (Alexander, Murphy, Woods, Duhon, & Parker, 1997).

*Expository text* is defined as non-fiction reading material in which the intent is to inform or explain (Williams, Stafford, Lauer, Hall, & Pollini, 2009).

*Learning* is a multidimensional process that incorporates a target of what is to be learned, the nature of the learner, the situation in which change or learning occurs, and the temporal nature of when change or learning takes place (Alexander et al., 2009).

*Levels of processing* is defined as the relative proportion of deep and surface processes employed during a cognitive task.

*Open-ended outcomes* are defined as outcomes in which there may be many possible goal states and many possible solution path (Frederickson, 1990; Simon, 1978).

*Persuasive text* is defined as text in which an author argues a point of view in order to change a reader's knowledge, beliefs, or interest (Kamalski et al., 2002; Murphy et al., 2003).

*Passage recall* is the ability to reproduce the text verbatim, in paraphrase, or by summarizing it (Kintsch, 1998).

*Surface processing* is the use of strategic and monitoring behavior related to the encoding of textual content (Alexander et al., 2004; Nolen & Haladyna, 1990).

*Topic knowledge* is defined as depth of knowledge about specific concepts related to a domain (Alexander et al., 1997).

## CHAPTER II

### REVIEW OF RELEVANT LITERATURE

*Agreement between theory and observation should count for nothing unless the theory is testable, and unless the agreement is found as a result of serious attempts to test it. (Popper, 1994, p. 89)*

In Popper's 1986 lectures at Stanford, he described the method of science as a method of critical discussion. According to Popper, critical discussion in scientific inquiry entails the testing of data-model fit and the development of alternate models to challenge existing ones. One such area in need of critical discussion is the investigation of deep and surface processing (or levels of processing). In this chapter, the need for critical discussion is outlined, four suggested areas to explore are delineated, and one alternative model (an interaction model) for testing depth of processing and performance is proposed.

The calls for this critical examination deep and surface processing are becoming increasingly numerous (e.g., Block, 2009; Heijne-Penninga, Kuks, Hofman, & Cohen-Schotanus, 2008). The rationale for embarking on this line of inquiry is the problem of ambiguity in how deep and surface processing are conceptualized and the inconsistency between models and studies in the extant literature (e.g., Block, 2009; Cano, 2007; Justicia, Pichardo, Cano, Berbén, & De la Fuente, 2008; Phan, 2009). In other words, there seems to be tension with the assumption by some that deep processing promotes better learning outcomes, while surface processing promotes weaker learning outcomes, however, this assumption only holds in some studies of the relation between depth of processing and performance. Therefore, the purpose of this chapter is to question the

prevailing wisdom as to the additive effects of deep processing (i.e., that deep processing should always promote better learning outcomes) in part, by exploring four areas identified as possible reasons why the prevailing wisdom and subsequent data-model fit differ. Some of these reasons come from previous studies examining the effects of deep and surface processing on performance, namely those that included situational factors (e.g., You & Jia, 2008). Other reasons for these mixed findings in the literature come from various areas outside of the depth of processing literature and include considerations of conceptualization, operationalization, and model specification (e.g., Dinsmore et al., 2008, Murphy & Alexander, 2002).

Although the focus of this review is on the conceptualization, operationalization, situational factors, and model specification of deep and surface processing (i.e. levels of processing) in the empirical literature, the implications of this examination reach far beyond research. Issues directly related to research can help inform constituents (e.g., students, parents, teachers, or administrators) about the role levels of processing may play in learning; specifically, how one's level of processing aids or hinders learning in various situations. Given the mixed findings in the literature thus far, it is doubtful that a *best* approach for students exists (i.e., a deep processing approach). Rather, there needs to be an examination of when surface processing, deep processing, or a combination of the two can be used to learn more effectively. However, the inconsistency of the findings in research thus far may be hampering our ability to uncover the complex relations that may exist.

The purpose of this literature review is to examine empirical research on deep and surface processing in four areas: conceptualizations (i.e., definitions) of deep and surface

processing; operationalization (i.e., measurement) of deep and surface processing; the situation or environment in which deep and surface processing is considered; and the exogenous and endogenous variables (i.e., model specification) of deep and surface processing. To facilitate a critical examination of these issues, this review of deep and surface processing is divided into four parts. First, the four areas listed previously (i.e., conceptualization, measurement, situation, and model specification) will be discussed in terms of how they may or may not be expected to contribute to the inconsistency of results between deep and surface processing and learning outcomes. Second, search and coding strategies will be specified for how studies were identified for inclusion in this review and how they were coded according to the four areas (i.e., conceptualization, operationalization, situational factors, and model specification). Third, the results of the systematic review will be presented. Finally, an exemplar of these issues will be given for text processing.

### **Four Areas for Systematic Investigation**

First, four areas for systematic investigation will be identified based on how each of these areas may explain the inconsistency of results in the literature between deep and surface processing and learning outcomes. These four areas were identified from past reviews in the educational psychology literature and from studies on deep and surface processing and frame the four guiding questions for this investigation.

#### **Conceptualization**

The definition of a construct proffered, such as deep and surface processing, reveals how it is being conceptualized in a study and how the study as a whole is framed. Clearly stated definitions provide internal consistency among the operationalization of

the construct, inclusion of relevant situational factors, and model specification within a study. However, reviews of other areas of educational psychology have demonstrated that explicit definitions are not the norm. In the motivation literature Murphy and Alexander (2002) found that key constructs such as individual interest and self-efficacy were typically not explicitly defined. Similarly, in the strategic processing literature Dinsmore et al. (2008) found that metacognition, self-regulation, and self-regulated learning were also not always explicitly defined.

An explicit conceptualization of deep and surface processing is the lynchpin for identifying or developing valid measures of the construct and ultimately for interpreting any resulting data collected. It is quite possible that the problems plaguing the motivation and strategic processing literature identified previously may be causing the inconsistent relations seen between deep and surface processing and learning outcomes.

### **Operationalization**

The operationalization of a construct refers to how the construct is enacted or measured. The operationalization of deep and surface processing can have an impact on the inconsistent results in the literature in terms of both the measure chosen and the validity of that measure.

The type of measurement chosen to quantify the relative prevalence of one's deep and surface processing may partially explain the inconsistent results. For example, measures and measurements of metacognition, self-regulation, and self-regulated learning varied widely in the literature from self-report questionnaires to codings of think-aloud protocols, with a particular reliance on self-report questionnaires (Dinsmore et al., 2008). The overreliance on self-report questionnaires in the strategic processing literature was

troubling because these measures require the participants to be meta-metacognitive – in other words, they have to think about their thinking of their own thinking. Problems such as this are potential causes of measurement error or bias. Since deep and surface processing is also a *process* measure, similar issues of self-report questionnaires may be contributing to the inconsistency of the results in the literature. A caveat to this previous point is that some form of self report may be necessary to elicit what participants are actually doing, however, care needs to be taken that an argument can be made for their validity.

Second, as stated earlier, the conceptual definitions should be the lynchpin to creating a measure or measurement of a construct. Evidence of a measure's validity should be provided by authors (e.g., Crocker & Algina, 1986; Schunk, 2008). In prior investigations of strategic processing, reliability was generally well established, but a clear link with the conceptual definition was often not well established and was difficult to determine (Dinsmore et al., 2008). Cronbach (1971) defined the process of validation as “the accuracy of a specific prediction or inference from a test score” (p. 443). Further, Cronbach outlines three types of validation procedures: content, criterion-related, and construct validation. [For an extended description of these three types of validation processes see Crocker and Algina, 1986, Chapter 10.] Content validation refers to the process of using a small number of items on a test to infer to a larger domain of interest. Criterion-related validation refers to inferences from a test score to some other external behavioral variable. And, finally, construct validation refers to inferring from a test score to a psychological construct. Cronbach points out that what is being validated is not the

test itself, rather the *interpretation* of the resulting data. Thus, if interpretations of the resulting data are not valid, inconsistent results are *likely* to occur.

### **Situational Considerations**

The next factor that could contribute to inconsistency of results between deep and surface processing and learning outcomes is the research context. This consideration could be an issue whether it is research conducted in a laboratory or in a more naturalistic setting such as a classroom. In Alexander, Scallert, and Reynolds's (2009) topography of learning they refer to situation as the *where* of learning and postulate it as a critical dimension in learning. Learning always occurs in some type of situation, whether one is referring to a physical context (e.g., a laboratory versus a classroom), a social context (e.g., working alone versus a group project), or a cultural context (e.g., a school versus a museum).

Situational considerations could be an important factor in challenging the assumption that deep processing is always good and surface processing is always bad and explain *why* the inconsistencies in the literature are occurring. For instance, the academic domain of an investigation has been a focal point in number of reviews in regards to motivation and strategic processing (e.g., Dinsmore et al., 2008; Murphy & Alexander, 2002). Specifically, some argue in regards to self-regulation and self-regulated learning that domain *does* matter (Alexander, Dinsmore, Parkinson, & Winters, 2011).

Additionally, influential frameworks of deep and surface processing, such as Biggs' approaches to learning place a major emphasis on the situational factors and these researchers question research designs that do not include relevant situational factors (e.g., Biggs, 1993). Situational considerations may be an important piece of the puzzle in



explaining why the relations between deep and surface processing and learning outcomes have been inconsistent.

### **Model Specification**

The fourth issue that may be a potential cause for the inconsistent result between deep and surface processing and learning outcomes is how researchers are specifying models of deep and surface processing in their investigations. This issue is important because misspecified models (e.g., models that omit relevant variables or include irrelevant variables) bias the estimation of coefficients in a model. [For a complete discussion of model misspecification see Pedhazur (1997) or Cohen, Cohen, West and Aiken (2003).] Both the exogenous variables (i.e., predictors) and endogenous variables (i.e., outcomes) of deep and surface processing chosen in a model of deep and surface processing has the potential to change the relation between deep and surface processing and a specified outcome (e.g., recall versus comprehension).

### **Guiding Questions**

To investigate the nature of these four issues in the extant literature four guiding questions were developed. To investigate these four questions, a systematic review of the literature on deep and surface processing was undertaken. The four issues described in the preceding section are the basis for the four questions chosen to guide a systematic review of the literature. These four questions are:

How are deep and surface processing conceptualized in the literature?

How have deep and surface processing been measured and what validity evidence has been provided for these measures?

What type of situations or environments have deep and surface processing been investigated in?

How have models of deep and surface processing been specified in regard to the antecedents/consequents of deep and surface processing?

To investigate these four questions, a search strategy and coding scheme were developed which is described next.

### **Systematic Review Procedure**

#### **Search Criteria**

In order to consider the state of deep and surface processing in regards to the aforementioned four questions, a pool of studies was created. This pool was created by identifying articles within the Psycinfo database that contained one or more of the following terms in the article abstract: “deep processing,” “deep processors,” “shallow processing,” “shallow processors,” “surface processing,” “surface processors,” “deep learning,” “deep learner,” “shallow learning,” “shallow learner,” “surface learning,” “surface learner,” “depth of processing,” “depth of learning,” “deep strategies,” and “surface strategies.” By searching for the terms in the article abstract instead of the whole document, this strategy ensures that deep and surface processing are key constructs in the studies pooled. The search was further delimited to those studies that were in peer-reviewed journals, were empirical, used human populations, and were in English. There were no limitations placed on the date in which the article was published.

This search strategy initially yielded 222 articles from the PsycInfo database from the years 1983 to 2009. One article identified in the Psycinfo search was discarded from the pool since it did not mention deep processing, surface processing, or levels of

processing (i.e., Jetton, 1994). The remaining studies were organized into a comprehensive table (see Appendix A). The table is organized first by year of publication and then alphabetically by first author. Articles were tabled and analyzed according to the following dimensions: theoretical frame, definitions of the constructs, measure(s) of deep or surface processing, explicitly stated validity evidence, stated predictors and/or outcomes of deep and surface processing, the task participants completed during the study (if any), and the academic domain under investigation.

### **A Priori Coding Schemes**

An explicit a priori coding scheme was developed for five dimensions related to the four issues described above: clarity of definitions, type of measure, validity evidence, presence of a research task, and academic domain. An initial attempt was made to code a sixth dimension, theoretical frame, but was dropped due to issues with the identification of the theoretical framework used. To check the reliability of the resulting five coding schemes, a second rater was used to code a randomly-selected five percent of studies in the pool with a high level of interrater agreement (percent agreement = 0.89). Any discrepancies were ameliorated in conference.

**Clarity of definitions.** The clarity of definition code was modeled from two previous systematic literature reviews (Dinsmore et al., 2008; Murphy & Alexander, 2000). Three broad definition categories were identified: explicit, implicit, and absent. *Explicit definitions* codes were applied when an author specifically and clearly stated a definition (coded as “E” in the table). For example, Jay, Caldwell-Harris, and King (2008) explicitly defined deep and shallow levels of processing by stating, “A shallow level of processing is one in which only superficial or physical aspects are encoded. A

deeper level of processing takes more time and effort to activate the semantic meaning of the stimulus” (p. 85).

Three different codes were used to describe the three forms of implicit definitions. *Conceptual definition* codes (coded as “C” in the table) were applied when the authors did not expressly state a definition but where words or phrases appeared in the text that alluded to meaning. For example, Blom and Severiens (2008) suggested that surface learning was rehearsal and that deep learning was elaboration and critical thinking, but did not explicitly state either of these attributes. *Referential definition* codes (coded as “R” in the table) were indicated when key references were applied as proxies for definitions. For example, Al-Emadi (2001) referenced the work of Elliot, McGregor, and Gable (1999) as a proxy for the author’s own conceptualization of deep and surface processing. An *absent* definition code (coded as “A” in the table) was used when there was neither an implicit or explicit definition provided. For example, in a study by Jundt and Hinsz (2002), a definition was not offered for deep, surface, or shallow processing.

**Measure type.** Five codes were used to describe the measure types. These codes were also adapted from a previous systematic literature review (Dinsmore et al., 2008). The first code indicated a self-report questionnaire (coded as SR), which are listed in the table by either the name of the measure (e.g., the MSLQ) or as a researcher developed questionnaire (labeled as “RDQ” for each study; e.g., Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004). The second code was used for studies that examined deep and surface processing by condition (coded as “CON” in the table). For example, McKelvie and Pullara (1988) tried to induce deep processing by having participants rate words based on pleasantness and tried to induce shallow processing by having participants rate

words based on visual complexity. The third code was applied to measures that used some physiological response to measure deep and surface processing (coded as “PHY” in the table). For example, She and Chen (2009) used eye tracking (i.e., physiological eye movements) to measure deep and surface processing. The fourth code was applied to measures that used some sort of coding scheme to measure deep and surface processing *during* a task (marked as “CODE” in the table). For instance, Fergusson-Hessler and de Jong (1990) coded participants think alouds while reading a physics task. The fifth code was used for studies where there was no identifiable measure of deep or surface processing (coded as “NONE” in the table).

**Validity evidence.** Validity evidence for each of the studies was coded using Cronbach’s (1971) description of the three major types of validity: content, criterion-related, and construct validation. Studies that reported validity evidence that related to content validation (i.e., inferences from a test score to a larger domain of similar items) were coded in the table as *content*. Procedures for assessing content validation include: defining the domain of interest, selecting a panel of qualified experts, and matching items to the performance domain (Crocker & Algina, 1986). For example, Henry, Van Lunen, Udermann, and Onate (2009) reported using expert validation for their researcher-designed questionnaire of deep and surface processing.

Studies that reported validity evidence that related to criterion-related validation (i.e., inferences from a test score to other behavioral variables) were coded in the table as *criterion*. Procedures for assessing criterion validity include: identifying a suitable criterion, identifying an appropriate sample, and determining the strength of the relation between the desired measure and the criterion performance (Crocker & Algina, 1986).

For instance, Camp, Markley, and Kramer (1983) correlated their measure of deep and surface processing (retrospective reports) with a criterion performance measure (strategy use they observed during the task). Finally, studies that reported validity evidence related to construct validation (i.e., inferences from a test score to the label of a particular psychological construct) were coded in the table as *construct*. Widely used approaches to construct validation include: differentiation between groups, factor analysis, and the multitrait-multimethod matrix. For example, Yamauchi and Miki (2000) used factor analysis on their *Learning Strategies Scale* to see if the structure identified in their sample matched their theorized structure. The last two categories were studies that either did not report any validity evidence (coded in the table as *none*) or only referenced previous studies to support their claims of validation (coded in the table as *ref*).

During the course of tabling, one additional type of validity evidence was explicitly mentioned, face validity (i.e., items which appear to measure a meaningful construct to laypersons; Crocker & Algina, 1986). For example, Hancock, Stock, Kulhavy, and Swindell (1996) explicitly provided evidence of face validity for their *Critical Incidence Questionnaire (CDQ)*. Since face validity was not one of the a priori codes, face validity was coded as *content*. This decision was made because while face validity and content validity are not exactly equivalent, they are highly related (Crocker & Algina, 1986).

**Presence of a task and academic domain.** The final a priori coding schemes examined the presence of a research task and what academic domain was under investigation. In regards to the research task, studies were coded as containing a research task (RT) or not containing a research task (NT). For example, Burgess and Weaver

(2003) studied deep and surface processing by having participants do a face recognition task and this study was coded as RT. On the other hand, Klinger (2006) did not include a research task, but instead correlated a self-report questionnaire with the type of perceived learning environment as was therefore coded as NT.

If the study indicated a particular academic domain either in the research task (for studies coded as RT in the previous scheme) or the measure of deep and surface processing (for studies coded as NT in the previous scheme), then one of the following academic domain codes was used: physical and life sciences (PLS); social science (SS); language arts (LA); mathematics (MM); business (BS); philosophical or religious studies (PRS); performance arts and athletics (PAA); and computer science (CS). For example, in regards to the academic domain of a research task Segalowitz, Cohen, Chan, and Prieur (2001) asked participants to recall elements from a musical score and was coded as *PAA*. In regards to the academic domains referenced by items in a scale, Cassidy and Eachus (2000) asked participants to answer items on the Approaches and Study Skills Inventory for Students (*ASSIST*) in terms of their performance in a psychology class and was coded in the table as *SS*. If no academic domain was indicated either in the research task or measure of deep and surface processing, then it was coded as domain general (*DG*). Finally, studies that indicated more than one academic as part of a research tasks or tasks or measures of deep and surface processing were coded as multiple domains (*MD*).

## **Findings from the Empirical Literature**

Data from the systematic review will be presented and discussed in the following sections for each of the four guiding questions. First, however, the attempt to code the theoretical frameworks for the tabled studies will be discussed.

### **The Theoretical Frameworks and Their Impact on the Four Guiding Questions**

As the selected studies were initially tabled, an attempt was made to table the theoretical framework used for each study. The difficulty in this endeavor was that only in rare cases was the theoretical frame explicitly mentioned. For example, You and Jia (2008) specifically invoked Approaches to Learning (AL) as their framework, Murphy and Alexander (2002) specifically built on the Model of Domain Learning (MDL) as their framework, and N'Kaoua, Véron, Lespinet, Claverie, and Sztark (2002) specifically referenced Craik and Lockhart's (1972) framework, an Information Processing Theory (IPT) model. However, the norm for these studies was either only a reference to some framework (such as Craik & Lockhart's model; e.g., Gadzella, Stephens, & Baloglu, 2002) or an introduction that alluded to a particular framework. While the coding scheme employed for the definitions (i.e., explicit, implicit-referential, implicit-conceptual, implicit-measure, and absent) was attempted for the conceptual frames, this was deemed too unreliable of a task between raters. In the rare cases where a framework was explicitly mentioned this was not a problem, however, it was very difficult to determine in cases where references or conceptual allusions to a framework were made if the authors intended to use the referenced or alluded to models or theories as the frameworks for the study.



Nonetheless, this initial foray into coding the framework did reveal three different types of frameworks invoked for deep and surface processing that serve as illustrative examples for the conceptualizations, operationalizations, situational considerations, and model specification discussed consequently. The three frameworks discussed here include: Approaches to Learning (AL; e.g., Biggs, 1978), Information Processing Theory (IPT) models (e.g., Craik & Lockhart, 1972), and the Model of Domain Learning (MDL; e.g., Alexander, 1997). It is important to acknowledge the theoretical frameworks because these frameworks are what allow researchers to systematically test the relations between deep and surface processing and learning outcomes and organize the findings. In other words, differences found in the relations between levels of processing and learning outcomes may be explained by the different interpretations placed upon the data by the theoretical frame employed. By specifying the theoretical frameworks explicitly, the task of critical discussion as described by Karl Popper becomes more straightforward. In order to pave the way toward this end, each of these three theoretical frames (AL, IPT, and the MDL) will be described briefly, followed by a discussion of how these different frames may impact the conceptualization, operationalization, situational consideration, and model specification of deep and surface processing.

**Intentionality and approaches to learning.** Perhaps the earliest cogent theoretical model of deep and surface processing was articulated by Marton and Säljö (1976a). In their writing, the authors discussed their attempts to investigate qualitative differences in the outcome and process of learning. In regard to the process, derived mostly through semi-structured interviews, they identified two different levels of learner engagement while participants read text, surface-level and deep-level processing, which

they reported to be clearly distinguishable. These levels referred to the aspects of the learning material on which the learner focused. Surface-level processing indicated that students had a reproductive conception of learning and directed their attention toward the features of the text itself. Deep-level processing indicated that students directed their intentions toward comprehending what the author had to say, or what was signified by the features of the text.

Marton and Säljö (1976b) expanded this initial framework to investigate whether or not students' intentionality could be changed given two assessment conditions. The surface-level condition asked students to recall specific points from a text, while the deep-level condition asked students to state the main points of the text. Hence, they were investigating whether they could induce deep-level or surface-level processing.

Biggs (1978), building on the work of Marton and Säljö, took a similar approach in terms of intentionality (termed Approaches to Learning; AL). Instead of the two approaches identified by Marton and Säljö, Biggs theorized there were three, consisting of reproducing, internalizing, and organizing approaches. The internalizing approach was analogous to deep-level processing, while reproducing and organizing were roughly analogous to surface-level processing. Additionally, Biggs placed a heavy emphasis on an individual's stable personality traits and to a lesser degree one's environment in determining their approach to studying (i.e., their process of learning). This model was developed based on data collected from a self-report instrument, the Study Process Questionnaire (SPQ) in which motivational (e.g., intrinsic motivation) and personality factors (e.g., test anxiety) were used to predict their study process and resulting academic outcomes (e.g., course grades).

**An information processing approach.** Approaches such as Craik and Lockhart's (1972) levels of processing approach consider deep and surface processing in terms of memory. This approach situates it within information processing theory (IPT) because of its consideration of encoding operations and retrieval of information. Craik and Lockhart postulated that the persistence or durability of a memory trace was not related to a specific memory store (e.g., short or long term memory), but rather was instead related to the type of processing during which the information was initially encoded (e.g., deep or shallow processing), hence the term levels of processing approach. The attempts here were to distinguish between what Craik and Lockhart termed qualitatively different domains of processing. For example, processing the phonemic aspect of words in a list was posited to have a qualitatively different encoding process than processing the semantic aspect of words in a list. These different processes were hypothesized to affect the retention of these words in memory and hence produce different rates of recall. While there were numerous IPT approaches present, Craik and Lockhart's (1972) levels of processing framework typified these approaches.

**A developmental approach.** Unlike a frame in which approaches to tasks are stable such as IPT and AL, a developmental approach would model some change over time in how learners engage in a task. One such developmental model that examines deep and surface processing is Alexander's Model of Domain Learning (MDL; 1997). The MDL examines a learner's path in an academic domain (e.g., mathematics) through three different stages of expertise: acclimation, competence, and proficiency. Individuals' development (or lack thereof) in these stages is guided by the interaction of three forces: knowledge, interest, and strategies. Further, the MDL distinguishes between

surface-level strategies (e.g., initial apprehension or deciphering of text) and deep-level strategies (e.g., personalization or transformation of text).

In acclimation, learners depend primarily on surface-level strategies to build subject-matter knowledge (Murphy & Alexander, 2002), which encompasses both domain knowledge (i.e., breadth of knowledge one possesses of a target domain) and topic knowledge (i.e., the depth of knowledge about a domain-specific concept; Alexander, 1997). Whereas an approach to learning framework would conceptualize processing as a stable trait (e.g., a personality trait) or IPT as a type of processing to encode, the MDL specifies a shift in processing as an individual develops expertise (i.e., moves towards competence or proficiency). Hence, according to the MDL, one's use of deep- or surface-level processing depends on individual characteristics within a given domain that are likely to change over time.

**Impact on the four guiding questions.** The theoretical framing of a study should guide the conceptualization, operationalization, situational considerations, and model specification. These three frameworks differ in a number of ways related to these four issues. One of these differences will be highlighted here as an exemplar and will be discussed, along with other differences between the frameworks, more thoroughly in the subsequent treatment of the four guiding questions.

These three frameworks would be expected to differ in regards to their treatment of situational considerations. In general, IPT frameworks such as Craik and Lockhart's (1972) levels of processing do not assume that situational factors (e.g., characteristics of the task or situation) affect the relation between processing and memory retrieval beyond the task instructions given (e.g., to process the words in a list semantically). On the other

hand, both AL and the MDL conjecture that the academic domain of the task under investigation (a characteristic of the task) *will* affect the processing of the task (Biggs, 1978; Alexander, 1997). In the case of IPT, the relation between levels of processing and learning outcomes is invariant across situations, whereas for AL and the MDL this relation is thought to vary across situations. For example, the MDL predicts an individual's use of deep and surface-processing strategies to be different when performing tasks in domains (e.g. mathematics versus history) in which s/he has varying levels of expertise. Further, the MDL differentiates itself from AL by considering not only the role of situational factors such as academic domains, but how factors such as these will change over time.

### **Conceptualization**

With regard to conceptualization, two specific issues were investigated. First, the explicitness of the definitions was examined, and second, what definitional keywords appeared most often in the definitions proffered (i.e., the content of the definitions).

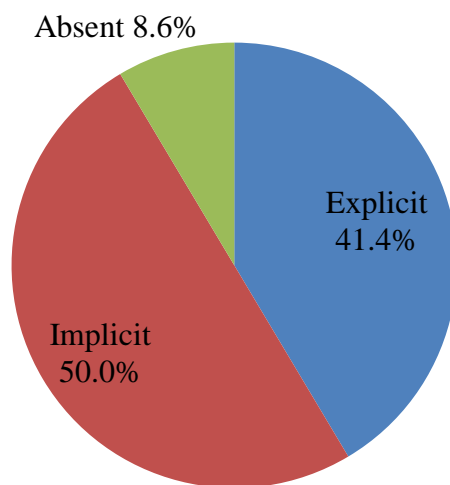
**Clarity of definitions.** Figure 1 shows the percentage of each definition code (i.e., explicit, implicit, and absent definitions. It was clear from this analysis that almost half of the studies examined failed to offer an explicit definition of deep processing, surface processing, or levels of processing. Further, for the studies that provided implicit definitions, 36.0% of those were conceptually defined (e.g., Freed, Marshall, & Phillips, 1998), 12.6% of those were referentially defined (e.g., Graham & Golan, 1991), and 50.4% of those were defined by measure (e.g., Cassidy, 2007). The low number of studies that explicitly defined deep and surface processing and the high number of studies

that defined deep and surface processing by measure and by condition are especially troubling.

First, the low percentage of studies that explicitly defined the construct of deep and surface processing (41.4%) is somewhat in line with what Murphy and Alexander (2000) found in the motivation literature (38%) and with what Dinsmore et al. (2008) found with regard to metacognition, self-regulation and self-regulated learning (49%). Dinsmore et al. (2008) suggested that the longevity of metacognition (back to the early 1970s) may have explained the relatively low percentage of explicit definition for metacognition during the five-year time period from which they tabled their studies (i.e., 2003-2007).

Figure 1

*Percentages of Each Definition Category Coded from the Tabled Studies*



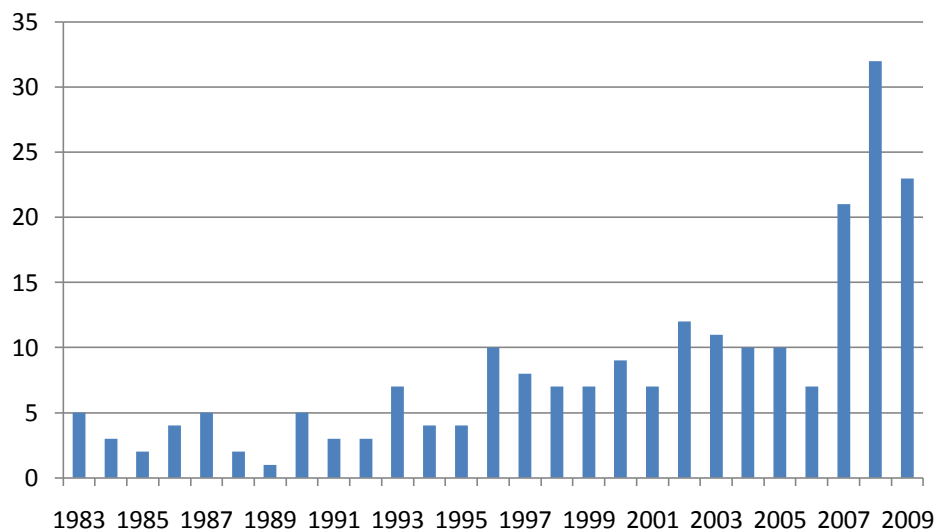
The relative longevity of a construct may also explain the lack of explicit definitions for deep and surface processing, particularly the absent definitions. A perusal of the table

demonstrates that, prior to 2000, only one study was coded as *absent* in contrast to 18 studies coded as *absent* after that date. Perhaps it is the case that as deep and processing became more common in the literature, the perceived need to provide a definition became less apparent (as Figure 2 shows, the prevalence of deep and surface processing has been increasing since its inception).

As Murphy and Alexander (2000) pointed out, implicit or absent definitions place a heavy burden on the reader, particularly those readers with less prior experience in the field. Hence, it is difficult to understand the interpretations placed on the resulting data if it is not clear exactly how the underlying constructs are conceptualized, thereby exacerbating the problem of interpreting and organizing research findings across studies. This is certainly a possibility for the ambiguity of the relation between deep and surface processing and learning outcomes.

Figure 2

*Studies of Deep and Surface Processing by Year*



Second, over half of the studies only defined deep and surface processing by the measure used or by the experimental condition described. As with the longevity of the construct itself, it is possible that as some of the measures are used more often (e.g., the Strategy Process Questionnaire; Biggs, 1979), these measure are deemed a proxy for a conceptual definition. While this is certainly better than no definition, it is still incumbent upon the researcher to demonstrate that there is a match between the conceptual definition of the construct and the measure chosen (e.g., Schunk, 2008). Similarly, the experimental conditions used as proxies for definitions (e.g., focusing participants on the physical features of a word for surface processing and semantic features of word for deep processing; Toppino, Fearnow-Kenney, Kiepert, & Teremula, 2009), although common, demonstrate no such conceptual-operational link. Thus, not only is it likely that the lack of explicit definitions is leading to ambiguity, but also the heavy reliance on defining the construct by measure may also be contributing to the ambiguous results.

