AFFECTIVE AND COGNITIVE COMPONENTS OF STATISTICS COURSE PERFORMANCE

BY

KENNETH C. BESSANT

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY IN

EDUCATION

Department of Curriculum: Mathematics and Natural Sciences
Faculty of Education
University of Manitoba
Winnipeg, Manitoba

0-612-51629-6

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

AFFECTIVE AND COGNITIVE COMPONENTS OF STATISTICS COURSE PERFORMANCE

BY

KENNETH C. BESSANT

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree

of

DOCTOR OF PHILOSOPHY

ABSTRACT

The primary goal of this thesis was to investigate the individual and combined effects of affective and cognitive–learning variables on university students' final grades in introductory-level statistics. The study was *exploratory* in nature, and it addressed the following general questions:

1. What is the nature of the relationship between statistics course performance and statistics anxiety and attitudes toward statistics?
2. To what extent do differential learning strategies (e.g., surface-instrumental versus deep-conceptual) and metacognitive problem-solving procedures impact final grade outcome?
3. Are reported levels of cognitive interference, distractibility, or cognitive concern (i.e., worry) related to achievement scores?
4. In what ways do measures of statistics anxiety, attitudes toward statistics, learning strategies, metacognitive problem-solving procedures, and cognitive interference interact with statistics course performance?

A multi-modal research design was devised to investigate these objectives, which combined both quantitative and qualitative methods of data collection and analysis. A two-stage sample comprising 435 Brandon University students provided questionnaire data for the study: (1) an existing data set (n = 358) based on "in-class" surveys administered between 1994 and 1996 and (2) a new sample (n = 77) collected in 1999. In addition, semi-structured interviews were conducted with nine students in order to gather first-hand accounts of learning experiences and affective responses to course demands.

The initial aggregate of cases was randomly partitioned into two subgroups for the purpose of internal (double) cross-validation. Beginning with the data in Subgroup 1, multiple regression and principal components analysis were used to identify a total of 15 predictors: seven affective and eight cognitive–learning

variables. The complete set of regressors accounts for approximately 40% of the criterion variance in either subgroup (i.e., R^2s of .3989 and .4245). The squared cross-validity coefficients obtained when weights derived in Subgroups 1 and 2 were applied to the data in opposing partitions (i.e., .3108 and .2948) affirm the stability of the predictive equations. Similar results were found when the two sets of regression coefficients were used separately to predict statistics achievement in the new sample (Subgroup 3): .3481 and .2839.

Commonality analysis revealed that 6 of the 15 predictors make significant ($ps < .05$) *unique* contributions to the explanation of final course performance: expected final grade (EXPGRADE), attentional focus (FOCUSATT), metacognitive problem solving (METAPROB), cognitive interference (INTRFERE), surface-disintegrated learning (SURFMIPS), and procedural learning (PROCMIPS). Only one of these six regressors, EXPGRADE, is aligned with the affective domain, but it proved to be the most powerful predictor of students' final grades either alone or in conjunction with other affective or cognitive factors. Furthermore, five of the eight cognitive–learning factors used to predict final grades significantly increment the R^2.

The findings of this study suggest the need for (1) more refined methods of investigating the nature of statistics affect (e.g., anxieties and attitudes), (2) greater research attention to metacognitive, cognitive-attentional, and learning-strategy variables, and (3) closer examination of the interactive role of affective and cognitive factors in statistics learning, problem solving, and achievement.

ACKNOWLEDGMENTS

The pursuit of an advanced degree is a career goal that demands not only personal sacrifice but the support and cooperation of advisors, family, and friends. I would like to extend my sincere appreciation to all of the members of my advisory committee and to my good friend and colleague, Dr. Erasmus D. Monu (Brandon University). I consider myself privileged to have had the opportunity to study with individuals who have committed their academic lives to the advancement of mathematics teaching and learning. I was particularly honoured to have Dr. Eric R. Muller (Brock University) serve as my external examiner. I dedicate this thesis to:

Dr. Eric D. MacPherson, who knew I could do it,

Dr. Lars C. Jansson, who guided me through it,

Dr. Jeffrey G. Williams, who worked with me to start it,

Dr. Erasmus D. Monu, who insisted that I try it, and

Ms. Josephine Scaletta Bessant, my wife, who inspired me to finish it.

Finally, I would like to acknowledge the support of the Brandon University Research Committee for a portion of the data used in this thesis.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

LIST OF APPENDIX TABLES

CHAPTER I

INTRODUCTION

The Emergence of Statistics Education

It has only been within the last century that statistics has become recognized as a separate field or discipline, but its beginnings date back to the early centuries of the first millennium A. D. (e.g., censuses conducted by Augustus in the Roman Empire). A historical survey of the emergence of statistical and probabilistic reasoning reveals a diverse collection of contributors from mathematics and from the empirical and social sciences. The meaning of statistics has evolved along with expanding conceptions and applications of quantitative technology in areas such as astronomy, heredity, and demography. Hald (1990) traces the word "statistics" to sixteenth century Italian origins, that is, the collection of information and facts of interest to a statesman (*statista*) or pertaining to the state (*stato*). Indications of a more modern use of statistics can be found in John Graunt's (1662) descriptive analysis of plague mortality rates entitled *Natural and Political Observations made upon the Bills of Mortality* (as cited in Hald, 1990).

Stigler (1986) has remarked that "modern statistics" arose out of the interplay between mathematical concepts and the needs of various applied sciences; "it is a logic and methodology for the measurement of uncertainty and for an examination of the consequences of that uncertainty in the planning and interpretation of experimentation and observation" (p. 1). This point is illustrated by the rise of positivistic sociology in the nineteenth century, and the pursuit of so-called "laws" of human nature. John Stuart Mill (1872) discussed the notion of a "science" (albeit inexact) of human actions, which he contended was able to furnish only "approximate generalizations" or "the lowest kind of empirical laws" (p. 434). As such, this science could not provide a basis for predicting, with "scientific accuracy," an individual's thoughts, feelings, or actions. However, Mill (1872) states that "The very events which in their own nature appear most capricious and uncertain, and which in any individual case no attainable degree of knowledge would enable us to foresee, occur, when considerable numbers are taken into the account, with a degree of regularity approaching to mathematical" (pp. 532-533).

On a related matter, Lambert Adolphe Jacques Quetelet proposed what he termed "social physics." Stigler (1986) observes that "Although the works of Bernoulli and Laplace foreshadowed the application of probability to the measurement of uncertainty in the social sciences, the works of Quetelet represent the first steps toward making this wish a practical reality" (p. 161). Notable among Quetelet's advances in statistical analysis is his notion of "the average man" (*l'homme moyen*), which was intended as an analogue to the physical concept of the centre of gravity; it represented a device for eliminating random variations in order to reveal societal regularities. Quetelet (1835/1969) proffered "the fundamental principle, that *the greater the number of individuals observed, the more do individual peculiarities, whether physical or moral, become effaced, and leave in a prominent point of view the general facts, by virtue of which society exists and is preserved*" (p. 6).

As the underpinnings of statistical and probabilistic methods have become more firmly established, increased attention has been devoted to cross-disciplinary applications. Expanded academic, research, and professional interests in statistics have in turn contributed to their inclusion not only in university- and college-level courses but also in school mathematics. Curricular attention to statistics and probability topics, across all grade levels, is rooted in suggestions proffered by committees and commissions on school mathematics established in the early to mid-twentieth century (e.g., *The Reorganization of Mathematics in Secondary Education*, 1923). Shulte and Smart (1981) have remarked that "All major curriculum groups in this century—including the NCTM in its recommendations for the curriculum of the 1980s—have stressed the importance of statistics and probability" (p. ix). Indeed, considerable discussion has been focused on rationalizing and justifying the inclusion of statistics and probability topics in school mathematics. Recently, the National Council of Teachers of Mathematics (NCTM, 1989) and Manitoba Education and Training (1995, 1996, 1997) curriculum guidelines have devoted *standards* or *strands* explicitly to the coverage of statistics and probability content at all grade levels.

Increased interest in statistics education stems partially from the widening exposure to statistical methods and statistics curriculum content across a diverse audience of academics, professionals, and students (Feinberg & Halperin, 1978). Undergraduate degrees in the social and behavioral sciences, as well as in pre-professional programs, now routinely require the completion of one or more statistics courses. Statistics educators, in various disciplines, recognize the stress that such requirements can evoke in students who have tended to avoid mathematics-related courses because of limited prerequisite skills, poor prior performance, low self-confidence, or negative attitudes (Blalock, 1987). Such experiences have raised researchers' awareness of and curiosity over the involvement of not only cognitive (e.g., information-processing) skills but also affective factors (e.g., attitudinal dispositions and anxiety responsiveness) in learning statistics. As a consequence, there has been a recent expansion in statistics education research, and some of it parallels investigative trends in mathematics teaching, learning, and assessment.

Background of the Study

Research on statistics anxiety has to some extent grown out of the exploration of affective and cognitive dimensions of mathematics learning and problem solving (e.g., McLeod, 1989a, 1989b). For example, early studies concerning the nature of mathematics attitudes have influenced current research on attitudes toward statistics and performance outcomes. Based on a meta-analysis of 151 studies, Hembree (1990) asserts that "Positive attitudes toward mathematics consistently related to lower mathematics anxiety, with strong inverse relations observed for an enjoyment of mathematics and self-confidence in the subject" (p. 38).[1] Furthermore, theoretical and empirical analyses of linkages among test anxiety, cognitive interference (Sarason, 1984), and ineffective study skills (Kirkland & Hollandsworth, 1980) suggest the potential involvement of

---

[1] Hembree (1990) reports *mean* correlations of – .37, – .47, and – .65 between mathematics anxiety and three "attitudinal constructs": usefulness of, enjoyment of, and self-confidence in mathematics, respectively. The latter two coefficients are based on post-secondary data.

cognitive-attentional factors in statistics achievement. It seems plausible that the investigation of selected topics in statistics education may benefit from research on affective and cognitive components of mathematics education.

Much has been written about the impact of cognitive and affective factors on mathematics learning and performance outcomes. However, researchers have tended to investigate these two broad classes of variables or influences separately, that is, without attention to their individual and joint (interactive) effects on achievement scores. Despite several decades of study, terminology within both the affective and cognitive domains is fraught with classificatory problems, most notably with regard to mathematics affect (e.g., anxiety, attitudes, and emotions). Some three decades ago, May (1969) characterized the affective objectives of the day as "notoriously fuzzy and difficult to interpret" (p. 35), and he did not expect these matters to improve with the publication of Krathwohl, Bloom, and Masia's (1964) *Affective Domain Taxonomy*.

Several mathematics education researchers have suggested that the lack of consistency in the definitions and uses of various affective concepts stems from an inadequate theoretical base (Hart, 1989; Kulm, 1980; Mandler 1989a; McLeod, 1988). Mandler (1989b) remarks that "absent some theoretical agreement, a search for consensus may well be futile" (p. 237). Because of conceptual confusion and imprecise or often overlapping definitions, a portion of the present study will endeavor to clarify and to refine operational definitions of the central variables (see Chapter IV, Analysis of the Questionnaire Data), followed by the exploration of statistically significant relationships in the data.

Statement of Problem and Research Questions

Research that analyzes the involvement of affective and cognitive factors in mathematics learning and problem-solving processes is at best *suggestive* concerning key research questions or hypotheses in the field of statistics education. These matters are further obscured by the relative dearth of literature bearing on linkages between statistics course performance and measures of statistics anxiety, attitudes toward statistics, and preferred or habituated learning practices. Harvey, Plake, and Wise (1985) have noted the importance of examining "a variety of both cognitive and affective variables" because "neither component should be isolated when attempting to predict performance or when developing remediation programs" (p. 4). Therefore, the primary research goal here is to investigate the *individual* and *combined* effects of affective and cognitive variables on the final grade performance of university-level students enrolled in introductory statistics courses. This research is *exploratory* in nature, and it examines the following general questions:

1. What is the nature of the relationship between statistics course performance and statistics anxiety and attitudes toward statistics?
2. To what extent do differential learning strategies (i.e., surface-instrumental versus deep-conceptual) and metacognitive problem-solving procedures impact final grade outcome?
3. Is there empirical support for the suggested theoretical linkage between cognitive interference, anxiety, and performance?

Limitations and Potential Contributions

One of the main sources of data in the present study involves students' responses to a series of self-report instruments. Forced-choice scales cannot fully represent (or detect) nuances in students' affective responses to learning and evaluative situations in statistics courses (e.g., intensity or magnitude of statistics anxiety). It is for this and other related methodological reasons that this research incorporates a qualitative dimension through the collection and analysis of semi-structured interviews. The study sample is drawn from a subject pool of students enrolled in introductory-level statistics courses at Brandon University. Although the overall design is longitudinal in nature, the results are not necessarily generalizable to the experiences of students of other post-secondary institutions.