***Definitional content.*** In addition to the explicitness of definitions, the actual content of the definition was examined in cases where one was explicitly provided. This examination was done by searching each explicit definition (where provided) for recurring keywords in the definitions. From an initial examination of these definitions certain keywords appeared multiple times. These keywords included: *intention, approach, style, aim, focus, effort, orientation, motivation, adoption, memory, encoding, strategies, deep-level strategies, and surface-level strategies*. Of the 92 studies that explicitly defined deep and surface processing, the percentages of definitions that contained these keywords are shown in Table 2. This table demonstrates the higher



frequency of the terms *intention*, *approach*, *focus*, *motivation*, *adopt*, *memory*, and *strategies*. The prevalence of these terms will be discussed in regard to the three theoretical frames previously identified: AL, the MDL, and information processing theory (IPT).

Given that AL, along with Marton and Saljo's (1976a; 1976b) intentionality to learn paper, espouse deep and surface processing as an approach or intention to learn, it was not surprising that the terms *intention* and *approach* appeared so frequently in the definitions. Since AL tends to assume that an individual's approach or intention is relatively stable, the higher percentage of the terms *focus* and *adopt* also supports the idea that AL is common framework in these studies. Quite common in these definitions was the deep *or* surface processing distinction.

Table 2  
*Definitional Keywords in Explicit Studies by Percentage*

Keyword	Percentage*
intention	17.4
approach	34.8
style	2.2
aim	4.3
focus	17.4
effort	5.4
orientation	3.3
motivation	19.6
adopt	15.2
memory	12.0
encode	6.5
strategies	23.9
deep-level strategies	3.3
surface-level strategies	3.3

*Note.* \*n = 92 studies.

For example, Chamorro-Premuzic, Furnham, and Lewis (2007) defined deep and surface processing as an *approach*:

A deep approach to learning is characterized by intrinsic motivation, engagement with the subject matter, and the desire to know everything about a given topic. Conversely, students who opt for a surface approach to learning are not interested in the task per se, but aim at learning the minimum amount of material required to pass. (p. 242)

In other words, these definitions dichotomize deep and surface processing so that learners are at least stable in their orientation to tasks in general.

Also not surprising was the occurrence of the term *strategy*. Given that models, such as the MDL, differentiate strategies based on levels of processing, this was expected. However, in examining some of the framing more closely, it was revealed that studies that cited AL as a framework (either explicitly or by reference) often used the term *strategy*. For example, Brown, Aoshima, Bolen, Chia, and Kohyama (2007) define surface learners as those who “typically utilize rote memory as a learning strategy” (p. 593). Here, the term *strategy* is being used in a very broad and less precise sense than in the strategy literature. On the other hand, Alexander (1997) defines strategies in the MDL as processes “purposefully invoked when a learner wishes to maximize performance or to circumvent problem in understanding or learning” (p. 223). In other words, a strategy is *not* an overall approach, rather a very deliberate process to achieve some end. After closer examination of the theoretical frames used (i.e., AL, the MDL, and IPT), it was apparent that the terms *deep-level strategies* and *surface-level strategies* used in the definitions were more indicative of theoretical frames like the MDL. As the

table shows, these percentages represented a relatively small number of the total studies (about 3%).

The last term that emerged quite often in these explicit definitions was *memory*. Upon closer examination of these studies, it was evident that information processing theory (IPT), particularly the work of Craik and colleagues (e.g., Craik & Lockhart, 1972), framed many of the studies in the table. For example, Osorio, Ballesteros, Fay, and Pouthas (2009) using Craik and Lockhart's (1972) framing defined deep and levels of processing as deep and shallow encoding this way:

...shallow encoding (processing words based on their orthographic or phonemic components) leads to a fragile memory trace susceptible to rapid forgetting. In contrast, deep encoding (semantic or meaning based processing) leads to a more durable memory trace. (p. 57)

As with AL, there is again a dichotomization of approach here, but it is more narrowly constrained to encoding of information.

Given the differences in the three frameworks identified here (i.e., AL, the MDL, and IPT) and the subsequent differences in the explicit definitions proffered, the need for explicitness of definitions is exacerbated. Researchers and readers alike *cannot* simply assume that deep and surface processing (or a term like *strategies* for that matter) means only one thing. That being said, there should also be some agreement on what the terms mean and these labels need to be used in very precise ways.

Not only is it a problem that these definitions vary so widely, there may also be a problem in *how* they vary. As demonstrated earlier in some of the definitions used in the AL framing, there is the presumption that a learner adopts one approach or the other (i.e.,

a surface or deep approach). On the other hand, according to the MDL, there is no such presumption that learners would only use deep-level strategies or surface-level strategies. In fact, a learner would use a combination of these strategies to accomplish a task. This distinction between the two conceptualizations is nontrivial and the dangers of dichotomizing deep and surface processing parallel those dangers that Thorndike and Gates (1929) discussed in making distinct categories:

...it should be noted that the children of a given age or in a given class do not fall into two or more separate groups in any ability. When the teacher says: "I have two kinds of workers, quick and careless, average workers, and the slow and sure" or when the psychologist, after applying an intelligence test, divides the kindergarten children into idiots, imbeciles, morons, the dull, the average, the superior, and the geniuses, they are misstating the facts. These divisions, which do not exist in nature, are artificial classes—mere matters of definition. They are justified by convenience of expression when one fully realizes the fact that in each trait or combination of traits there is only one type—the average—from which individuals vary by steps imperceptibly minute. (pp. 218-219)

While researchers and teachers may talk about students who typically use more or less deep-level strategies, the conceptualization of deep and surface processing must be more precise, otherwise there is a risk of deleterious effects on the measurement of deep and surface processing which will be discussed next.

## **Operationalization**

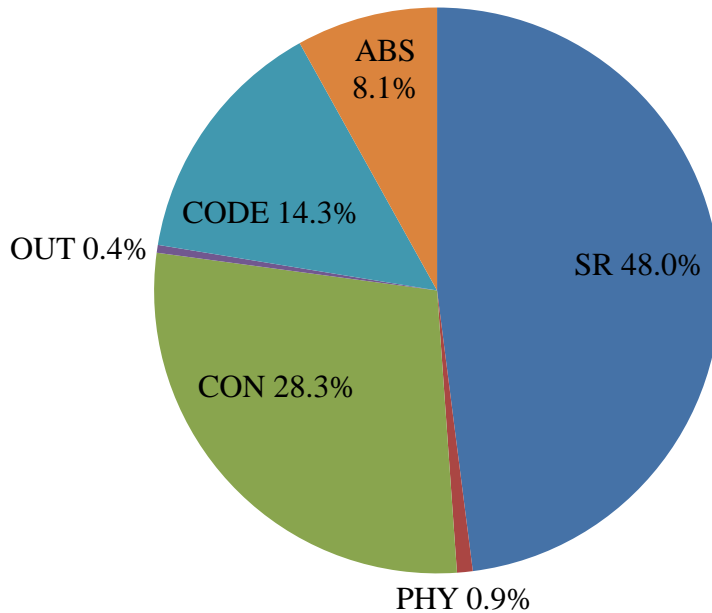
The conceptualizations described previously should relate to the measurement used (e.g., APA, 2010; Schunk, 2008). The types of measures and the validity evidence provided by these measures were investigated to see if there might be any potential for these issues to impact the inconsistent findings between deep and surface processing and learning outcomes.

*Types of measures.* In addition to the coding categories outlined in the a priori coding scheme section (i.e., self-report questionnaire, by condition, coding scheme, physiological response, or no measure), a sixth category was needed, which was labeled *outcome*. Here deep or surface processing was inferred from some outcome measure (i.e., the product of learning) instead of from a process. For example, Sanchez, Garcia-Rodicio, and Acuna (2009) used reading comprehension outcomes only to assess depth of processing. The percentages for each of these five measures in the pooled studies are presented in Figure 3.

Not surprisingly, and in line with reviews of other constructs (e.g., Dinsmore et al. 2008), self-report questionnaires accounted for nearly half of the all the measurements used. This category encompassed numerous self-report measures of deep and surface processing. These included both established measures, such as the Inventory of Learning Processes (ILP; e.g., Gadzella & Masten, 1998), the Strategic Processing Questionnaire (SPQ; e.g., Furnham, Christopher, Gardwood, & Martin, 2008), and the Motivated Strategies for Learning

Figure 3

*Percentages of Each Measure Category Coded from the Tabled Studies*



*Note.* SR = self-report; CON = by condition; CODE = coding scheme; ABS = no operationalization; PHY = physiological response; and OUT = outcome measure.

Questionnaire (MSLQ; e.g., Lau, Liem, & Nie, 2008), as well as numerous researcher-developed questionnaires (e.g., Senko & Miles, 2008). It is likely that the heavy reliance on self-report questionnaires is due to the relative ease of collecting such data from a large amount of participants over some of the other types of measurements subsequently discussed.

The second most common category was measurement of deep and surface processing by condition. In these studies, the instructions to the participants or the task itself were assumed to induce deep or surface processing. For instance, Ally et al. (2009) measured the effects of deep and surface processing on word recall by instructing participants to indicate if a word had more or less than three syllables in a surface

processing condition and whether they liked or disliked a word in a deep processing condition. The high occurrence of method to investigate deep and surface processing seems to be linked back to the frame used, as these studies tended to cite IPT (such as Craik and colleagues) as their theoretical frame (e.g., Gadzella, Ginther, & Williamson, 1986; Hinojosa, Martín-Loeches, Muñoz, Casado, & Pozo, 2004).

Less occurring measures included coding schemes, physiological measures, and absent measures. The coding scheme category encompassed studies that examined both concurrent reporting (i.e., verbal reports *during* a task) and retrospective reporting (i.e., verbal reports *following* a task). For example, Ferm and Johansen (2008) coded retrospective interviews of music students' experiences over the past semester. An example of concurrent coding can be found in Fergusson-Hessler and de Jong's (1990) investigation that examined the reading of physics texts using a concurrent think-aloud protocol. The two smallest categories were those that measured deep and surface processing using a physiological task (e.g., eye tracking; Klein, Zwickel, Prinz, & Frith, 2009) or by using an outcome task such as a reading comprehension measure (e.g., Sanchez et al., 2009). Lastly, there were some studies where there was no apparent operationalization of deep and surface processing (e.g., Pearce & Lee, 2009).

The heavy reliance on self-report questionnaires to measure cognitive processing presents two problems. First, as Dinsmore et al. (2008) point out in regards to metacognition, self-regulation, and self-regulated learning, reporting one's own metacognition actually requires one to be meta-metacognitive. In other words, students have to think about their thinking of their thinking. In the case of deep and surface processing, accurately reporting one's strategic processing requires that an individual is

metacognitive. Participants who do not monitor their use of strategies well are not likely to accurately report which strategies they used and which strategies they did not use. Since errors in measurement attenuate relations between variables (e.g., Pedhazur, 1997), errors of measurement caused by the inaccurate measurement of deep and surface processing would be likely to attenuate the relation between deep and surface processing and learning outcomes.

However, the self-report of strategies can be useful in measuring processing. At this point in time, uncovering covert mental processing (as well as attitudes) can *only* be uncovered using self-reports. At issue here is how the self-report is elicited, particularly the metacognitive awareness required to assess one's own cognition and metacognition with these existing self-report questionnaires. Measures such as concurrent reporting and retrospective reporting (another form of self-report) are situation specific and do not require an extra layer of metacognitive awareness of the participant. Further, concurrent reporting has one major advantage over retrospective reporting in that it is not affected by the memory ability of the person who is doing the reporting.

Second, while errors in measurement may attenuate the relations between levels of processing and learning outcomes, dependence of observations may be artificially increasing the relation between levels of processing and learning outcomes. For example, Bandalos, Finney, and Geske (2003) correlated a self-report questionnaire of deep learning strategies and participants' learning goals, also measured by self-report questionnaire. Since both self-report questionnaires were completed by the same individual, these two measurements do not meet the assumption of independence, an assumption for regression (Pedhazur, 1997).



The condition category has a similar advantage to self-report, in that varying deep and surface processing by giving participants different instructions is relatively easy to carry out. Unfortunately, this approach makes two major assumptions. The first assumption is that the participants follow the given instructions (e.g., process words in a word list given their like or dislike of a word). The second, and the more troubling assumption, is whether these instructions *actually* induce deeper or shallower processing. Without a process measure of some sort, it is impossible to know if the directions in each of the conditions had the intended effect. In fact, this issue has been controversial for a long period of time (e.g., Lockhart & Craik, 1990; Nelson, 1977).

The last issue identified with the measures and measurement of deep and surface processing relates back to the previous discussion of dichotomization in the conceptualization of the construct. It is no surprise then that these measures are often used to place participants in a certain category (i.e., a deep processor or a surface processor), particularly those studies defining deep and surface processing using the AL identified terms. For instance, Campbell, Smith, Boulton-Lewis, Brownlee, Burnett, Carrington, and Purdie (2001) used a self-report questionnaire to do just that:

Selection of students was based on LPQ responses, such that, for each class, two students with high deep and low surface approaches (hereafter termed a “deep” approach to learning), and two students with high surface and low deep responses (hereafter termed a “surface” approach to learning), were chosen [for further analysis]. (p. 176)

For one, the selection of participants to groups based on extreme scores makes the research design vulnerable to the internal validity threat of regression (see Campbell and Stanley, 1963, for an extended discussion of the threat of regression).

The bigger issue here, however, was the dichotomization, whether through extreme scores or even a median split of scores on an outcome measure. One possible explanation for this strategy was the match to the conceptual definitions prevalent in these studies, particularly the AL frameworks. Although this match between conception and operationalization was good on one hand, if the conception was somehow flawed, so was the measurement strategy. Another explanation, and one that is complementary to the first, was that this strategy was chosen in order to fit an analysis of variance (ANOVA) test, such as a t or F test. This analysis strategy is disadvantageous in these situations because it, in essence, takes a continuous variable and turns it onto a categorical one, thereby reducing the amount of variance in the variable (Pedhazur, 1997). Since ANOVA is simply a special case of regression (Pedhazur, 1997), it makes more sense to keep the data continuous and analyze it as such.

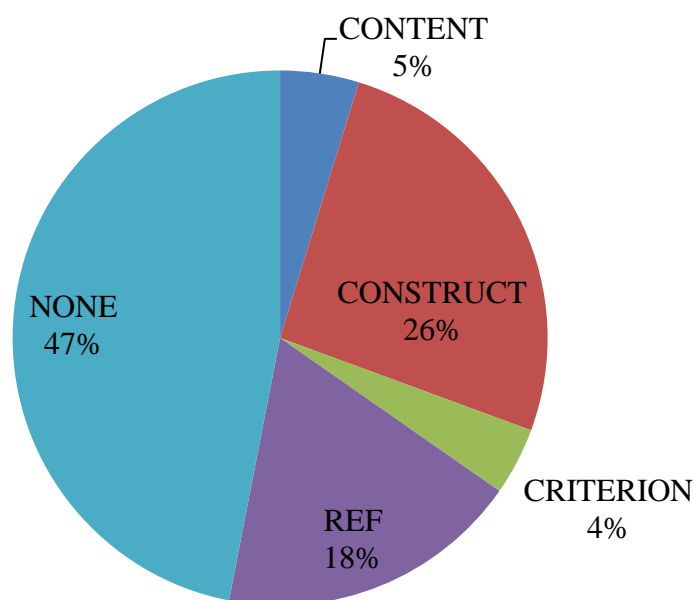
**Validity evidence.** Similar to the theoretical frames used and the explicitness of definitions, validity evidence was not often explicitly stated. In fact, for studies that included a processing measure (this excludes studies that manipulated processing by condition) the most notable aspect of this coding scheme (Figure 4) was that over 65% of the studies (whether explicit or implicit) did not provide any of the three forms of validity evidence for the current study (this includes references to previous studies).

The most common type of validity evidence when present was construct validity. This was not surprising as Messick (1981) called construct validity the heart of validity.

Typically, construct validation was undertaken using factor analysis (e.g., Nolen & Haladyna, 1990), which is a common approach for providing construct validation (Crocker & Algina, 1986). Both criterion and construct validity evidence occurred much less frequently.

Figure 4

*Percentages of Validity Evidence Coded from the Tabled Studies*



*Note.* REF = references to prior studies.

Multiple types of validity evidence are preferred (Crocker & Algina, 1986), but one type of evidence is certainly better than none. However, even when validity evidence was present (such as a factor analysis), it was often difficult to determine whether the evidence presented was *meant* to be validity evidence. As Schunk (2008) has pointed out in regard to metacognition, self-regulation, and self-regulated learning “it is incumbent

upon researchers to clearly explain how their measures are reliable and valid indicators of the variables they are attempting to study” (p. 466). In all likelihood, the paucity of validity evidence found in this review only adds to the problems associated with the measurement issues discussed in the previous section.

**Situational consideration.** The next issue considered is the research context, both in terms of whether there was a research task present and the academic domain in which deep and surface processing was being investigated.

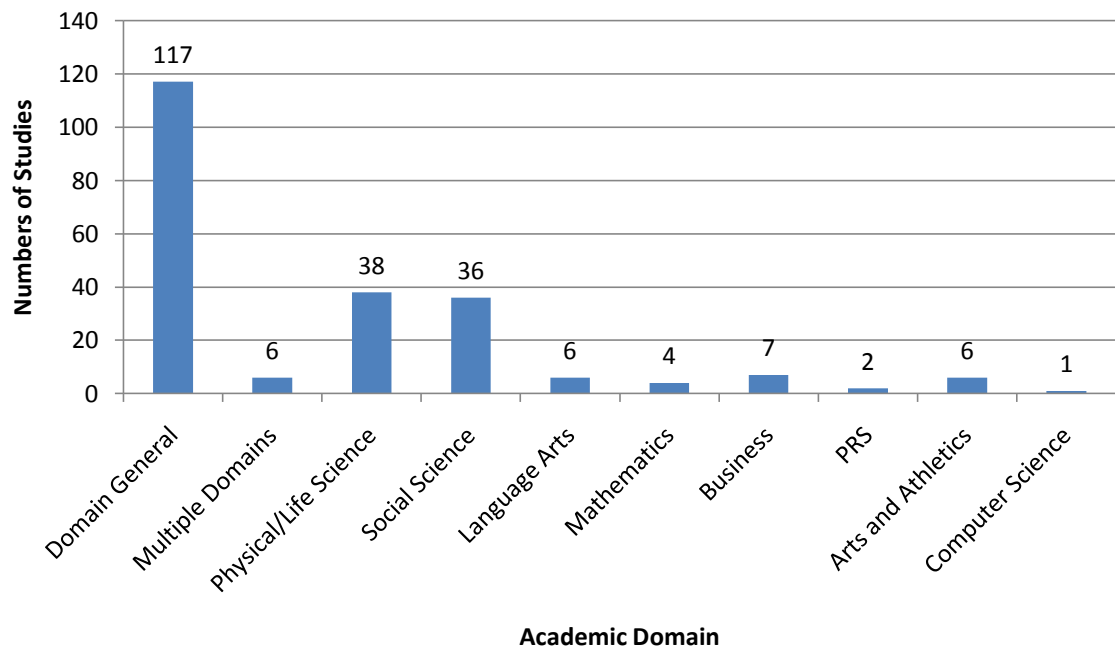
Nearly 40% of the studies tabled did not include a research task (39.5%; coded as NT in Appendix A). In these studies, the situation (or research task) either was considered through instructions in a questionnaire (e.g., Rodriguez, 2009) or was not considered at all (e.g., Coutinho & Neuman, 2008). Either way, the role of the specific situational variables during deep and surface processing was underrepresented. There is a reasonable expectation that deep and surface processing should differ depending on the situation, as learning in general is expected to differ depending on the characteristics of that situation (e.g., Alexander et al., 2009).

Figure 5 presents the number of studies in each domain category (i.e., domain general, multiple domains, physical/life science, social science, language arts, mathematics, business, philosophy and religious studies, arts and athletics, and computer science). Over half of the tabled studies investigated deep and surface processing with no specific academic domain (i.e., domain general). Of those studies that did examine deep and surface processing in a specific domain, the physical and life sciences and social sciences were the most frequently occurring.

Given that both AL and the MDL take into account academic domain (Alexander, 1997; Biggs, 1978), it was surprising that there were so many studies that were domain general. This could be explained by the high number of studies in the table that define or identify their framework as some form of IPT, as IPT does not often consider the situation (e.g., Craik & Lockhart, 1972). However, given the high frequency of AL studies, it was surprising that this percentage was as high as it was (i.e., greater than 50 percent of the total studies). Another

Figure 5

*Percentages of Task Type Coded from the Tabled Studies*



*Note.* PRS = philosophy/religious studies.

explanation for this high percentage could be a high number of studies that correlated deep and surface processing with *general* academic performance (e.g., Cano, 2007). If,

in fact, the situation and academic domain *do* make a difference, studies that do not take these factors into account may be investigating a relation that does not exist (i.e., a relation between levels of process and general academic performance). The relations from these studies between levels of processing and learning outcomes may in fact be obscuring a more complex relation in which the relation between levels of processing and learning outcomes are variant in different situations.

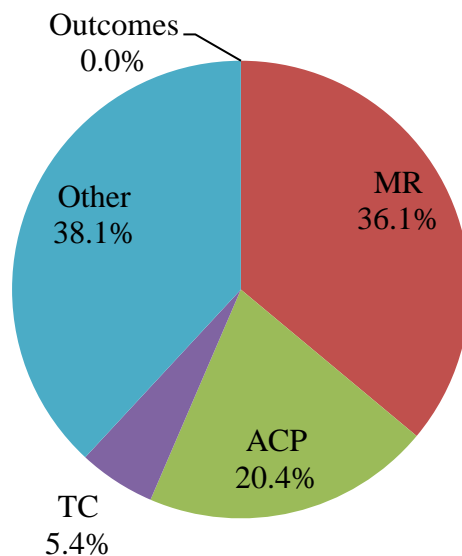
**Model specification.** The last of the four guiding questions concerned the endogenous (i.e., outcomes) and exogenous (i.e., predictors) related to deep and surface processing in the tabled studies. The outcomes and predictors for each of the tabled studies were identified and categories were developed. Figures 6 and 7 present the categories for the outcomes and predictors (respectively) that were over five percent, with the rest listed in Appendix A. In regards to outcomes related to deep and surface processing, memory/recall and academic/course performance had the largest percentages. The high frequency of memory and recall may be partially explained by the higher frequency of IPT frameworks (e.g., Toppino et al., 2009). Although, there are examples in the table of frames other than IPT that investigate outcomes related to memory and recall as well (e.g., Nolen, 1988). The emphasis on memory and recall may also be due to the fact these measures are seen to be relatively more objective than other outcome measures, such as text summaries (e.g., Sanchez et al., 2009).

In regards to the predictors included in the models of the tabled studies, the most numerous were achievement goals. Two possible explanations can be forwarded for this finding. First, since AL has been a rather enduring framework, achievement goals, particularly those that are seen as a stable orientation, match Biggs's (1978) model that

uses personality factors as predictors. Second, the high percentage may also coincide with the increase in interest in self-regulated learning beginning in the 1990s (Dinsmore et al., 2008). Particularly influential may have been Pintrich's (2000) chapter in the *Handbook of self-regulation* entitled "The Role of Goal Orientation in Self-Regulated Learning." The modest incidence of self efficacy as a predictor also may lend some support for this explanation as well (e.g., Meier, McCarthy, & Schmeck, 1984).

Figure 6

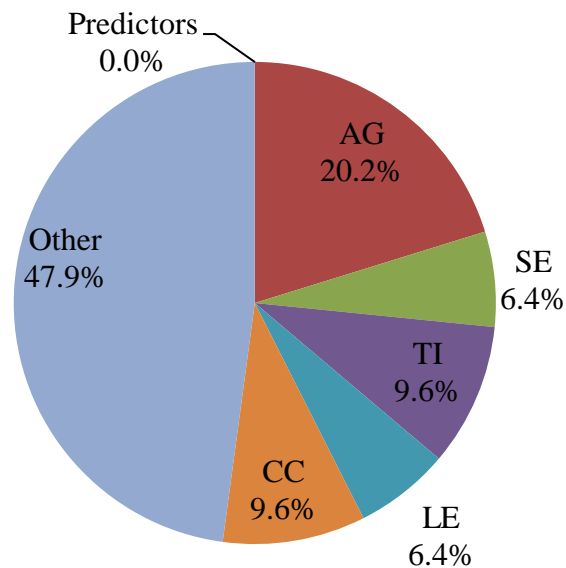
*Percentages of Outcome Categories from the Tabled Studies*



*Note.* MR = memory/recall; ACP = academic/course performance; TC = text comprehension.

Figure 7

*Percentages of Predictor Categories from the Tabled Studies*



*Note.* AG = achievement goals; SE = self efficacy; TI = teaching/instructional method; LE = learning environment; CC = cultural characteristics.

In light of the differences in both the predictors and the outcomes of deep and surface processing specified for these studies, it seems reasonable that the relation between levels of processing and learning outcomes has been consistent. However, one must wonder why in many cases motivation was chosen as the sole predictor. While motivation almost certainly plays a role in determining deep- versus surface-level processing, one must wonder about the role of cognitive processes, such as metacognitive processes. Further, the specification in a large number of these studies on memory begs the question; can we equate memory and meaningful learning? If the learning outcomes in the table are not equitable, then it should be no surprise that the relations found in these studies are inconsistent.



### **Concluding Remarks and Indications for Future Research**

Before drawing some conclusions across these four issues that were examined and suggesting future directions for research, it is worth reiterating that this review is meant to identify broad issues across the literature. There was no intention to rebuke individuals. Rather the codings were meant to identify reasons why the relation between deep and surface processing and learning outcomes has been ambiguous and inconsistent. To this end, one of the issues identified in the review, namely the precision and explicitness of definition and description, was also a difficulty in the coding process. For example, when validity evidence seemed to be present but was not made explicit, a judgment had to be made whether or not the evidence *was* in fact validity evidence. Using multiple raters with 90% agreement on a subset of these studies indicates that while we can be relatively confident about the overall findings, the subjectivity of this process may lead to some disagreement in several particular instances with a few of these studies.

Despite these difficulties, this review does serve as a broad overview to some issues that may be contributing to the ambiguity and inconsistency in the extant literature on levels of processing thus far. First, the inconsistencies in the relation between depth of processing and performance may be attributable to lack of conceptual clarity. If the definitions are not well specified, the measures of the construct and resulting interpretations are questionable. Also, since the definitions that were explicitly provided differed greatly from study to study, it is also reasonable to assume that the measures and interpretations across these studies *should* differ. Second, in regards to measurement of deep and surface processing two problems arose; the heavy reliance and problem of using

self-report questionnaires and the paucity of validity evidence provided for these measures. Third, situational considerations may have played a role in the inconsistencies. Differences were found across these studies in terms of whether or not a research task was employed (i.e., the presence of some research task) and which, if any, academic domain was under investigation. Lastly, the predictors and outcomes specified in the models of deep and surface processing varied widely in these studies. For instance, in terms of learning outcomes, differences range from simple recall of a face recognition task (e.g., Block, 2009) to more complex learning outcomes such as assessment portfolios (Baeten, Dochy, & Struyven, 2008).

Taken separately, any of these four issues may be seen as possible contributors to this inconsistency in the relation between levels of processing and learning outcomes in the extant literature. Taken together, it should not be surprising that there are inconsistencies in this relation between processing level and cognitive performance. However, it is also possible that these identified issues may not be problems at all. Instead, the problem may be the question that is being asked. The focus of much of this research seems to be on finding if deep processing or surface processing leads to more meaningful learning. Implicit here is the expectation that deep and surface processing are invariant over *who* is doing the processing (i.e., the characteristics of the learner), *where* the processing is taking place (i.e., the situation), *when* the processing occurs (i.e., the temporal aspect), and to *what* end they are employed for (i.e., the target of learning). These four dimensions proposed by Alexander et al. (2009), the *who*, *where*, *when*, and *what*, present a complex topographical perspective of learning in which each of these four dimensions (and their resulting interactions) play a pivotal role in learning.

If it is the case that these four dimensions *do* matter, as Alexander et al. (2009) and others contend, it changes the nature of the questions that should be asked. The inconsistent findings thus far may in fact be evidence that the relation between levels of processing and performance is *not* invariant across these four dimensions. In other words, individual characteristics (e.g., prior knowledge) and situation characteristics (e.g., type of text) may be moderating the relation between the use of deep- and surface-processing strategies and learning outcomes. This changes the question that should be asked from “what is the relation between levels of processing and learning outcomes,” to “what is the relation between levels of processing and learning outcomes for whom, at what point in development, in what situations, and for what end.” In other words, notions that deep processing is “good” and surface processing is “bad” are naive at best, and misleading at worst.

For research on deep and surface processing to have any bearing on practice, a number of issues need to be remedied. Most important to this aim is a strong theoretical rationale for the conceptualization and measurement of deep and surface processing. Definitions of deep- and surface-processing strategies must be explicitly defined, so either congruity in definitions can be reached, or at the very least there is acknowledgment of how the construct in question is being defined. This explicit conceptualization and the proceeding theoretical frame must be specified in such a way as to not obscure variability that does exist. In other words, there must be an acknowledgment that deep- *and* surface-processing strategies can coincide and that this may be advantageous to the learner in many situations.

The measurement and validity evidence provided must also be better specified. When multiple forms of validity evidence are possible, these multiple forms should be provided (Crocker & Algina, 1986). Validity evidence should be used to inform the choice of measure used. Further, the type of measurement (e.g., self-report questionnaire, concurrent reporting) needs to be carefully considered. For example, concurrent reporting has the advantage of not asking participants to be aware of the types of strategies they are using (such as a self-report questionnaire administered prior to a task) or relying on their memory (such as a self-report questionnaire administered after a task or retrospective reports such as semi-structured interviews). However, even concurrent reporting has certain limitations, such as the range of strategies that participants may be able to report. It is quite possible that participants may not be able to report complex strategies they are using because of heavier task demands these complex strategies may create. Nonetheless, these issues need to be acknowledged and weighed carefully when choosing a measure of deep- and surface-processing strategies.

Second, deep and surface processing should be systematically investigated in terms of the four dimensions of learning (i.e., the *who*, *where*, *when*, and *what*) to better determine the impact each of these four dimensions have on deep and surface processing or vice versa. The decision of what to systematically investigate in terms of these four dimensions needs to be informed by theory. For example, in terms of the *who*, models such as the MDL can guide which characteristics of the learner may affect their levels of processing (i.e., knowledge and interest). The work of Muis (2007) can be used to investigate if epistemic beliefs impact the use of deep and surface processing. In terms of the situation, there are many issues which need to be addressed here. Both AL (e.g.,

Biggs, 1978) and the MDL (e.g., Alexander, 1997) indicate that domain *does* make a difference. There should be systematic investigations of these differences both within and across domains. The cross-domain investigations of those in the table (e.g., Lehman & Schraw, 2002) are a good start. In terms of development, models should take into account change in both the learner and situation over time. This is where developmental models such as the MDL may be particularly useful. And finally, in terms of learning outcomes, there should be systematic investigations of what types of learning outcomes and assessments can aptly capture principled learning, not just encoding or recall of information. There is a great need for simultaneous investigation of learning outcomes and if in fact levels of processing are more or less advantageous for different types of outcomes. Whichever theoretical frames are chosen, it is their use and particularly the justification of their use (through critical discussion of the theoretical frames being tested) that will ultimately provide the necessary guidance to conceptualize deep and surface processing, measure deep and surface processing validly, and interpret the data from these studies to inform practice.

The implications that this research agenda has on practice are numerous. These findings will undoubtedly guide decisions of what strategies to model (e.g., deep versus surface), when to model them, where to model them, and which outcomes to model them for. It is not enough to say that deep processing or deep-processing strategies are better, and therefore those should be the only ones taught or encouraged. Rather, there needs to be an effort to understand the situations in which one or a combination of both types of strategies is useful. Through this effort instruction can focus on guiding students to use particular strategies in *appropriate* situations. This type of instruction is more likely to

encourage the type of thinking that Dewey (1910/1997) promoted in *How We Think*. Here, Dewey advocated that it is one's working habits, methods of inquiry, and reasoning appropriate to the various situations students may encounter that should be the focus of education.

### **Instantiating an Investigation of Deep and Surface Processing with Text**

As discussed throughout this chapter, this literature review was intended to probe the evidence in the extant literature about the relation between depth of processing and performance; why there may be confounding evidence in the extant literature as to the relation of processing level and performance, and what other gaps may exist. The following section illustrates one such study that builds on previous research and investigates gaps in the current literature. This empirical study is described in further detail in subsequent chapters and is but one study in a program of study needed to develop sufficient evidence as to the nature of the relation between depth of processing and performance. The following sections specify how the current study builds on existing research and explores gaps in the extant literature on deep and surface processing in terms of the theoretical framing and measures of deep and surface processing.

### **Theoretical Framing**

As alluded to earlier, any investigation of deep and surface processing should be situated within a strong theoretical framework. The theoretical framework acts as both a lens from which to view the construct and can offer guidance as to hypothesized relations that may exist in the hypothetical model. In the case of the current study, three guiding frames served these functions: a topographical perspective of learning (Alexander, et al, 2009), the Model of Domain Learning (MDL; Alexander, 1997), and the Construction-

Integration Model of Text Comprehension (CIM; Kintsch, 2004, Kintsch & van Dijk, 1978).

**The topographical perspective.** The first of these frames, the topographical perspective, serves as one possible guidepost to the dimensions of learning that should concern us. In other words, this perspective helps us to specify models that investigate the relation between depth of processing and performance. Alexander, et al (2009) postulated four dimensions of learning: the *who*, *what*, *where*, and *when*. In the current study, the *who*, *what*, and *where* will be addressed. Although the *when* is not addressed in the current study, broad programs of research should include the temporal nature of learning as well.

The *who* refers to the individual characteristics of the learner (e.g., subject-matter knowledge, interest, and epistemic beliefs). Including the individual characteristics of the learner are essential for exploring learning because they allow a look at individual differences that are often ignored (Pedhazur, 1997). The range of individual differences that could potentially play a role in moderating the relation between depth of processing and performance are many. Based on previous research (e.g., Alexander, Sperl, Buehl, Fives, & Chiu, 2004; Murphy & Alexander, 2002), prior knowledge was chosen as a possible mediator in the relation between depth of processing and performance (research question 2). Further explication of this choice is described in the MDL section to follow.

The *what* refers to the object of learning, in other words, what is to be learned. Past research that examined the relation between depth of processing and performance primarily looked at the object of learning (i.e., the *what*) as recall (e.g., Ally et al., 2009; Burgess & Weaver, 2003). This was also the case with studies of text processing in

which the object of learning was passage recall (i.e., ability to reproduce or recall text; e.g. Murphy & Alexander, 2002). While, there were certainly a few studies that examined open-ended tasks (e.g., Boekaerts 2002; Sullivan & Cantwell, 1999), only one study reviewed considered recall tasks and open-ended tasks simultaneously (Lehman & Schraw, 2002), although these tasks were confounded with measures of processing in that particular study. Hence, this study seeks to directly compare the relation between two different target of learning, both recall and an open-ended task (research question 1).

The *where* refers to the situation in which learning takes place (Alexander et al., 2009). The texts presented in this study are an aspect of the *where* because they are part of the physical environment that influences learning (specifically tested in the current study by manipulating the type of text). As mentioned, a program of research is necessary to investigate how and to what extent the relation between depth of processing and performance differs as a consequence of individual and situational factors, since the range of possible tasks is endless. This study investigates one area where such an interaction may occur, namely, the type of text.

Learning from text was chosen for this study because it is one of the primary ways that students are expected to learn over the course of a semester. However, investigations of course material text comprehension suggest that students do not comprehend typical course texts very well (e.g., Fox, Dinsmore, Maggioni, & Alexander, 2009). One promising avenue to help readers process text more deeply is the use of persuasive text (e.g., Dinsmore et al, 2009). This investigation seeks to examine whether the situation (i.e., expository or persuasive text) moderates the relation between depth of processing and performance (research question 3).



**The Model of Domain Learning (MDL).** Situated within the topographical perspective, the MDL (Alexander, 1997, 2004) serves as a useful framework to guide selection of variables that constitute important individual characteristics and hypothesize the relation between knowledge and strategic processing. The MDL is a developmental model of expertise that describes three forces (i.e., knowledge, interest, and strategic processing) and three stages (i.e., acclimation, competence, and proficiency) of expertise within an academic domain.

The first of these forces, knowledge, is composed of two types of knowledge, domain and topic knowledge. Domain knowledge refers to the breadth and scope of subject-matter knowledge, while topic knowledge refers to the depth of knowledge about specific concepts related to a domain (Alexander et al., 1997). Similarly, interest is composed of two types as well, individual and situational interest. Individual interest refers to “an enduring predisposition to reengage particular contents over time” (Hidi & Renninger, 2006, p. 111), while situational interest refers to “focused attention and the affective reaction that is triggered in the moment by environmental stimuli, which may or may not last over time” (Hidi & Renninger, p. 113).

Finally, strategic processing is likewise composed of two types, surface-level strategies and deep-level strategies. Surface processing is the use of strategic and monitoring behavior related to the encoding of textual content. By contrast, *deep processing* is the use of strategic and monitoring behavior that involve a more extensive manipulation of a task or text (Alexander et al., 2004; Nolen & Haladyna, 1990).

These three forces are hypothesized to change over the course of expertise development within an academic domain. In acclimation, individuals are hypothesized to

have low levels of domain knowledge and rely primarily on surface-level strategies. As individuals progress through competence to proficiency, they build domain knowledge (as well as topic knowledge) and begin to rely more on deep-level strategies. This relation between knowledge and processing has particular implications for the relation between depth of processing and task performance. Specifically, it is likely that individuals must leverage their prior knowledge in order for deeper-level strategies to be effective for tasks in a domain.

Subject-matter knowledge is not meant to be completely representative of the “who,” rather, one indicator of the who dimension. Subject-matter knowledge is hypothesized to be one individual difference variable that may moderate the relation between levels of processing and learning outcomes. Additionally, the MDL is a useful framework for the studying the relation between levels of processing and outcomes because the MDL specifically addresses how levels of processing interact with subject matter knowledge. The MDL hypothesizes the co-occurrence of lower levels of subject matter knowledge with a greater reliance on surface-level processes (i.e., a relatively high proportion of surface-level processes to deep-level processes) in acclimation. Conversely, those individuals in proficiency with higher levels of subject-matter knowledge would be expected to rely on a larger proportion of deep-level processes during a task.

**The Construction-Integration Model (CIM).** Since learning in the current study uses a reading task, there must also be some model to guide how strategies are defined at the text comprehension level. The CIM is helpful in differentiating deep- and surface-level strategies during reading comprehension. The CIM distinguishes between

two levels of representation: the textbase and the situation model. The textbase, defined as the semantic representation of a text, includes both the microstructure (i.e., propositions in the text that represent meaning) and the macrostructure (i.e., global organization of the microstructure in the text; Kintsch, 2004). On the other hand, the situation model represents the integration of the text with a reader's prior knowledge (Kintsch, 2004).

The CIM is a useful framework in distinguishing surface- and deep-level strategies during text comprehension because the definitions of the textbase and situation model parallel those of general deep- and surface processing strategies defined previously (i.e., surface processing-strategies as the use of strategic and monitoring behaviors related to the encoding of textual content and deep-processing strategies as the use of strategic and monitoring behavior that involve a more extensive manipulation of a task or text). In fact, Kintsch (2004) even refers to the mental representations a reader forms from the textbase as "surface-level memory" (p. 1273). These distinctions can aid in discerning what one classifies as deep-level strategies and surface-level strategies. Further, the CIM was chosen because it delineates two types of processes: a bottom-up process (representation of the textbase) that aligns with the conceptual definition of surface-level processes and a top-down process (construction of the situation model) that aligns to the conceptual definition of deep-level processes.

### **Measuring Deep and Surface Processing**

The other main issue discussed in the review was how deep and surface processing was measured and what validity evidence was provided for that measure. Of the measures and measurements reviewed, there is certainly no perfect measure of

strategic processing, however, a stronger validity argument can be made for verbal protocols of reading. Think-aloud protocols (e.g., Ericsson, 2006) and more specifically verbal protocols of reading (VPR; e.g., Pressley & Afflerbach, 1995) elicit cognitive and metacognitive processes. Since these processes are covert, it is necessary in some way to have individuals make these processes overt. VPR elicits utterances of what readers are doing concurrently, that is *while* reading.

Most of the studies reviewed (particularly those that aligned with approaches to learning) used questionnaires prior to the task. In this case, concurrent VPRs align more closely to the definition proffered here in that depth of processing represents deep-level and surface-level strategies used during a task, providing some construct validity to this investigation. Concurrent VPRs are also advantageous in this instance to strategic checklists, questionnaires, or interviews conducted subsequent to a task because these measures and measurements rely on an individual's memory of their task performance, which may not reflect what they actually did (Ericsson, 2006). In Pressley and Afflerbach's (1995) view, they stated three advantages of VPs, namely:

first, it can provide data on cognitive processes and reader responses that otherwise could be investigated only indirectly; second, verbal reports sometimes can provide access to the reasoning processes underlying sophisticated cognition, response, and decision making; third, verbal reports allow of the analysis of affective processes of reading in addition to (or in relation to) cognitive processes. (p. 4).

Additionally, VPRs allow for different coding schemes at different grain sizes. In this particular investigation, the 30 reading behaviors described by Pressley and Afflerbach

(1995) were used along with Kintsch's (2004) construction-integration model to lend more construct validity to this measurement. The CIM could be used to broaden the grain size of the investigation to three types of strategies: deep level, surface level, and other. This procedure is described in greater detail in chapter three.

However, the use of concurrent VPRs is not without their limitations. First, VPRs may elicit only a certain range of strategic processes. It is possible that VPRs may underreport some surface-level strategies and some complex deep-level strategies. Second, although coding schemes can be carefully constructed, there is undoubtedly some subjectivity in coding. The need to build reliability during the coding process does take time and may limit wider applications of VPRs. However, in this investigation, which is concerned with the interaction of subject-matter knowledge and type of text on the relation between depth of processing and performance, the complexity of coding was favored since it was considered a more valid measurement.

## **CHAPTER III**

### **METHODOLOGY**

This section describes the participants, materials (i.e., text passages), measures (i.e., demographics, subject-matter knowledge, passage recall, and beliefs about ETCs), and measurements (i.e., concurrent think-aloud protocol and open-ended responses) that were the basis for this study of the relation between depth of processing and performance. These are followed by the procedures employed in conducting the study, as well an overview of the data analyses plan that was pursued.

#### **Participants**

The participants for this study consisted of 151 (109 female) undergraduates from a large mid-Atlantic university. The sample consisted of 27 freshman, 14 sophomores, 59 juniors, 45 seniors, and 6 participants who reported more than four years of undergraduate education with an average age of 20.55 ( $SD = 2.38$ ). These undergraduates were ethnically diverse (61.6% Caucasian, 18.5% Asian, 12.6% African-American, 2% Hispanic, .7% American Indian, and 4.6% Other) and reported a wide variety of academic majors. These majors consisted of primarily physical and life sciences and the social sciences (34.4% and 57.0% respectively). Students reported their grade point average (GPA) which had a mean of 3.31 ( $SD = .42$ ) for this sample. Participants were recruited from numerous courses on campus. All participants consented to participate in the study and all participants but one consented to be audiotaped. Participants were either offered extra course credit or compensated \$10 (US) for their participation.

## Materials

The materials for this study consisted of two text passages (Appendix B); an expository text meant to inform and a two-sided refutational text meant to persuade. A persuasive text can be defined as a message in which the author's intent is to change how an individual views an issue (Buehl et al., 2001; Kamalski et al., 2008). The passages were adapted from Stephen Webb's (2002) book, *Where is Everybody? Fifty Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life*. This topic was selected because individuals with and without prior knowledge of the subject would be able to respond meaningfully to the prompt, since it was likely that participants would at the very least have some folk knowledge about the topic and at least a moderate level of situational interest for the topic. Situational interest has been described a momentary affective reaction stimulated by the situation which may or may not last over time (Hidi & Renninger, 2006).

Both texts addressed the possibility of the existence of intelligent extraterrestrial civilizations (ETCs) in the galaxy. The main difference between the expository and refutational texts was the addition of the persuasive content in the two-sided refutational text. The two-sided refutational text contained 10 sentences that were intended to persuade the reader into adopting the view that ETCs do not exist and refute the other two views (i.e., that ETCs exist and are here, or that ETCs exist but have not yet communicated). For example, in support of the notion that species might extinguish themselves through nuclear war or another catastrophe, the passage stated, "the rate at which human activity is wiping out other species supports the idea that we may indeed extinguish ourselves before we are able to communicate with other ETCs." The passage

was specifically designed to persuade the reader that ETCs do not exist because this was the least prevalent view as a majority of Americans believe that ETCs do exist *and* that they have or are currently visiting Earth (Webb, 2002).

In order to ensure that the two texts were equivalent with regard to length and difficulty, additional sentences were added to the expository passage. The 10 additional sentences, while relevant to some part of the passage, did not change the nature of the ideas or concepts presented. For example, in the expository passage after the sentence, “Perhaps ETCs are using other types of signals such as gravitational signals, particle signals like neutrinos, or hypothetical tachyon signals that we may not be able to detect or interpret,” one added sentence read, “Tachyons are any theoretical particles that have an imaginary mass.” The added sentence gave an additional detail about tachyons that did not affect the overall premise or argument that the passage made. These sentences were added to make each of the passages as comparable as possible while differentiating the genre of each passage (i.e., expository and persuasive). Table 3 presents data that demonstrate that the passages are equivalent in terms of length and difficulty.

The same existing text was used for the expository and persuasive texts to ensure they shared common ideas between them. The only difference between the texts was that sentences were added to the persuasive text to create an explicit argument. To keep the length and difficulty consistent between the passages the same number of sentences was added to the expository text. These sentences were expository in nature and did not add new ideas, only elaborated on existing ones.

Reading experts ( $n = 3$ ) and content experts (astronomy and physics;  $n = 2$ ) were asked to judge the equivalence and suitability of the passages and their contents. These



external evaluators determined that: (a) the texts were equivalent in regard to length and difficulty, (b) the premise in the expository text was a description of three views about the Fermi paradox, (c) there was an argument in the two-sided refutational text that the view that ETCs do not exist is the most likely view, and (d) the passages were accurate with a minor modification about the number of planets in other solar systems discovered due to the continuing new discoveries by the Kepler spacecraft.

Table 3

*Length and Difficulty Data for the Expository and Persuasive Text Passages*

Measure	Expository	Two-sided refutational
Word count	1799.00	1850.00
Number of sentences	75.00	75.00
Flesch reading ease	47.00	45.80
Flesch-Kincaid grade level	11.80	12.20

### Measures

The measures for the study consisted of: a demographics measure; a subject-matter knowledge test; a prior beliefs about intelligent Extraterrestrial Civilizations (ETCs) measure; a post-passage measure beliefs about intelligent Extraterrestrial Civilizations measure, and a passage recall measure.

### Demographics

The demographics measure (Appendix C) consisted of questions pertaining to participants' gender, age, ethnicity, year in school (e.g., freshman), grade point average (GPA), and academic major.

### **Subject-Matter Knowledge**

The subject-matter knowledge measure (Appendix D) was designed to quantify participants' domain and topic knowledge in astronomy and astrobiology, respectively. This measure consisted of 16 total multiple-choice items, ten multiple-choice items measuring their domain knowledge of astronomy (i.e., their breadth of knowledge about astronomy) and six multiple-choice items measuring their topic knowledge of extraterrestrial life (i.e., their depth of knowledge about the study of the origin, evolution, distribution, and future of life in the universe; NASA, 2010).

This measure was constructed by sampling concepts from textbooks about astronomy and astrobiology (Jones, 2004; Lunine, 2005; Shostak, 2003). Domain knowledge items developed for inclusion in the measure were concepts about astronomy common across the astronomy textbooks. These items measured the breadth of individuals' knowledge about astronomy. Topic knowledge items developed for inclusion in the measure were concepts common across chapters in the astronomy textbooks about astrobiology and concepts found in astrobiology textbooks. These items measured the depth of individuals' knowledge about astrobiology (a topic within the domain of astronomy).

The response model for the multiple-choice items in this measure was a graduated response model (Alexander, Murphy, & Kulikowich, 1998). A graduated response model was used in order to discriminate between individuals' levels of knowledge about the targeted domain or topic (i.e., astronomy and astrobiology). A sample domain knowledge item with the associated points presented in parentheses was:

The lowest-energy photons of all electromagnetic radiation are:

- a. radio waves (4) – domain correct (astronomy)

- b. gamma rays (2) – domain incorrect (astronomy)
- c. tidal waves (1) – non-astronomy science incorrect (oceanography)
- d. stun rays (0) – misconception or popular lore (science fiction)

Answer choice (a) was the correct response and was the response choice most likely to be selected by someone with high levels of domain knowledge. Answer choice (b) was an incorrect response but was still within the domain of astronomy (they are higher energy photons than radio waves) and was likely to be selected by someone with at least some domain knowledge of astronomy. Answer choice (c) was an incorrect response from a different domain of science, in this case from oceanography. Answer choice (d) was an incorrect response and was either a common misconception held by someone who was a layperson in the domain or popular lore. In the example, answer choice (d) was from science fiction. All items on this measure followed the same distracter model.

The six topic knowledge items were in a similar graduated-response format. A topic correct response was scored a four, a topic incorrect response was scored a two, a non-astronomy science incorrect response was scored a one, and a common misconception or popular lore response was scored a zero. All item responses for both the domain knowledge and topic knowledge questions were randomly ordered.

Content validity was established for this measure using the four steps suggested by Crocker and Algina (1986). First, the performance domains (i.e., astronomy and astrobiology) were defined. Second, items were sent to experts ( $n = 3$ ) in the domains of astronomy, physics, and biochemistry. Third, these experts were asked to: (a) make sure each item matched the intended domain or topic (i.e., astronomy or astrobiology), and (b) judge each response in terms of the response model (i.e., domain correct, domain incorrect, non-astronomy science incorrect, or common misconception/popular lore).

Fourth, these data were collected and either items that did not match the performance domain or responses that did not match the response model were either revised or deleted. Specifically, three items were reworded and six answer choices were reworded or changed.

To correct for measurement error in these 16 items ( $\alpha = .73$ ), the items were subjected to an exploratory factor analysis (EFA) and a factor score was calculated for each individual. The EFA indicated a one-factor solution that explained 22.00% of the total variance. These factor scores have a mean of 0 and a standard deviation of 1.

### **Beliefs about ETCs**

Participants' beliefs about the existence or non-existence of ETCs were measured using a two-item questionnaire before and after reading either the expository or two-sided refutational text passage (Appendix E). For this study, the participants' beliefs about ETCs were used to aid in coding their open-ended response. The first item asked participants to choose, "which of the following most closely describes your beliefs about the existence of *intelligent Extraterrestrial Civilizations* (ETCs)?" Three of the four options appeared in the text passage about ETCs: ETCs exist *and* they are here or have visited Earth in the past; ETCs exist *but* they have not yet communicated with us or visited Earth; and, ETCs do not exist *and* we are alone in the universe. A fourth possible option, the existence or nonexistence of ETCs is unknowable, was added to make the possible response choices exhaustive. The second item asked participants how confident they were in their response to the previous item. They indicated their confidence by making a slash mark on a 100-millimeter line from *not confident* to *very confident*. Scores for their confidence ratings ranged from 0 to 100.

### **Passage Recall Measure**

The passage recall measure (Appendix H) was a 10-item multiple-choice measure assessing participants' recall of the text passages. Like the domain and topic knowledge measures, the passage recall measure also used a graduated response model. The levels of the graduated response model for the passage recall questions were: correct response from the passage (scored 4); incorrect response from the passage (scored 2); incorrect response from the domain of astronomy but not from the passage (scored 1); and, an incorrect response not from the domain of astronomy (scored 0). A sample item from this measure is:

The Voyager probe took 21 years to reach:

- a. Pluto (4) – correct response from the passage
- b. Proxima Centauri (2) – incorrect response from the passage
- c. Saturn (1) – incorrect response from the domain of astronomy
- d. Cato Neimoidia (0) – incorrect response not from the domain of astronomy

The passage recall measure followed the same validity procedure (content validity) as the domain and topic knowledge measures. All items were judged by the experts ( $n=3$ ) to match the topic of the passage. The alpha coefficient was low for the passage recall measure ( $\alpha = .53$ ). However, since these items were used as indicators for a latent variable in the resulting model, the structural relations in the model were corrected for any measurement error in the original indicators (Byrne, 1994).

### **Measurements**

In addition to the quantitative measures described above, both think-aloud data and open-ended passage outcome data were collected then quantified using the described coding schemes.

### **Think-aloud Data**

While reading the text passage, participants were asked to think aloud. The specific instructions for the verbal protocol are included in Appendix G. These think-aloud verbalizations were recorded using an Olympus WS-110 digital voice recorder. The recordings were then transcribed into text documents by the author.

Using an initial set of 30 codes developed in preceding studies (e.g., Fox et al., 2007, 2008) and based on the set of possible behaviors seen in verbal protocols of reading in Pressley and Afflerbach's summative overview (1995), Kintsch's construction-integration model (CIM; Kintsch, 2004) was used to code these participants' utterances as either a surface-processing strategy (S; i.e., a strategy aimed at constructing the textbase), a deep-processing strategy (D; i.e., a strategy aimed the integrating the textbase with their prior knowledge), or as some other type of strategy (O; e.g., a regulatory strategy). Further, if the surface-processing or deep-processing strategies did not appear to be helpful, a minus sign was included with the code (i.e., an S- or D-).

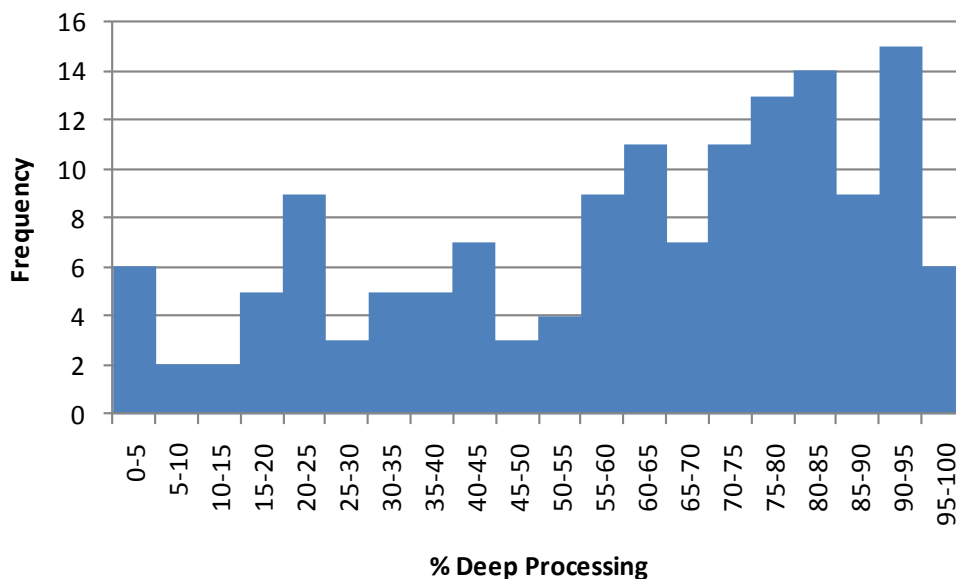
Since the helpfulness of each strategy was exploratory, care was taken to be conservative in assigning a minus to a strategy only when it was readily apparent it was not a helpful strategic move. For instance, with if a deep-processing strategy such as an elaboration was completely irrelevant to the text, it was coded as a D-. If a surface-processing strategy, such as a local restatement was incorrect, it was coded an S-. The list of codes and examples of coded utterances from this study appear in Appendix H. To build interrater agreement, the think-alouds were coded independently by the author and another trained rater until acceptable agreement was reached. For 25 of the 150 participants (16.67%) the kappa coefficient for the codes consisting of S, S-, D, and D-

(409 total utterances) was .85. This agreement was considered substantial and the remaining transcripts were coded by the author. All differences in codes were resolved by discussion before the remaining transcripts were coded.

Level of processing was calculated by dividing an individual's total number of deep-processing strategies by the sum of his or her deep- and surface-processing strategies. The resulting proportion for depth of processing scores ranged from .00 (no deep-processing strategies) to 1.00 (all deep processing strategies;  $M = .61$ ,  $SD = .28$ ). The distribution for these scores by percentage is presented in Figure 8.

Figure 8

*Distribution of Level of Processing Percentages*



### Open-ended Response

Participants were asked to respond to an open-ended question about the text passage. This question immediately followed a posttest administration of their beliefs about ETCs in which question one asked about their beliefs about the existence of ETCs

(are here, exist but have not communicated, or do not exist). The open-ended question asked participants to, “justify your answer to Question 1 based on evidence from the passage or from your background knowledge.”

Their written responses were typed and saved as an electronic document.

Participants’ responses to the open-ended question were quantified using the Structured Outcome of the Learning Observation (SOLO) taxonomy (Biggs & Collis, 1982). The SOLO taxonomy evaluates the structure of the participant responses based on capacity (working memory and attention span), relating operations (how question and response interrelate), consistency and closure (amount of data utilized and openness of conclusion), and overall structure of the response. Using these elements, the taxonomy classifies responses as: prestructural, unidimensional, multistructural, relational, or extended abstract. The rubric (from Biggs & Collis, 1982) as well as the score values for each classification is included in Table 4.

The scoring rubric also allowed for transitional responses (i.e., responses that fall between levels of the taxonomy), resulting in a scale that ranged from zero (prestructural) to four (extended abstract) with the possibility of a response being between two of the levels (e.g., a 3.5 if the response fell between relational and extended abstract). The resulting distribution of scores on the open-ended response ( $M = 1.57$ ,  $SD = 1.06$ ) is presented in Figure 9.



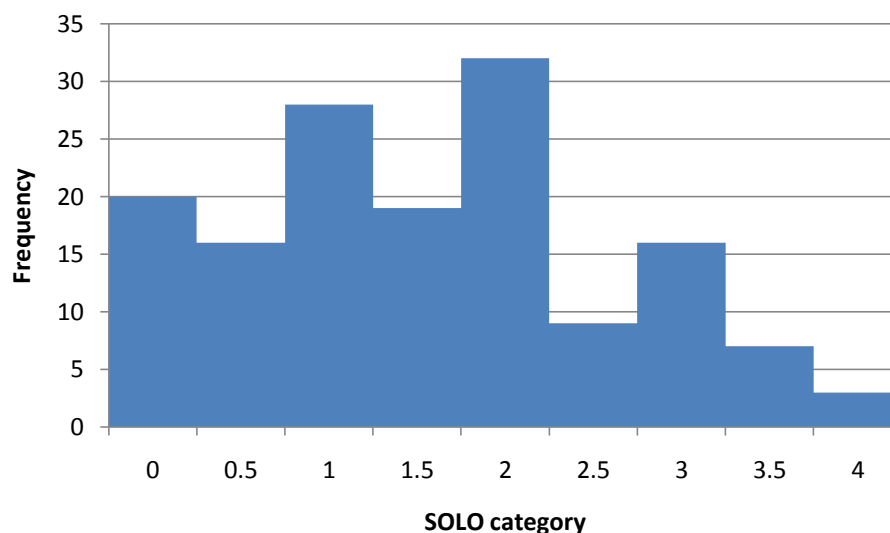
Table 4

*SOLO Taxonomy for the Problem-Solving Task Outcome*

Taxonomy Level	Score	Response characteristics
Prestructural	0	Cue and response undifferentiated No logical interrelation for cue and response High closure or low consistency Cue linked with irrelevant feature(s)
Unidimensional	1	Relate question with one piece of relevant data with a logical operation Drawing a conclusion from a particular instance Responses equally correct but inconsistent with each other One relevant feature to link question and response
Multistructural	2	Two or more relevant concepts or data Uses several features but does not link them Closure but lack of consistency Several relevant features link question and response
Relational	3	Response which interrelates multiple concepts Overall concept or principle accounting for data presented Waits for all aspects before interrelating to make coherent whole Definite overgeneralized answer tied to concrete experience Uses relevant data in a conceptual scheme
Extended abstract	4	Give information comprehended in relevance to an overriding abstract principle True logical deduction Heavily qualifies set out principle to application in given situations Question left relatively open Relevant data with interrelations under hypothetical abstract structure with alternative outcomes and no definite closure
Transitional responses	.5; 1.5; 2.5; or 3.5	At a level of the taxonomy but marked by confusion or inconsistency Handles more information than able to cope with Loses track of the argument Forced to give up before reaching next SOLO level

*Note.* From Bigg and Collis (1982).