Notwithstanding these cautionary remarks, this thesis represents an effort to explore the involvement and the interaction of cognitive and affective variables in statistics course performance. Treatment of the study data incorporates aspects of both internal (double) cross-validation and external replication. Prior to the widespread acceptance of qualitative methods, some educational researchers rationalized such studies as building a foundation for quantitative-empirical designs. In a somewhat parallel but reverse sense, it is hoped that the research questions and findings associated with this thesis will prompt more in-depth (qualitative) investigations of the complex linkages between cognitive and affective processes, particularly with regard to their effects on statistics learning outcomes. On a more practical note, such research is aimed at improving our understanding of how statistics learning strategies are associated with affective responses, which may in turn facilitate improved instruction and curriculum design.

CHAPTER II

THEORETICAL CONSIDERATIONS

Statistics Anxiety

Research on what is termed "statistics anxiety" is relatively recent (e.g., Zeidner, 1991). Prior to the consideration of statistics anxiety as a distinct topic of investigation, researchers surveyed students enrolled in undergraduate and graduate statistics courses with instruments designed to measure "mathematics anxiety." As a result, some analyses of students' affective responses while learning statistics content mirror aspects of mathematics anxiety research.

Interest in mathematics anxiety can be traced to Gough's (1954) early discussion of "mathemaphobia" and Dreger and Aiken's (1957) research on "number anxiety." Also, MacPherson (1966) used Cattell's IPAT Anxiety Scale (Institute for Personality and Ability Testing, 1957) to investigate correlates of anxiety in learning programmed mathematics. More than four decades have passed since Dreger and Aiken (1957) expressed the concern that "almost no controlled research has been attempted in the realm of emotional problems associated with arithmetic and mathematics" (p. 344).

In the late 1970s and throughout the 1980s, a considerable body of literature developed on what is commonly referred to as mathematics anxiety (e.g., Brush, 1978; Rounds & Hendel, 1980). These studies frequently analyzed the factorial structure of the Mathematics Anxiety Rating Scale (MARS; Suinn, Edie, Nicoletti, & Spinelli, 1972) and correlated extracted dimensions of the MARS with a range of mathematics affect and performance variables. Hembree (1990) reports inverse (mean) correlations between mathematics anxiety and college-level performance variables such as course grades and aptitude/achievement measures of computation, problem solving, and abstract reasoning (rs = – .27, – .25, – .27, and – .40, respectively).[2]

Of late, there has been some research on statistics anxiety, a term which shares several conceptual attributes with general anxiety and mathematics anxiety (i.e., worry, emotionality, nervousness, tension, and physiological arousal). It is, therefore, not surprising to find much of the substance and the language of Richardson and Suinn's (1972) early and often-cited definition of mathematics anxiety reflected in Zeidner's (1991) conception of statistics anxiety. Zeidner (1991) remarks that statistics anxiety "is commonly claimed to debilitate performance in a wide variety of academic situations by interfering with the manipulation of statistics data and solution of statistics problems" (p. 319).

Results of empirical studies, although mixed, generally suggest an inverse relationship between mathematics-related anxiety and statistics performance (e.g., Harvey et al., 1985). Zeidner (1991) finds a negative correlation between levels of statistics anxiety[3] and final course grades (i.e., $r = -.13$, $p < .05$). Lester and Hand (1989) similarly observe that students' examination scores in a psychological statistics course are negatively and significantly related to "anxiety about mathematics evaluation," as measured by a modified version of the MARS

---

[2] Hembree (1990) reports *mean* correlations for differing sample sizes and levels of education. The coefficients for computation and problem solving were based on studies of school-age (i.e., grades 7 and grades 9 to 12) *and* post-secondary students, whereas the correlations pertaining to course grades and abstract reasoning were restricted to the latter subject pool.

[3] Zeidner (1991) operationalizes statistics anxiety with a set of items adapted from the MARS; this inventory is discussed in the *Study Questionnaire: Instruments* section of Chapter III.

(i.e., $r = -.47, p < .001$). Recently, Birenbaum and Eylath (1994) have reported a non-significant correlation of $-.11$ between course grades and statistics anxiety, based on a single self-rating scale ranging from 1, "No anxiety at all" to 10, "A very high level of anxiety" (p. 94). Zimmer and Fuller (1996) suggest that discrepant empirical findings may be due to differences in operational measures of statistics anxiety, that is, single items versus multi-item instruments.

It is important to note that mathematics anxiety and, by implication, statistics anxiety, lack conceptual and operational clarity. Research on mathematics-related anxiety has been (is) confronted by a range of theoretical, conceptual, and measurement issues. In lieu of explicit conceptual definitions, some researchers have tended to focus attention on operational measures of mathematics attitudes and anxieties. Hart (1989) contends that some instruments include items dealing with both anxiety and attitudes toward mathematics, a matter of some concern to Kulm (1980), who cautions against "combining in meaningless ways [affective] characteristics that ought to be considered separately" (p. 365). Hart (1989) argues that the absence of clear and concise meanings for many affective constructs has impaired valuable research on the role of affect in mathematics learning.

Attitudes Toward Statistics

Some time ago, Rokeach (1968) noted confusion over the term 'attitude,' but he recommended "continued critical analysis with the aim of giving it a more precise conceptual and operational meaning" (p. 111). Research on what are broadly termed "attitudes toward mathematics" is fairly recent. The potential impact of affective factors on mathematics learning is of particular concern to current researchers, educators, and curriculum constructors. The National Assessment of Educational Progress, in conjunction with the NCTM, has published a series of monographs since 1978 that report information on school mathematics performance and students' affective responses to mathematics (e.g., Kenney & Silver, 1997). The investigation of attitudes toward mathematics has gone beyond descriptive profiles; some researchers have inquired into the dynamic of attitude–performance interaction. On this point, McLeod (1992) notes that several major evaluation studies (both national and international) have reported positive correlations between mathematics attitudes and achievement at various grade levels.

Although a variety of measurement techniques have been devised to explore attitudes toward mathematics, self-report paper-and-pencil instruments are common in the literature (e.g., Thurstone- and Likert-type scales, inventories, check lists, semantic differential scales, and projective techniques). Aiken (1974) and Fennema and Sherman (1986) have constructed a series of multi-item scales that focus on attitudinal dimensions such as mathematics enjoyment, valuation, and self-confidence. Researchers interested in examining the impact of (self-rated) attitudinal dispositions on statistics learning must therefore consider whether they will use existing measures of mathematics affect or construct new (content-specific) instruments. Feinberg and Halperin (1978), for instance, reported a statistically significant correlation ($r = .35$) between statistics performance and a 14-item scale dealing with *attitudes toward quantitative concepts*. Although such findings suggest that measures of attitudes toward mathematics may have some explanatory value in specialized quantitative situations, a number of researchers have perceived a need to develop instruments for use in statistics education.

Bendig and Hughes (1954) conducted one of the earliest studies concerning the relationship between students' "emotional attitude" and achievement in an introductory (psychological) statistics course. The authors solicited students' views and attitudes about taking a statistics course and, from these, constructed a Likert-format Statistics Course Attitude Scale (50 items). More recently, Roberts and Bilderback (1980) developed the Statistics Attitude Survey (SAS) "that was designed to be more relevant than other affective measures in the prediction of performance in statistics" (p. 236). The SAS is composed of 33 items couched in "statistical jargon," which cover themes such as beliefs about statistics, affective responsiveness, problem-solving competence, and perceived usefulness of statistics.

Wise (1985) devised the Attitudes Toward Statistics (ATS) scale in response to several perceived limitations of the SAS. His primary criticism of the SAS was that many of the items, perhaps one-third, deal with students' success in solving problems and understanding concepts, that is, matters of achievement rather than attitudes. Wise designed the ATS to measure students' attitudes toward (1) the statistics course in which they are currently enrolled and (2) the usefulness of statistics in their chosen field of study. Of these two dimensions of the ATS, the attitude toward course subscale has been found to be a stronger predictor of statistics achievement (Elmore, Lewis, & Bay, 1993; Harvey et al., 1985).

Notwithstanding Wise's (1985) remarks, Roberts and Reese (1987) have suggested that "the SAS and ATS are measuring the same characteristic" (p. 763). The SAS and the ATS follow the general format and phrasing of mathematics attitude scales, but constituent items describe learning experiences that the authors deem pertinent to the circumstances of statistics education.

Negative attitudes toward statistics have been linked to fear of failure (or evaluation), low self-confidence, and anxiety. Feinberg and Halperin (1978) report significant associations between mathematics background, attitudes toward quantitative concepts, perceived mathematical ability, and achievement in introductory statistics. Benson's (1989) findings indicate that the (indirect) effects of prior mathematics courses on statistics test anxiety are mediated by mathematics self-concept and achievement. Furthermore, students' misconceptions about mathematical skill requirements and their perceptions of the usefulness of statistics are thought to reinforce negative attitudes that interfere with instruction and learning (Gal & Ginsburg, 1994).

Mandler's (1989a, 1989b) constructivist views on emotions and mathematical problem solving provide a theoretical basis for interpreting the interaction of affective and cognitive processes in statistics learning. He suggested that values about mathematics, in conjunction with other situational factors, influence the intensity of emotions. Mandler (1989a) states that "The well-known phenomenon of *number shock* demonstrates an underlying value that mathematical manipulation is difficult, complex, and potentially frightening" (p. 7). It follows that belief systems and attitudes toward statistics (e.g., usefulness) may predispose individuals to respond differently to statistics learning and instruction.

Procedural versus Conceptual Learning of Statistics Content

A number of authors have referred to a theoretical distinction between *procedural* and *conceptual* learning (e.g., Greeno, 1980; Wertheimer, 1959). Notwithstanding variation in terminology, there is some consensus that students can learn, understand, or be taught mathematics concepts and procedures in quite dissimilar fashions. Skemp (1978, 1986) has distinguished *relational* and *instrumental* understanding, that is, knowing what to do and why, as contrasted with simply applying a rule (i.e., "rules without reasons"). Skemp (1978) suggests that these two modes of teaching and learning are so entrenched that "*there are two effectively different subjects being taught under the same name, 'mathematics'*" (p. 11).

Similar theoretical discussions have arisen in the field of educational psychology regarding contrasting approaches to learning: "one described as deep, meaning-oriented, transformational, or internalising; the other as surface-oriented, reproducing, or memorising" (Speth & Brown, 1988, p. 247). Biggs (1985) has outlined a model of student learning in which three learning approaches (i.e., *deep*, *surface*, and *achieving*) mediate the impacts of personal (e.g., previous experiences, cognitive styles, and abilities) and situational factors (e.g., subject matter, mode of instruction, and task demands) on performance. According to Biggs, a "deep" approach is characterized by intrinsic interest in the subject matter, integrative-associative learning strategies, and a focus on maximizing understanding, whereas a "surface" orientation is associated with extrinsic motivation, rote memorization, and the investment of minimal effort to meet task requirements. Biggs (1993) further suggests that an "achieving" approach emphasizes ego-enhancement and the pursuit of high grades through organization, time management, and efficient allocation of energy.

Characteristic differences in the type or mode of learning, of the sort proposed by mathematics education researchers and educational psychologists, are relevant to research on statistics course performance. Variation in statistics learning strategies and motives may be linked to attentional focus (i.e., susceptibility to cognitive interference) and to levels of processing in memory. Wachtel (1967) notes that "Concepts of breadth of attention have been prominent in research on cognitive styles, much of which can be regarded as the study of attention deployment" (p. 417). Craik and Lockhart (1972) have proposed a framework in which retention is linked to the depth and type of encoding, that is, "trace persistence is a function of depth of analysis, with deeper levels of analysis associated with more elaborate, longer lasting, and stronger traces" (p. 675). These authors discuss levels of processing in terms of a continuum ranging from transient, sensory analyses (shallow) to more elaborate, semantic-associative (deep) operations. Variability in the depth of processing used to encode and retrieve information is pivotal to learning.

Empirical studies have been conducted to explore the effects of students' preferred or habituated modes of learning on proneness to mathematics-related anxiety. Reece and Todd (1989) have reported a moderate inverse relationship between a formal-deductive style of thinking and anxiety over working with numbers. These findings are consistent with other researchers' contentions that a surface approach is linked to anxiety (Marton & Säljö, 1984) and negative affect (Biggs, 1985). On a related topic, Hudak and Anderson (1990) find that success in introductory statistics is positively related to level of formal operational ability, whereas "the concrete style of learning, marked by lacking the use of theory and inference, is particularly maladaptive in statistics and computer science courses" (p. 233).

Metacognitive Problem Solving

A number of researchers have deemed metacognitive knowledge and skills germane to mathematics learning and problem solving (Lester, Garofalo, & Kroll, 1989; Schoenfeld, 1987; Silver, 1987). Garofalo and Lester (1985) have noted the important role of executive management processes in mathematical problem solving: using cognitive strategies to improve one's understanding of a task or problem, formulating a plan, selecting appropriate strategies, monitoring the execution of strategies and activities, evaluating outcomes, and revising unproductive plans and procedures. Garofalo and Lester (1985) suggest that it is the *self-monitoring* function that distinguishes metacognitive from cognitive activities; "cognition is involved in doing, whereas metacognition is involved in choosing and planning what to do and monitoring what is being done" (p. 164). Briars (1983) agrees that process-oriented aspects of metacognition are highly relevant to strategy selection and problem solution.

Further, McLeod (1989a) has remarked that affective variables can have a substantial effect on the cognitive and metacognitive processes involved in mathematical problem solving. Mandler (1989a) states that "emotional experiences are frequently not conducive to the full utilization of the cognitive apparatus; thought may become simplified (i.e., stereotyped and canalized) and tend to revert to simpler modes of problem solving" (p. 9). The relationship between affective intensity and metacognition may be influenced (perhaps mediated) by the beliefs that students bring to the learning environment, that is, beliefs about self, mathematics, and problem-solving processes. Schoenfeld (1987) has argued convincingly that beliefs and metacognition can have significant effects on mathematical behavior. Mandler (1989a) has pointed to the role of students' beliefs in explaining the inhibiting effects of affective arousal on metacognitive experiences. According to Lester, Garofalo, and Kroll (1989), "Beliefs often interact with and, at times, shape attitudes and emotions, and beliefs influence the decisions made during problem solving" (p. 77).

Metacognitive problem solving implies greater control over the direction of attention, systematic exploration of contextually relevant information, goal-directed behavior, planful action, and ongoing evaluation. Further, metacognitive awareness and skills have been linked to deep-level learning motives, strategies, and intentionality, as well as more positive (less-anxious) affective dispositions toward learning (Biggs, 1985). Metacognitive functioning is highly pertinent to the analysis of statistics learning and performance insofar as it is theoretically related to depth of information processing, learning strategies, and affective responsiveness while solving problems.

Cognitive Interference, Attentional Focus, and Cue Utilization

Some researchers (e.g., Sarason, 1984; Tobias, 1985; Wine, 1980) have discussed the cognitive-attentional components of test anxiety. In this approach, individuals are presumed to have access to a limited cognitive capacity for carrying out tasks at any given time. Heightened anxiety is thought to impair performance by increasing negative self-focusing tendencies and diverting attention away from task-completion behavior. Sarason (1984) has termed this process *cognitive interference*, that is, "intrusive thoughts that keep the individual from directing full attention to the task at hand" (p. 932).

Easterbrook's (1959) arousal-cue-utilization hypothesis is similar to Sarason's (1984) notion of cognitive interference insofar as it predicts that increased anxiety level can impact cognition by restricting attentional focus to a narrower or smaller set of central cues. Attention is devoted to essential cues, while peripheral stimuli are progressively eliminated from consideration. Easterbrook (1959) has proposed "that, when the direction of behavior is constant, increase in drive is associated with a reduction in the range of cue use" (p. 183). Wachtel (1967) uses the analogy of a beam of light to conceptually distinguish types of broad and narrow attention. Attentional *scanning* is represented by the extent of the beam's movement around the perceptual field, whereas attentional *focus* is characterized as the overall width of the beam. Metaphorically speaking, the attention of the very anxious person can be likened to "a narrow beam which roams all over the field" (Wachtel, 1967, p. 421). On a related matter, Korchin (1964) has suggested that, under conditions of extreme stress, attention may become diffuse or poorly controlled "as part of the general breakdown of organized behavior" (p. 71).

Cognitive-attentional, cue-utilization, and interference models of test anxiety assume that emotional arousal mediates the effects of cognition on performance. Increased anxiety is thought to excite task-irrelevant cognitions that impair concentration, cognitive efficiency, information retrieval, and performance. However, a recursive effect is also possible; diminished cognitive functioning or cognitive confusion may generate (or accentuate) emotional reactivity under conditions of evaluative stress. In other words, if affective responsiveness to learning tasks or evaluation can generate cognitive interference, it is also plausible that certain modes of cognitive processing may make individuals more or less prone to anxiety.

An Information-Processing Model of Cognition, Anxiety, and Performance

Several concepts associated with information-processing models of cognition (i.e., encoding, storage, organization, and retrieval) have been used to analyze test anxiety (e.g., Mueller, 1980). Information-processing models of test anxiety commonly suggest that increased anxiety impairs cognitive functioning through its effects on attentional structure, cue utilization, and memory processes. Tobias (1979) has outlined an information-processing model of "the effects of anxiety on learning from instruction" (p. 575). This model purports that anxiety affects learning indirectly through its impact on cognitive processes, particularly "preprocessing," "processing," or "postprocessing."

Anxiety is thought to divert attention to task-irrelevant cognition, thereby interfering with (1) the initial presentation of material, (2) the cognitive operations students use to store, organize, and manipulate information, and (3) the retrieval of previously-learned content. Tobias (1985) has proposed "a limited cognitive processing capacity formulation" (p. 138) of the effects of anxiety and study skills on academic performance. Low test-anxious students who employ competent study methods are considered well-equipped to deal with learning demands, whereas high anxiety and weak skills are thought to reduce available cognitive capacity.

Benjamin, McKeachie, Lin, and Holinger (1981) have examined an information-processing model of how test anxiety affects performance. They find support for both the "encoding deficit hypothesis" (weak study skills) and the "retrieval deficit hypothesis" (interference). Poor performance is attributed to retrieval problems in the testing situation and to cognitive processes involved in learning new information. The results of a study by Naveh-Benjamin, McKeachie, and Lin (1987) are consistent with an information-processing framework insofar as high-anxious students exhibited weaker cognitive organization of course materials and concepts. Lower achievement levels of high test-anxious students have been explained in terms of deficiencies in the encoding and organization of information, as well as retrieval problems in evaluative settings (Naveh-Benjamin, 1991).

Summary of Salient Relationships Among Study Variables

To the extent that this study is exploratory in nature, the foregoing review of literature suggests relationships among variables that warrant particular analytical attention. The substantive issues raised in the discussion of relevant research have been used to formulate six questions that expand on the more general goals identified in Chapter I. However, this statement of research objectives does not preclude the investigation of other relationships or combinations of variables implied in the larger framework of analysis depicted in Figure 1.

1. Is statistics course performance related to self-reported attitudes toward statistics, as defined in terms of students' perceptions of their current statistics course or the usefulness of statistics?
2. Is the level of statistics anxiety, that is, students' self-reported responsiveness to learning and evaluative processes, related to final grade outcome?
3. Are deep-conceptual versus surface-procedural learning strategies (or learning preferences) differentially related to achievement scores?
4. Is the level of self-reported use of metacognitive problem-solving strategies associated with statistics course performance?
5. Are reported levels of cognitive interference, distractibility, or cognitive concern (i.e., worry) related to final grades?
6. In what ways do measures of statistics anxiety, attitudes toward statistics, learning strategies, metacognitive problem-solving procedures, and cognitive interference interact with statistics course performance?

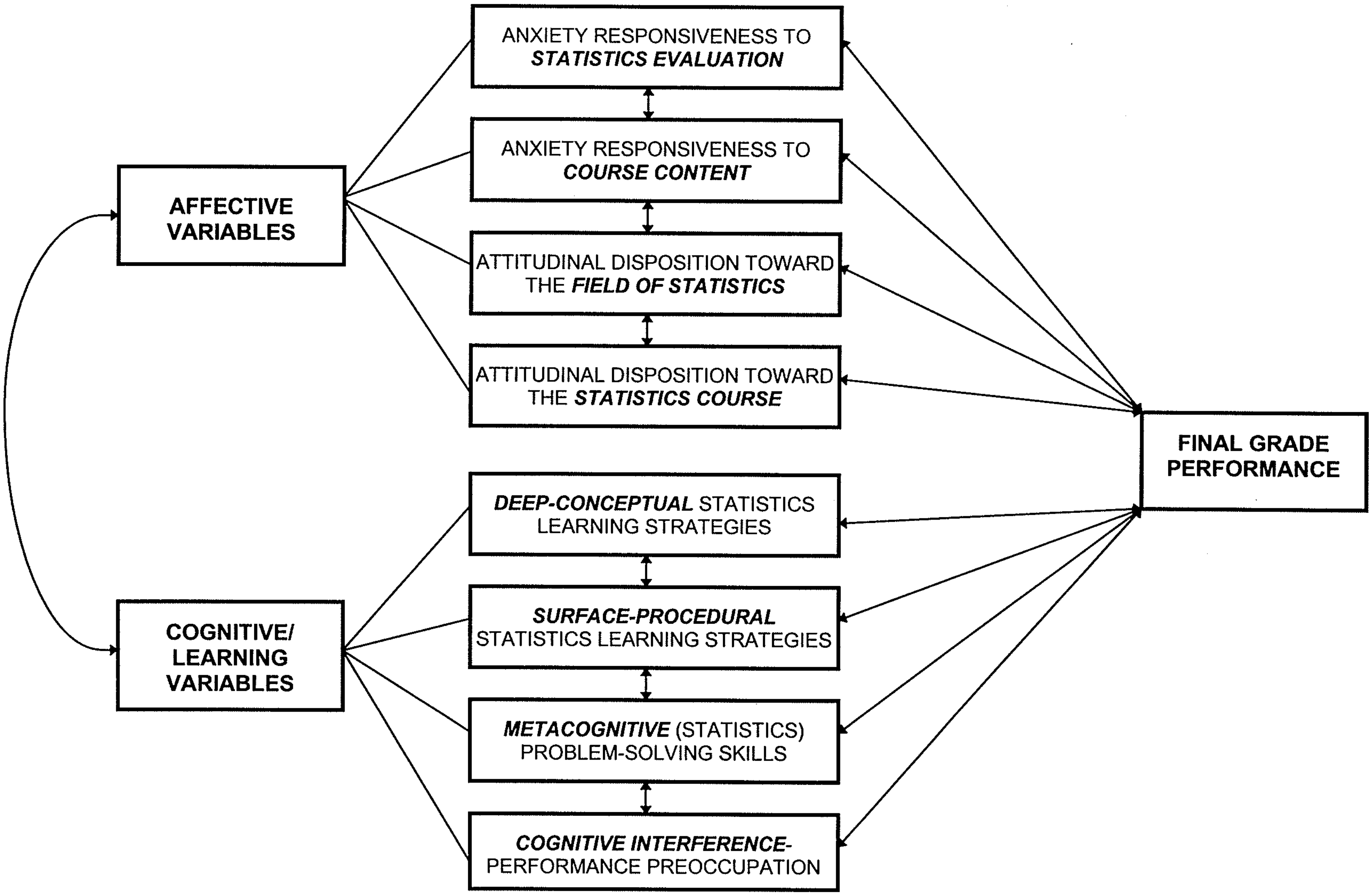


Figure 1. A Pictorial Summary of Relationships Among Affective, Cognitive, and Performance Variables.

CHAPTER III

METHOD

Sample and Procedures

This study is based on a combination of questionnaire and personal interview data collected in three interconnected processes: (1) an existing data set (n = 358) consisting of "in-class" survey responses of Brandon University students, (2) a subsequent administration of a refined set of the original study instruments, and (3) supplementary interviews with consenting students.

Phase one of data collection. The initial data set was collected between 1994 and 1996. At the time these study questionnaires were administered, the respondents were enrolled in one of three introductory-level statistics courses offered by the departments of mathematics, psychology, or sociology. The inclusion of the three "departmental" samples permitted access to students from diverse mathematics backgrounds, faculties, and degree programs. The three courses share similar curricula, for example, measures of central tendency and dispersion, probability, sampling distributions, confidence intervals, and hypothesis testing.

All of the instructors who cooperated with phase one of this study provided class time in which to administer the instruments. After an explanation of ethical considerations and item response categories, students were asked to complete four self-rating scales in the following order: the Statistics Anxiety Scale (SAS), the Mathematics Information Processing Scale (MIPS), the Attitudes Toward

Statistics (ATS) scale, and the Study Process Questionnaire (SPQ).[4] Most students required approximately 50 minutes to respond to the items and to provide basic demographic information. In all instances, these data were collected midway through the fall semester of each respective academic year. This procedure ensured that the respondents were exposed to statistics content, instruction, and evaluation processes prior to participating in the study.