Figure 9

*Distribution of Open-ended Response Scores*

To build interrater agreement for the open-ended responses, the following procedure was undertaken. First, the author and another rater scored five random open-ended responses together. A score and rationale for the score was tabled. Next, each rater independently scored another five random open-ended responses marking a score and a rationale for the score. The author and other rater compared both the scores and rationale for the scores. Differences were rectified in conference. The author and rater then independently scored another ten responses. For the 15 responses independently coded (10.0% of the total responses) the kappa coefficient for the exact ratings was .69. Further, the average deviation of ratings for both raters for these 15 responses was .03 points. This kappa coefficient was considered good according to Fleiss's (1981) guidelines and the average deviation of the rater score were considered well within tolerable limits. All differences in codes were resolved by discussion before the remaining transcripts were coded. The remaining transcripts were coded by the author.

## Procedures

Participants completed the measures and measurements in a research laboratory at the university. Participants first read and signed the consent form. Participants then completed the demographics measure, the subject-matter knowledge measure, and beliefs about ETCs measure.

Next, participants were given a practice passage to read while thinking aloud. The practice passage is about mosquitoes and was adapted from a popularly written science article by Marston Bates (1975). Once participants were comfortable with the think-aloud protocol, they either read the expository or persuasive text. The text condition (i.e., expository or persuasive) was randomly assigned to obtain roughly the same number of participants in each condition ( $n_{\text{expository}} = 76$ ,  $n_{\text{persuasive}} = 75$ ).

After reading the passage and thinking aloud participants again completed the passage recall measure, again completed the beliefs about ETCs questionnaire, and the open-ended response. During the passage recall measure participants did not have access to the text, but did have access to the text during the open-ended response. The average participant took approximately 30 minutes to complete these tasks.

## Overview of Analyses

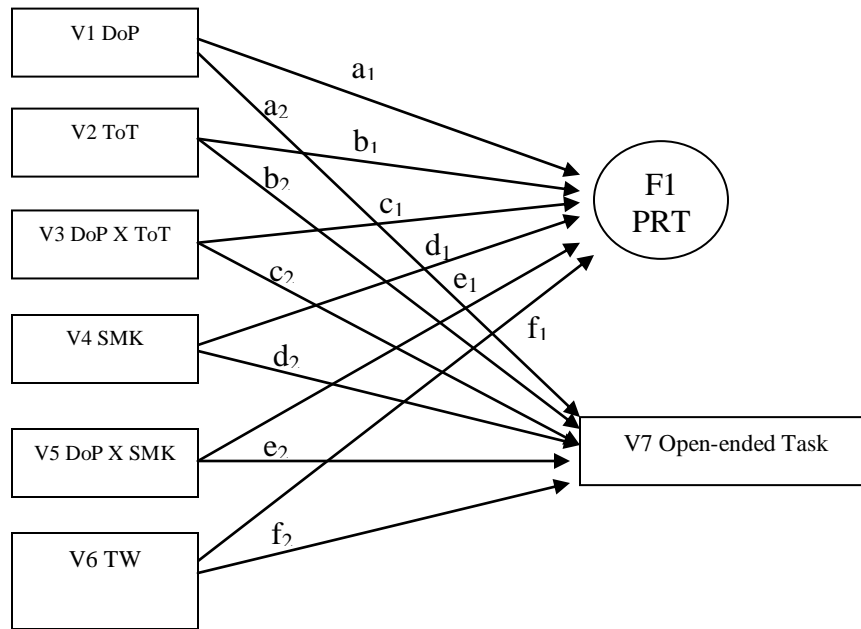
Structural equation modeling (SEM; Figure 10) was used to provide the statistical tests for these three questions. Question 1 was a test for the moderation of type of text on the relations between depth of processing and the two outcome measures (i.e., the passage recall task and the open-ended task;  $H_0: c_1 = 0$  and  $H_0: c_2 = 0$ ). The parameters of the model tested for this question were the direct effects of the interaction of depth of processing and type of text on the passage recall task scores and open-ended task scores

(paths  $c_1$  and  $c_2$  in Figure 1 respectively). Specifically, the null hypothesis was that the direct effects of this interaction on each of the outcomes in the model were 0.

Question 2 was a test for the moderation of subject-matter knowledge on the relations between depth of processing and the two outcome measures (i.e., the passage recall task and the open-ended task;  $H_0: e_1 = 0$  and  $H_0: e_2 = 0$ ). The parameters of model tested for this question were the direct effects of the interaction of depth of processing and subject-matter knowledge on the passage recall task scores and open-ended task scores (paths  $e_1$  and  $e_2$  in Figure 1 respectively). Specifically, the null hypothesis was that the direct effects of this interaction on each of the outcomes in the model were 0.

Question 3 was a test of the difference between three pairs of paths. The first of these pairs were the paths between depth of processing on the passage recall task and depth of processing on the open-ended task (paths  $a_1$  and  $a_2$  in Figure 1 respectively). The second of these pairs were the paths between the interaction of depth of processing and type of text on the passage recall task and open-ended task respectively (paths  $c_1$  and  $c_2$  in Figure 1). The third of these pairs were the paths between the interaction of depth of processing and subject-matter knowledge on the passage recall task and open-ended task respectively (paths  $e_1$  and  $e_2$  in Figure 1). Specifically, the null hypotheses for these tests was that the estimated parameters for these paths were equal ( $H_0: a_1 = a_2$ ,  $H_0: c_1 = c_2$ , and  $H_0: e_1 = e_2$ ). The only difference in the analysis for question 3 was that the passage recall task and open-ended task were standardized (i.e., transformed to have a mean of zero and standard deviation of 1) to allow for a direct comparison of these paths using a t test.

Figure 10

*Structural Model for Analyses*

*Note:* DoP = depth of processing; ToT = Type of text; SMK = Subject-matter knowledge; TW = Three way interaction (DoP X ToT X SMK); PRT = Passage recall task.

Prior to conducting the study, an a priori power analysis (i.e., sample size determination) was conducted for testing parameters within a model, according to the procedure outlined by Hancock (2006). The first step was to decide on the desired level of power ( $\pi = .80$ ) and alpha level ( $\alpha = .05$ ) for the focal parameter tests (i.e.,  $H_0: a_1 \neq a_2$ ,  $H_0: c_1 \neq c_2$ ;  $H_0: e_1 \neq e_2$ ;  $H_0: c_1 \neq 0$ ,  $H_0: c_2 \neq 0$ ,  $H_0: e_1 \neq 0$ , and  $H_0: e_2 \neq 0$ ), which corresponded to a noncentrality parameter of  $\lambda = 7.85$ . Next, numerical values were selected for all parameters (standardized) of the full model as well as for the standardized focal parameters (Table 5).

Table 5

*Standardized focal and peripheral parameters values for the a priori power analysis*

Parameter estimate	Standardized coefficient
Focal parameters	
c <sub>1</sub>	.20
c <sub>2</sub>	.20
e <sub>1</sub>	.20
e <sub>2</sub>	.20
a <sub>2</sub> -a <sub>1</sub>	.15
c <sub>2</sub> -c <sub>1</sub>	.15
e <sub>2</sub> -e <sub>1</sub>	.15
Peripheral parameters	
a <sub>1</sub>	.20
a <sub>2</sub>	.20
b <sub>1</sub>	.30
b <sub>2</sub>	.30
d <sub>1</sub>	.40
d <sub>2</sub>	.40
f <sub>1</sub>	.10
f <sub>2</sub>	.10
V <sub>1</sub> V <sub>2</sub>	.20
V <sub>1</sub> V <sub>3</sub>	.20
V <sub>2</sub> V <sub>3</sub>	.00

*Note.* All other peripheral parameters were estimated to be .01.

Using EQS v.6.1 (Bentler, 2006) moment information was determined for population. This information was used to determine population data-model fit for a reduced model. An arbitrary sample size was set and EQS was again used to determine a model fit function value for the reduced model. Necessary sample sizes were derived for each of the focal parameters of interest. The largest of the sample sizes needed for testing  $H_0: a_1 = a_2$  with  $\pi = .80$  was  $n = 150$ . The sample size for testing  $H_0: c_1 = c_2$  and  $H_0: e_1 = e_2$  was deemed too large for the scope of this study, so while it will still be tested, it should be noted that these were not tested with  $\pi = .80$ .

To evaluate the hypotheses in questions one and two, the measurement equations and standard errors were evaluated as part of the EQS output. For these questions, the  $z$  value reported in EQS was evaluated at  $\alpha = .05$ . For question three, the differences in the path values were divided by the standard error of the difference (i.e., the square root of the sums of each parameters standard error minus two times the covariance of each pairs parameter estimate) to obtain a  $t$  value that was also evaluated at  $\alpha = .05$ . To make the paths comparable to the outcome variables, each outcome variable was standardized prior to making these comparisons (i.e., they both had means of zero and standard deviations of 1). Factor scores were computed using Principle Components Analysis (PCA).

Following data collection and coding, a structural equation model was run using EQS v.6.1 (Bentler, 2006). According to the procedure suggested by Aiken and West (1991), exogenous variables were first centered then multiplied to derive each of the product terms using the centered exogenous variables (i.e., depth of processing, type of text, and subject-matter knowledge). The measurement model was run first with error terms for the passage recall measures left uncorrelated. The measurement model demonstrated satisfactory fit and any attempts to constrain error terms in the measurement model lead to convergence issues. Standardized loadings for the passage recall factor in the measurement model ranged from .17 to .52. Construct reliability for the factor was calculated using coefficient H (Hancock & Mueller, 2001). Coefficient H for this factor was .62 which indicates moderate construct reliability.

The structural model was then run using full-information maximum likelihood (FIML; five participants' depth of processing and one participant's open-ended response

were missing). Correlations between exogenous variables were estimated along with the covariance between the error and disturbance of the endogenous variables.

To evaluate the hypotheses in questions one and two, the measurement equations and standard errors were evaluated as part of the EQS output. For these questions, the z value reported in EQS was evaluated at  $\alpha = .05$ . For question three, the differences in the path values were divided by the standard error of the difference (i.e., the square root of the sums of each parameters standard error) to obtain a t value that was also evaluated at  $\alpha = .05$ .



## **CHAPTER IV**

### **RESULTS AND DISCUSSION**

The purpose of this study was to investigate interactions of type of text and subject-matter knowledge on the relation between depth of processing and task performance as well differences in the relations between passage recall and an open-ended question. Structural equation modeling (SEM) was employed to analyze data relevant to the specific research questions in the study. This chapter consists of the results and discussion for the preliminary data analyses as well as each of the three research questions.

#### **Preliminary Data Analyses**

The purpose of the preliminary data analyses was to ensure that the resulting data were appropriate for the SEM analysis. In regards to normality, these data met Finney and DiStefano's (2006) suggestion that multivariate kurtosis (Mardia-based kappa) should be less than three ( $\kappa = 0.16$ ). Additionally, raw data of the cases with the largest contribution to normalized multivariate kurtosis were examined and not found to be problematic. Lastly, as suggested by Bentler (2006) the distribution of standardized residuals were symmetric and centered closely around zero. Table 6 presents the model covariance matrix for the measured and latent variables along with the means for each of the measured variables. Additionally, overall model fit was good according to the joint criteria suggested by Hu and Bentler (1999): CFI = .99, SRMR = .065, and RMSEA = .048 (90% CI: .023, .068). These values indicate that the suggested model is one plausible model for these data. The standardized path coefficients are displayed in Figure 11 with significant paths bolded.

Table 6

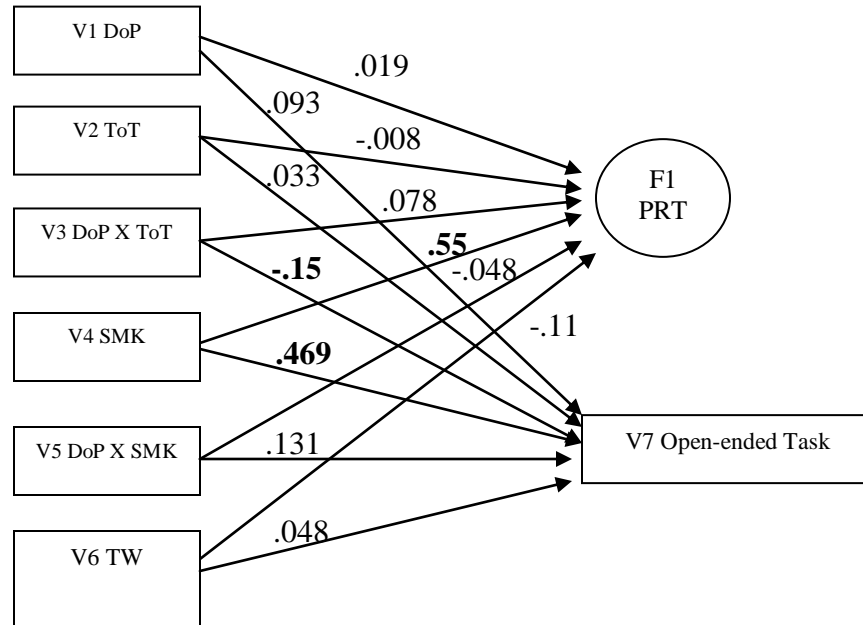
*Model covariance and mean matrix for measured and latent variables*

	DoP	ToT	DopX Tot	SMK	DoPX SMK	TW	OET	PR1	PR2	PR3	PR4	PR5	PR6	PR7	PR8	PR9	PR10	PRT	
DoP	.076																		
Tot	.000	.250																	
DoPXTot	.002	.000	.019																
SMK	.047	-.027	.009	.993															
DoPXSМК	.000	.008	-.005	.034	.076														
TW	-.004	.011	.000	-.009	.002	.019													
OET	.047	.012	-.018	.512	.062	.003	1.117												
PR1	.007	-.005	.003	.102	.000	-.004	.066	.400											
PR2	.013	-.011	.006	.205	.000	-.008	.134	.068	1.057										
PR3	.009	-.007	.004	.141	.000	-.005	.092	.047	.094	.293									
PR4	.012	-.010	.006	.190	.000	-.007	.124	.063	.127	.087	1.361								
PR5	.018	-.014	.008	.276	.000	-.011	.180	.091	.184	.127	.171	1.824							
PR6	.025	-.020	.011	.383	.000	-.015	.250	.127	.256	.176	.237	.345	1.805						
PR7	.008	-.006	.004	.121	.000	-.005	.079	.040	.081	.056	.075	.109	.151	1.683					
PR8	.012	-.009	.005	.183	.000	-.007	.119	.061	.122	.084	.113	.165	.229	.072	1.262				
PR9	.015	-.012	.007	.237	.000	-.009	.155	.079	.159	.109	.147	.214	.297	.094	.142	.848			
PR10	.013	-.010	.006	.196	.000	-.008	.128	.065	.131	.090	.122	.177	.246	.078	.117	.152	2.151		
PRT	.007	-.005	.003	.102	.000	-.004	.066	.034	.068	.047	.063	.091	.127	.040	.061	.079	.065	.034	
Means	.000	.000	.000	.000	.044	.009	.000	3.77	3.52	3.84	3.17	2.87	3.13	2.14	3.13	3.66	2.58		

*Note.* DoP = Depth of Processing; Tot = type of text; DoP X Tot = interaction of depth of processing and type of text; SMK = subject-matter knowledge; DoP X SMK = interaction of depth of processing and subject-matter knowledge; TW = three way interaction of depth of processing, type of text, and subject-matter knowledge; OET = open-ended task; PR = passage recall items; PRT = passage recall task factor

Figure 11

*Standardized path coefficients for the structural model*



*Note.* DoP = Depth of Processing; Tot = type of text; DoP X Tot = interaction of depth of processing and type of text; SMK = subject-matter knowledge; DoP X SMK = interaction of depth of processing and subject-matter knowledge; TW = three way interaction of depth of processing, type of text, and subject-matter knowledge; PRT = passage recall task factor.

Significant paths ( $p < .05$ ) appear in bold.

### **Interaction of Depth of Processing and Type of Text on Performance**

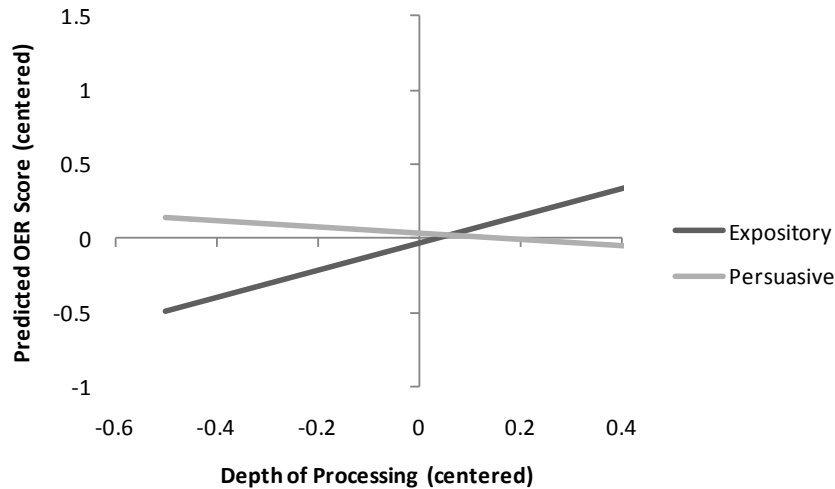
The first research question addressed how the relation between depth of processing and performance (a passage recall task and an open-ended task) would be moderated by type of text (i.e., expository versus persuasive). For the passage recall task and open-ended task, this was done by testing the null hypotheses  $H_0: c_1 = 0$  and  $H_0: c_2 = 0$  (respectively) from Figure 11. This was a test if the direct effects of the interaction of depth of processing and type of text on each of the outcomes were significantly different

than zero. Results from the path estimates in the model revealed that for there was a significant interaction of type of text on the relation between depth of processing and the open-ended task ( $p_{V7V3} = -1.13, p = .039$ ) but not for the passage recall task ( $p_{F1V3} = .11, p = .42$ ). In other words, the direct effect of the interaction between depth of processing and type of text on the open-ended task was significantly different than zero (in the negative direction), while the direct effect of the interaction between depth of processing and type of text on the passage recall tasks was not significantly different than zero.

The significant path between the interaction of depth of processing and type of text with the open-ended responses indicates that those individuals with shallower processing (lower DoP percentage) had higher open-ended responses in the persuasive condition than those in the expository condition. However, individuals with deeper processing (higher DoP percentages) had higher open-ended responses in the expository condition than the persuasive condition. This interaction is displayed graphically in Figure 12. The regression equations for the simple slopes were calculated (Cohen, Cohen, West, & Aiken, 2003) for the two text conditions:  $z_{\text{exp}}: \hat{Y} = .92x - .035$  and  $z_{\text{per}}: \hat{Y} = -.21x + .035$ .

Figure 12

*Interaction graph for type of text on the relation between depth of processing and open-ended response*



Further, the actual slopes (i.e., those calculated using the scores from both the expository and persuasive conditions) was significantly different than zero for the expository text condition ( $b=1.17$ ,  $t=2.60$ ,  $p=.011$ ), but not for the persuasive condition ( $b=.13$ ,  $t=.29$ ,  $p=.77$ ).

Two explanations are offered for this conditional effect. The first of these explanations centers on an individual's prior beliefs. For the expository condition, there was *no* explicit argument as to which of the three explanations for extraterrestrial civilizations (ETCs) was most scientifically testable (i.e., they are here, they exist but have not communicated, or they do not exist). However, it may be the case that by juxtaposing these three viewpoints into one document just meant to expose individuals to these three views, some individuals (in many cases those with high levels of subject-matter knowledge) projected an argument onto the text or noticed that there is in fact an

implicit argument present. Thus, it may have prompted the engagement in deeper levels of processing as a means to justify their position in the open-ended task.

However, for those in the persuasive condition, deeper levels of processing were not necessarily helpful in responding to the open-ended question. It may be the case that the explicit argument present in the refutational text interfered with their processing in some way. This may be particularly salient with this text since the text argued in favor of the position (i.e., they do not exist) not widely held in the population (Webb, 2002) or in this sample (13.2% before reading the text and 12.7% after reading the text). One participant remarked, “I think we don’t want to be alone I think it’s hard to believe, to think that we are the very few or the only existence, then why us, you know?” Underlying this may be an affective response by participants summed up by this response, “I think it’s absolutely ridiculous to think we are alone in the universe. Statistically that’s impossible. Also, it would be incredibly depressing.” In other words, this issue was clearly emotionally charged for some of these participants, which may not have helped them justify their responses on their open-ended response.

The second of these explanations centers on the usefulness of the strategies employed (i.e., deep or surface). As part of the coding scheme unhelpful strategies were coded with an S- (unhelpful surface-level strategy) or a D- (unhelpful deep-level strategy). Although, there was good interrater agreement for these codes, how helpful a strategy was for a participant was difficult to determine upon examining the transcripts. To this end, both raters were very conservative in this coding and caution should be observed when interpreting these data. To this end, however, the correlation between the helpfulness of the strategies employed (proportion of D- and S- to the total number of

strategies) and the open-ended responses was significant ( $r = .27$ ,  $t = 2.36$ ,  $df = 144$ ,  $p = .020$ ). It is possible that the explicit argument in the persuasive text prompted participants to engage in unhelpful deep-processing strategies that the expository text did not elicit.

### **Interaction of Depth of Processing and Subject-Matter Knowledge on Performance**

The second research question addressed how the relation between depth of processing and performance (a passage recall task and an open-ended task) would be moderated by subject-matter knowledge. For the passage recall task and open-ended tasks, this was done by testing the hypotheses  $H_0: e_1 = 0$  and  $H_0: e_2 = 0$  (respectively) from Figure 11. This was a test if the direct effects of the interaction of depth of processing and subject-matter knowledge on each of the outcomes were significantly different than zero. Results from the path estimates in the model revealed no significant interactions of subject-matter knowledge on the relation between depth of processing and the open ended tasks ( $p_{V7V5} = .503$ ,  $p = .067$ ) or the passage recall task ( $p_{F1V5} = -.032$ ,  $p = .62$ ). In other words, the direct effect of the interaction between depth of processing and type of text on both outcomes were not significantly different than zero.

The lack of significance between the path testing the interaction of depth of processing and subject-matter knowledge to the open-ended response was a bit surprising. This could potentially be a power issue, since the parameter estimates used in the a priori power analyses were somewhat different the actual parameter estimates in the model. For example, the peripheral parameter estimates for depth of processing to the passage recall task and for depth of processing to the open-ended response were thought to be about .20. However, the actual model parameters fell far below those at .019 and



.093 respectively. It may be the case that a larger sample size would be necessary for path significance, rather than there not being a significant relation in the population.

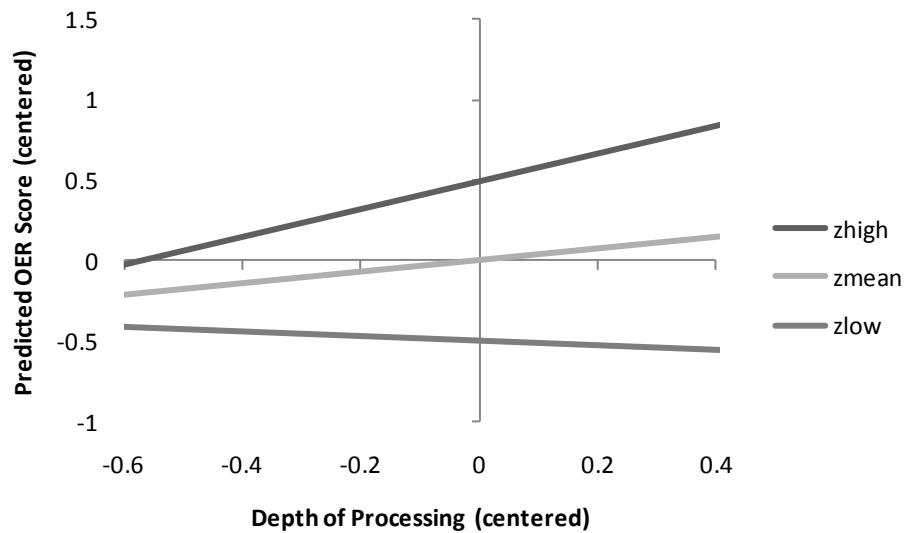
The failure to reject the null hypothesis (i.e., that there was no interaction effect between depth of processing and subject-matter knowledge on the open-ended task) in the typical null hypothesis testing (NHT) framework precludes generalization of these findings. However, the close proximity of the test statistic to the region of rejection warrants further examination. In other words, retaining the null hypothesis does not rule out that there are interaction effects, rather it signifies an inability to confidently state that the effect is significantly different than zero in a population according to a chosen decision rule ( $\alpha = .05$ ). One possible reason for the inability to reject the null hypothesis may be a lack of power (given the differences in the parameter estimates in the power analysis and those estimated in the subsequent model used for these analyses). In the ensuing discussion of the depth of processing and subject-matter interaction, caution should be exercised in interpreting these effects beyond this particular sample.

To meaningfully interpret the interactions of two continuous variables, Cohen et al. (2003) suggested examining the regression of the criterion (scores on the open-ended task) on one predictor (depth of processing) at each of several values of the other predictor (subject-matter knowledge). The regression line of the criterion on one of the predictors (depth of processing) at one value of the other predictor (subject-matter knowledge) is termed a *simple regression line*. In accordance with Cohen et al.'s (2003) suggestion, the three values chosen for subject-matter knowledge in calculating the simple slopes were a subject-matter knowledge value at the mean (labeled  $z_{\text{mean}}$ ), a subject-matter knowledge value one standard deviation above the mean (labeled  $z_{\text{high}}$ ),

and a subject-matter knowledge value one standard deviation below the mean (labeled  $z_{low}$ ). Since the factor scores for subject-matter knowledge had a mean of zero and a standard deviation of one, these three values for subject-matter knowledge were 0, 1, and -1 respectively. The equations for these simple slopes were:  $z_{low}: \hat{Y} = -.15x - .070$ ;  $z_{mean}: \hat{Y} = .36x$ ; and,  $z_{high}: \hat{Y} = .86x + .070$  and are graphically displayed in Figure 13.

Figure 13

*Interaction graph for subject-matter knowledge on the relation between depth of processing and open-ended response*



*Note.*  $z_{high}$  = regression slope for a score one standard deviation above the mean;  $z_{mean}$  = regression slope for a score at the mean;  $z_{low}$  = regression slope for a score one standard deviation below the mean.

While parallel lines would indicate no interaction, the pattern of the slopes in these simple regression lines indicates an interaction effect for this sample (again, this

should not be generalized beyond this sample). The positive interaction term in this sample indicates that higher proportions of deep-processing strategies more often lead to higher scores on the open-ended response for those with higher and mean amounts of subject-matter knowledge. However for those with lower values of subject-matter knowledge, a higher proportion of deeper-processing strategies more often lead to lower scores on the open-ended response.

Similar to the proposed interaction for type of text above, the helpfulness of the strategies employed may be one of the reasons for the presence of the subject-matter knowledge and depth of processing interaction for this sample. In fact, the correlation between the proportion of helpful strategies (i.e., S- and D- divided by the total number of strategies) and subject-matter knowledge was significant ( $r = .39$ ,  $t = 5.11$ ,  $df = 144$ ,  $p < .001$ ). This strong relation lends credence to the argument that for this sample deep processing may only have been effective if one possessed the requisite subject-matter knowledge to deploy the strategic process in a helpful way. This conclusion is consistent with the relations between strategic processing and knowledge forwarded by Alexander and Judy (1988).

### **Differences Between the Passage Recall Task and Open-Ended Task**

The third question investigated differences in the paths between three of the variables of interest (i.e., depth of processing, the interaction of depth of processing and type of text, and the interaction of depth of processing and subject-matter knowledge) and the two outcomes (i.e., passage recall task and the open-ended task). An initial examination of these paths reveal only one significant path (from the interaction of type of text and depth of processing to the open-ended task). The path from the interaction of

type of text and depth of processing to the open-ended task was *not* significantly different from zero. The other paths (the interaction of depth of processing and subject-matter knowledge to both outcomes and depth of processing to both outcomes) were all nonsignificant. This pattern of relations would seem to indicate only possible significant differences in the path from the interaction of type of text and depth of processing to the open-ended task and the path from the interaction of type of text and depth of processing to the passage recall task. Since the scale of outcome variables (the open-ended and passage recall tasks) could not be assumed to be equivalent which can be problematic in comparisons of the unstandardized paths (Kwan & Chan, 2010), Hotchkiss (1976) suggested standardizing each of the outcome variables for comparison. After both outcome variables were standardized, a comparison of the path between the interaction of type of text and depth of processing to passage recall and the interaction of type of text and depth of processing to the open-ended task were significant using this method ( $t = 2.15$ ,  $df = 141$ ,  $p = .033$ ).

Furthermore, in this sample the standardized effects were larger in general for the open-ended response than for the passage recall task (except for the subject-matter knowledge and the three way interaction in the model). Comparing these values should be interpreted with caution as these paths did not attain significance at the .05 level. The differences in these effects, particularly the significant depth of processing and type of text interaction between these outcomes, may be due to the demands of these two tasks and the participants' familiarity with these types of tasks. First, the passage recall task and open-ended response place different cognitive demands on participants. While the passage recall task only requires surface-level memory (e.g., Kintsch, 2004) to reproduce

a correct answer, the open-ended tasks requires justification of a position, a more higher-order task (e.g., Alexander et al., 2011; Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). Unlike mere reproduction of information relying on surface-level memory, the justification task would likely require readers to construct their own situation model during and after reading. In other words, the open-ended task is a production task rather than a reproduction task.

Second, with regard to the participants' familiarity with the tasks, the passage recall task is a more typical school-type task. Passage recall is not only commonplace in the classroom, but is a typical task for the standardized testing to which most of these individuals have been exposed. However, since justification of knowledge is not as commonplace in schools, it could be that this task was something unfamiliar or novel to many. These two issues (the task demands and familiarity of the participants with them) could have contributed to the significant interaction of depth of processing and type of text as well the significant difference in the two outcomes.

## **CHAPTER V**

### **CONCLUSIONS**

This study was designed to test the interactions of type of text and subject-matter knowledge on the relation between depth of processing and two outcomes, a passage recall task and an open-ended task. Overall, the findings of this study support the contention that a consideration of the relation between levels of processing and performance without a consideration of moderating factors may be contributing to the mixed findings in the literature. There were significant differences in the interaction of depth of processing and type of text on the open-ended tasks, as well as differences in this interaction between the open-ended task and the passage recall task. Although there was no conclusive evidence of a significant interaction between depth of processing and subject-matter knowledge on the open-ended response, this should be further investigated, and suggestions for this are given in subsequent sections. This chapter will proceed by situating the findings of this study in regard to the three theoretical frames and the extant literature on deep and surface processing. Following that will be a discussion of future programs of research in deep and surface processing, as well as implications for practice.