Phase two of data collection. Following a review of the original questionnaire, a revised set (or booklet) of instruments was administered in the second phase of data collection. Chapter IV, Analysis of the Questionnaire Data, includes a detailed discussion of the statistical procedures used to analyze the existing data set in order to identify and to group selected items for the analysis of phase-two data. The students who agreed to complete the follow-up questionnaire were also invited to participate in a semi-structured interview concerning their experiences while learning and taking tests in introductory statistics courses. The primary goal of the interview component of this study was to ask students supplementary questions concerning their affective responses to statistics course content and evaluative processes. It was anticipated that the information extracted from the interviews would facilitate the interpretation of questionnaire data, focus attention on factors not included in the study instruments, and provide direction for future research. It should be noted that the interview data were *not* included in formal statistical analyses or tests of relationships.

---

[4] Although Biggs' (1985) SPQ is included in the study questionnaire, the accompanying data are not analyzed in this study. However, theoretical and conceptual discussions of the SPQ (e.g., surface versus deep approaches to learning) bear some relevance to *conceptual* versus *procedural* learning and, as such, are incorporated into the analysis of the MIPS.

The second phase of data collection proceeded as follows. As a preliminary step, two statistics instructors at Brandon University were contacted regarding their potential (and voluntary) involvement in the project. At this time, the instructors were informed of the general nature of the research, the administration of the questionnaires, and the estimated amount of class time needed to complete the instruments. Both instructors agreed to participate and arrangements were then made to visit their classrooms. After a brief introduction, I proceeded to describe the study and to address matters of informed consent, confidentiality of information, and voluntary participation. Students were asked to read the cover letter prior to answering any questions, in order to familiarize them with research goals, ethical considerations, and related instructions.

The final page of the booklet included a two-part consent form that requested students' (1) informed (and signed) permission to access their final statistics course grades (i.e., the dependent variable in this study) and (2) voluntary participation in a brief interview concerning their experiences while learning and taking tests in the current statistics course. The questionnaire consent form, in conjunction with the cover letter, assured students that they were under no obligation to release grades or to participate in the interview, but if they agreed, all collected survey responses would be held in the strictest confidence. The consent form also explained the conditions under which grade information would be requested and stored (i.e., coded for entry) in a computer data base so as to safeguard the confidentiality of the data.

The second part of the consent form asked students who had completed the instruments if they would be willing to participate in a short, approximately 30-minute interview. As noted earlier, the central aim of the interviews was to further investigate the nature of students' affective responses (e.g., attitudes, anxieties, and self-confidence) while learning content and taking tests in statistics courses. Also, the interview component was included in the overall research design to complement the analysis of the questionnaire data. Prior to conducting the interviews, however, students were asked to read an "Interview Consent Form" describing the general nature of the study and the interview process. Interviews were conducted with a total of nine students, and the results are summarized and discussed in Chapter V, Analysis of the Interview Data.

The interviews were semi-structured in nature, which principally involved informal discussions of students' affective and learning experiences. Structured instruments were incorporated into the interview process on a discretionary basis to explore students' affective responsiveness (i.e., direction and intensity) and self-confidence in statistics courses, for example, the Positive and Negative Affect Schedule[5] (PANAS; Watson, Clark, & Tellegen, 1988) and the Fennema-Sherman Mathematics Attitudes Scales (Fennema & Sherman, 1986). The administration of the PANAS typically involves asking respondents to rate a series of "mood descriptors" in terms of the *extent* to which or the *frequency* with which they experience the following affective states:

---

[5] Copyright 1988 by the American Psychological Association. Reprinted (or adapted) by permission.

Enthusiastic	Scared
Interested	Afraid
Determined	Upset
Excited	Distressed
Inspired	Jittery
Alert	Nervous
Active	Ashamed
Strong	Guilty
Proud	Irritable
Attentive	Hostile

In constructing the PANAS, Watson et al. (1988) identified only "high-pole" adjective markers representing relatively extreme levels of either positive or negative emotional intensity. Mossholder, Kemery, Harris, Armenakis, and McGrath (1994) recommend the inclusion of "low-pole" measures of affectivity such as "at ease," "calm," "relaxed," "drowsy," or "sluggish." Therefore, the PANAS was used, in conjunction with a number of low-pole adjectives, as a basis for interviewing students concerning their affective responsiveness to statistics learning tasks, most notably, preparing for and taking tests. Over the course of the interviews, students were asked to indicate the extent to which they experienced each of the various affective states using a 4-point scale: 1 (*definitely do not feel*), 2 (*do not feel*), 3 (*slightly feel*), and 4 (*definitely feel*).[6] Asking students to consider high- and low-pole descriptors of negative and positive affectivity offers an alternative method of gauging their affective responsiveness (i.e., direction and intensity) to statistics learning and testing procedures.

---

[6] See Russell (1979) for a comparative discussion of 4- versus 5-point rating scales in research on bipolar affective space.

The interviews also addressed students' perceptions of self-confidence while learning and taking tests in statistics courses. These discussions focused on, but were not restricted to, dimensions of the Fennema-Sherman Mathematics Attitudes Scales (Fennema & Sherman, 1986), for example, the Confidence in Learning Mathematics Scale and the Effectance Motivation in Mathematics Scale (see Appendix A). Students were encouraged to comment broadly about matters of self-confidence as opposed to simply rating scale items.

Study Questionnaire: Instruments

Statistics Anxiety Scale.[7] A modified version of the MARS,[8] entitled the Statistics Anxiety Scale (SAS), was used to measure level of statistics anxiety. Forty-two of the original 98 MARS items dealing with mathematics learning, instruction, and evaluation processes were reworded to reflect *statistics* content. In responding to the SAS, students consider statements describing a variety of course-related learning and evaluative situations, each of which they rate in terms of how much they would be made anxious on a 5-point scale ranging from 1 (*not at all*) to 5 (*very much*). Items defining learning or instructional contexts include "Watching a teacher work out a statistics problem...," "Walking into a statistics class," and "Starting a new chapter...." The evaluation items describe situations such as "Taking a final examination...," "Being given a 'pop' quiz...," "Thinking about an upcoming statistics test...," and "Waiting to have a statistics test returned."

---

[7] Permission for this scale from R. M. Suinn. Any reproduction is prohibited without written permission from: R. M. Suinn, Department of Psychology, Colorado State University, Ft. Collins, CO 80523.

[8] Richardson and Suinn (1972) report an internal reliability coefficient alpha of .97 for the MARS ($n = 397$) and a test-retest (7 weeks later) Pearson product-moment coefficient of .85 ($n = 35$).

The SAS is patterned after instruments developed by Richardson and Woolfolk (1980) and Zeidner (1991). Richardson and Woolfolk (1980) selected 40 MARS items involving academic contexts or applications of mathematics that exhibited highest item-to-total correlations. They presumed that the 40-item MARS was "as reliable, stable, and valid as the original AMRS [*sic*]" (Richardson & Woolfolk, 1980, p. 274). Zeidner (1991) constructed the Statistics Anxiety Inventory (SAI) by revising the items included in Richardson and Woolfolk's shortened version of the MARS so as to depict experiences associated with learning statistics. Zeidner (1991) has identified two empirical dimensions in the SAI that reflect *statistics content anxiety* and *statistics test anxiety*, with corresponding internal reliability estimates of .94 and .92.

Attitudes Toward Statistics scale.[9] An instrument developed by Wise (1985) was employed to assess students' attitudes toward their introductory statistics course(s). The Attitudes Toward Statistics (ATS) scale is composed of 29 items that can be separated into two subscales: *Attitude Toward Field of Statistics* (20 items) and *Attitude Toward Course* (9 items). The *Field* subscale includes statements such as "I feel that statistics will be useful to me in my profession" and "Statistics is a worthwhile part of my professional training," whereas the *Course* subscale contains items like "I get upset at the thought of enrolling in another statistics course" and "Dealing with numbers makes me uneasy." Students rate each item on a Likert-format scale of 1 (*strongly disagree*) to 5 (*strongly agree*).

---

[9] Permission from S. L. Wise, Department of Educational Psychology, University of Nebraska, Lincoln, NE 68588-0345.

The direction of the weighting procedure generates scales on which higher scores reflect more positive attitudes toward statistics. Wise (1985) reports alpha coefficients of .92 and .90 for the *Field* and *Course* dimensions of the ATS, followed by two-week, re-test reliability coefficients of .82 and .91, respectively. Other writers have confirmed the reliability (Perney & Ravid, 1990) and factorial validity (Woehlke, 1991) of the ATS.

Differential statistics learning strategies. Differences in students' learning and study strategies while dealing with statistics course content were gauged with an instrument developed by the author entitled The Mathematics Information Processing Scale[10] (MIPS; Bessant, 1997). The 87 statements that comprise the MIPS (see Appendix I) can be separated into three distinct sets of items pertaining to students': (1) deep-conceptual (or relational) versus surface-procedural (or instrumental) learning of statistics content (Items 1 to 52), (2) metacognitive problem-solving strategies (Items 53a to 53i), and (3) attentional deployment during tests and examinations (Items 54a to 54z). Respondents use a 5-point scale of 1 (*not at all typical*) to 5 (*very typical*) to rate statements dealing with their experiences while learning, being instructed, and taking tests in statistics courses.

The initial 52 items of the MIPS are aimed at identifying fundamental differences in the ways that students study, learn, and understand statistics content, for example, "I prepare for tests by looking for associations and relationships between ideas" (relational or deep), "I am unsure what test questions mean or what they are asking me to do" (superficial or surface), and "I do not care

[10] Copyright 1997 by Sage Publications, Inc. Reprinted by permission of Sage Publications.

if I know what a statistic is used for, as long as I can perform the calculations" (procedural or instrumental). This group of 52 items describes a wide range of learning tasks such as listening to lectures, reading textbooks, completing assignments, developing problem-solving procedures, and preparing for tests and examinations. The wording of these statements was designed to reflect distinctions between deep-conceptual versus surface-procedural learning of statistics content.

Metacognitive problem-solving procedures. A set of nine MIPS items probes the extent to which students use (metacognitive) problem-solving procedures such as establishing a general framework to interpret the question, identifying a major goal, determining relevant information, developing a plan of action, as well as selecting and evaluating solution strategies. Again, students are asked to indicate, on a 5-point scale, the extent to which they make use of the stated problem-solving strategies. The content of these statements is based on Garofalo and Lester's (1985) discussion of metacognition and mathematical problem-solving processes.

Cognitive-attentional focus. The MIPS contains 26 items intended to gauge the nature of students' attentional deployment, range of cue use, cognitive interference, distractibility, and performance preoccupation in evaluative settings. A number of scale elements are concerned with the breadth or narrowness of attention (in testing situations), that is, the consideration of central and peripheral problem cues, attentional focus and scanning, and cognitive (dis)organization. Statements dealing with cognitive interference include external distractions (e.g., noise) and lack of focus due to heightened arousal. Several items are patterned after Liebert and Morris' (1967) concept of *cognitive concern*. Exemplars include "I spend a lot of time waiting for the answer to come to me" (cognitive disorganization), "I have difficulty determining what information in the question or problem is crucial..." (cue deafness), "I restrict my analysis to what I think is the most important information" (attentional focus), and "I am preoccupied with what others will think of my performance" (cognitive concern).

Procedures of Quantitative Data Analysis

The reproducibility of results is an important but often-neglected aspect of social and behavioral science research (Shaver, 1993). This issue is made salient by the realization that *all* classical parametric methods are correlational (Knapp, 1978) and, as such, capitalize on sampling error. On a related matter, there has been extensive debate over the relative merits of tests of statistical significance, measures of effect size, and estimates of replicability for evaluating result importance.[11] Ang (1998) remarks that "Researchers have all too often ignored the replicability of results because they overly rely on significance testing" (p. 1143). Replication is a widely recognized method of investigating the robustness of reported relationships and of addressing the threat to generalizability posed by sampling bias (Carver, 1993).

In extending this discussion to multiple regression, a distinction should be made between *explanatory* and *predictive* research. Used in the first sense,

---

[11] For a discussion of key arguments, refer to the special issue of the *Journal of Experimental Education* entitled "Statistical Significance Testing in Contemporary Practice: Some Proposed Alternatives with Comments from Journal Editors" (Thompson, 1993).

regression analysis estimates the strength of relationship between a set of independent variables and a dependent variable. The optimal linear composite of variables is determined so as to maximize the multiple correlation in the population (Huberty & Mourad, 1980). Sample regression weights are treated as approximations of their population counterparts, but the R^2 will almost always be overestimated with regard to both the population (i.e., the coefficient of determination, ρ^2) and a new sample (i.e., the squared coefficient of cross-validation, ρ_c^2).[12]

In predictive regression, ρ_c^2 is of interest insofar as "the purpose is to estimate the effectiveness of a specific prediction equation when the equation is applied to data other than those in which the equation was derived" (Kromrey & Hines, 1995, p. 902). The selection of regressor variables to be included in the final equation is determined empirically so as to optimize the prediction of the criterion. As noted above, when a set of weights derived in one sample is used to predict the criterion scores in another sample, the resulting R^2 is typically smaller than that obtained in the data used to calculate the regression coefficients. Pedhazur (1982, 1997) refers to this phenomenon as *shrinkage*, which results from optimizing the weights to fit the idiosyncrasies of the sample under analysis and treating zero-order correlations as "error-free." Because of capitalization on chance, the R is biased upwards and the regression equation is *un*likely to fit any other data as well as the original sample (Mosteller & Tukey, 1977).

---

[12] According to Kromrey and Hines (1995), $\rho_c^2 \leq \rho^2$.

Pedhazur (1982) asserts that cross-validation is perhaps "the best method for estimating the degree of shrinkage" (p. 149). As an alternative to carrying out "external" replication studies, some researchers and statisticians have discussed the merits of "data splitting" (e.g., Picard & Berk, 1990), that is, partitioning available data into two subgroups for cross-validatory purposes. In this manner, "the future can be constructed by reserving part of the present, available data" (Picard & Cook, 1984, p. 576). Ang (1998) refers to this procedure as an "internal" method of replication. Although the cross-validation procedure[13] has been used for some time now to assess the validity of regression equations (e.g., Kurtz, 1948; Mosier, 1951), Barnard (1974) remarks that it "may perhaps be said to be one of the most seriously neglected ideas in statistics" (p. 133).

In *simple* cross-validation, a predictive equation is based on the so-called *derivation*[14] sample, for which the R^2 and the regression coefficients are calculated. The constituent weights are then applied to the predictor variables of a second (i.e., *validation*) sample to yield a Y' for each subject. The correlation ($r_{yy'}$), calculated in the validation sample, between *observed* values of the criterion variable (Y_i) and the *predicted* scores (Y_i') generated with weights from the derivation sample, is termed the cross-validity coefficient. In *double* cross-validation, the regression coefficients obtained in *each* of two samples (or *each* of two partitions of a larger set of data) are carried over to the *other* sample and two

[13] Other empirical computer-intensive approaches to predictive assessment include multi-cross-validation, jackknife, and bootstrap methods.

[14] The terminology used to differentiate between the sample in which regression weights are first developed versus that in which weights are later evaluated for predictive validity is variable and sometimes contradictory. For example, the expression "calibration" sample has been used in both senses.

separate $r_{yy'}$s are calculated. If the difference (i.e., shrinkage) between the two corresponding sets of squared multiple correlations (R^2s) and squared cross-validity coefficients ($R_{CV}{}^2$s) is small, the researcher may choose to base future predictions on the larger, and presumably more stable, combined sample (Pedhazur, 1997). In either simple or double cross-validation procedures, coefficients of cross-validation can be used to evaluate the predictive validity of regression equations.

This study incorporates several cross-validation procedures of data analysis:

1. The existing data set (n = 358), collected between 1994 and 1996, was first randomly partitioned into two pools: the derivation and the validation samples (alternately termed Subgroups 1 and 2).
2. Regression analysis was then used in the derivation sample to explore, choose, and organize variables so as to maximize the R^2. Stepwise and backward elimination methods of regression analysis were used, in conjunction with principal components analysis, to select predictor variables from larger sets of scale items, to be retained and combined with other similarly identified variables.
3. Weights from the regression equation obtained in the derivation sample were then applied to X_i scores in the validation sample to predict statistics course performance (Y_i'). The resulting Y' values were then correlated with observed criterion scores (Y) to yield a coefficient of cross-validation ($r_{yy'}$).
4. Based on the internal double cross-validation procedure, and using a uniform set of regressor variables, multiple Rs and regression coefficients were calculated in each of the two pools of data specified in Step 1. The application of these two regression equations to the predictor scores in the corresponding (validation) samples yielded separate coefficients of cross-validation for each partition. This method permitted comparisons of the actual R^2s and $R_{CV}{}^2$s between the two subgroups.
5. A new (external) sample was collected in 1999 (n = 77) to further evaluate the regression equations obtained in the two partitions of existing data. This involved independently carrying the weights generated in Subgroups 1 and 2 over to this third pool of data (Subgroup 3) for the purposes of calculating two additional coefficients of cross-validation.

In cross-validation designs of the type outlined above, it is acceptable to use a series of predictor-selection procedures (e.g., multiple regression and factor analysis) to analyze the data in the derivation sample (Subgroup 1). Two key objectives in this process are parsimony with regard to the number of predictors and maximization of the multiple R^2. Once the selection of regressors has been finalized, the goal is to evaluate the predictive power of the derived weights in other data, that is, in the validation sample (Subgroup 2), in the context of data-splitting, or in new data (Subgroup 3), in the case of external replication.

Data Reduction

It should again be mentioned that the data under examination here take the form of student responses to three instruments: the Statistics Anxiety Scale (SAS), the Attitudes Toward Statistics (ATS) scale, and the Mathematics Information Processing Scale (MIPS). At the outset of this study, it was noted that there are a number of conceptual and measurement issues bearing on the analysis of these or other similarly collected data, most notably within the affective domain. Persistent definitional ambiguity concerning anxiety and attitude constructs suggests the need for a careful examination of the SAS and the ATS scale.

Previous research indicates the presence of two or more empirical dimensions within each of these two "affective" instruments, that is, the SAS and the ATS (refer to discussion in Chapter II). For this reason, principal components analysis was used as a data-reduction tool to assist the exploration and identification of clusters of salient scale items to be included in the calculation of regression coefficients. As part of the process of optimizing predictive equations, the items in all three scales were analyzed in terms of item-to-criterion correlations, factorial structure (both within and across scale boundaries), content validity, item-to-total(scale) correlations, and coefficients of internal consistency (i.e., Cronbach's alpha).

CHAPTER IV

ANALYSIS OF THE QUESTIONNAIRE DATA

This chapter discusses the results of predictor-selection procedures, as well as internal and external cross-validation. Chapter III, Method, provided a brief overview of the major cross-validatory features included in this investigation of statistics course performance. However, Figures 2 and 3 below present more detailed representations of the constituent steps and methods of data analysis. The directional and connective lines in these two figures show linkages among various aspects of the data-analysis process, while the numbering of the boxes (which is synchronized between Figures 2 and 3) indicates the order or sequence of analysis. The numbers corresponding to the elements depicted in the two flow charts are incorporated into the textual discussions of this chapter as parenthetical notations (e.g., *Figure 2, No. 1*).

Figure 2 outlines the general process of predictor selection in phase-one data ($n = 358$), for example, the partitioning of these data for the purposes of cross-validation. As is indicated in Figure 2, multiple regression and factor analysis were used to select a smaller (i.e., more parsimonious) set of scale items from the original instruments, to serve as predictors of statistics performance. Again, it is important to note that, in this type of cross-validation design, the methods outlined for the treatment or handling of the data in the derivation sample are legitimate procedures for establishing a set of regressors that best predicts the criterion. In simple cross-validation, the weights derived in one sample are applied to other data in order to evaluate the predictive power and stability of the

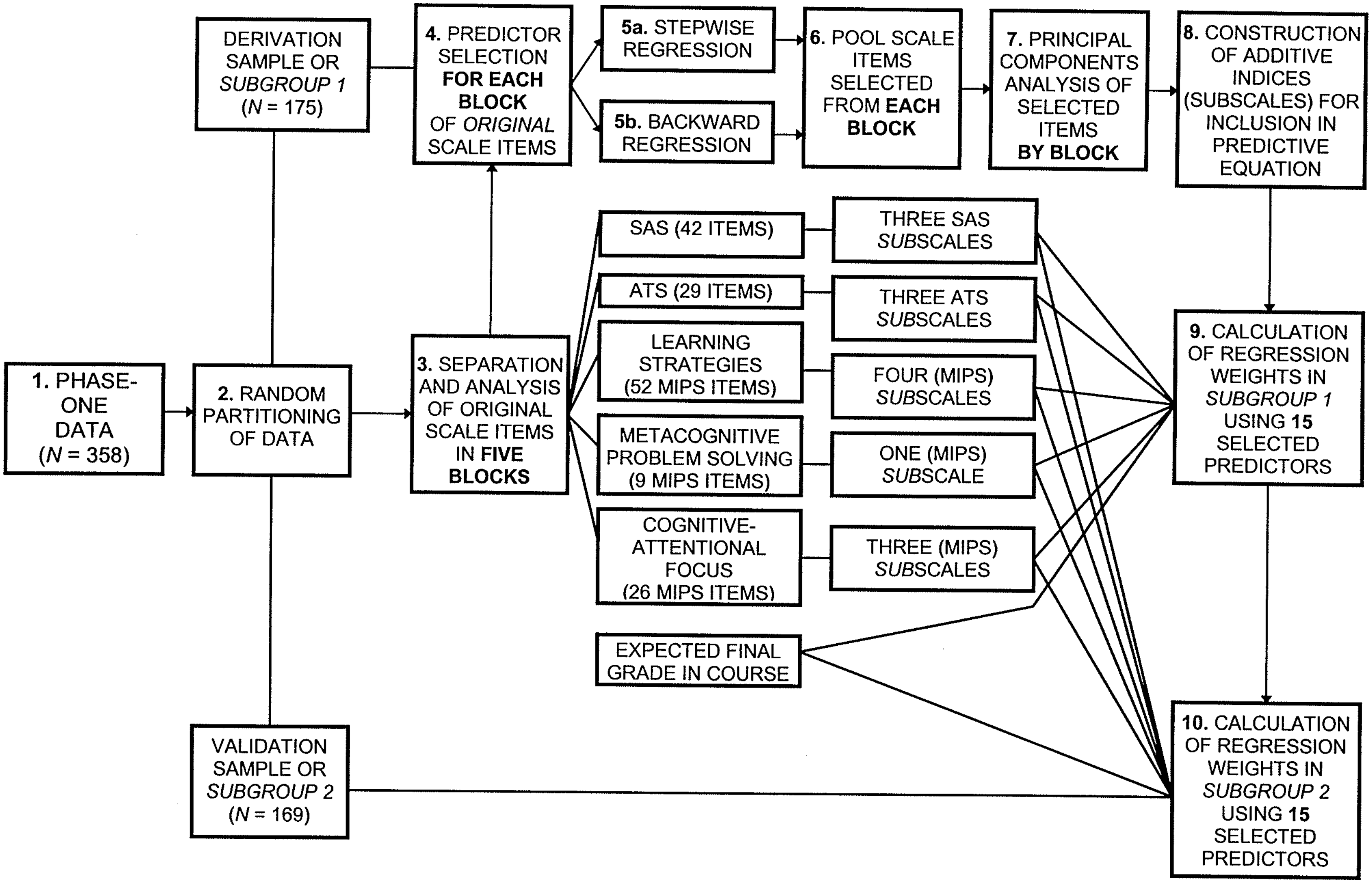


Figure 2. A Flow Chart of Regression Predictor-Selection Procedures for Phase-One Data.