#### **The Three Theoretical Frames**

Although there was good statistical model-data fit, the findings in this study represent just one of many plausible models to explain the relation between these variables. Further, the interpretation of the results of this study should be considered in regard to the theoretical frames in which the model was situated; namely, the Construction-Integration Model (CIM; Kintsch, 2004), the Model of Domain Learning

(MDL; Alexander, 1997), and the Topographical Perspective on Learning (Alexander et al, 2009). This section will discuss the results in terms of these theoretical frames followed by a more general discussion of what these results mean for research on depth of processing generally.

### **The Construction-Integration Model**

For this study, the CIM provided a framework in which deep- and surface-processing strategies could be both conceptualized and operationalized. Hence, any interpretations of the results of study must be made relative to this framework. While there are certainly other models that could have been used for this purpose (e.g., the Landscape Model; van den Broek, Young, Tzeng, & Linderholm, 1999), the CIM is advantageous for depth of processing because it addresses two processes of learning from text, namely strategies aimed at the textbase versus strategies aimed the situation model. These two processes helped to operationalize deep- and surface-processing strategies during text comprehension.

Coding of the think-aloud utterances using this framework functioned well as evidenced by the high interrater agreement achieved. Using the 30 codes developed from previous studies (e.g., Fox et al., 2009), the CIM was useful in reducing the codes to three; namely, surface-processing strategies (i.e., those aimed at the textbase), deep-processing strategies (i.e., those aimed at the situation model), and other codes (self-regulation and evaluation). This larger grain size not only made coding more efficient, but was also quite useful to distinguish between these the two types of processes here (i.e., deep and surface).

Kintsch (2004) postulated that learning from text requires the construction of a situation model (deep processing) and failure to do so leads to encapsulated knowledge. Further, Kintsch argued that this encapsulated knowledge was likely to only be retrieved through specific episodic text memory (such as passage recall) and would not be useful in situations beyond reproduction of the text (e.g., justification of a position in the open-ended response). The results of this particular study did not reveal a direct relation between depth of processing (the percentage of strategies aimed at the situation model) and the two outcomes, but subsequent discussion of the results in light of the other theoretical frames demonstrate that while there was no direct relation, depth of processing does play a role in learning from text when conditional effects are considered (i.e., interactions).

### **The Model of Domain Learning**

The MDL was a critical frame in terms of modeling individual differences that may exist between learners, specifically in this study the interaction of subject-matter knowledge and depth of processing. The MDL posits that as an individual progresses from acclimation (the first stage of expertise in the MDL) to proficiency (the final stage of expertise the MDL) there are changes in both individuals' subject-matter knowledge and strategic processing. Empirical evidence using the MDL has suggested that there is an increase in both domain and topic knowledge through these stages of expertise development (e.g., Alexander, Jetton, & Kulikowich, 1995). Additionally, the MDL hypothesizes and empirical evidence has demonstrated that at acclimation the learner is relying primarily on surface-processing strategies, while individuals in proficiency rely primarily on deep-processing strategies (e.g., Alexander, 1997; Alexander et al., 2004).



It is the concurrent rise of subject-matter knowledge and increase in deep-processing strategies in the MDL that was of particular interest for this study. The suggestion by some that only deep processing should be encouraged and surface processing should be discouraged (e.g., Tian, 2007) seems at odds with the developmental trajectory suggested by the MDL in which an increase in subject-matter knowledge occurs jointly with an increased reliance on deeper-processing strategies. Nonetheless, surface-level strategies serve an important and continuing role across the stages of academic development, although in differing degrees at each stage. Is it the case that the rise of subject-matter knowledge is necessary for one to rely primarily on deep-level processing? Or, would encouraging individuals' with lower levels of subject-matter knowledge to engage primarily in deep-processing strategies also be effective?

The preliminary evidence provided here tends to support the prior notion that subject-matter knowledge may be necessary in order to effectively engage in primarily deeper levels of processing, in other words, even in the stages of competence and proficiency it appears that surface-level processing plays a key role. This preliminary evidence, however, comes with two important caveats. First, although the interaction graph of depth of processing and subject-matter knowledge on the open-ended response demonstrated that on average a high percentage of deep-processing strategies only resulted in increased open-ended tasks scores for those with high and mean levels of subject-matter knowledge, this interaction path was not statistically significant at  $\alpha = .05$ . However, as will be discussed later this interaction may still merit considerations and further investigation. Nonetheless, given the nature of the interaction patterns between subject-matter knowledge and depth of processing interaction on the open-ended

response in this sample (i.e., that only higher or mean levels of subject-matter knowledge coincided with increased performance on the open-ended task), it is critical that future investigations consider these interactions.

The second caveat to be acknowledged in terms of the subject-matter knowledge and depth of processing interaction is that neither of these variables had been experimentally manipulated. In designs where neither variable has been manipulated, the results are, “generally more ambiguous and more complex” (Pedhazur, 1997, p. 607).

### **Topographical Perspective on Learning**

Turning now to the broadest of these theoretical frames, the Topographical Perspective on Learning certainly incorporates both the processing and individual differences addressed in the CIM and MDL (e.g., individual differences such as subject-matter knowledge; the *who*), but also extends the scope of the study by considering the situation in which learning takes place (the *where*) as well as the target of learning (the *what*). This topographical perspective helps to frame the possible dimensions that might influence learning beyond the individual and can be particularly helpful for explicating reasons why the relations between depth of processing and performance have been mixed in the extant literature. Differences between situation and the target of learning will be addressed.

First, in regard to the situation, the differences between expository and persuasive text were experimentally tested by randomly assigning participants to either condition. While there were no significant main effects of either type of text or depth of processing, there was a significant depth of processing by type of text interaction for the open-ended task. In other words, the relation between depth of processing and performance on the

open-ended task may be conditional depending on the situation (in this case type of text). Unlike the interaction between the two non-manipulated variables discussed previously (i.e. depth of processing and subject-matter knowledge); the interaction between the manipulated variable (i.e., type of text) and depth of processing is less ambiguous (Pedhazur, 1997). Hence, the interaction effects for type of text and depth of processing can more confidently be attributed to the manipulation of type of text (i.e., expository versus persuasive).

The situational difference found here is but one example within the topographical perspective of variables that may affect the relation between processing and performance. With regard to text, many such variables may potentially change the relation between depth of processing and performance. Text difficulty, text cohesion, text coherence, the density of low frequency words, and text structure have been found to affect text comprehension (e.g., Anderson & Freebody, 1981; Ozuru, Dempsey, & McNamara, 2009). And, similar to type of text, it is certainly plausible that these variables may also interact with depth of processing on learning outcomes. These situational elements need not be constrained to text either when considering tasks outside of reading comprehension, as individuals may learn from a wide variety of tasks from attending a class lecture to complex problem-solving tasks.

Second, in regard to the target of learning, the conditional nature of type of text on the relation between depth of processing and performance needs to be qualified further. That is, this relation differed significantly between the passage recall task and the open-ended task. This additional condition further qualifies what can be said about the effects of depth of processing and type of text on performance; namely, that the

relation between this interaction and performance changes depending on how an individual is asked to demonstrate learning (i.e., what the performance is). In this case, differences occurred in tasks that required either reproduction of the text (i.e., passage recall) or the justification of a position (i.e., open-ended task).

As with the type of text interaction, the difference between these two outcome tasks is but one example of a task difference that might affect the relation between depth of processing and performance. The two outcome tasks administered here were chosen to represent two ends of a continuum (from reproducing text to producing a justification), but one can certainly imagine many other outcomes of learning that could be investigated. For example, these outcome tasks may include text summaries or a more open-ended problem-solving task in which individuals design a gym bag.

### **The Extant Literature on Deep and Surface Processing**

Overall, the results of this investigation provide some evidence that the relation between depth of processing and performance is conditional. The evidence in regard to the situation (i.e., type of text) and the target of learning (i.e., the passage recall and open-ended tasks) having a conditional effect was statistically significant. However, the evidence in regard to individual differences (i.e., subject-matter knowledge) should be further investigated before drawing more substantive conclusions. Despite the non-significant finding in regard to subject-matter knowledge, it *is* clear that considering depth of processing and performance in isolation may hamper attempts to fully understand this complex relation.

The findings of the current study provide evidence for some of the possibilities forwarded for the inconsistent findings between depth of processing and performance in

the literature. These four issues were: conceptualization of deep and surface processing, operationalization of deep and surface processing, situational factors, and model specification. These four issues will be discussed starting with conceptualization and operationalization, followed by situational factors and model specification.

**Conceptualization and operationalization.** In the current study, the conceptualization and operationalization of variables were carefully framed using the relevant theoretical frames discussed previously. So, when the results of the current study are juxtaposed those of previous studies, differences between the relations of variables can at least be partially explained by how the variables were conceptualized and measured. For example, the conceptualization of deep- and surface-processing strategies while reading were conceptualized in the current study as strategies aimed at the textbase and situation model respectively while reading, may generate different relations between processing and performance than an otherwise similar study that defines deep and surface processing as prior intent to memorize the text or learn the text (e.g., *Approaches to Learning*; e.g., Biggs, 1978; Heijne-Penninga et al., 2008). Likewise, choosing to measure deep and surface processing using a think-aloud protocol may also produce a different pattern of relations from an otherwise similar study using a self-report checklist due to memory constraints and other factors. It is important to account for the advantages and disadvantages of each measure or measurement when comparing the relations in these variables across studies.

**Model specification and situational considerations.** Next, the related issues of model specification and situational considerations were specifically investigated in this study. The purpose was to demonstrate that the situation (in this case the type of text)

*does* matter. In other words, investigations of deep and surface processing need to be interpreted in regard to the situation or situations in which participants are exposed to. With the wide variety of situations to which students are exposed, it then becomes difficult to correctly specify a model of deep and surface processing. As Box and Draper (1987) said, “Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful” (p. 74). In other words, which of these conditional effects need to be modeled, and which of these conditional effects can be ignored? While the data-model fit was satisfactory for this model, it is certainly plausible that competing models can and will further explicate these complex relations and replace the current one.

### **Limitations**

Although the proposed study forwarded the literature on deep and surface processing, there were a few limitations present. First, as with any measure, the concurrent think-aloud protocol was likely to elicit only a certain range of processes (Pressley & Afflerbach, 1995). It is probable that processes that are overly complex or overly simple (skillful reading) were not reported by participants. Therefore, it is likely that the think-aloud protocol elicited a range of strategic processes but certainly did not capture all strategic processes or skillful reading. Specifically, the number of utterances may not necessarily correspond to the number of strategies a participant uses. They may be more likely to report deep strategies (because they have better access).

Second, processing in this investigation was measured for one instance with one passage. Certainly there would be an expectation that one’s depth of processing may differ from situation to situation depending on the properties of task at hand. Lastly, the learning outcome, as with any experimental task was necessarily constrained beyond that

of what individuals might do in their everyday lives. For instance, individuals did not have access to materials that they might normally use to complete a task (e.g., other resource materials such as textbooks). Here, they were purposefully being limited to their prior knowledge and the text they read in their given text condition.

Last, the use of the CIM, while helpful in coding deep-level and surface-level processes does not address metacognitive or self-regulatory processes. In this study, these processes were coded as “other” and not used in the calculation of levels of processing. It is likely given the previous literature on metacognitive and self-regulatory processes that these might be important to include. Ways to include these processes are discussed in the directions for future research.

Additionally, although the interrater agreement for coding the think-aloud utterances was high, there were instances where coding an utterance was difficult. For instance, the utterance, “I’m guessing from this sentence that interstellar mean between stars,” was very easy to code as a surface-level process (i.e., guessing the meaning of a word in context.” However, there were instances where coding was not quite so easy. For example, the utterance, “That doesn’t make sense to me,” is harder to code as either a metacognitive process (i.e., monitoring comprehension) or as an evaluation of agreement. In cases such as this, the context of their remaining utterances had to be relied upon to aid in the coding process.

### **Future Research on Deep and Surface Processing**

The current study is merely a first step in modeling the complex interactions of depth of processing and performance. The limitations of the current study (some of which have also been discussed previously) measures and modeling of deep and surface

processing as well as the inclusion of other relevant variables. Despite these limitations, and given that future models of depth of processing *should* include interaction effects, what directions should programs of research in deep and surface processing ultimately take? Directions for future research in regards to measuring and modeling deep and surface processing, particularly taking into account the limitations of the current study, will be discussed.

### **Measuring Deep and Surface Processing**

Three recommendations for measuring deep and surface processing will be discussed including: measurement of depth of processing with multiple indicators; the need for multi-trait multi-method studies; and, the avoidance of dichotomization of deep and surface processing. First, in regard to the measurement of depth of processing, it may be the case that a single indicator (such as the proportion of deep- versus surface-processing strategies) may not be adequate to represent the construct of depth of processing. It is quite possible that strategies aimed at the textbase (surface-processing strategies) and those aimed at the situation model (deep-processing strategies) may not be adequate indicators of depth of processing. For example other indicators, such as monitoring and control processing, may play an integral part in a larger latent construct.

In addition to more fully representing the construct of depth of processing, including multiple indicators of depth of processing may also provide some convergent validity evidence (i.e., whether theoretically related indicators are actually related to each other in practice; Campbell & Fiske, 1959) for depth of processing. One preliminary study to provide convergent validity evidence would be to design and carry out a multi-trait multi-method (MTMM; Campbell & Fiske, 1959). MTMM designs work by



measuring several traits with several different measures (or methods). By examining the resulting correlation matrix, convergent validity can be determined by examining the validity diagonal in each method block (i.e., the measures should correlate highly with each other for the same trait) and divergent validity can be determined by examining the heterotrait-monomethod triangles (i.e., the measures should not correlate highly with each other for different traits) and the heterotrait-heteromethod triangles (i.e., different measures with different traits should also not correlate highly with one another). In addition to measured variable MTMM designs first described by Campbell and Fiske (1959), latent variable MTMM designs are also possible (e.g., Marsh, 1989).

Third, dichotomization of depth of processing (i.e., designating an individual as a deep processor or a surface processor) should be avoided. In order to legitimately dichotomize depth of processing, one would have to conceptually claim that individuals use primarily deep-processing strategies or primarily surface-processing strategies. This runs counter to the models presented here (both the CIM and the MDL). Additionally, evidence from this study (see Figure 8 in particular) suggests that there were many participants whose proportion of deep-level strategies was neither near 0% or 100%. Therefore, the evidence presented here suggests that the dichotomization of depth of processing is both theoretically and empirically unjustified.

### **Modeling Deep and Surface Processing**

Beyond further refinements to the measures and measurements of deep and surface processing, statistically modeling the relation between depth of processing and performance also needs to be a continuing consideration. These modeling considerations

comprise: including other pertinent individual and situational variables; the latent modeling of these variables; and inclusion of the usefulness of strategies employed.

First, while the current study provided evidence that an additive model (i.e., one without interactions) is likely not sufficient to model the relation between deep and surface processing, much remains to be seen as to what other individual or situational variables may also play a key role in the relation between depth of processing and performance. In related areas of investigation (such as metacognition and self regulation) other variables that have played key roles in theoretical models include interest and epistemic beliefs. Interest (i.e., the feeling- and value-related attributes toward an object; Schiefele, 1999), which is also one of the forces considered in the MDL, may also potentially moderate the relation between depth of processing and performance. Epistemic beliefs, while not a force modeled in the MDL, have been theorized to play a key role in self-regulated learning and other theoretical frameworks (e.g., Muis & Franco, 2009). Additionally, situational variables discussed previously, such as text coherence and text difficulty, should also be systematically investigated for interaction effects in future models.

Second, dealing with interactions in the relation between depth of processing and performance (e.g., interaction effects and multiple outcomes) brings with it many analytic challenges. Casual modeling (i.e., SEM) can be a helpful approach in this endeavor for a variety of reasons. These types of models can address “theory-driven causal research questions” at the latent variable level (Hancock & Mueller, 2006, p. 6). Additionally, the latent variable approach is designed to estimate the effects of latent constructs (e.g., depth of processing), rather than a single indicator (e.g., proportion of deep-processing

strategies; Cohen et al., 2003; Pedhazur, 1997) and measurement errors in the model can be somewhat accounted for (Byrne, 1994). Utilizing a multivariate framework, these models are flexible in that they can consider multiple outcomes (Pedhazur, 1997) and are capable of estimating the interactions of latent variables (e.g., Marsh, Wen, Hau, Little, Bovaird, & Widamin, 2007). Lastly, these models move away from simple rule based procedures to model building that relies on a scientific epistemology (Rodgers, 2010). For example, in the current study the decision to retain the null hypothesis in terms of the subject-matter and depth of processing interaction is rather difficult to make, since the  $p$  value is so close to the cutoff value (i.e.,  $\alpha = .05$ ). While, these models typically require large sample sizes, the benefits of such models make the effort worthwhile.

One particularly helpful outcome of the current study is that parameter estimates from the current model can assist with future modeling. Specifically, the current model may serve another indicator of parameter estimates for a priori power analyses, whether for testing particular parameters as in this study or for model fit as a whole. However, as was the issue in this case, focal and peripheral parameters values were chosen from studies using different conceptualizations and operationalizations of deep and surface processing, so if different measures are used or different variables are modeled, these estimates will need to be continually revised.

Third, one of the more difficult considerations in modeling the relation between deep depth of processing and performance is how the usefulness of the strategies employed (whether deep or surface) plays into the overall relation between depth of processing and performance. For instance, if individuals use deep-processing strategies, but these strategies are ineffective (e.g., an unhelpful elaboration about an event unrelated

to the text), should these strategies really be counted as deep-processing strategies?

While this sounds like a reasonable approach, preliminary attempts to ascertain whether strategies were helpful or unhelpful from participants' think-aloud utterances were difficult to make. One possible avenue to investigate this issue may rely on a multi-pronged approach using both think-aloud protocols and semi-structured interviews. For example, it may be helpful to use the think-aloud transcript as a way to query participants' reasons for engaging in and their judgments of the success of a particular strategy.

Both the recommendations for measurement and model building will clearly not be addressed by a single study. These issues will require a systematic program of research to investigate valid measures of depth of processing, appropriate interaction terms for the model, and performance outcomes that assess learning differently. Hopefully programs of research from a variety of different theoretical positions can shed new light on how depth of processing relates to task performance.

### **Implications for Practice**

Given the laboratory nature of this study and the nature of the think-aloud data used, specific classroom implications would be ill advised. However, the nature of the interactions demonstrated in this study may be important to consider for more overarching standards suggested for classroom practice. One such set of standards that has been proposed and has been gaining increased attention is the Common Core State Standards Initiative (CCSSI; [www.corestandards.org](http://www.corestandards.org)). Currently 43 states and the District of Columbia have adopted these standards.

One of the key design considerations for these core standards (specifically those for English and Language Arts; ELA) is on the results and not the means by which these results are obtained:

By emphasizing required achievements, the Standards leave room for teachers, curriculum developers, and states to determine how those goals should be reached and what additional topics should be addressed. Thus, the Standards do not mandate such things as a particular writing process or the full range of metacognitive strategies that students may need to monitor and direct their thinking and learning. Teachers are thus free to provide students with whatever tools and knowledge their professional judgment and experience identify as most helpful for meeting the goals set out in the Standards. (p. 4)

The results of this study corroborate the notion that students should be free to flexibly use a variety of skills and strategies during reading. In the current, study different levels of processing were found to be related to better outcomes for individuals depending on what type of text the individual was reading as well the amount of prior subject-matter knowledge they possessed. This would be further enhanced by the suggestion in the standards that a variety of texts be used (e.g., p. 31).

However, while the standards maintain the notion that particular processes should not be mandated, as standards they do dictate outcomes of learning. For example, one of the standards states that students should be able to, “determine a theme or central idea of a text and analyze in detail its development over the course of the text, including how it emerges and is shaped and refined by specific details; provide an objective summary of the text,” (p. 40). The findings of this study indicated that the significant interactions occurred with the justification outcomes, and not with passage recall. In many cases (including the texts used for the current study), providing an “objective summary of the

text” would be difficult given the implicit or explicit argument made in the text in conjunction with an individual’s prior beliefs about the topic. By not allowing students to argue from their point of view (whether or not it is in agreement with the author’s point of view) may have consequences for learning. For example, in this study, simply the presence of an explicit argument had a significant effect on those students who read the persuasive text.

It is precisely the differences found with regards to these two outcomes that highlight the importance of considering the type of learning outcome used to gauge student learning. If objective assessments continued to be relied upon, the relative influence of the depth of one’s processing or the particular strategies used is likely to make little difference in their learning outcomes. However, if more subjective assessments are used (such as the justification task used in the current study), constructs such as levels of processing and subject-matter knowledge will play a critical role in how well students can perform.

## Appendix A

Summary Table of Studies for the Systematic Literature Review

Citation	Definition	Definition Code	Measure Code	Validity Code	Outcome Code	Predictor Code	Domain Code	Task Code
Ally, Gold, & Budson (2009)	Yonelinas et al. (1998)	R	CON		MR		DG	Y
Anmarkrud & Braten (2009)	Weinstein and Mayer (1986)	R	RDQ	CONSTRUCT	PO		SS	Y
Arteche, Chamorro-Premuzic, Ackerman, & Furnham (2009)	<i>Learning approaches define the individual's preferred style for learning and can be classified into three categories: deep (intrinsic motivation, engagement with the subject matter, and desire to know everything about the topic); surface (aim at learning the minimal amount necessary to pass); and achievement (goal-orientated study strategies)</i>	E	SPQ	NONE	IE	IE	DG	N
Aycicegi-Dinn & Caldwell-Harris (2009)	deep processing occurs with more novel words; shallow processing encompasses visual display	C	CON		MR		DG	Y
Bentley, Driver, & Dolan (2009)	Craik and Tulving, 1975; Baddeley, 1990	R	CON		MR	PI	DG	Y
Block (2009)	Craik & Lockhart, 1972	R	N		PO		DG	Y
Boatright-Horowitz, Langley, & Gunnip (2009)	Craik & Lockhart, 1972	R	CON		MR		DG	Y

Chamorro-Premuzic & Furnham (2009)	Biggs, 1992	R	SPQ	REF		PF	DG	N
Corson, Verrier, & Bucic (2009)		A	N				DG	Y
Gadzella, Masten, & Zascavage (2009)	shallow (physical analysis to stimuli); deep (semantic analysis)	C	ILP	REF		ST	DG	N
Henry, Van Lunen, Udermann, & Onate (2009)		A	RDQ	CONTENT	PS		PAA	N
Klein, Zwickel, Prinz, & Frith (2009)	Rayner, 1998	R	PHY		MSA	ToM	DG	Y
Marinkovic, Oscar-Berman, Urban, O'Reilly, Howard, Sawyer, & Harris (2009)	deep (semantic); shallow (perceptual)	C	CON		MR	PI	DG	Y
Meijer, de Groot, van Gerven, van Boxtel, & Jolles (2009)	deeper processing - activation of semantic information; shallow processing - word is only visually processed	C	CON		RT		DG	Y



Osorio, Ballesteros, Fay, & Pouthas (2009)	<i>Shallow encoding (processing words based on their orthographic or phonemic components) leads to a fragile memory trace susceptible to rapid forgetting. In contrast, deep encoding (semantic or meaning based processing) leads to a more durable memory trace</i>	E	CON	NONE		A	DG	Y
Pearce & Lee (2009)		A	N				BS	N
Phan (2009a)	<i>students may adopt a deep approach to learning with an intention to understand the authors' meaning and linking it to their prior knowledge and personal experience (Phan, 2006; Phan &amp; Deo, 2007). In contrast, students may also adopt a surface learning approach where the main emphasis is on studying merely for the intention of reproducing information without any further analysis (Murphy &amp; Tyler, 2005)</i>	E	RDQ	NONE	AP	AG	SS	N
Phan (2009b)	<i>students adopt a deep approach to learning when the main intention is to seek understanding from the author's meaning and relate this to prior knowledge and personal experience. A surface approach to learning, in contrast, involves engaging in learning simply for the intention of reproducing information without any detailed or further analysis</i>	E	RDQ	NONE		AG	SS	N

Phan (2009c)	<i>According to this theoretical framework, students may adopt a deep approach to learning, with the intention of understanding the author's meaning and linking it to their prior knowledge and personal experience (Phan &amp; Deo, 2007). In contrast, students may also adopt a surface learning approach, whereby their main emphasis is on studying merely for the intention of reproducing information without any further analysis (Murphy &amp; Tyler, 2005).</i>	E	RDQ	NONE	AP		DG	N
Rodriguez (2009)	<i>A deep approach suggests an intention to understand the material induced by the subject matter.; A surface approach reflects the student's intentino to reproduce the learning material through adopting routine procedures.</i>	E	ASSIST	CRITERION	ACP	AG	BS	N
Sanchez, Garcia-Rodicio, & Acuna (2009)		A	OUT	NONE	TC		PLS	Y
She & Chen (2009)		A	PHY	REF	EF	EF	PLS	Y
Shen & Dillard (2009)		M	CODE			MF	DG	Y

Toppino, Fearnow-Kenney, Kiepert, & Teremula (2009)	shallow (physical analysis); deep (semantic analysis)	C	CON		MR		DG	Y
Vogel, Kennedy, & Kwok (2009)	Biggs (2003)	R	CODE	REF	ACP		DG	N
Barnhardt & Geraci (2008)		M	CON		MR		DG	Y
Baeten, Dochy, & Struyven (2008)	<i>The deep approach to learning is associated with student intention to understand and to distil meaning from the content to be learned; The surface approach is characterised by a student's intention to cope with course requirements</i>	E	SPQ	NONE	AP	TI	BS	Y
Black & Plowright (2008)		A	CODE	CONSTRUCT			PLS	N
Blom & Severiens (2008)	rehearsal (surface learning); elaboration and critical thinking (deep learning)	C	MSLQ	CONSTRU CT	TP	CC	MM	Y

Cerdán & Vidal-Abarca (2008).	<p><i>Thus, whereas students answering high-level questions reviewed and connected more relevant units of information (i.e., a review-and-integrate search pattern), students answering low-level questions concentrated on locating isolated textual units (i.e., a locate-and-memorize search pattern). Hence, although high-level questions required the students to devote more effort to the task, they helped the students integrate text information, which is apparent when measuring deep but not superficial levels of comprehension.</i></p>	C	CODE	NONE	TI	PLS	Y	
Chamorro-Premuzic & Furnham (2008)	<p><i>Deep learners are intrinsically motivated and enjoy exploring the subject matter as much as they can. On the other hand, achieving students are extrinsically motivated and want to do well because of the rewards attached to high performance. Finally, surface learners are interested in learning the indispensable facts only and expend minimum effort to achieve this (Chamorro-Premuzic, Furnham, &amp; Lewis, 2007)</i></p>	E	RDQ	REF	TP	IQ	DG	N
Cheng, Lin, & Tsai (2008)	<p>Blaxton, 1989; Challis, Velichkovsky, &amp; Craik, 1996; Jacoby, 1983; Jacoby &amp; Dallas, 1981; Shimamura, 1986; Tulving &amp; Schacter, 1990; Weldon &amp; Roediger, 1987; Winnick &amp; Daniel, 1970</p>	R	CON		MR	DG	Y	

Coutinho & Neuman  
(2008)

*Deep processing, considered to be the most successful approach to learning, has been labelled elaborative processing or critical thinking (Weinstein and Mayer 1986). People using a deep-processing learning style challenge the authenticity of new information and focus on the content of the information in pursuit of comprehension. In contrast, people using a surfaceprocessing learning style fail to understand the true nature of the information. They adopt repetitive rehearsal and rote memorisation of information in order to learn and focus on verbatim recall of texts and facts. The last type of learning style is disorganisation, which refers to the learner's failure to process information due to inability to establish or maintain a structured, organised and orderly approach to learning.*

E

SSS

NONE

MC

AG

DG

N

Cutica & Bucciarelli (2008)	<p><i>deep learning involves the construction and manipulation of mental representations that reproduce the state of affairs described. The listener constructs such mental representations on the basis of the semantic and pragmatic information contained in the text, together with his or her prior knowledge, and any inferences that are drawn; they do not generally contain surface information (the linguistic form of sentences)</i></p>	E	CON	G	DG	Y	
Ferm & Johansen (2008)	<p><i>'Surface learning' denotes minimalist strategies for memorising and reproducing along with an intentional focus on passing assessments and is possibly connected to negative attitudes to learning. As opposed to this, deep learning is characterised by strategies for achieving meaningful learning, an intention to develop wider understanding and a positive attitude to learning. By meaningful learning we mean understanding the potential of what is learned for its utilisation within practices of music teaching, communication with peers and professors and strengthening of theoretical insights. It is characterised by 'the negotiation of new meanings – not just the acquisition of new skills or information' (Wenger, 2006: 12)</i></p>	E	CODE	NONE	TI	PAA	N

Froger, Taconnat, Landré, Beigneux, & Isingrini (2008)	M	CON	MR	DG	Y
Furnham, Christopher, Gardwood, & Martin (2008)	E	SPQ	REF	AP	DG N
<p><i>A student using a predominantly surface approach is supposedly goal-oriented rather than focused on deriving any intrinsic meaning from the task, and learns in a superficial manner (rote fashion) with the aim of achieving the minimum requirements. A student taking a deep approach is interested in reaching a meaningful understanding through extensive reading and research. A student with an achieving approach is highly committed to gaining good grades and would take a systematic approach to studying.</i></p>					
Gallo, Meadow, Johnson, & Foster (2008)	E	CON	MR	DG	Y
<p><i>The levels-of-processing effect refers to the finding that memory for a list of words is better when the meaning or semantics of the words is encoded (deep processing), relative to focusing on more superficial aspects of the words (shallow processing) such as their perceptual, phonological, or orthographic characteristics (e.g., Craik &amp; Tulving, 1975)</i></p>					

Heijne-Penninga, Kuks, Hofman, & Cohen-Schotanus (2008)	<p><i>Students who focus on rote learning with the intention of reproducing the learning material are using a surface approach. By contrast, deep learning can be described as focusing on understanding by applying and comparing ideas; In summary, deep learning seems to be characterised by three dimensions: 1 understanding: trying to understand the learning material by gaining an overview and creating outlines and structure; 2 elaboration: relating the learning material to other sources and personal ideas, and questioning and using evidence critically, and 3 analysis: trying to clarify the learning material by searching for its major points, finding reasons for what is being said and arriving at conclusions.</i></p>	E	DIP	REF	PT	PLS	Y
Jay, Caldwell-Harris, & King (2008)	<p><i>A shallow level of processing is one in which only superficial or physical aspects are encoded. A deeper level of processing takes more time and effort to activate the semantic meaning of the stimulus.</i></p>	E	CON	MR	DG	Y	
Justicia, Pichardo, Cano, Berbén, & De la Fuente (2008)		M	SPQ	CONSTRUCT		DG	N
Lau, Liem, & Nie (2008)	<p>deep: [self-regulatory and metacognitive]</p>	C	MSLQ	CONSTRUCT	AG	DG	N



Leelawong & Biswas (2008)	deep [clear conceptual organization of the materials]	C	CON		TP	FB	PLS	Y
Liem, Lau, & Nie (2008)	<i>Deep learning is characterized by such strategies as elaborating ideas, thinking critically, and linking as well as integrating one concept with another (Biggs, 1987). In comparison, surface learning is characterized by such strategies as memorization and reproduction of the learning materials (Biggs, 1987)</i>	E	RDQ	CONSTRUCT		AG	LA	N
Major & Horton (2008)		M	CON		MR		DG	Y
Nairne, Pandeirada, & Thompson (2008)		M	CON		MR		DG	Y
Nijhuis, Segers, & Gijsselaers (2008)	deep learning [relating, structuring, and critical processing]; surface learning [memorizing and analyzing]	C	SPQ	NONE		PLE	DG	N

Papinczak, Young, Groves, Haynes (2008)	<p><i>In the surface approach, students focus on recall of facts using strategies such as rote learning. They are usually motivated by external factors, such as fear of failure, which are extrinsic to the task (Biggs &amp; Moore, 1993); Students who adopt deeper approaches are task focussed and motivated by intrinsic factors, actively seeking to extract understanding and meaning from their learning (Entwistle, 2000), and utilising their metacognitive capabilities to monitor and self-regulate their learning (Blumberg, 2000)</i></p>	E	ASSIST	CONSTRU CT	SE	SE	PLS	Y
Phan (2008)	<p><i>students may adopt a deep approach to learning, with the intention of understanding the author's meaning and linking it to their prior knowledge and personal experience (Phan, 2006; Phan &amp; Deo, 2007). In contrast, students may also adopt a surface learning approach, whereby their main emphasis is on studying merely for the intention of reproducing information without any further analysis (Murphy &amp; Tyler, 2005)</i></p>	E	N		ACP	AG	SS	N
Powell, Tindal, & Millwood (2008)		A	N			PLE	DG	N

Senko & Miles (2008)	deep: probe the course material, such as connecting it to other concepts, generating personal examples, formulating opinions about it, or questioning its veracity; “surface” learning strategies that focus on memorizing and reciting terms, facts, and the broad essentials of a theory	E	RDQ	NONE		AG	SS	Y
Shell & Husman (2008)		M	SPOCK	NONE	AG	AG	DG	N
Shen, Hiltz, & Bieber (2008)	<i>A deep learning approach is consistent with a search for knowledge and understanding, whereas a surface learner is concerned only with passing exams by memorizing facts.</i>	E	RDQ	REF		CE	U	Y
Sins, van Joolingen, Savelsbergh, & van Hout-Wolters (2008)	<i>Deep cognitive processing, as described in the work of Marton and Sa’ljo” (1976, 1997), Ramsden (1992), and Entwistle (1988, 2001), involves active learning processes, such as relating ideas, looking for patterns and principles and attempting to integrate new information with prior knowledge and experience. Surface cognitive processing, in contrast, entails processes without much reflecting and involves treating the learning material as more or less unrelated bits of information. Surface processing does not implicate elaboration of the learning material and leads to more restricted learning processes.</i>	E	CODE	CONSTRUCT	GP	SE	PLS	Y

Sithiworachart & Joy (2008)	<i>Some students see learning as a matter of memorizing and comprehending knowledge only to cope with course requirements, and these are strategies in surface learning (Entwistle 2001); Others see learning as a way to satisfy their own requirements to develop new skills by relating previous knowledge with experiences, and these are strategies in deep learning (Entwistle 2001)</i>	E	N	CA	CS	Y
Weinstein, Bugg, & Roediger (2008)	Craig & Lockhart (1972)	R	CON	MR	DG	Y
You & Jia (2008)	<i>The deep learning approach is generally associated with meaning-based learning strategies (deep strategy (DS)) for the purpose of sensemaking and personal growth (deep motivation (DM)) while the surface-learning construct concerns the rote memorization of the narrow content (surface strategy) for the sake of passing the exams only (surface motivation) (Biggs et al., 2001; Scouller, 1998)</i>	E	SPQ	REF	SS	N

Brown, Aoshima, Bolen, Chia, & Kohyama (2007)	<p><i>surface learners' motivation is to repeat what they have learned without a primary concern toward understanding the material. Surface learners typically utilize rote memory as a learning strategy. This student tends to take a narrow view and concentrate on detail and thus is unable to distinguish principles from examples or generalize their knowledge; deep learners have the intention of understanding, engaging and valuing the subject. The student actively seeks to understand the material and relates ideas to one another. Deep learners tend to read and study beyond the course requirements.</i></p>	E	SPQ	NONE	CC	DG	N
Bryceson (2007)	personal reflective mechanism	C	N			MD	N

Cano (2007)	<p><i>The motivation of those learners deploying a surface approach is generally extrinsic (in order to fulfill course requirements and avoid failure with the least personal effort and involvement). They tend to resort to a repetitive strategy (memorising specific facts and accurately reproducing them). By contrast, the motivation of those deploying a deep approach tends to be intrinsic (they strive to understand the author's intent and seek self-fulfilment from the material). Their strategy is more meaningful (searching for meaning, integrating normal knowledge with personal experience, and relating facts to conclusions).</i></p>	E	LPQ	CONSTRUCT	ACP	CC	DG	N
Cassidy (2007)	<p><i>A deep approach to learning is characterized by intrinsic motivation, engagement with the subject matter, and the desire to know everything about a given topic. Conversely, students who opt for a surface approach to learning are not interested in the task per se, but aim at learning the minimum amount of material required to pass.</i></p>	M	ASSIST	NONE	ACP	PLS	N	
Chamorro-Premuzic, Furnham, & Lewis (2007)		E	SPQ	REF	PF	PF	PLS	N
Cotton & Gresty (2007)		M	CODE	CONTENT	RT	PLS	Y	

Coyne, McCoach, & Kapp (2007)	<i>Activities that encourage deep processing challenge students to move beyond memorizing simple dictionary definitions to understanding words at a richer, more complex level by, for example, describing how they relate to other words and to their own experiences (McKeown &amp; Beck, 2003).</i>	E	CON	NONE	VG		DG	Y
Craig et al (2007)		A	N	NONE	MR		DG	Y
Evans & Gibbons (2007)		M	CON			CSI	PLS	Y

Fenollar, Román, & Cuestas (2007)

*Marton and Saljo (1976) used the distinction between deep and surface learning to describe the qualitative differences in the processes students engage in while carrying out a prose-reading task, while Entwistle developed and clarified the distinction in terms of processing strategies (Entwistle, 1997; Tait & Entwistle, 1996). Deep learners actively try to understand meaning by working out relationships between concepts, relating new material to previously known information and adopting a critical attitude to information (Weinstein & Mayer, 1986). Surface learners on the other hand focus on memory strategies (i.e. rote processing) which emphasizes the retention of knowledge (Boyle, Duffy, & Dunleavy, 2003).*

E

RDQ

CRITERIO  
N

ACP

AG

DG

N



Furnham, Christopher, Garwood, & Martin (2007)	Surface Extrinsic/instrumental: Views the task as a demand to be met to reach a goal. Avoids personal meaning in the task Reproducing: rote learning relies on memorisation to a large extent; Deep Intrinsic: a search for meaning in text studies. A concern for greater understanding in specific areas. Motivated by a desire for a sense of deep personal satisfaction Meaningful: reading widely, extensive research. Attempts to theorise. Links past with present information. Attempts to build a global picture	E	SPQ	NONE	GK	DG	N	
Matthews, Lietz, & Darmawan (2007)	<i>Those students using a surface approach were unable to explain the central message of the article read and only recalled fragments of the material. However, those students adopting a deep approach were able to show a more global understanding of the author's intentions and even recalled extracts from the text.</i>	E	SPQ	NONE	TV	DG	N	
Neuville, Frenay, & Bourgeois (2007)	<i>The first two could be categorised as surface-level strategies [rehearsal and organization] while the other two involve more deep-processing [relating and critical thinking] (Entwistle &amp; Ramsden, 1983).</i>	E	RDQ	CONSTRU CT	ACP	TV	DG	N

Phan (2007)	Students may adopt a deep approach, which entails an intention to understand meaning and link it to prior knowledge and personal experience. In contrast, students may take a surface approach, where the main emphasis is on studying merely for the intention of reproducing information without any further analysis (Murphy & Tyler, 2005)	E	SPQ	NONE	SE		DG	N
Smith, Clegg, Lawrence, & Todd (2007)	<i>This dual reflection encourages deep learning where the student attempts to make sense of what is to be learnt (Gibbs, 1992, p. 2).</i>	C	CODE	NONE		RF	DG	N
Stewart, Holler, & Kidd (2007)	deep processing, such as when the identity of the pronominal referent is explicitly probed (e.g., through comprehension questions)	E	CON		TP		DG	Y
Thomas & Gadbois (2007)	<i>Students who have a surface learning approach are extrinsically motivated and work to avoid failure. They favour superficial learning strategies like rote memorization. In contrast, students who adopt a deep-learning approach are intrinsically motivated and strive to understand and integrate content with already learned information.</i>	E	SPQ; MSLQ	NONE	SH		SS	Y
Tian (2007)	<i>By contrast, assignment essay is said to focus on critical learning skills, to assess the ability of students to resolve problems themselves, and to encourage deep learning because it requires 'a response in continuous prose to a specific question, for which the student has received advance notice, which is to be prepared in the student's own time' (Biggs, 1988, p. 185)</i>	C	SPQ	NONE	ACP	LM	LA	Y
Wecker, Kohnle, & Fischer (2007)		A	CODE	NONE			PLS	Y
Whiteford & McAllister (2007)		A	N			CC	PLS	N
Wong & Lam (2007)	<i>Biggs (1987) differentiated three types of learning approaches in higher education: the deep approach, the achieving approach, and the surface approach. The deep</i>	E	SPQ	NONE	TP		SS	N

Abraham, Kamath, Upadhyaya, & Ramnarayan (2006)	<p><i>approach involves learning motives and strategies that are aimed at learning for learning's sake, and it connotes an abstract level of conceptualization and an internal drive to learn. The achieving approach involves learning motives and strategies that maximize the chance of obtaining good grades or improved career prospects. The surface approach involves the reproduction of that which is taught to meet the minimum external requirements of assessment. Students who adopt a surface approach memorise facts without intending to understand the subject, whereas students who adopt a deep approach intend to understand the material and relate it to personal experiences.</i></p>	E	RDQ	NONE	TI	PLS	N
Aharony (2006)	<p><i>deep approach: the ability to relate new information to previously acquired knowledge; to study different aspects of the material in order to obtain the entire picture; to search for a relevant meaning and a connecting point between the learning material and daily life and personal experiences. Other characteristics of this approach include the students' tendency to use metacognitive skills, to develop learning materials that create a basis for new ideas, to offer other solutions from an inquisitive-critical perspective, and from there, to search and discover their 'inner self' (Beishuizen &amp; Stoutjesdijk, 1999; Biggs, 1993; Entwistle, 1977); surface approach: a student's tendency to choose the quickest way to accomplish the task; to acquire the learning material without asking in-depth questions, to study the material in a linear manner; to relate to minimal aspects of material or to a problem without showing interest; or the need to understand it in its entirety; to learn by rote by relying on memory and not on comprehension; and to be concerned with the time needed to fulfil the learning task (Biggs, 1993). This learning approach, which focuses on memorizing the main elements, has almost no use for or expression of meta-cognitive skills.</i></p>	E	CODE	CONSTRUCT	CC	LA	Y
Igo, Riccomini, Bruning, & Pope	<p><i>mental processes required to create summaries and paraphrases (or to identify</i></p>	E	CODE	CONSTRUCT	MR	PLS	Y

(2006)	<i>and note main ideas) are deeper than those required to record verbatim notes</i>							
Klinger (2006)	Entwistle (2001) defines deeper learning as students actively transforming content and visualizing the conceptual connectedness. Deep learners: relate ideas to previous knowledge and experience, look for patterns and underlying principles, check evidence and relate it to conclusions, examine logic and argument cautiously and critically, are aware of the understanding that develops while learning, and become actively interested in the course content.	E	SPQ	NONE		LE	SS	N
Ragland et al (2006)	[deep is semantic]	C	CON		MR		DG	Y
Upton & Adams (2006)		A	RDQ	NONE	ACP		SS	Y
Zhou, Hu, Sun, & Huang (2006)		M	CON	NONE	MR		DG	Y
Biswas, Leelawong, Schwartz, & Vye (2005)		A	N				PLS	Y
Clump (2005)	[behaviors they employ to process the material, such as critically evaluating it, rewording class information and connecting it to their lives, focusing on facts and details, or using commonly prescribed study methods]	C	ILP	REF			DG	N
Goldman (2005)	<i>deep: when they seek the meaning of what they are learning and intentionally relate it to their existing knowledge</i>	E	CODE	CONTENT	SS		SS	Y

Groves (2005)	<i>intrinsic interest in the nature of the problem motivates students to develop a comprehensive understanding of all of the elements required for its solution</i>	C	SPQ	NONE	RS	RS	PLS	N
Henson, Hornberger, & Rugg (2005)	semantic decision (“deep” task) or orthographic decision (“shallow” task)	C	CON		MR		DG	Y
Minasian-Batmanian, Lingard, & Prosser (2005)	Prosser and Trigwell, 1999; Ramsden, 2003	R	CODE	REF			PLS	N
Nijhuis, Segers, & Gijsselaers (2005)	<i>These qualities [deep processing] are the ability to appropriately engage with, and respond to, the professional situations which they encounter, to understand the structural complexity of the task and the rationale behind facts, and to seek meanings.</i>	E	SPQ	NONE		TI	BS	N
Oulasvirta, Kärkkäinen, & Laarni (2005)		A	N				DG	Y
Pierce, Sullivan, Schacter, & Budson (2005)	Smith et al (2001)	R	CON		MR		DG	Y

Salovaara (2005)	<i>They [deep-processing strategies] include strategic processes such as explaining, reasoning, organizing, summarizing and elaborating (Weinstein &amp; Meyer 1986; Pintrich &amp; Schrauben 1992). Another class of strategies is surface-processing strategies, which include, for example, memorization and rehearsal strategies (Weinstein &amp; Meyer 1986).</i>	E	CODE	CONSTRUCT	LE	LA	Y
Abd-El-Khalick & Akerson (2004)		A	N		ACP	PLS	N
Boldini, Russo, & Avons (2004)	[elaborativeness of processing]	C	CON	MR		DG	Y
Gibson (2004)	<i>In the examination of expertise, for example, novices often transfer solutions when shallow primes (i.e., primes based on perceptual, surface features, such as the same word.s) are present: Novice solvers are more likely to solve a problem when the prime (the first problem solved) shares more surface characteristics than it does conceptual characteristics with the subsequent target problem (e.g., Nuvick, 1988).</i>	E	CON	MR		DG	Y
Hinojosa, Martín-Loeches, Muñoz, Casado, & Pozo (2004)		M	CON	MR		DG	Y
McParland, Noble, &	<i>Students who use surface learning</i>	E	SPQ	REF		PLS	Y

Livingston (2004)	<i>commonly learn by rote, with the aim of simply reproducing the material. In contrast, students using deep learning try to understand the meaning of the material being studied and relate it to previous knowledge and personal experiences. Students using strategic learning focus on achieving high grades, and at any particular time might elect to use either a surface or a deep approach depending on which they believe will be most successful.</i>							
Royet, Koenig, Paugam-Moisy, Puzenat, & Chasse (2004)	semantic descriptions or associations; accessing of stimulus related knowledge	C	CON	NONE	MR		DG	Y
Simon, Fagley, & Halleran (2004)	Maheswaran & Meyers-Levy, 1990	R	CON		RS		MM	Y
Vansteenkiste, Simons, Lens, Sheldon, & Deci (2004)		M	RDQ	CONSTRUCT		IER	BS	Y
Westman & Alexander (2004)	looking for conceptual understanding	C	LSS	CONSTRUCT			PRS	N

Yan & Kember (2004)	<p><i>a student who adopts a deep approach: . is interested in the academic task and derives enjoyment from carrying it out; . searches for the meaning inherent in the task (if a prose passage, the intention of the author); . personalises the task, making it meaningful to own experience and to the real world; . integrates aspects or parts of task into a whole (for instance, relates evidence to a conclusion), sees relationships between this whole and previous knowledge; and . tries to theorise about the task, forms hypothesis. And, a student who adopts a surface approach: . sees the task as a demand to be met, a necessary imposition if some other goal is to be reached (a qualification for instance); . sees the aspects or parts of the task as discrete and unrelated either to each other or to other tasks; . is worried about the time the task is taking; . avoids personal or other meanings the task may have; and . relies on memorisation, attempting to reproduce the surface aspects of the task (the words used, for example, or a diagram or mnemonic)</i></p>	E	CODE	NONE	LE	MD	Y
---------------------	--	---	------	------	----	----	---



Bandalos, Finney, & Geske (2003)	<i>Deep-processing strategies were defined as "challenging the veracity of information encountered and attempting to integrate new information with prior knowledge and experience" (p. 549), whereas surface processing included such things as rehearsal and memorization. Disorganization was defined as "difficulty in establishing or maintaining a structured, organized approach to studying" (p. 549)</i>	E	MSLQ	CONSTRUCT	ACP	SE	DG	N
Burgess & Weaver (2003)	deep: more holistic judgment; shallow: one physical feature	E	CON		MR		DG	Y
Carrier (2003)	deep: studying lecture notes, making chapter notes, highlighting, underlining, going to office hours; surface rereading textbook chapter	E	RDQ	NONE	ACP		SS	Y
Clump & Skogsberg (2003)		M	ILP	CONSTRUCT			DG	N
Daselaar, Veltman, Rombouts, Raaijmakers, & Jonker (2003)	[deep is semantic]	C	CON	NONE		PI	DG	Y
Elander (2003)		A	N			TI	SS	N

Gadzella & Baloglu (2003)	<i>deep - critically evaluates, conceptually organizes, and compares and contrasts information under study; elaborative processing - translates new information into his/her own terminology; fact retention - processing of specific factual information; methodical study - uses systematic techniques recommended in "how-to-study" manuals</i>	E	ILP	REF	ACP		SS	N
Mills (2003)	Freeman, 1995	R	N			TI	PLS	N
Nathanson & Botta (2003)	[how much one processes (quantity)]	C	CODE	CONSTRUCT	IP		DG	N
Watkins, McInerney, Akande, & Lee (2003)	<i>...surface or deep approaches to learning. In the former case, students typically tried to memorize details or key terms to be able to answer subsequent questions, whereas in the latter case the students tried to understand the message that the passage was trying to convey</i>	E	LPQ	CONSTRUCT			DG	N
Yamauchi & Miki (2003)		M	RDQ	NONE		AG	DG	N
Boekaerts (2002)		M	RDQ	CONSTRUCT		BL	SS	Y
Chang, Sung, & Chen (2002)	[deep processing is construction whereas shallow is passive]	C	N		TC		PLS	Y

Hill & Woodland (2002)	<i>Deep learning is the acquisition of higher order skills such as analysing, interpreting and evaluating information rather than simply amassing, reproducing and describing it. Deep learning is holistic, defined by an integration of facts to produce understanding, rather than atomistic, characterised by the accumulation of disparate facts (Entwistle &amp; Ramsden, 1983; Marton et al., 1984; Ramsden, 1988)</i>	E	CODE	NONE	KC	SS	Y
Fujii et al (2002)	deep (semantic processing); shallow (phonological processing)	C	CON		MR	DG	Y
Gadzella, Stephens, & Baloglu (2002)	<i>The Deep Processing scale measures high-order type of items which include analyses, evaluations, and comparisons of information; whereas, the Fact Retention scale measures low-order type of items which include factual information and memory recall.</i>	E	ILP	NONE	ACP	SS	Y
Jundt & Hinsz (2002)		A	N		AF	DG	Y

Lehman & Schraw (2002)	<i>shallow text processing, such as recognition and recall of text facts, though not necessarily deeper processing (Vidal-Abarca, Martinez, &amp; Gilabert, 2000); The effects of coherence on deeper processing are less consistent. Coherence sometimes increases the number of inferences while reading and the extent to which those inferences are linked within an integrated representation of the text in memory (Graesser et al., 1994; Kintsch, 1998; van den Broek, Risden, &amp; Husebye-Hartmann, 1995)</i>	E	CODE	NONE	TC	R	MD	Y
McLaughlin & Mandin (2002)		M	CODE	CONSTRUCT	ACP		PLS	N
Murphy & Alexander (2002)	<i>Surface-level reading strategies refer to techniques, such as rereading or omitting unfamiliar words, that facilitate the initial apprehension or deciphering of a written text. By comparison, deep-level processing strategies, which include such procedures as relating the text to prior knowledge or building a mental image, entail a personalization or transformation of the printed message.</i>	E	RDQ	CONSTRUCT	TC		SS	Y
N'Kaoua, Véron, Lespinet, Claverie, & Sztark (2002)	<i>Depth of processing varies along a continuum from shallow (sensory) processing to deep (semantic) processing.</i>	E	CON		MR		DG	Y

Toichi & Kamio (2002)	<i>Semantic processing (the use of semantic memory in encoding) of verbal materials facilitates LTM for the materials better than 'shallow' (phonological or perceptual) processing, which is known as the levels-of-processing effect</i>	E	CON		MR		DG	Y
Worthington (2002)		M	RDQ	NONE			SS	N
Al-Emadi (2001)	(Elliot et al, 1999)	R	SSS	CONSTRUCT	ACP	AG	DG	N
Barab, Thomas, & Merrill (2001)		A	CODE	NONE		LE	SS	N
Campbell, Smith, Boulton-Lewis, Brownlee, Burnett, Carrington, & Purdie (2001)	<i>The deep approach involves an intention to gain personal understanding, with corresponding strategies to gain meaning from the learning task. This 'meaning' aspect is complemented by a structural aspect whereby information is organised by integrating the whole and the parts. The surface approach describes an intention to avoid failure with corresponding strategies that facilitate the memorisation of facts without meaning and organisation (Ramsden, 1992)</i>	E	LPQ	REF	AT		DG	N
Segalowitz, Cohen, Chan, & Prieur (2001)	<i>We identify an aesthetic focus on the music with deep-level processing, analogous to deep-level (meaning) processing in studies of verbal memory. This deep-level processing can be contrasted with processing that focuses on the surface form of the music, such as the individual notes and their characteristics taken without</i>	E	CON		MR		PAA	Y

Sinha & Kumari (2001)	<i>reference to the ir contribution to the aesthetic dimension of the music (analogous to focusing on the letters or sounds that make up a word without reference to its meaning).</i>	M	RDQ	NONE		CC	DG	N
Smith, Tindell, Pierce, Gilliland, & Gerken (2001)		M	CON		MR		DG	Y
Vermetten, Lodewijks, & Vermunt (2001)	covering four learning components: cognitive processing, metacognitive regulation, mental learning models, and learning orientations	C	ILS	NONE	SS		DG	N
Cassidy & Eachus (2000)	<i>A deep approach is associated with intrinsic motivation where education is seen as a means of personal development and where there is learning for learning sake. A surface approach is associated with outcome goals rather than learning goals, that is, education to obtain qualiŽ cations (extrinsic motivation) not understanding, and a fear of failure rather than a need for achievement (Biggs, 1978; Entwistle &amp; Ramsden, 1983)</i>	E	ASSIST	NONE	ACP	SE	SS	Y
deVries & Downie (2000)		A	N				DG	N

Hara, Bonk, & Angeli (2000)	<i>depth of processing which differentiates surface level processing (e.g., repeating of what was already stated) from deep level processing (e.g., providing the advantages and disadvantages of a situation)</i>	E	CODE	CONSTRUCT			SS	Y
Martin, Stark, & Jolly (2000)		M	LSI	REF	CE		PLS	N
Moreno & Mayer (2000)		M	CODE	NONE	TP	PM	PLS	Y
Sauzéon, N'kaoua, Lespinet, Guillem, & Claverie (2000)	<i>in this context, it is possible to distinguish between different levels of quality within conceptual processes such as phonetic processing and deeper semantic levels of processing</i>	C	CON		MR		DG	Y
Sedikides & Green (2000)	[deep is elaboration]	C	CON		MR		DG	Y
Verma (2000)		M	ILS	REF	SE	SE	DG	N
Yamauchi & Miki (2000)		M	LSS	CONSTRUCT		AG	DG	N
Aaron & Skakun (1999)	<i>The surface (or instrumental learner memorizes lists of superficial knowledge, the strategic (or achieving) learner focuses on the requirements of assessment, and the deep learner searches for understanding and meaning.</i>	E	ASSIST	CONSTRUCT		ACP	PLS	N

Beishuizen &  
Stoutjesdijk (1999)

*Deep processing is characterised by actively relating various parts of the text to one another and to prior knowledge, by organising separate topics into a whole to gain insight into the subject matter, and by a critical approach. Surface processing, in contrast, is characterised by step-by-step analysis of the material and reading of each part of the text thoroughly, with great attention to detail and focus on factual information, by an attempt to remember as much as possible, and by the use of study strategies like rehearsing and memorising.*

E

ILS

NONE

SS

SS

Y



Elliot, McGregor, & Gable (1999)	<p><i>Deep processing (also labeled elaboration or critical thinking; e.g., Weinstein &amp; Mayer, 1986) involves challenging the veracity of information encountered and attempting to integrate new information with prior knowledge and experience, whereas surface processing (also labeled rehearsal or memorization; e.g., Zimmerman &amp; Pons, 1986) involves the repetitive rehearsal and rote memorization of information (Entwistle &amp; Ramsden, 1983). The deep-surface processing distinction is widely regarded in study strategy literature as having conceptual and predictive utility. Disorganization refers to the learner's difficulty in establishing or maintaining a structured, organized approach to studying (Entwistle, 1988).</i></p>	E	RDQ	CONSTRU CT	ACP	AG	SS	Y
Gibbs (1999)		M	ASSIST	NONE		LE	PRS	Y
Grady, McIntosh, Rajah, Beig, & Craik (1999)	<p><i>deep encoding, such as semantic processing of items; shallow encoding, such as purely perceptual judgements</i></p>	E	CON		MR		DG	Y
O'Brien & Hart (1999)	<p><i>surface learning they suggest is where students attempt to 'spot facts' in, say, a given text, particularly those they perceive as a ready focus for questions or assessment by teachers. This 'fact spotting' is followed by committing the facts to memory, focusing on the surface level of the material under study; any attempt to understand what the author of a given text might be saying, to focus on underlying meaning and seeking to integrate the components of the text is seen as evidence of 'deep learning'</i></p>	E	N				BS	N

Sullivan & Cantwell (1999)	<i>The deep approach, for example, is marked by an intention to understand content in a structurally complex way, while the surface approach is marked by a desire to simplify structural complexity to a point of minimally acceptable competence (e.g., Biggs, 1979; Cantwell, 1997; Marton and Saljo, 1976).</i>	E	SPQ	CONSTRUCT	SS		PAA	Y
Bentin, Moscovitch, & Nirhod (1998)	deep: [semantic processing]	C	CON	CONSTRUCT	MR		DG	Y
Freed, Marshall, & Phillips (1998)	deep: [semantic processing]; shallow: [phonetic processing]	C	CON		MR		DG	Y
Gadzella & Masten (1998)		M	ILP	REF		AM	DG	N
Gadzella, Masten, & Stacks (1998)		M	ILP	REF	ST	ST	DG	N
James (1998)		M	CODE	NONE	TP		DG	Y
Onion & Bartzokas (1998)	<i>superficial cognitive processing of the information by recipients (shallow, sensory appraisals); deep processing of information where messages are related to previous knowledge and experience</i>	E	CON		AT		PLS	Y
Vu, Van Der Vleuten, & Lacombe (1998)	<i>[with a deep approach] students were more involved in meaningful learning and in reaching understanding, and were motivated by an interest in the subject matter or vocational relevance.</i>	E	ILP	NONE			PLS	N
Alexander, Murphy, Woods, Duhon, & Parker (1997)	<i>text-based strategies (e.g., rereading); deeper-processing strategies (e.g., building a mental representation)</i>	E	RDQ	CONSTRUCT	TP		SS	Y

Brébion, Amador, Smith, & Gorman (1997)		M	CON		MR	O	DG	Y
Evans & Honour (1997)	<i>Students who demonstrated a deep approach to learning start with the intention of understanding the meaning of the article, questioned the author's arguments, and related them both to previous knowledge and to personal experience, and tried to determine the extent to which the author's conclusions seemed to be justified by the evidence presented. Surface-level approaches were characterised by those students who were intent upon memorising those parts of the article which they considered to be important in view of the types of questions they anticipated afterwards. Their focus of attention was thus limited to the specific facts or pieces of information which was rote learned. (Marton &amp; Saljo, 1976a)</i>	E	CODE	NONE		OL	SS	Y
Gadzella, Ginther, & Bryant (1997)		M	ILP	NONE	ACP		SS	Y
Gadzella, Ginther, Masten, & Guthrie (1997)	deep: evaluate, conceptualize, organize, and compare information; shallow: rote learning and memorization of facts	E	ILP	NONE	ST		DG	N

Sandberg & Barnard (1997)	<i>Deep knowledge can be described in terms of four distinct aspects: Knowledge of multiple models and multiple viewpoints of the domain, such as structural, behavioural, causal and teleological (Kamsteeg, 1992; White &amp; Frederiksen, 1990). Knowledge about the relations between different models and viewpoints. Knowledge of reasoning procedures to solve quantitative and qualitative problems. Knowledge of first principles and knowledge to reason on their basis in order to solve novel or unfamiliar problems (Chi et al., 1980)</i>	E	CODE	NONE		IP	PLS	Y
Tehan, Fallon, & Randall (1997)	<i>rating items for pleasantness is a prototypic deep-level process</i>	E	CON		MR		DG	Y
Weldon & Bellinger (1997)		M	CON		MR		DG	Y
Greene & Miller (1996)	deep: [meaningful cognitive engagement]	C	RDQ		ACP	AG	SS	N
Grosschalk & Gregg (1996)	<i>shallow processing invokes mainly data driven processing; deep tasks invoke conceptually driven processing</i>	E	CON			HY	DG	Y
Hancock, Stock, Kulhavy, & Swindell (1996)		M	RDQ	CONTENT		GN	DG	N
Meehan & Pilotti (1996)		M	CON		MR	AP	DG	Y
Nichols (1996)		M	RDQ	NONE		LE	MM	Y
Niles (1996)	surace - [rote]	C	SPQ	CONSTRUCT		CC	DG	N