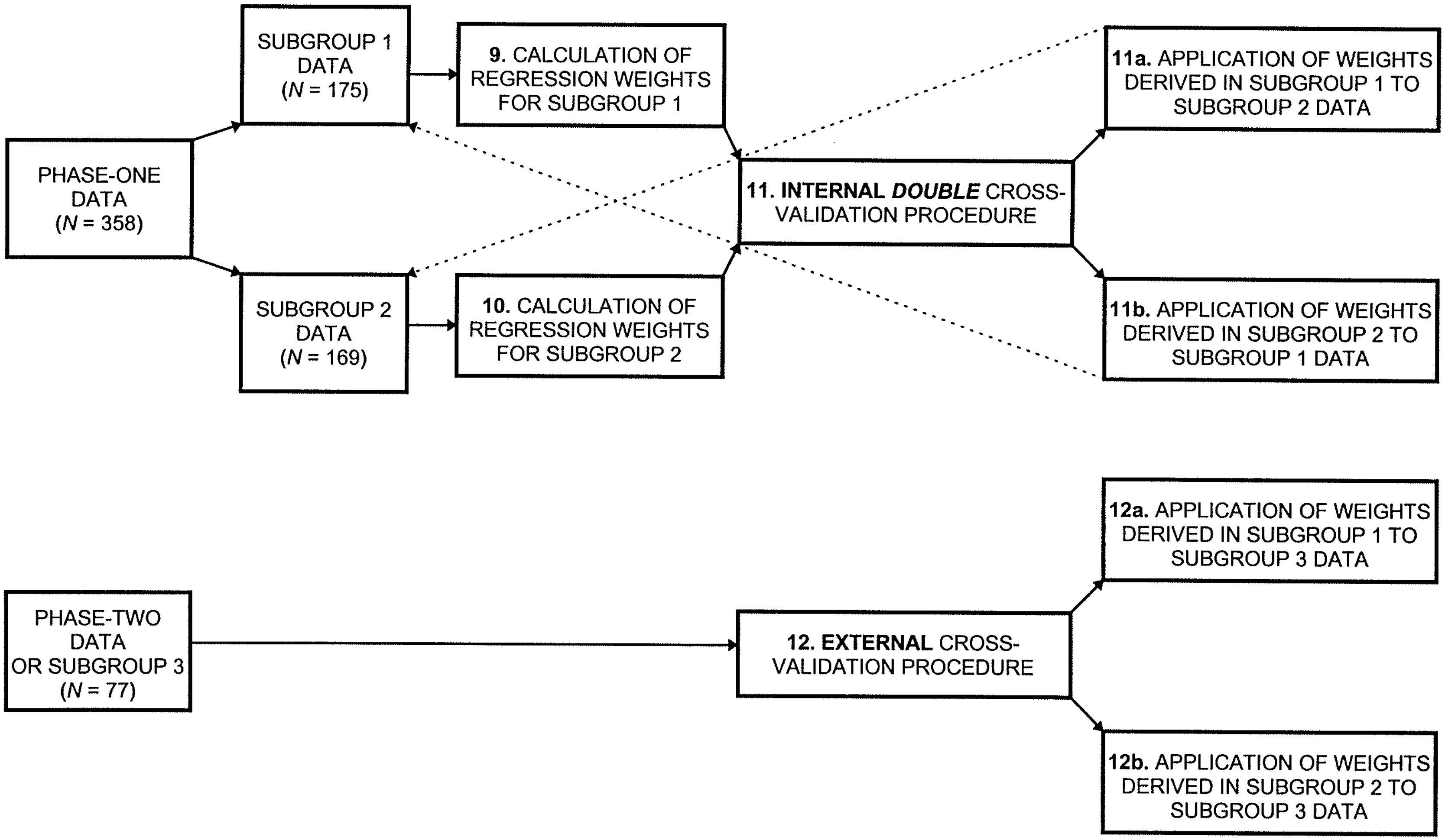


Figure 3. A Flow Chart of Internal and External Cross-Validation Procedures: Phase-One and Phase-Two Data.

regression equation. However, this study incorporates aspects of both internal (double) cross-validation and external replication, as outlined in Figures 2 and 3.

Selection of Regression Predictors

The existing data set of 358 cases (*Figure 2, No. 1*) was first randomly partitioned into two samples (*Figure 2, No. 2)* using a computerized random-generator procedure (SPSS, 1993): the derivation sample (or Subgroup 1) and the validation sample (or Subgroup 2).[15] The selection of variables to be included in the final predictive regression equation was carried out with the derivation sample data. Predictor selection proceeded according to the following five blocks or subsets of scale items (*Figure 2, No. 3 & 4*).

1. Statistics Anxiety Scale (SAS) - 42 items.
2. Attitudes Toward Statistics (ATS) scale - 29 items.
3. Learning strategies component of the Mathematics Information Processing Scale (MIPS) - 52 items.
4. Metacognitive problem-solving procedures - 9 (MIPS) items.
5. Cognitive-attentional focus in evaluative contexts - 26 (MIPS) items.

Each of these groups of variables was first analyzed with a stepwise selection method of multiple regression (*Figure 2, No. 5a*). Pedhazur (1997) has discussed the relative merits of several regression-based procedures of selecting predictors, but he remarks that such decisions ultimately reflect "the researcher's specific aims, resources, and frame of reference" (p. 211). The primary goal is to identify a set of predictors that is nearly as efficient as the entire pool. This process began by submitting each of the five blocks of scale items (as noted above) to separate regression procedures. For each of these five *sub*-analyses, the

[15] Due to missing values for the EXPGRADE variable, the final sample sizes for the derivation and validation samples are 175 and 169 cases, respectively.

variables retained in the initial stepwise solution were removed from the original group of scale items, and the remaining predictors were re-submitted to subsequent regression solutions. This procedure was continued until no further items met the criteria for inclusion.[16] A similar process was followed for each of the five blocks of variables but with a backward elimination method of regression analysis (*Figure 2, No. 5b*).

The two methods of regression analysis yielded somewhat different, but largely overlapping, subsets of items for potential inclusion in the final predictive equation. The two respective lists of variables were compared and compiled, in conjunction with an examination of item-to-criterion correlations, for all variables in each of the five blocks of scale items (*Figure 2, No. 6*). The resulting five subsets of predictors were then submitted separately to principal components analysis, with varimax rotation, for further data reduction (*Figure 2, No. 7*). The examination of rotated factor coefficients provided information concerning the deletion of variables due either to high levels of factorial complexity or to weak relationships with factors. Factor analysis was used primarily to explore empirical dimensions among the various subsets of scale items, in order to assist the construction of additive indices needed to calculate the final predictive regression equation. The following subsections discuss the results of factor analyzing predictor variables selected from the five blocks of original scale items. Rotated

---

[16] The default criteria for inclusion and removal in the stepwise solution are 0.05 (Probability of F-to-enter) and 0.10 (Probability of F-to-remove).

factor matrices for the associated principal components analyses are presented in Appendix B.

Statistics Anxiety Scale regressors. The investigation of statistics anxiety is commonly based on paper-and-pencil instruments adapted from earlier mathematics anxiety research. Although it has been suggested that mathematics and statistics anxiety represent content-specific forms of a more general concern over evaluation, it is unclear to what extent subject matter issues contribute to students' affective responsiveness to statistics learning. Some of the central arguments against treating statistics anxiety as a specialized form of test anxiety stem from the mathematical and problem-solving features of statistics content.

As was noted in Chapter III, Method, the Statistics Anxiety Scale (SAS) used in this study comprises 42 items adapted from the MARS (Suinn et al., 1972). The SAS items depict multiple facets of two general dimensions of statistics anxiety, namely students' feelings of apprehensiveness while (1) carrying out basic learning activities and (2) preparing for or taking tests in a current course. Zeidner (1991) conducted research on statistics anxiety using a similarly designed Statistics Anxiety Inventory (SAI), within which he identifies two factorial dimensions labelled *Statistics Content Anxiety* and *Statistics Test Anxiety*.

Based on the results of regressing statistics course performance (i.e., final grades) on all of the SAS items, 21 scale elements were retained for orthogonal principal components analysis. Zeidner's (1991) work in this area suggests a possible two-factor structure in the SAS. In this case, however, a decision was made to retain three factors after comparing the predictive efficiencies of the two- and three-factor solutions for the SAS. The results of these exploratory extraction procedures indicated the presence of two substantive factors and one smaller, nontrivial factor. The three-factor rotated extraction accounts for 58.2% of the total variance, with eigenvalues of 7.56, 3.47, and 1.19, respectively (see Table B1 of Appendix B for rotated factor matrix).

The first of the three principal components, Factor I, *Statistics Evaluation Anxiety*, seems to tap students' self-reported responsiveness to evaluative processes. This factor accounts for 36% of the variance in the 21 items selected from the original SAS, which represents more than one-half of the variance explained by all three factors. Ten of the SAS items included in this analysis exhibit highest correlations with Factor I, with coefficients ranging from .872 to .469. A prominent feature of the evaluation dimension involves students' responsiveness while preparing for, thinking about, and awaiting the results of statistics tests or examinations. Although the SAS includes items describing a number of evaluative contexts, the results of factor analysis emphasize situations pertaining to pre-test stress and study processes. Performance anxiety is also evident in several of the item markers,[17] for example, completing difficult homework assignments and carrying out problem-solving activities. Three of the 10 items with largest coefficients on Factor I deal with statistical problem solving (e.g., "Not knowing the formula needed to solve a problem").

---

[17] The expression "item markers" is used here and elsewhere in this chapter to refer to scale items with highest correlations on respective factors. For example, there are 10 item markers associated with Factor I, three of which pertain to statistical problem solving.

Factor II of the SAS, *Statistics Learning Anxiety*, resembles a more general form of apprehension associated with learning statistical concepts, formulas, and procedures. Nine SAS items are most highly associated with this factor. The corresponding coefficients range between .797 and .522. Exemplary item markers describe situations in which students *anticipate* having to read chapters, apply formulas, attempt problems, and start assignments. As diverse as these items may appear, they all revolve around students' needs to address the specialized demands of acquiring statistical content. Factor II draws attention to the kinds of experiences that evoke feelings of apprehensiveness in students as they begin to confront statistics curricula through classroom instruction, textbook study, and applied tasks.

The final factor extracted from the SAS, *Statistics Course Anxiety*, appears to represent a subdimension of the statistics learning theme. The two items most highly related to this factor address students' self-reported feelings of anxiety while walking "to" and "into" a statistics class. It would seem that both of these elements of the SAS capture feelings of anticipation about what will be covered in the upcoming class period. In this sense, Factor III appears to blend aspects of the two more substantive dimensions of the SAS, that is, students' responsivity both to statistics content and to prospective evaluations (e.g., tests or problem-solving exercises).

Attitudes Toward Statistics regressors. The investigation of attitude structures in statistics is at least partially an outgrowth of similar research in mathematics education. Specialized instruments, such as Wise's (1985) Attitudes

Toward Statistics (ATS) scale, have been developed to gauge students' enjoyment or valuation of statistics learning. The ATS is intended to probe students' general attitudes toward statistics courses and their perceptions of the usefulness of statistics in professional and everyday life. Analyses of the ATS scale (Wise, 1985) commonly identify a two-factor structure, that is, attitude toward the field of statistics versus attitude toward course content.

Prior to the exploration of the factorial structure of the ATS, the complete set of 29 scale items was analyzed with stepwise and backward selection methods of multiple regression to identify salient predictors of the criterion variable. Thirteen ATS items were selected and subsequently submitted to principal components analysis. As with the SAS, two factorial dimensions were expected, but the three-factor solution yielded somewhat improved predictive efficiency. These three factors together account for 60.3% of the total variance in the 13 ATS items with respective eigenvalues of 3.97, 2.76, and 1.11 (see Table B2 of Appendix B for factor coefficient matrix).

The first factor, *Attitude Toward the Course*, seems to depict students' self-reported concerns over enrolling in a statistics course, as well as dealing with numbers and formulas. The four items with highest coefficients on this factor (i.e., ranging from .896 to .807) imply discomfort, nervousness, and confusion while taking a statistics course. Because of the potential overlap between the attitude and anxiety constructs, all items selected from both the SAS and ATS scales were jointly submitted to principal components analysis. The results reveal a clear separation between the ATS and SAS factors and their respective item markers (see Table B3 of Appendix B for rotated matrix).

Factor II, *Attitude Toward the Field*, appears to emphasize students' perceived valuation of statistics in their particular field of study, as defined by research needs, professional activities, and performance or training requirements. Six ATS items exhibit highest coefficients on Factor II, with corresponding values ranging from .766 to .572. In this three-factor solution, three of the items that Wise (1985) includes in the Attitude Toward Field subscale split away to form a separate factor labelled *Attitude Toward Statistics in Everyday Life*. This latter factor highlights a subdimension of students' attitudes toward the field of statistics, which is concerned with the usefulness of statistics in everyday life. As such, the ATS items associated with Factor III can be empirically distinguished from those dealing with the perceived relevance of statistics to professional or disciplinary applications. Otherwise, the factor analysis results reported here are consistent with Wise's (1985) interpretation of the dimensionality of the ATS.

Statistics learning strategy regressors. Research dealing with the impact of cognitive factors on mathematics anxiety and performance has often been restricted to high-school achievement levels or basic mathematics skills. There are many other cognitive variables that may influence statistics learning outcomes. In contrast to content-specific cognitive skills, the Mathematics Information Processing Scale (MIPS, see Appendix I) probes learning-related factors such as study method, knowledge acquisition, information processing, problem solving, and attentional focus. Regression and dimensional analysis of the MIPS proceeded after separating the 87 scale items into three groups dealing with (1) statistics learning strategies (Items 1 to 52), (2) metacognitive problem-solving procedures (Items 53a to 53i), and (3) attentional deployment in evaluative contexts (Items 54a to 54z).