Pitts & Heap (1996)	M	CODE	NONE	MR		DG	Y
Shum, Sweeper, & Murray (1996)	M	CON		MR		DG	Y
Verheij, Stoutjesdijk, & Beishuizen (1996)	E	ILS	CRITERIO N	LS	LS	SS	Y
			<i>Vermunt and Van Rijswijk (1988) distinguish between deep processing and surface processing of text. Deep processors draw relationships between various unrelated parts of the text by activating prior knowledge. They organize separate topics into a whole and look at the issues in a critical way. Surface processors, in contrast, analyze the material step by step, with much attention to details. They focus on factual information and try to remember as much as possible, using study activities like rehearsing and memorizing.</i>				
Watson (1996)	A	CODE	NONE	TP		BS	N

Hilliard (1995)	<i>Students who adopt a surface approach are predominantly motivated either by a desire to simply complete the course or by a fear of failure. Their intention is to fulfill the course requirements by memorizing and reproducing specific facts or pieces of disconnected information for examinations, they tend to be anxiously aware of assessment requirements and prefer to restrict learning to a defined syllabus and specified tasks.; students who adopt the deep approach are predominantly motivated by an interest in learning for its own sake and an interest in the subject material, they attempt to understand the underlying structure and meaning. Students examine evidence critically, use it cautiously, and actively relate new information to previous knowledge.</i>	E	SPQ	CONTENT	ACP	CC	DG	Y
Muthukrishna & Borkowski (1995)		M	RDQ	NONE	TP	TI	MM	Y
Thill & Brunel (1995)	sensory (shallow); semantic (deep)	C	RDQ	CONSTRUCT		AG	PAA	Y
Wautier & Westman (1995)	deep: emphasizing underlying concepts	E	ILP	NONE	TP		DG	Y
Albaili (1994)		M	ILP	REF		GPA	DG	N

Beishuizen, Stoutjesdijk, & van Putten (1994)	deep: relating & structuring, actively processing information critically; surface: memorizing and rehearsal, analyzing of small units step by step	E	ILP	NONE	TC	SS	Y	
Cantwell & Millard (1994)	<i>deep- high levels of intrinsic motivation coupled with the utilisation of a sufficiently broad strategic repertoire to enable the motivational goals of learning to be advanced; mindless (surface) - less effortful and less purposeful application of routine procedures across all domains of learning</i>	E	LPQ	REF	SS	PAA	Y	
Verma (1994)		M	ILP	NONE		DG	N	
Albaili (1993)	deep: efforts to organize conceptually and evaluate information critically	E	ILP	CONSTRUCT	ACP	DG	N	
Barton & Sanford (1993)		M	CON		TC	DG	Y	
Jakoubek & Swenson (1993)	deep [beyond rote learning]	C	ILP	REF	ACP	CC	DG	N
Richardson & Fergus (1993)		M	ILP	REF		PAP	DG	N
Sandman (1993)		M	CON		MR	DG	Y	
Terry (1993)	deep processing (as in judging traits); shallow processing (as in focusing on physical features)	C	CON		MR	DG	Y	
Westman (1993)		M	ILP	NONE	ACP	MD	N	
Kozminsky & Kaufman (1992)		M	ILP	CONSTRUCT	ACP	MD	N	

Volet & Chalmers (1992)		M	RDQ	NONE	ACP	AG	SS	N
Wilding & Hayes (1992)	<i>surface learning concentrates on rote learning of facts without elaboration or connection with existing knowledge for the learner, and is motivated mainly by fear of failure; Deep learning attempts to understand and interconnect material and is motivated by interest in the material for its own sake</i>	E	SPQ	NONE	SS		SS	Y
Graham & Golan (1991)	Craik & Tulving (1975)	R	CON		MR	AG	DG	Y
Kirby & Pedwell (1991)	<i>Deep learners say that they use meaningful learning strategies, and are intrinsically motivated; Surface learners report rote learning strategies and extrinsic motivation.</i>	E	CON	NONE	TC		SS	Y
Sporer (1991)	deep: [more elaborated]	C	CON		FR		DG	Y
Fergusson-Hessler & de Jong (1990)	superficial processing - reading the text, comparing symbols in text and figures; deep is divided into: integrating - bringing structure into new knowledge and connecting - relating new knowledge to previous knowledge	E	CODE	CONTENT	TP		PLS	Y
Grodsky & Giambra (1990)	shallow [simple stimuli]; deep [semantic]	C	CON		TP		MD	Y
Nolen & Haladyna (1990a)		M	RDQ	CONSTRUCT		TO	PLS	N



Nolen & Haladyna (1990b)	deep - integration of new information and prior knowledge [by three processes elaborating, monitoring, and organizing]; surface - rote learning of facts and formulas, reading the whole chapter, and trying to memorize as much information as possible	E	RDQ	CONSTRUCT	SS	TO	PLS	N
Skaggs et al (1990)		M	ILP	NONE	MR		PLS	Y
Tooth, Tonge, & McManus (1989)	surface - emphasizes rote-learning of facts motivated by fear of failure; deep - stresses understanding of underlying concepts, motivated by interest and a personal need for understanding	E	SPQ	NONE	ACP		PLS	N
McKelvie & Pullara (1988)		M	CON		MR		DG	Y
Nolen (1988)	<i>deep processing strategies include discriminating important information from unimportant information, trying to figure out how new information fits with what one already knows, and monitoring comprehension; surface level strategies include simply reading a whole passage over and over, memorizing all the new words, and rehearsing information</i>	E	RDQ	CRITERION	MR	AG	PLS	Y
Gadzella, Ginther, & Williamson (1987)	deep: students who developed a thorough understanding of the material studied; shallow: devoted more time to rote learning of specific knowledge	E	ILP	CRITERION	ACP		DG	N

Koivukari (1987)	surface: mere memorization of signifiers; comprehension [deep]: all aspects of the linguistic sign have been assimilated	E	CODE	CONTENT	TI	SS	Y
Maylor, Rabbitt, & Kingstone (1987)		M	CON	MR		DG	Y
Miller, Alway, & McKinley (1987)	<i>deep processing style is characterized by looking for differences and similarities among topics, encoding main ideas and supporting details, comparing and contrasting ideas, detailed and systematic organization of information, and critical analysis of relationships; reproductive style invovles writing down exact words</i>	E	ILP	REF	ACP	DG	N
Vaid (1987)		M	CODE	NONE	LP	DG	Y
Arlin (1986)	quantity (high or low) of information processed	C	CON	TPR		DG	Y
Boyd & Ellis (1986)	<i>deep processing relies on an associative network</i>	C	CON	MR		DG	Y
Gadzella, Ginther, & Williamson (1986)		M	ILP	CONSTRU CT	ACP	DG	N
Schommer & Surber (1986)		M	CON	TC		SS	Y
Geva & Ryan (1985)		M	CON	TC		PLS	Y
Meier & Schmeck (1985)		M	ILP	REF	BO	DG	N

Alesandrini, Langstaff, & Wittrock (1984)	<i>depth implies a greater degree of semantic or cognitive analysis</i>	E	ILP	NONE	TP		PLS	Y
Dixon & von Eye (1984)	deep [semantic]; shallow [nonsemantic]	C	CON		MR		DG	Y
Meier, McCarthy, & Schmeck (1984)		M	ILP	REF	TP	SE	LA	Y
Camp, Markley, & Kramer (1983)		M	CODE	CRITERION	MR		DG	Y
Lockhart & Schmeck (1983)	deep: conceptual, spending more time categorizing, critical evaluating appropriateness of categorization, and contrasting categorizations with one another	E	ILP	CONSTRUCT	ACP		SS	N
McCarthy & Meier (1983)	deep - searching for meaning, comparing and contrasting ideas and elaborating concepts; shallow - rote memorization in its given form	E	ILP	REF	TP		LA	N
McClain (1983)	[progressively deeper from graphemic encoding to semantic encoding]	C	CON		MR		DG	Y
Silberman, Weingartner, Laraia, Byrnes, & Post (1983)		M	CODE	NONE	MR		DG	Y

## Appendix B

### Expository and Refutational Text Passages

#### Expository Text

**Adapted from: Stephen Webb, “Where is Everybody?: Fifty Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life.” Copernicus Books, 2002.**

While working at Los Alamos National Laboratory in the 1950s, the physicist Enrico Fermi was chatting with a few of his colleagues about the likelihood of interstellar travel in the next ten years when he asked, “Where is everybody?” His question addressed the contradiction that we (Earthlings) would *expect* to find some form of intelligent life on another planet, yet there is no direct evidence that extraterrestrial civilizations (ETCs) exist anywhere in the universe (beyond a few suspect pictures of UFOs). This simple yet profound question became known as the Fermi paradox and has sparked heated debate over the past 60 years.

How do we begin to address the Fermi paradox? Many scientists suggest that the Drake equation can be used as a helpful tool to estimate the number of intelligent civilizations in the universe, whether it is just one (i.e., Earthlings) or many ETCs in addition to us. This Drake equation is written as:

$$N = R \times f_p \times n_e \times f_l \times f_i \times f_c \times L$$

where N represents the number of intelligent communicating civilizations in the galaxy; R is the yearly rate at which stars form in the galaxy;  $f_p$  represents the fraction of stars that possess planets;  $n_e$  represents the number of planets with habitable environments suitable for life within those stars with planets;  $f_l$  represents the fraction of suitable planets on which life actually develops;  $f_i$  represents the fraction of planet on which life

develops intelligence;  $f_c$  represents the fraction of intelligent life-forms that develop a culture capable of interstellar communication; and finally,  $L$  represents the time in years that such a culture will devote to interstellar communication.

Over the past 60 years three different resolutions of the Fermi paradox have emerged. First, some claim that “they” (ETCs) are already here or have been here and left evidence of their existence (meaning  $N$  in the Drake equation would be greater than one). Second, some claim that ETCs exist but have not yet communicated (meaning  $N$  in the Drake equation would be greater than one). Finally, some claim that ETCs do not exist and we are alone in the universe (meaning  $N$  in the Drake equation would equal one).

#### **“They” are here or “they” were already here**

The first of these positions is that ETCs have already visited Earth or are here now. These views make values in the Drake equation high and assume that there are many ETCs in the galaxy. One such view is that aliens have been observing us by sending small unmanned probes, which we have not yet detected, rather than a whole fleet of ships. After all, the solar system is large and there are plenty of places from which ETCs could observe Earth without being detected. One set of places are the Earth-Moon Lagrangian points. Lagrangian points are places around two orbiting masses (such as the Earth and Moon) where a third smaller body can maintain a fixed distance from the larger masses using their gravitational pull. In fact, NASA has used these Lagrangian points with quite a few American spacecraft. While we have certainly scanned these areas for such probes and none have been found, it is important to remember that these are very large expanses of space.

Another response that fits into this category is panspermia, which Crick and Orgel forwarded in 1973. Panspermia refers to the idea that life originated elsewhere and was somehow transported to Earth. This implies that our DNA is extraterrestrial in origin and our galactic ancestors are likely somewhere out in the galaxy, making  $N$  in the Drake equation greater than one. In this view we are *also* aliens! Numerous studies of genetics demonstrate that all life on Earth is highly related, which some use as evidence that our DNA came from a single origin outside of Earth. For example, some studies indicate that 75% of human genes or close variants exist in worms, that we share half our genes with the banana, and that 98.4% of human DNA is similar to that of ape DNA. Specifically, Crick and Orgel discussed the idea of *directed panspermia* in which they claimed ETCs may have deliberately seeded planets such as Earth by aiming spores full of DNA at many planets.

A third response in this category is the so-called zoo scenario forwarded by Ball in 1973. This view specifically addresses *why* ETCs would be observing Earth. This scenario assumes that ETCs with advanced technology are common and that these advanced ETCs are in control of the universe, while less advanced civilizations will be destroyed, tamed, or assimilated. The zoo scenario argues that Earth could be a “wilderness” area that is not to be disturbed and has been set aside by a more advanced ETC so that it can develop naturally, similar to what humans do in wildlife sanctuaries and zoos. In this scenario, we are the creatures behind the iron bars with ETCs watching us. In other words, an ETC or multiple ETCs have set aside Earth in order to let mankind evolve to a state of perfection.

**They exist but have not yet communicated**

The second possible resolution of the Fermi paradox is that ETCs exist, but that they have not yet been able to communicate with us or we have been unable to listen to their communications, meaning as in the previous scenarios that  $N$  in the Drake equation is greater than one. There have been a number of explanations as to why such communication has not occurred. First, the distances between stars that would possess planets capable of supporting life are *very* great, too great to permit interstellar travel. For example, the Voyager probe took 21 years to reach Pluto and will not reach the closest star for tens of thousands of years. The closest star to Earth is Proxima Centauri. Even if one could travel at the speed of light (theorized as the “universal speed limit”), which for now seems unlikely, it would still take four years to reach the nearest star, and there are no guarantees that there are habitable planets around that star. The amount of time required for interstellar travel at sub-light speeds make it very impracticable.

Additionally, it has been forwarded that ETCs are out there and are signaling, but we simply do not know how to listen for them. For instance, Scientists at the SETI (Search for Extra Terrestrial Intelligence) Institute "listen" primarily for electromagnetic signals such as visible light, radio waves, and infrared waves. Perhaps ETCs are using other types of signals such as gravitational signals, particle signals like neutrinos, or hypothetical tachyon signals that we may not be able to detect or interpret. Tachyons are any theoretical particles that have an imaginary mass. This inability to listen may explain the “great silence” in the universe.

A third explanation under this category is that planetary catastrophes ( $n_e$  in the Drake equation) prevent ETCs from communicating because they were destroyed before they develop the technology to do so. One explanation is that it may be common for

meteors to destroy intelligent life on planets *before* they are able to achieve interstellar travel or communication. Hundreds of small meteors hit the Earth every day. Others have suggested that perhaps the urge for nuclear war is great and as civilizations keep becoming advanced they keep annihilating themselves. Another explanation is that overpopulation destroys intelligent life before it becomes able to colonize other planets. Each of these catastrophes may preclude *both* interstellar travel and communication.

### **They do not exist**

The third class of explanations is that intelligent ETCs do not exist and we are alone in the galaxy. Without any intelligent ETCs, the value for N in the Drake equation is one; namely Earthlings. The first of these explanations is that complex life requires a planet like Earth to originate and evolve. However, these planetary systems may be very rare. Although we have not yet seen planets like Earth in other solar systems directly, we have already inferred their existence by examining the gravitational pull that they exert on their parent stars. This gravitational pull creates a “wobble” in the parent star’s visible light emissions. This wobble is referred to as the Doppler effect.

A second explanation as to why we are alone in the universe is that Earth has an optimal “pump of evolution.” One of the key mechanisms in evolution is survival of the fittest. In other words, there are certain events that must occur to move evolution forward. For example, Jupiter’s gravitational pull infrequently pulls asteroids across the path of the Earth. As discussed, larger asteroids do infrequently hit planets such as Earth, but perhaps these infrequent and well-timed larger planetary strikes can be beneficial in the long run. For example, if the meteor impact of 65 million years ago had not happened, Earth might still be home to the dinosaurs so that mammals could not evolve as they have. Without



these infrequent crashes new more intelligent species may not be able to evolve, thus infrequent asteroid strikes provide an “optimal pump of evolution” conducive to intelligent life that other planetary systems may not have.

Finally, planetary systems may just be very dangerous places. Unintelligent life on other planets may have to deal with many threats both from space and on their own planet that do not allow sufficient time for intelligence to develop. These threats could be from meteor impacts or supervolcano eruptions. One such supervolcano on Earth is in Yellowstone National Park. The corresponding mass extinction events may make it impossible for life to hold on very long. In fact, there are data from the Holocene epoch (the past 10,000 years to the present) that depict a mass extinction event; this one due to human activity. The rate at which species are becoming extinct is estimated to be 120,000 times the “normal” or “background rate.” Perhaps as life starts to become intelligent, it extinguishes itself much more quickly than the responses in the second category of responses.

Each of these three responses to the Fermi paradox represents different values associated with the Drake equation. The first two responses (i.e., they are here *or* that they are out there and just have not communicated), assume  $N$  in the Drake equation to be greater than one. Some assume that not only is  $N$  greater than one, but that ETCs are plentiful in the galaxy. On the other hand, the last response (i.e., we are alone), assumes that some or all of the variables on the right side of the Drake equation are very small, making  $N$  equal to one, with Earth possessing the only intelligent species interested in and capable of interstellar travel and communication in the universe.

### **Persuasive Text**

**Adapted from: Stephen Webb, “Where is Everybody?: Fifty Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life.” Copernicus Books, 2002.**

While working at Los Alamos National Laboratory in the 1950s, the physicist Enrico Fermi was chatting with a few of his colleagues about the likelihood of interstellar travel in the next ten years when he asked, “Where is everybody?” His question addressed the contradiction that we (Earthlings) would *expect* to find some form of intelligent life on another planet, yet there is no direct evidence that extraterrestrial civilizations (ETCs) exist anywhere in the universe (beyond a few suspect pictures of UFOs). This simple yet profound question became known as the Fermi paradox and has sparked heated debate over the past 60 years.

How do we begin to address the Fermi paradox? Many scientists suggest that the Drake equation can be used as a helpful tool to estimate the number of intelligent civilizations in the universe, whether it is just one (i.e., Earthlings) or many ETCs in addition to us. This Drake equation is written as:

$$N = R \times f_p \times n_e \times f_l \times f_i \times f_c \times L$$

where  $N$  represents the number of intelligent communicating civilizations in the galaxy;  $R$  is the yearly rate at which stars form in the galaxy;  $f_p$  represents the fraction of stars that possess planets;  $n_e$  represents the number of planets with habitable environments suitable for life within those stars with planets;  $f_l$  represents the fraction of suitable planets on which life actually develops;  $f_i$  represents the fraction of planet on which life develops intelligence;  $f_c$  represents the fraction of intelligent life-forms that develop a culture capable of interstellar communication; and finally,  $L$  represents the time in years that such a culture will devote to interstellar communication.

Over the past 60 years three different resolutions of the Fermi paradox have emerged. First, some claim that “they” (ETCs) are already here or have been here and left evidence of their existence (meaning  $N$  in the Drake equation would be greater than one). Second, some claim that ETCs exist but have not yet communicated (meaning  $N$  in the Drake equation would be greater than one). Finally, some claim that ETCs do not exist and we are alone in the universe (meaning  $N$  in the Drake equation would equal one).

**“They” are here or “they” were already here**

The first of these positions is that ETCs have already visited Earth or are here now. These views make values in the Drake equation high and assume that there are many ETCs in the galaxy. One such view is that aliens have been observing us by sending small unmanned probes, which we have not yet detected, rather than a whole fleet of ships. After all, the solar system is large and there are plenty of places from which ETCs could observe Earth without being detected. One set of places are the Earth-Moon Lagrangian points. Lagrangian points are places around two orbiting masses (such as the Earth and Moon) where a third smaller body can maintain a fixed distance from the larger masses using their gravitational pull. While we have certainly scanned these areas for such probes and none have been found, it is important to remember that these are very large expanses of space. Keeping this vast expanse of space in mind, there have been no plausible reports of extraterrestrial spacecraft in these regions (or any others) to date.

Another response that fits into this category is panspermia, which Crick and Orgel forwarded in 1973. Panspermia refers to the idea that life originated elsewhere and was somehow transported to Earth. This implies that our DNA is extraterrestrial in origin and our galactic ancestors are likely somewhere out in the galaxy, making  $N$  in the Drake

equation greater than one. In this view we are *also* aliens! Numerous studies of genetics demonstrate that all life on Earth is highly related, which some use as evidence that our DNA came from a single origin outside of Earth. Specifically, Crick and Orgel discussed the idea of *directed panspermia* in which they claimed ETCs may have deliberately seeded planets such as Earth by aiming spores full of DNA at many planets. One problem with this particular solution is that it is very difficult to test empirically.

A third response in this category is the so-called zoo scenario forwarded by Ball in 1973. This view specifically addresses *why* ETCs would be observing Earth. This scenario assumes that ETCs with advanced technology are common and that these advanced ETCs are in control of the universe, while less advanced civilizations will be destroyed, tamed, or assimilated. The zoo scenario argues that Earth could be a “wilderness” area that is not to be disturbed and has been set aside by a more advanced ETC so that it can develop naturally, similar to what humans do in wildlife sanctuaries and zoos. In other words, an ETC or multiple ETCs have set aside Earth in order to let mankind evolve to a state of perfection. A major problem with this scenario is that there is no way to empirically test it.

### **They exist but have not yet communicated**

The second possible resolution of the Fermi paradox is that ETCs exist, but that they have not yet been able to communicate with us or we have been unable to listen to their communications, meaning as in the previous scenarios that  $N$  in the Drake equation is greater than one. There have been a number of explanations as to why such communication has not occurred. First, the distances between stars that would possess planets capable of supporting life are *very* great, too great to permit interstellar travel. For

example, the Voyager probe took 21 years to reach Pluto and will not reach the closest star for tens of thousands of years. Even if one could travel at the speed of light (theorized as the “universal speed limit”), which for now seems unlikely, it would still take four years to reach the nearest star, and there are no guarantees that there are habitable planets around that star. The amount of time required for interstellar travel at sub-light speeds make it very impracticable. While this may explain why we have not *seen* ETCs, it does not explain why they have not communicated, since interstellar communication should be easier than interstellar travel.

Additionally, it has been forwarded that ETCs are out there and are signaling, but we simply do not know how to listen for them. For instance, Scientists at the SETI (Search for Extra Terrestrial Intelligence) Institute "listen" primarily for electromagnetic signals such as visible light, radio waves, and infrared waves. Perhaps ETCs are using other types of signals such as gravitational signals, particle signals like neutrinos, or hypothetical tachyon signals that we may not be able to detect or interpret. This inability to listen may explain the “great silence” in the universe. While this response is testable, it is problematic given the wide range of communication options ETCs could employ and it would be rather difficult to disprove.

A third explanation under this category is that planetary catastrophes ( $n_e$  in the Drake equation) prevent ETCs from communicating because they were destroyed before they develop the technology to do so. One explanation is that it may be common for meteors to destroy intelligent life on planets *before* they are able to achieve interstellar travel or communication. Others have suggested that perhaps the urge for nuclear war is great and as civilizations keep becoming advanced they keep annihilating themselves.

Another explanation is that overpopulation destroys intelligent life before it becomes able to colonize other planets. Each of these catastrophes may preclude *both* interstellar travel and communication. However, if this view posits that ETCs are common, it seems unlikely that *every* ETC have faced crises such as these.

### **They do not exist**

The third class of explanations is that intelligent ETCs do not exist and we are alone in the galaxy. Without any intelligent ETCs, the value for N in the Drake equation is one; namely Earthlings. The first of these explanations is that complex life requires a planet like Earth to originate and evolve. However, these planetary systems may be very rare. Although we have not yet seen planets like the Earth in other solar systems directly, we have already inferred their existence by examining the gravitational pull that they exert on their parent stars. This gravitational pull creates a “wobble” in the parent star’s visible light emissions. While new planets are being found each day, few of these planets appear to be in habitable zones which could support life, lending credence to the view that we may be alone in the universe.

A second explanation as to why we are alone in the universe is that Earth has an optimal “pump of evolution.” In other words, there are certain events that must occur to move evolution forward. For example, Jupiter’s gravitational pull infrequently pulls asteroids across the path of the Earth. As discussed, larger asteroids do infrequently hit planets such as Earth, but perhaps these infrequent and well-timed larger planetary strikes can be beneficial in the long run. For example, if the meteor impact of 65 million years ago had not happened, Earth might still be home to the dinosaurs so that mammals could not evolve as they have. Without these infrequent crashes new more intelligent species

may not be able to evolve, thus infrequent asteroid strikes provide an “optimal pump of evolution” conducive to intelligent life that other planetary systems may not have. Fossil records, such as those from 65 million years ago provide some evidence that the asteroid that destroyed the dinosaurs allowed humans to survive and evolve.

Finally, planetary systems may just be very dangerous places. Unintelligent life on other planets may have to deal with many threats both from space and on their own planet that do not allow sufficient time for intelligence to develop. These threats could be from meteor impacts or supervolcano eruptions. The corresponding mass extinction events may make it impossible for life to hold on very long. In fact, there are data from the Holocene epoch (the past 10,000 years to the present) that depict a mass extinction event; this one due to human activity. The rate at which species are becoming extinct is estimated to be 120,000 times the “normal” or “background rate.” Perhaps as life starts to become intelligent, it extinguishes itself much more quickly than the responses in the second category of responses. The rate at which human activity is wiping out other species supports the idea that we may indeed extinguish ourselves before we are able to communicate with other ETCs.

These three responses to the Fermi paradox vary in the types and levels of evidence available to support them. The first two responses (i.e., that they are here or that they are out there and just have not communicated) tend to be more difficult to support or refute as hypotheses, and would rely on relatively weak evidence. On the other hand, the last response (i.e., that we are alone) is based on hypotheses that tend to be more testable, and which would be based on stronger scientific evidence.

## Appendix C

## Demographics Measure

**Part A (university)****DIRECTIONS:** Please circle or fill in the appropriate responses.**Gender:**      Male                  Female**Age:** \_\_\_\_\_**Ethnicity (check all categories that apply):**

\_\_\_\_\_ Caucasian/Non-Hispanic White

\_\_\_\_\_ Hispanic

\_\_\_\_\_ African American/Black

\_\_\_\_\_ American Indian

\_\_\_\_\_ Asian/Pacific Islander

\_\_\_\_\_ Other (Please specify: \_\_\_\_\_)

**Major:** \_\_\_\_\_**Year in school:** \_\_\_\_\_**GPA:** \_\_\_\_\_**Are you a native English speaker?:**                  Yes                  No



## Appendix D

## Subject-matter Knowledge Measure

**DIRECTIONS:** Circle the letter of the response that *best* answers the question. Please answer all of the questions. If you do not know the answer, please make your best guess.

1. The four fundamental physical forces are gravity, strong nuclear force, weak nuclear force, and:
  - a. dark matter
  - b. electromagnetism
  - c. magnitude
  - d. centrifugal force
  
2. The two basic forms of energy are kinetic and:
  - a. potential
  - b. physical
  - c. electrical
  - d. eukaryotic
  
3. A chemical element number in the periodic table is determined by:
  - a. their abundance in the universe
  - b. the number of protons in the nucleus
  - c. the number of electrons in the nucleus
  - d. the order in which they were discovered
  
4. The inverse of wavelength is:
  - a. wave number
  - b. amplitude
  - c. absorption rate
  - d. wave height
  
5. Argon is a:
  - a. gas giant
  - b. natural gas
  - c. noble gas
  - d. transition element

6. Distance divided by time is referred to as:
  - a. acceleration
  - b. speed
  - c. space-time
  - d. force
  
7. The attraction objects have with a force proportional to the objects' masses and inversely proportional to the square root of the distance between their centers is referred to as:
  - a. kinetic energy
  - b. black hole
  - c. gravitational force
  - d. magnetic force
  
8. Electrons occupy spaces in an atom best described as:
  - a. circular orbits
  - b. uncertainty fields
  - c. wave functions
  - d. covalent bonds
  
9. The unit of magnetic field strength is a(n):
  - a. Tesla
  - b. Angstrom
  - c. Ohm
  - d. Newton
  
10. The lowest-energy photons of all electromagnetic radiation are:
  - a. radio waves
  - b. gamma rays
  - c. tidal waves
  - d. stun rays
  
11. Stars generate energy in their interior through the process of:
  - a. fusion
  - b. internal combustion
  - c. fission
  - d. solar flares
  
12. The specific icy and rocky area beyond 30 astronomical units (AUs) from the sun is known as the:

- a. Kuiper belt
  - b. Cirrus cloud
  - c. Oort cloud
  - d. Asteroid belt
13. The notion that Earth was seeded from life elsewhere in the cosmos is referred to as the:
- a. big bang theory
  - b. panspermia model
  - c. model of evolution
  - d. terrestrial model
14. By what means was the first planet revolving around a star like our sun detected?
- a. Hubble telescope
  - b. spectroscopy
  - c. radio waves
  - d. astrometry
15. The oxygen-rich atmosphere of the Earth is a direct result of the long-term oxidation of Earth's crust by:
- a. tectonic plates
  - b. anaerobic bacteria
  - c. increased methane
  - d. photosynthesis
16. Lifeforms on Earth are based on the element:
- a. carbon
  - b. oxygen
  - c. water
  - d. proteins

## Appendix E

## Beliefs About the Existence of ETCs

**DIRECTIONS:** Please check or make a slash for the appropriate response.

1. Which of the following most closely describes your beliefs about the existence of *intelligent* **Extra-Terrestrial Civilizations** (ETCs)? (check one)

ETCs exist *and* they are here or have visited Earth in the past.

ETCs exist *but* they have not yet communicated with us or visited Earth.

ETCs do not exist *and* we are alone in the universe.

The existence or nonexistence of ETCs is unknowable.

2. Indicate how confident you are in your response to question 1 by making a slash anywhere on the line (or ends) from *not confident* to *very confident*.


  
 not confident very confident

## Appendix F

## Passage Recall Measure

**DIRECTIONS:** Circle the letter of the response that *best* answers the question. Please answer all of the questions. If you do not know the answer, please make your best guess.

1.  $N = R \times f_p \times n_e \times f_l \times f_i \times f_c \times L$  is referred to as the:
  - a. Fermi paradox
  - b. Drake equation
  - c. Neugebauer equation
  - d. Newton's law of gravity
  
2. The areas where a third smaller body can maintain a fixed distance from two larger objects are called:
  - a. Gravitational singularities
  - b. Holocene epochs
  - c. Golden means
  - d. Lagrangian points
  
3. The idea that life originated elsewhere and was somehow transported to Earth is referred to as:
  - a. prime directive
  - b. panspermia
  - c. fertilization
  - d. interstellar travel
  
4. The Voyager probe took 21 years to reach:
  - a. Proxima Centauri
  - b. Saturn
  - c. Cato Neimoidia
  - d. Pluto
  
5. Scientists at the SETI institute listen for electromagnetic signals such as:
  - a. capillary waves
  - b. infrared waves
  - c. tachyon particles
  - d. subatomic particles

6. Planets outside our solar system are inferred to exist by their gravitational pull on:
  - a. black holes
  - b. radio waves
  - c. parent stars
  - d. beta-cell mass
7. The only types of planets astronomers have found outside our solar system are:
  - a. mesoplanets
  - b. exoskeletons
  - c. lagrangians
  - d. gas giants
8. Enrico Fermi worked at the:
  - a. Search for Extra Terrestrial Intelligence Institute
  - b. Max Planck Institute for Extraterrestrial Physics
  - c. Los Alamos National Laboratory
  - d. Department of Energy
9. The extinction of the dinosaurs 65 million years ago was likely caused by:
  - a. a meteor impact
  - b. a supervolcano
  - c. human presence
  - d. change in axial tilt
10. The largest mass extinction event occurred during the:
  - a. Paleozoic era
  - b. Baroque period
  - c. Holocene epoch
  - d. Fermi paradox

## Appendix G

### Instructions for the Verbal Protocol Analysis

In this investigation, we are interested in what you think and do while you read a text. What we want you to do is say what you are thinking and doing out loud. You can decide for yourself whether you would like to read the text silently or out loud, or do some of both. Do whatever feels most natural to you. We are only interested in what you are thinking and doing as you read. For example, if you are going back to reread, please say that's what you are doing. If something in the text reminds you of prior experiences or things you already know, please say that. If you are thinking that you don't understand something, please say that, too. There are no right or wrong things to say here, just whatever is going through your head as you read. If you are quiet for a period of time, I'll ask you to say what you're thinking. Do you have any questions?