The initial 52 elements of the MIPS were designed to provide indications of students' preferred or habituated study orientations to statistical material. These items address a number of theoretical issues associated with the distinction made between deep-conceptual and surface-instrumental learning:

- comprehensive (understanding) versus procedural-instrumental learning
- integrative versus reproductive (rote memorization) study techniques
- independent-mastery versus instructor-dependent learning
- pattern recognition versus random (spontaneous) recall
- structured versus disorganized study
- hierarchical versus serial storage and retrieval processes

This original set of 52 scale elements dealing with statistics learning strategies was reduced to 31 variables based on the results of item-to-criterion correlations, regression analysis, and estimates of internal consistency. The exploration of the dimensionality of selected MIPS items began with a number of trial factor solutions. Substantive factors representing deep-conceptual and surface-instrumental features of statistics learning were highly stable across analyses extracting 2 to 5 components, with some variation in the configuration of peripheral factors. The final rotated solution was restricted to a four-factor structure, based on explained variances and eigenvalues. These four factors account for 41.2% of the total variance in the selected MIPS items, with eigenvalues of 5.16, 4.12, 1.91, and 1.58 (see Table B4 in Appendix B for rotated coefficient matrix).

Factor I, *Surface-Disintegrated Learning*, suggests lack of structure in study processes and general uncertainty over how or what to learn. This latter issue is illustrated by item markers that describe students' inability to grasp the meaning of test questions and persistent unproductive problem-solving experiences. Students who espouse surface learning strategies tend to be more easily confused and lack confidence in their ability to manipulate formulas or to solve statistical problems, whereas students who employ integrative study methods are more apt to be deliberate and orderly in their acquisition of statistical material. The factor coefficients corresponding to the 10 scale elements exhibiting highest associations with this factor range between .763 and .360.

Factor II, *Deep-Conceptual Learning*, seems to depict an organized orientation to learning statistics that involves looking for associations between ideas and probing deeper meanings in the content. Exemplary items describe study strategies that parallel Biggs' (1985) deep approach to learning, for example, identifying central ideas, clustering course materials, examining theoretical issues, and preparing study objectives. Eleven of the MIPS items manifest highest correlations with this factor, with coefficients ranging from .657 to .398.

Factor III, *Strategic Learning*, resembles aspects of Biggs' (1985) achieving approach to learning and Entwistle's (1988) achieving orientation. Both of these conceptions of learning imply an efficient, though not necessarily extensive, use of time and effort to perform tasks. The term *strategic* here connotes deliberate or planned use of learning-related resources. Entwistle (1988) describes the achieving orientation as "a combination of careful planning, systematic study methods, positive attitudes, and conscientiousness which… [depend] on the student's individual conceptions of learning and purposes in studying" (p. 43). The seven items most highly associated with Factor III appear to reflect a strategic orientation to statistics learning that emphasizes independent study, metacognitive problem solving, mastery learning, and selective use of memorization tactics for test preparation. The rotated factor coefficients for corresponding item markers range from .673 to .470.

Finally, Factor IV, *Procedural Learning*, seems to represent a subtle but important departure from the first three factors. Of the MIPS items selected by regression analysis, only three show strong associations with Procedural Learning, and of these, two are moderately cross-correlated with Surface-Disintegrated Learning. Factor IV is reminiscent of Skemp's (1978) conception of *instrumental* understanding, that is, performing tasks without knowledge of underlying principles. Despite the negative connotation sometimes associated with procedural learning, it need not impair course performance.

Metacognitive problem-solving regressors. The MIPS includes nine items dealing with students' self-reported use of metacognitive problem-solving procedures. The combined results of stepwise and backward regression analyses of these items identified five metacognitive strategies for potential inclusion in the final equation (in descending order of predictive value): determining where the problem fits into the topics on the test (Item 54b), developing a general frame of understanding (Item 54a), selecting strategies to carry out a plan (Item 54g), identifying the major goal (Item 54c), and developing a plan of action (Item 54f). On closer examination, multiple regression analysis revealed that Item 54b was as efficient as (and often more efficient than) *any* additive combination of two or more of the five metacognitive items in predicting the criterion variable. Hence, only Item 54b was retained in the final regression equation.

Cognitive-attentional regressors. Earlier theoretical discussions in Chapter II indicated that increased levels of test anxiety may affect the operation of a number of cognitive-attentional processes. Twenty-six items in the MIPS probe students' perceptions of how (or to what extent) evaluation procedures impact their attentional focus, concentration levels, retrieval processes, self-preoccupied thoughts, perceptions of self-doubt, and the like. Based on the results of multiple regression, 17 of these MIPS items were submitted to principal components analysis. The three-factor extraction yielded the most interpretable solution, with eigenvalues of 4.72, 1.70, and 1.43, respectively, and a total explained variance of 46.2% (see Table B5 in Appendix B for factor matrix).

Factor I, *Distractibility*, appears to represent a darting and unfocused attentional structure evoked by evaluative conditions. Six MIPS items exhibit highest coefficients on this factor, with values ranging from .773 to .497. Exemplary markers describe attentional disruptions (i.e., distractibility) associated with increased levels of emotional reactivity, difficulties determining central cues, and examining irrelevant or peripheral information. As a consequence, attention moves across a broad range of cues in a haphazard and non-strategic manner.

The second component is labelled *Cognitive Interference* based on Sarason's (1984) discussions of how internal (e.g., worry) and external (e.g., noise) factors may obstruct or disrupt attentional focus on the task at hand. Seven of the 17 MIPS items included in this factor solution share highest associations with this factor, and the corresponding coefficients range between .730 and .428. This factor is related to a variety of self-focusing tendencies such as thinking about grades and worrying about others' perceptions of one's performance. Concern over poor performance is involved in this factor, which potentially involves aspects of negative self-preoccupation.

Aspects of Factors I and II are consistent with Korchin's (1964) characterization of high-anxious individuals as "unable to concentrate, hyperresponsive, and hyperdistractible" (p. 71). Although the attentional field is generally thought to narrow as anxiety increases, Korchin (1964) suggests that, at extreme levels, there is a tendency toward cognitive disorganization or regression. Furthermore, the progressive deployment of attention toward self-preoccupying and self-deprecating thoughts may account for the occurrence of "mental blocks" or interruptions in problem-solving processes.

Factor III, *Focused Attention*, resembles a more strategic and directed form of attentional deployment. Of the four items most highly related to this factor (i.e., coefficients of .706 to .447), two are concerned with the student's perceived ability to manage or to control his or her own affective responsiveness to evaluative stressors. Factor III combines this capacity for self-regulation with being able to shut out performance inhibitions. Concentration may then be focused on identifying central cues and finding strategies to answer questions, thereby preventing the absorption of thought processes in self-doubt. In general, the results of factor analyzing this block of MIPS items indicate the presence of two substantive and quite distinctive forms of attentional deployment in evaluative conditions, that is, highly focused attention as contrasted with distractibility and cognitive interference.

Construction of additive indices. Based on the results of multiple regression and principal components analysis, 83 out of a total of 158 original scale items included in the SAS, the ATS, and the MIPS (see p. 41) were used to generate 14 subscales. Table 1 below provides the names, variable labels, and numbers of items associated with the predictors selected for further analysis (*Figure 2, No. 8)*. Thirteen of these measures are multi-item additive indices, whereas metacognitive problem solving is operationalized with a single scale item.

Table 1 identifies a fifteenth predictor, Expected Final Course Grade (EXPGRADE), which was included in the final regression equation along with the 14 affective and cognitive–learning variables. The socio-demographic section of the study questionnaire included an item that asked "What final grade do you expect to receive in this course: (A+, A, A-, B+, B, B-, C+, C, D, F)?" Students' indications of their expected final letter grades in introductory statistics were converted into the numeric variable EXPGRADE, ranging from 0 ('F') to 9 ('A+').

Table 1
Names, Labels, and Numbers of Items Associated with Final Regressor Variables

Name of Index (Subscale)	Regressor Label	Number of Items
Affective Variables:		
Statistics Evaluation Anxiety	SASEVAL	11
Statistics Learning Anxiety	SASLEARN	8
Statistics Course Anxiety	SASCLASS	2
Attitude Toward Course	ATSCOURS	4
Attitude Toward Field	ATSFIELD	6
Attitude Toward Statistics in Everyday Life	ATSLIFE	3
Expected Final Course Grade	EXPGRADE	1
Cognitive–Learning Variables:		
Surface-Disintegrated Learning	SURFMIPS	10
Deep-Conceptual Learning	DEEPMIPS	11
Strategic Learning	ACHVMIPS	7
Procedural Learning	PROCMIPS	3
Metacognitive Problem Solving	METAPROB	1
Distractibility	DISTRACT	6
Cognitive Interference	INTRFERE	5
Focused Attention	FOCUSATT	6

As is indicated both in Table 1 and in Figure 4, these 15 predictors can be broadly grouped into *affective* and *cognitive–learning* variables.

EXPGRADE combines aspects of *self-attribution*, such as students' estimations of their own cognitive skills, feelings of self-confidence, and subjective interpretations of previous learning (and evaluative) experiences in statistics or other related courses. In this sense, students' expectations concerning their final grades may be understood to blend the effects of a multitude of interrelated affective factors. Insofar as EXPGRADE can be interpreted as

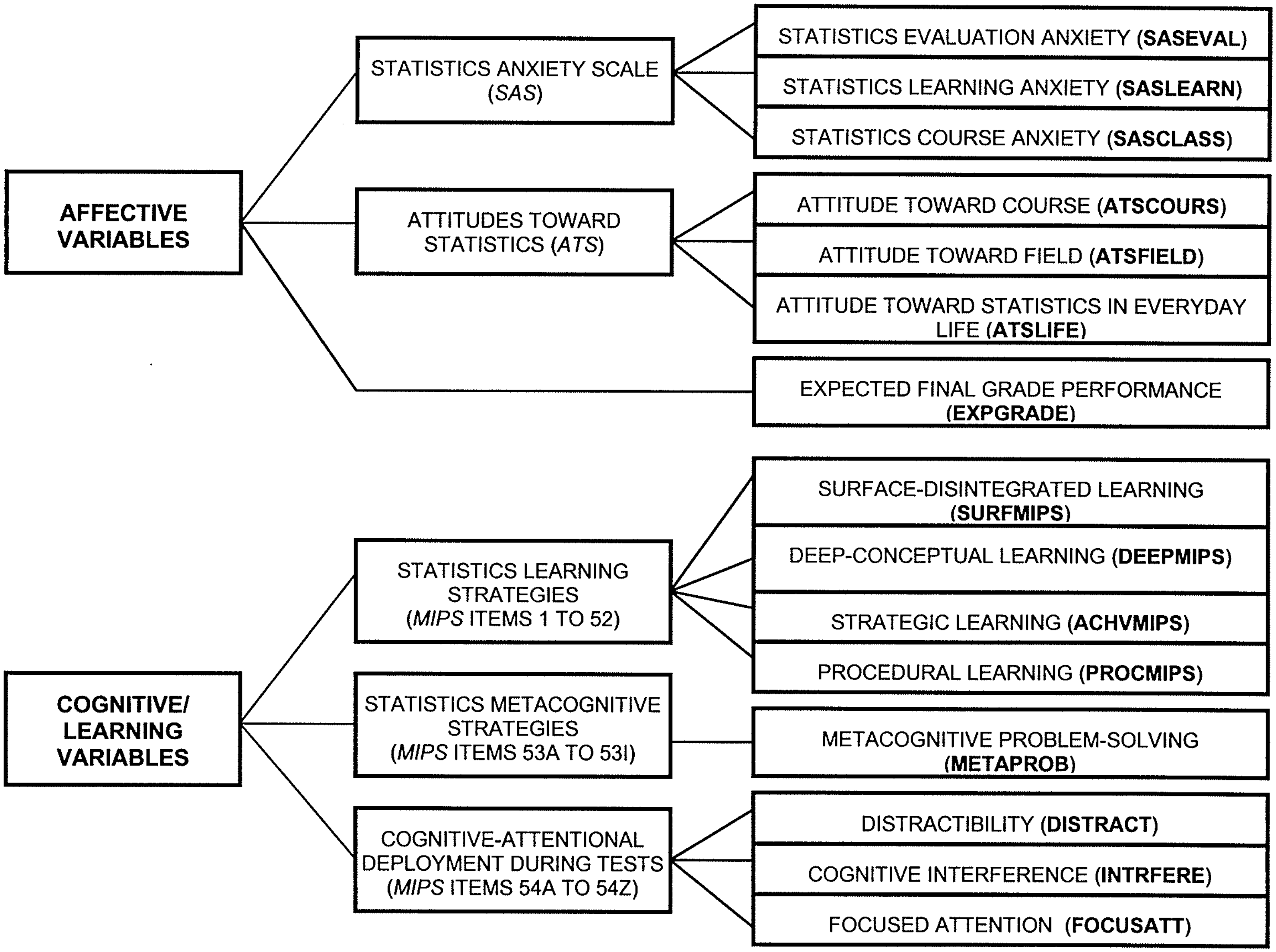


Figure 4. Classification of 15 Regression Predictors into Affective and Cognitive–Learning Variables.

representing aspects of statistics self-confidence, it is highly relevant to the prediction of course performance. Feinberg and Halperin (1978) and Goldstein and High (1992) report significant positive correlations (i.e., *r*s = .33 and .15) between expected or predicted grades and statistics performance.