## Appendix H

## Coding Scheme for the Verbal Protocol Analysis

*Surface-processing Strategies* (strategies aimed at the construction of the textbase)

S = use of surface-processing strategic behavior

S- = unsuccessful use of surface-processing strategic behavior

- reading aloud
  - S “Overpopulation destroys intelligent life before it becomes able to colonize intelligent planets.”
- rereading
  - S “Okay, rereading about the Fermi paradox.”
- skimming (skipping portions)
  - S “Now I’m skimming ahead.”
- guessing the meaning of a word in context
  - S “I’m guessing from this sentence that interstellar mean between stars.”
- predicting (about the microstructure or macrostructure)
  - S “Okay, here we go, this paragraph addresses why extraterrestrials would be observing Earth.”
- underlining or marking the text
  - S “And I’m underlining everything that seems important so it’s better retained.”
- using a text feature (such as a heading or table)
  - S “Um, looks like an equation.”
- rehearsing (repeating information to retain it in memory)
  - S “I’m going to say this a couple times in my head.”
- local restatement (paraphrasing or repeating text information at the microstructure level)
  - S “Okay, looks like we’re looking at different views of extraterrestrials.”
  - S- “I think this sentence is saying that the nearest star is the sun.”
- global restatement (paraphrasing or repeating text information at the macrostructure level)
  - S “Okay, so this is suggesting that there are extraterrestrials and they are trying to communicate to us but we either aren’t listening or we haven’t figured out how to listen.”
  - S- “So this is saying that there are four different views.”
- making connections to prior text
  - S “Oh, huh, they’re explaining it [the Drake equation].”
  - S- “I forgot what the Fermi paradox was already.”
- making connections to research task (a prior or subsequent measure)



- S “I think Panspermia was a response for a multiple choice question in the study.”
- evaluating text quality of the micro or macrostructure
  - S “Huh, they did not consistently put quotes around this phrase.”
  - S- “And in my opinion this information is really dense...”
- evaluation of the importance of the text
  - S “Um, I don’t know the point of this equation.”
- evaluation of task difficulty
  - S “This equation looks crazy.”
  - S- “There’s too many theories and I’m not going to be able to remember them all.”

*Deep-processing Strategies* (strategies aimed at integration into the situation model)

D = use of deep-processing strategic behavior

D- = unsuccessful use of deep-processing strategic behavior

- predicting (about the argument)
- questioning
  - D “Who knows if there is any unintelligent life on other planets.”
- arguing with text
  - D “Okay, just because numerous studies of genetics demonstrate that all life on Earth is highly related, some use as evidence that our DNA came from a single origin outside of Earth. Why couldn’t that single origin be within Earth? It seems plausible.”
  - D- “And I don’t understand how the Drake equation can exist if no one knows whether or not for sure whether there’s extraterrestrial life out in the universe.”
- global restatement (aimed at the argument)
  - D “So this is saying that it is more likely that we are alone in the galaxy.”
  - D- “It’s trying to say that none of these hypotheses are testable.”
- making connections to background knowledge
  - D “I remember hearing about the Doppler effect before.”
  - D- “I don’t know any of these, uh, symbols in this equation.”
- making connections to personal experience
  - D “I’ve never seen evidence of ETCs anywhere on this Earth.”
  - D- “When Pluto is mentioned it makes me feel like Pluto is still a planet.”
- interpreting (reasoning with information beyond the textbase, e.g., from the situation model, to continue integration from the textbase into the situation model)
  - D “And this is a recent discovery.”
  - D- “I feel the zoo scenario is the same thing we are doing, by looking for extraterrestrial communication.”

- elaborating (reasoning with information beyond the textbase, e.g., from the situation model, to build meaning tangential to the text)
  - D “The Earth is precariously placed as far as distance from the sun so it makes sense that would be really rare, extremely rare, that there would be another planet like Earth that could sustain life. At another star.”
  - D- “This reminds me of men in black.”
- evaluating agreement with the text
  - D “I also agree that there aren’t ETCs.”
- evaluating text quality about the argument
  - D “I wish these people would provide proof of where they’ve seen these extraterrestrial sightings.”
  - D- “This just seems like a silly article.”

*Regulatory strategies*

- evaluating comprehension
  - O “Not sure what they mean by the great silence.”
- moving on
  - O “Okay, moving on.”

*Other strategies or evaluations* (strategies the neither help construct the textbase or integrate text with the situation model)

- evaluation of interest (positive or negative)
  - O “Interesting the 98.4% of human DNA is similar to that of ape DNA.
- expression of empathy (sympathy or feelings felt or imputed to others)
  - O “I feel sorry for any aliens that try to come here.”
- expression of surprise
  - O “Wow, we share 50% of our DNA with bananas!”
- expression of amusement
  - O “We share half our genes with a banana. That’s funny.”
- No code (not enough information to categorize the utterance into a code category)
  - O “Turning the page.”
- adjusting reading rate (either speeding up or slowing down)
  - O “Okay, I’m going to go faster.”

## References

- Aiken, L.S., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, Ca: Sage.
- Al-Emadi, A. A. (2001). The relationships among achievement, goal-orientation, and study strategies. *Social Behavior and Personality*, 29, 823-832. doi: 10.2224/sbp.2001.29.8.823
- Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 213–250). Greenwich, CT: JAI Press.
- Alexander, P. A. (2004). A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 273-298). Mahwah, NJ: Lawrence Erlbaum Associates.
- Alexander, P. A., Dinsmore, D. L., Parkinson, M. M., & Winters, F. I. (2011). Self-regulated learning in academic domains. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance*. New York: Routledge.
- Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, 87, 559-575.

- Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404.
- Alexander, P. A., Murphy, P. K., & Kulikowich, J. M. (1998). What responses to domain-specific analogy problems reveal about emerging competence: A new perspective on an old acquaintance. *Journal of Educational Psychology, 90*, 397-406. doi: 10.1037/0022-0663.90.3.397
- Alexander, P. A., Murphy, P. K., Woods, B. S., Duhon, K. E., & Parker, D. (1997). College instruction and concomitant changes in students' knowledge, interest, and strategy use: A study of domain learning. *Contemporary Educational Psychology, 22*, 125-146. doi: 10.1006/ceps.1997.0927
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*, 315-343. doi: 10.2307/1170635
- Alexander, P. A., Schallert, D. L., & Reynolds, R. E. (2009). What is learning anyway? A topographical perspective considered. *Educational Psychologist, 44*, 176-192. doi: 10.1080/00461520903029006
- Alexander, P. A., Sperl, C. T., Buehl, M. M., Fives, H., & Chiu, S. (2004). Modeling domain learning: Profiles from the field of special education. *Journal of Educational Psychology, 96*, 545-557. 10.1037/0022-0663.96.3.545
- Ally, B. A., Gold, C. A., & Budson, A. E. (2009). An evaluation of recollection and familiarity in Alzheimer's disease and mild cognitive impairment using receiver

operating characteristics. *Brain and Cognition*, 69, 504-513. doi:  
10.1016/j.bandc.2008.11.003

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In: J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

APA (2010). *Publication manual of the American Psychological Association*. Washington, DC: American Psychological Association.

Arteche, A., Chamorro-Premuzic, T., Ackerman, P., & Furnham, A. (2009). Typical intellectual engagement as a byproduct of openness, learning approaches, and self-assessed intelligence. *Educational Psychology*, 29, 357-367. doi:  
10.1080/01443410902927833

Au, W. W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36, 258-267.

Au, W. W. (2008). Devising inequality: A Bernsteinian analysis of high-stakes testing and social reproduction in education. *British Journal of Sociology of Education*, 29, 639-651. doi: 10.1080/01425690802423312

Baeten, M., Dochy, F., & Struyven, K. (2008). Students' approaches to learning and assessment preferences in a portfolio-based learning environment. *Instructional Science*, 36, 359-374. 10.1007/s11251-008-9060-y

Bandalos, D. L., Finney, S. J., & Geske, J. A. (2003). A model of statistics performance based on achievement goal theory. *Journal of Educational Psychology*, 95, 604-616. doi: 10.1037/0022-0663.95.3.604

- Bates, M. (1975). The lady lives on blood. In A. Ternes (Ed.), *Ants, Indians, and little dinosaurs* (pp. 74-82). New York: Charles Scribner & Sons.
- Bentler, P. M. (2006). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software, Inc.
- Bentley, P., Driver, J., & Dolan, R. J. (2009). Modulation of fusiform cortex activity by cholinesterase inhibition predicts effects on subsequent memory. *Brain: A Journal of Neurology*, *132*, 2356-2371. doi: 10.1093/brain/awp176
- Biggs, J. B. (1978). Individual and group differences in study processes. *British Journal of Educational Psychology*, *48*(3), 266-279.
- Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, *63*(1), 3-19.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (structure of the observed learning outcome)*. New York: Academic Press.
- Block, R. A. (2009). Intent to remember briefly presented human faces and to her pictorial stimuli enhances recognition memory. *Memory & Cognition*, *37*, 667-678. doi: 10.3758/MC.37.5.667
- Blom, S., & Severiens, S. (2008). Engagement in self-regulated deep learning of successful immigrant and non-immigrant students in inner city schools. *European Journal of Psychology of Education*, *23*, 41-58. doi: 10.1007/BF03173139

- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Karthwohl D. R. (1956). *Taxonomy of educational objectives: the classification of educational goals; Handbook I: Cognitive Domain*. New York, Longmans.
- Boekaerts, M. (2002). Bringing about change in the classroom: Strengths and weaknesses of the self-regulated learning approach – EARLI Presidential Address, 2001. *Learning and Instruction*, 12, 589-604.
- Boldini, A., Russo, R., & Avons, S. E. (2004). One process is not enough! A speed-accuracy tradeoff study of recognition memory. *Psychonomic Bulletin & Review*, 11(2), 353-361.
- Box, G. E. P., & Draper, R. (1987). *Empirical model-building and response surfaces*. Oxford: John Wiley & Sons.
- Brown, M. B., Aoshima, M., Bolen, L. M., Chia, R., & Kohyama, T. (2007). Cross-cultural learning approaches in students from the USA, Japan and Taiwan. *School Psychology International*, 28, 592-604. doi: 10.1177/0143034307085660
- Buehl, M. M., Alexander, P. A., Murphy, P. K., & Sperl, C. T. (2001). Profiling persuasion: The role of beliefs, knowledge, and interest in the processing of persuasive texts that vary by argument structure. *Journal of Literacy Research*, 33, 269-301. doi: 10.1080/10862960109548112
- Burgess, M. C. R., & Weaver, G. E. (2003). Interest and attention in facial recognition. *Perceptual and Motor Skills*, 96, 467-480. doi: 10.2466/PMS.96.2.467-480
- Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Thousand Oaks, CA: Sage.

- Camp, C. J., Markley, R. P., & Kramer, J. J. (1983). Spontaneous use of mnemonics by elderly individuals. *Educational Gerontology, 9*, 57-71. doi: 10.1080/0380127830090106
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.
- Campbell, J., Smith, D., Boulton-Lewis, G., Brownlee, J., Burnett, P. C., Carrington, S., & Purdie, N. (2001). Students' perceptions of teaching and learning: The influence of students' approaches to learning and teachers' approaches to teaching. *Teachers and Teaching: Theory and Practice, 7*, 173-187. doi: 10.1080/13540600120054964
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Co.
- Cano, F. (2007). Approaches to learning and study orchestrations in high school students. *European Journal of Psychology of Education, 22*, 131-151. doi: 10.1007/BF03173518
- Carrell, P. L., & Connor, U. (1991). Reading and writing descriptive and persuasive texts. *The Modern Language Journal, 75*, 314-324. doi: 10.2307/328725
- Cassidy, S. (2007). Assessing 'inexperienced' students' ability to self-assess: Exploring links with learning style and academic personal control. *Assessment & Evaluation in Higher Education, 32*, 313-330. doi: 10.1080/02602930600896704
- Cassidy, S., & Eachus, P. (2000). Learning style, academic belief systems, self-report student proficiency and academic achievement in higher education. *Educational Psychology, 20*, 307-320. doi: 10.1080/713663740



- Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences, 44*, 1596-1603. doi: 10.1016/j.paid.2008.01.003
- Chamorro-Premuzic, T., Furnham, A., & Lewis, M. (2007). Personality and approaches to learning predict preference for different teaching methods. *Learning and Individual Differences, 17*, 241-250. doi: 10.1016/j.lindif.2006.12.001
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation for the behavioral sciences*. Hillsdale, NJ: Erlbaum and Associates.
- Coutinho, S. A., & Neuman, G. (2008). A model of metacognition, achievement goal orientation, learning style and self-efficacy. *Learning Environments Research, 11*, 131-151. DOI 10.1007/s10984-008-9042-7
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11*, 671-684. doi: 10.1016/S0022-5371(72)80001-X
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando: Holt, Rinehart and Winston.
- Cronbach, L. J. (1971). Validity. In R. L. Thorndike (Ed.), *Educational measurement (pp. 443-507)*. Washington DC: American Council on Education.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Oxford, England: Irvington.
- Dewey, J. (1910/1997). *How we think*. Mineola, NY: Dover Publications, Inc.

- Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20*, 391-409. doi: 10.1007/s10648-008-9083-6
- Dinsmore, D. L., Fox, E., & Parkinson, M. M. (2010, May). A deeper look at why readers succeed or fail. In D. McNamara (Chair), *Integration, depth, and complexity: Characterizing reader types through multidimensional profiling*. Symposium to be presented at the annual meeting of the American Educational Research Association, Denver.
- Dinsmore, D.L., Loughlin, S. M., & Parkinson, M. M. (2009, April). The effects of persuasive and informational text on metacognitive monitoring and control. In P. Alexander (Chair), *Meta what? Measuring Monitoring and Control*. Symposium conducted at the annual meeting of the American Educational Research Association, San Diego.
- Duffy, G. G., Roehler, L. R., & Herrmann, B. A. (1988). Modeling mental processes helps poor readers become strategic readers. *Reading Teacher, 41*, 762-767.
- Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91*, 549–563. doi: 10.1037/0022-0663.91.3.549
- Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 223-241). Cambridge: Cambridge University Press.

- Ferguson-Hessler, M. G., de Jong, T. (1990). Studying physics texts: Differences in study processes between good and poor performers. *Cognition and Instruction*, 7, 41-54. doi: 10.1207/s1532690xci0701\_2
- Ferm, C., & Johansen, G. (2008). Professors' and trainees' perceptions of educational quality as related to preconditions of deep learning in musikdidaktik. *British Journal of Music Education*, 25, 177-191. doi: 10.1017/S0265051708007912
- Finney, S. J. & DiStefano, C. (2006). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269-314). Greenwich CT: Information Age Publishing.
- Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in CALL. *Computer Assisted Language Learning*, 20, 409-442. doi: 10.1080/09588220701746013
- Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5, 105-112.
- Fox, E., Dinsmore, D. L., & Alexander, P. A. (in press). Reading competence, interest, and reading goals in three gifted young adolescent readers. [Special Issue on Motivation and giftedness] *High Ability Studies*.
- Fox, E., Dinsmore, D. L., & Alexander, P. A. (2007, April). Situational success at reading challenging texts: Exposing the fragile understanding of college students. In P. A. Alexander (Chair), *Fragile understanding: When good ideas go bad*. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago.

Fox, E., Dinsmore, D. L., Maggioni, L., & Alexander, P. A. (2008, March).

*Undergraduates' independent and scaffolded reading of course texts: Further evidence of fragile understanding.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Frederickson, N. (1990). Systems approaches in EP practice: A re-evaluation. In N. Jones & N. Frederickson (Eds.), *Refocusing educational psychology* (pp. 130-164). Oxford: Palmer Press.

Freed, D. B., Marshall, R. C., & Phillips, D. S. (1998). Comparison of semantically and phonemically based training procedures in an overlearned naming task. *Perceptual and Motor Skills*, 87(3), 795-800.

Furnham, A., Christopher, A., Garwood, J., & Martin, N. G. (2008). Ability, demography, learning style, and personality trait correlates of student preference for assessment method. *Educational Psychology*, 28, 15-27. doi: 10.1080/01443410701369138

Gadzella, B. M., Ginther, D. W., & Williamson, J. D. (1986). Differences in learning processes and academic achievement. *Perceptual and Motor Skills*, 62(1), 151-156.

Gadzella, B. M., & Masten, W.G. (1998). Critical thinking and learning processes for students in two major fields. *Journal of Instructional Psychology*, 25(4), 256-261.

Gadzella, B. M., Stephens, R., & Baloglu, M. (2002). Prediction of educational psychology course grades by age and learning style scores. *College Student Journal*, 36, 62-68.

- Good, T. L., Wiley, C. R. H., & Sabers, D. (2010). Accountability and educational reform: A critical analysis of four perspectives and considerations for enhancing reform efforts. *Educational Psychologist, 45*, 138-148. doi: 10.1080/00461521003720171
- Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology, 83*, 187-194. doi: 10.1037/0022-0663.83.2.187
- Gress, C. L. Z., Fior, M., Hadwin, A. F., & Winne, P. H. (2010). Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior, 26*, 806-814. doi: 10.1016/j.chb.2007.05.012
- H.R. Res. 1, 107<sup>th</sup> Cong., Cong. Rec. H9773-10052 (2001) (enacted).
- Hancock, G. R. (2006). Power analysis in covariance structure modeling. In: G.R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 69-115). Greenwich, Connecticut: Information Age Publishing.
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural Equation Modeling: Present and Future — A Festschrift in honor of Karl Jöreskog*. Lincolnwood, IL: Scientific Software International, Inc.
- Hancock, G. R., & Mueller, R. O. (2006). Introduction. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 1-12). Greenwich CT: Information Age Publishing.

- Hancock, T. E., Stock, W. A., Kulhavy, R. W., & Swindell, L. K. (1996). Gender and developmental differences in the academic study behaviors of elementary school children. *Journal of Experimental Education, 65*(1), 18-39.
- Heijne-Penninga, M., Kuks, J. B. M., Hofman, W. H. A., & Cohen-Schotanus, J. (2008). Influence of open- and closed-book tests on medical students' learning approaches. *Medical Education, 42*, 967-974. doi: 10.1111/j.1365-2923.2008.03125.x
- Henry, K. J., Van Lunen, B. L., Udermann, B., & Onate, J. A. (2009). Curricular satisfaction levels of national athletic trainers' association-accredited postprofessional athletic training graduates. *Journal of Athletic Training, 44*, 391-399. doi: 10.4085/1062-6050-44.4.391
- Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111-127. doi: 10.1207/s15326985ep4102\_4
- Hinojosa, J. A., Martín-Loeches, M., Muñoz, F., Casado, P., & Pozo, M. A. (2004). Electrophysiological evidence of automatic early semantic processing. *Brain and Language, 88*, 39-46. doi: 10.1016/S0093-934X(03)00158-5
- Hotchkiss, L. (1976). A technique for comparing path models between subgroups. *Sociological Methods & Research, 5*, 53-76.
- Hu, L-T, & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
- Jay, T., Caldwell-Harris, C., & King, K. (2008). Recalling taboo and nontaboo words. *American Journal of Psychology, 121*(1), 83-103.

- Jetton, T. L. (1994). Information-driven versus story-driven: What children remember when they are read informational stories. *Reading Psychology, 15*, 109-130. doi: 10.1080/0270271940150203
- Jones, B. W. (2004). *Life in the solar system and beyond*. New York: Springer.
- Jundt, D. K., & Hinsz, V. B. (2002). Influences of positive and negative affect on decisions involving judgmental biases. *Social Behavior and Personality, 30*, 45-52. doi: 10.2224/sbp.2002.30.1.45
- Justicia, F., Pichardo, M. C., Cano, F., Berbén, A. B. G., & De la Fuente, J. (2008). The Revised Two-Factor Study Process Questionnaire (R-SPQ-2F): Exploratory and confirmatory factor analyses at item level. *European Journal of Psychology of Education, 23*, 355-372. doi: 10.1007/BF03173004
- Kamalski, J. Sanders, T., & Lentz, L. (2008). Coherence marking, prior knowledge, and comprehension of informative and persuasive texts: Sorting things out. *Discourse Processes, 45*, 323-345. doi: 10.1080/01638530802145486
- Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology, 114*, 175-191. doi: 10.2307/1423513
- Kintsch, W. (1998). *Comprehension*. New York: Cambridge University Press.
- Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In: R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading (Fifth Edition; pp. 1270-1362)*. Newark, DE: International Reading Association.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

- Klein, A. M., Zwickel, J., Prinz, W., & Frith, U. (2009). Animated triangles: An eye tracking investigation. *The Quarterly Journal of Experimental Psychology*, *62*, 1189-1197. doi: 10.1080/17470210802384214
- Klinger, T. H. (2006). Learning approach, thinking style and critical inquiry: The online community. *Korean Journal of Thinking & Problem Solving*, *16*(1), 91-113.
- Kulikowich, J. M., & DeFranco, T. C. (2003). Philosophy's role in characterizing the nature of educational psychology and mathematics. *Educational Psychologist*, *38*, 147-156. doi: 10.1207/S15326985EP3803\_4
- Kwan, J. L. Y., & Chan W. (2011). Comparing standardized coefficients in structural equation modeling: A model reparameterization approach. *Behavior Research Methods*. doi: 10.3758/s13428-011-0088-6
- Lau, S., Liem, A. D., & Nie, Y. (2008). Task- and self-related pathways to deep learning: The mediating role of achievement goals, classroom attentiveness, and group participation. *British Journal of Educational Psychology*, *78*, 639-662. doi: 10.1348/000709907X270261
- Laxman, K. (2010). A conceptual framework mapping the application of information search strategies to well and ill-structured problem solving. *Computers & Education*, *55*, 513-526. doi: 10.1016/j.compedu.2010.02.014
- Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, *65*, 1-12.
- Lehman, S., & Schraw, G. (2002). Effects of coherence and relevance on shallow and deep text processing. *Journal of Educational Psychology*, *94*, 738-750. doi: 10.1037/0022-0663.94.4.738



- Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology, 44*, 87-112. doi: 10.1037/h0084237
- Lodewyk, K. R., Winne, P. H., & Jamieson-Noel, D. L. (2009). Implications of task structure on self-regulated learning and achievement. *Educational Psychology, 29*, 1-25. doi: 10.1080/01443410802447023
- Lunine, J. I. (2005). *Astrobiology: A multidisciplinary approach*. San Francisco: Pearson.
- Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and few solutions. *Applied Psychological Measurement, 13*, 335-361.
- Marsh, H. W., Wen, Z., Hau, K-T, Little, T. D., Bovaird, J. A., & Widamin, K. F. (2007). Unconstrained structural equation models of latent variables: Contrasting residual- and mean-centered approaches. *Structural Equation Modeling, 14*, 570-580.
- Marton, F., & Saljo, R. (1976a). On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology, 46*(1), 4-11.
- Marton, F. & Saljo, R. (1976b). On qualitative differences in learning: II. Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology, 46*(2), 115-127.
- McKelvie, S. J., & Pullara, M. (1988). Effects of hypnosis and level of processing on repeated recall of line drawings. *Journal of General Psychology, 115*(3), 315-329.
- Meier, S. T., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research, 8*, 107-120. doi: 10.1007/BF01173038

- Messick, S. (1981). Constructs and their vicissitudes in educational and psychological measurement. *Psychological Bulletin*, *89*, 575-588. doi: 10.1037/0033-2909.89.3.575
- Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist*, *42*(3), 173-190.
- Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology*, *34*, 306-318.
- Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, *25*, 3-53. doi: 10.1006/ceps.1999.1019
- Murphy, P. K., & Alexander, P. A. (2002). What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance. *Journal of Experimental Education*, *70*, 197-214. doi: 10.1080/00220970209599506
- Murphy, P. K., Long, J. F., Holleran, T. A., & Esterly, E. (2003). Persuasion online or on paper: A new take on an old issue. *Learning and Instruction*, *13*, 511-532. doi: 10.1016/S0959-4752(02)00041-5
- Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning & Verbal Behavior*, *16*, 151-171. doi: 10.1016/S0022-5371(77)80044-3
- N'Kaoua, B., Véron, A.-L., H., Lespinet, V. C., Claverie, B., & Sztark, F. (2002). Time course of cognitive recovery after propofol anaesthesia: A level of processing approach. *Journal of Clinical and Experimental Neuropsychology*, *24*, 713-719. doi: 10.1076/jcen.24.6.713.8401

- Nolen, S. B. (1988). Reasons for studying: Motivational orientations and study strategies. *Cognition and Instruction, 5*, 269-287. doi: 10.1207/s1532690xci0504\_2
- Nolen, S. B., & Haladyna, T. M. (1990). Personal and environmental influences on students' beliefs about effective study strategies. *Contemporary Educational Psychology, 15*, 116-130. doi: 10.1016/0361-476X(90)90011-O
- O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes, 43*, 121-152.
- Osorio, A., Ballesteros, S., Fay, S., & Pouthas, V. (2009). The effect of age on word-stem cued recall: A behavioral and electrophysiological study. *Brain Research, 1289*, 56-68. doi: 10.1016/j.brainres.2009.07.013
- Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19*, 228-242.
- Pearce, G., & Lee, G. (2009). Viva voce (oral examination) as an assessment method: Insights from marketing students. *Journal of Marketing Education, 31*, 120-130. doi: 10.1177/0273475309334050
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. New York: Holt, Rinehart, and Winston.
- Phan, H. P. (2008). Unifying different theories of learning: Theoretical framework and empirical evidence. *Educational Psychology, 28*, 325-340. doi: 10.1080/01443410701591392

- Phan, H. P. (2009). Amalgamation of future time orientation, epistemological beliefs, achievement goals and study strategies: Empirical evidence established. *British Journal of Educational Psychology*, 79, 155-173. doi: 10.1348/000709908X306864
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego: Academic Press.
- Popper, K. R. (1994). *The myth of the framework: In defence of science and rationality*. London: Routledge.
- Pressley, M. (1998). Comprehension strategies instruction. In: J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 113-127). NY: The Guilford Press.
- Pressley, M. & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. New York: Holt, Rinehart, and Winston.
- Rodriguez, C. M. (2009). The impact of academic self-concept, expectations and the choice of learning strategy on academic achievement: The case of business students. *Higher Education Research & Development*, 28, 523-539. doi: 10.1080/07294360903146841
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2009). Responses to issues raised about validity, bias, and fairness in high-stakes testing. *American Psychologist*, 64, 285-287. doi: 10.1037/a0015473

- Sanchez, E., Garcia-Rodicio, H., & Acuna, S. R. (2009). Are instructional explanations more effective in the context of an impasse? *Instructional Science, 37*, 537-563. doi: 10.1007/s11251-008-9074-5
- Schiefele, U. (1999). Interesting and learning from text. *Scientific Studies of Reading, 3*, 257-279.
- Schommer, M., & Surber, J. R. (1986). Comprehension-monitoring failure in skilled adult readers. *Journal of Educational Psychology, 78*, 353-357. doi: 10.1037/0022-0663.78.5.353
- Schraw, G. (2010). No school left behind. *Educational Psychologist, 45*, 71-75. doi: 10.1080/00461521003720189
- Schunk, D. H. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review, 20*, 463-467. doi: 10.1007/s10648-008-9086-3
- Segalowitz, N., Cohen, P., Chan, A., & Prieur, T. (2001). Musical recall memory: Contributions of elaboration and depth of processing. *Psychology of Music, 29*, 139-148. doi: 10.1177/0305735601292004
- Senko, C., & Miles, K. M. (2008). Pursuing their own learning agenda: How mastery-oriented students jeopardize their class performance. *Contemporary Educational Psychology, 33*, 561-583. doi: 10.1016/j.cedpsych.2007.12.001
- She, H.-C., & Chen, Y.-Z. (2009). The impact of multimedia effect on science learning: Evidence from eye movements. *Computers & Education, 53*, 1297-1307. doi: 10.1016/j.compedu.2009.06.012

- Shostak, G. S., & Barnett, A. (2003). *Cosmic company: The search for life in the universe*. Cambridge: Cambridge University Press.
- Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 271-295). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sullivan, Y. M., & Cantwell, R. H. (1999). The planning behaviours of musicians engaging traditional and non-traditional scores. *Psychology of Music*, 27, 245-266.
- Thomas, C. R., & Gadbois, S. A. (2007). Academic self-handicapping: The role of self-clarity and students' learning strategies. *British Journal of Educational Psychology*, 77, 101-119. doi: 10.1348/000709905X79644
- Thorndike, E. L., & Gates, A. I. (1929). *Elementary principles of education*. New York: The Macmillan Company.
- Tian, X. (2007). Do assessment methods matter? A sensitivity test. *Assessment & Evaluation in Higher Education*, 32, 387-401. doi: 10.1080/02602930600898460
- Toppino, T. C., Fearnow-Kenney, M. D., Kiepert, M. H., & Teremula, A. C. (2009). The spacing effect in intentional and incidental free recall by children and adults: Limits on the automaticity hypothesis. *Memory & Cognition*, 37, 316-325. doi: 10.3758/MC.37.3.316
- van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of memory representation. In: H. van Oostendorp, & S. R. Goldman (Eds.), *The construction of mental*

- representations during reading (pp. 71-98)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Van Houtte, M. (2005). Climate or culture? A plea for conceptual clarity in school effectiveness research. *School Effectiveness and School Improvement, 16*, 71-89. doi: 10.1080/09243450500113977
- Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology, 87*, 246-260. doi: 10.1037/0022-3514.87.2.246
- Webb, S. (2002). *If the universe is teeming with aliens...Where is everybody? Fifty solutions to the Fermi paradox and the problem of extraterrestrial life*. New York: Copernicus Books.
- William, D. (2010). Standardized testing and school accountability. *Educational Psychologist, 45*, 107-122. doi: 10.1080/00461521003703060
- Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology, 101*, 1-20. doi: 10.1037/a0013152
- Yamauchi, H., & Miki, K. (2000). Effects of students' perception of teachers' and parents' attitudes on their achievement goals and learning strategies. *Psychologia: An International Journal of Psychology in the Orient, 43*(3), 188-198.
- You, Z., & Jia, F. (2008). Do they learn differently? An investigation of the pre-service teachers from US and China. *Teaching and Teacher Education, 24*(4), 836-845.