The zero-order correlation between EXPGRADE and students' final grades, in the derivation sample, is .51. Due to the overall magnitude of this coefficient, EXPGRADE was entered in the last step of the regression analysis summarized in Table 2 in order to prevent this variable from dominating or otherwise obscuring the interpretation of the predictive equation. Put differently, EXPGRADE was stepped into the regression procedure *after* all other affective and cognitive–learning subscales had already been entered. This method of analyzing the data permitted a clearer understanding of how the variables combined (or interacted) in the prediction of final course performance.

Calculation of the Final Predictive Regression Equation (Figure 2, No. 9)

Multiple regression procedures routinely optimize coefficients to suit the idiosyncrasies of the data being analyzed. However, care was taken in this study not to "overfit" the equation to the derivation sample, so as not to unduly compromise its predictive efficiency in either the validation sample or newly collected data. In an article entitled "Underprediction from Overfitting: 45 Years of Shrinkage," Wherry (1975) discusses a series of issues surrounding the loss of predictive power when weights derived in one sample are applied to new data (e.g., .90 to .30). Wherry (1975) offers an important insight that has guided the general process of selecting variables for inclusion in the final regression analysis;

he states "I was still convinced, however, that the real basis of the problem [shrinkage] was simply the overfitting of error" (p. 8).

Factor analysis and multiple regression were used here to identify a set of affective and cognitive–learning variables that together yielded the most efficient regression equation predicting statistics course performance. The predictor-selection procedures were aimed at maximizing the R^2, while simultaneously minimizing the total number of regressor variables. Pedhazur (1997) has noted that the overall R^2 coefficient is invariant regardless of the order in which the predictors are entered into regression analysis. However, when the variables are intercorrelated, the proportion of variance attributed to each variable depends on the order or point of its inclusion. It is for this reason that some authors have argued against making efforts to ascertain "unique contribution to variance" (Darlington, 1968) or to determine the relative importance of intercorrelated predictors (Pedhazur, 1997).

No attempt is made in this study to apportion components of the R^2 to particular variables or sets thereof. Attention is focused instead on developing an efficient and stable regression equation to predict statistics course performance in available and new sample data. However, the results of commonality analysis (see Tables 3 and 4) permit some general remarks concerning the relative contributions of groups of variables (e.g., affective versus cognitive–learning) to the prediction of the criterion variable.

Table 2 displays the results of one of many possible "algebraically equivalent" formulas for R^2, that is, one of several different sequences in which the 15 predictors outlined in Table 1 can be entered into regression analysis.[18]

Table 2

Summary of Multiple Regression Analysis (Enter Method) for Variables Predicting Statistics Course Performance in the Derivation Sample

Predictor	R	R^2	R^2 Change	F Change
1. SASEVAL	.1851	.0343	.0343	6.14**
2. SASLEARN	.1953	.0381	.0039	0.69
3. SASCLASS	.2024	.0410	.0029	0.51
4. ATSCOURS	.2401	.0577	.0167	3.01*
5. ATSFIELD	.2654	.0704	.0128	2.32
6. ATSLIFE	.2918	.0851	.0147	2.70
7. SURFMIPS	.3346	.1120	.0268	5.05**
8. DEEPMIPS	.4075	.1660	.0541	10.76***
9. ACHVMIPS	.4091	.1674	.0014	0.27
10. PROCMIPS	.4377	.1916	.0242	4.90**
11. METAPROB	.4567	.2086	.0171	3.51*
12. DISTRACT	.4612	.2127	.0040	0.83
13. INTRFERE	.4729	.2236	.0109	2.27
14. FOCUSATT	.5204	.2708	.0472	10.35***
15. EXPGRADE	.6316	.3989	.1281	33.90***

Note. $N = 175$. SASEVAL = Statistics Evaluation Anxiety, SASLEARN = Statistics Learning Anxiety, SASCLASS = Statistics Course Anxiety, ATSCOURS = Attitude Toward Statistics Course, ATSFIELD = Attitude Toward Field of Statistics, ATSLIFE = Attitude Toward Statistics in Everyday Life, SURFMIPS = Surface-Disintegrated Learning, DEEPMIPS = Deep-Conceptual Learning, ACHVMIPS = Strategic Learning, PROCMIPS = Procedural Learning, METAPROB = Metacognitive Problem Solving, DISTRACT = Distractibility, INTRFERE = Cognitive Interference, FOCUSATT = Focused Attention, and EXPGRADE = Expected Final Course Grade.
* $p < .10$. ** $p < .05$. *** $p < .01$.

[18] More detailed results of multiple regression analysis, for both the derivation and validation samples, are presented in Tables C1 and C2 of Appendix C.

In the derivation sample, the R^2 associated with these 15 predictors is .3989, which represents slightly less than 40% of the total variance in the criterion variable. The first six variables can be termed affective measures (i.e., statistics anxieties and attitudes), followed by eight cognitive and learning-related factors (e.g., study strategies, metacognitive problem solving, and attentional deployment in evaluative contexts). For the reasons noted earlier, the EXPGRADE variable (i.e., self-attribution), which is closely aligned with the affective domain, was entered in the final step of the analysis. These same 15 variables were used to establish two separate sets of regression weights calculated in the derivation (*Figures 2 & 3, No. 9*) and the validation samples (*Figures 2 & 3, No. 10*), both of which were required for double cross-validation (see Figure 3).

Although varying the order of entry would yield different coefficient values for R, R^2, R^2 Change, and F Change, several broad comments are warranted. First, the three anxiety subscales account for 4.1% of the variance in the criterion variable, despite being entered in the initial three steps of this analysis. The three attitude subscales explain an additional 4.4%, bringing the total R^2 associated with these particular affective variables to 8.5%. In this sequence of predictors, the six affective measures, excluding EXPGRADE, are responsible for .0851/.3989 (or 21.3%) of the total variance explained by the entire set of 15 regressors.

Second, of the eight learning-related and cognitive-attentional variables entered *after* the inclusion of the anxiety and attitude subscales, four contribute significantly ($ps < .05$) to the increment of R^2: SURFMIPS, DEEPMIPS, PROCMIPS, and FOCUSATT. Taken together, the cognitive–learning predictors account for 18.6% of the variance in statistics course performance, over and above the combined effects of the affective variables. Third, EXPGRADE, the global measure of statistics self-confidence, explains an additional 12.8% of the variance in the criterion variable (after partialing out the effects of all other predictors). This brings the total R^2 to .3989.

Although the results of multiple regression analysis are not explicitly interpreted in terms of variance partitioning, a brief discussion of *commonality analysis* is relevant to the research questions outlined in Chapter I. Commonality analysis was "designed to identify the proportions of variance in the dependent variable that may be attributed uniquely to each of the independent variables, and the proportions of variance that are attributed to various combinations of independent variables" (Pedhazur, 1997, pp. 261-262). The unique contribution of each predictor, that is, its *uniqueness* (Pedhazur, 1997) or *usefulness* (Darlington, 1968), is designated by the increment in the explained variance attributed to it when it is entered in the last step of the regression procedure. This same logic or method of data analysis may be extended to blocks of variables (e.g., affective versus cognitive–learning factors), but with "the realization that *commonality analysis is useful and meaningful in predictive but not in explanatory research*" (Pedhazur, 1997, p. 269).

Initially, separate multiple regression procedures were performed to determine the increment in R^2 when each of the 15 predictors was entered last into the analysis. Six of the regressors contributed significantly (ps $< .05$) to the total explained variance in statistics course performance, after partialing out the effects of all other variables: EXPGRADE, FOCUSATT, METAPROB, INTRFERE, SURFMIPS, and PROCMIPS, in descending order of uniqueness. However, estimations of the uniqueness of particular variables based on commonality analysis should be interpreted carefully in light of the set of predictors under investigation. Commonality analysis indicates the relative predictive power of variables, both individually and jointly. Pedhazur (1997) amplifies this point by noting that the inclusion or deletion of regressors may have a dramatic effect on the proportion of variance (uniquely) attributed to each variable.

Commonality analysis was next applied to the derivation sample data after grouping all predictors except EXPGRADE into two broad classes of variables (see Figure 4):

1. *Affective Factors*—the three subscales of the Statistics Anxiety Scale (SASEVAL, SASLEARN, SASCLASS) and the three subscales of the Attitudes Toward Statistics scale (ATSCOURS, ATSFIELD, ATSLIFE).
2. *Cognitive–Learning Factors*—the eight subscales of the Mathematics Information Processing Scale (SURFMIPS, DEEPMIPS, ACHVMIPS, PROCMIPS, METAPROB, DISTRACT, INTRFERE, and FOCUSATT).

Table 3 shows the results of this procedure when EXPGRADE is maintained as a separate predictor apart from the affective and cognitive–learning variables. Two regression analyses were performed with EXPGRADE entered in the *first* step, followed alternately by the block of affective factors (*Analysis 1*) and then the cognitive–learning factors (*Analysis 2*). Two parallel regression solutions were calculated in which EXPGRADE was entered *last* into the equation. EXPGRADE was introduced in the first and then in the last steps of these two sets of regression analyses in order to shed light on the relative contributions of the affective and cognitive–learning predictors in combination with EXPGRADE. Commonality analysis for these four arrangements of regressor variables is presented below.

Table 3

Summary of Commonality Analysis for EXPGRADE, Affective, and Cognitive–Learning Variables Predicting Statistics Course Performance in the Derivation Sample

Predictor	R	R^2	R^2 Change	F Change
Analysis 1:				
EXPGRADE	.5098	.2599	.2599	60.75***
AFFECTIVE[a]	.5259	.2766	.0167	0.64
COGNITIVE[b]	.6316	.3989	.1223	4.04***
Analysis 2:				
EXPGRADE	.5098	.2599	.2599	60.75***
COGNITIVE	.6066	.3679	.1080	3.53***
AFFECTIVE	.6316	.3989	.0310	1.37
Analysis 3:				
COGNITIVE	.4698	.2207	.2207	5.88***
AFFECTIVE	.5204	.2708	.0501	1.83*
EXPGRADE	.6316	.3989	.1281	33.90***
Analysis 4:				
AFFECTIVE	.2918	.0851	.0851	2.61**
COGNITIVE	.5204	.2708	.1856	5.09***
EXPGRADE	.6316	.3989	.1281	33.90***

Note. $N = 175$.

[a] The AFFECTIVE block of variables includes the three subscales of the SAS and the three subscales of the ATS.

[b] The COGNITIVE block includes the eight subscales of the MIPS.

* $p < .10$. ** $p < .05$. *** $p < .01$.

As a final test of the relative predictive power of the two broad classes of variables, EXPGRADE was grouped with the affective measures. Statistics course performance was again regressed on the affective and cognitive–learning factors, after alternating the entry of each of these two blocks in the final step. Table 4 shows the results of commonality analysis concerning the relative contributions of affective versus cognitive–learning variables to the prediction of final grades.

Table 4

Summary of Commonality Analysis for Blocks of Affective (including EXPGRADE) and Cognitive–Learning Variables Predicting Statistics Course Performance in the Derivation Sample

Predictor	R	R^2	R^2 Change	F Change
Analysis 1:				
AFFECTIVE[a]	.5259	.2766	.2766	9.12***
COGNITIVE[b]	.6316	.3989	.1223	4.05***
Analysis 2:				
COGNITIVE	.4698	.2207	.2207	5.88***
AFFECTIVE	.6316	.3989	.1782	6.73***

Note. $N = 175$.
[a] The AFFECTIVE block of variables includes the three subscales of the SAS, the three subscales of the ATS, and EXPGRADE.
[b] The COGNITIVE block includes the eight subscales of the MIPS.
* $p < .10$. ** $p < .05$. *** $p < .01$.

Internal (Double) Cross-Validation: Phase-One Data

As was noted in Chapter III, Method, cross-validation has been used to empirically investigate the replicability of results. Table 5 presents *invariance* statistics for an internal double cross-validation procedure (*Figure 3, No. 11*), which involved randomly splitting an available data set into two subgroups. This method of analysis yielded separate multiple correlation coefficients (actual *R*s) for the two subgroups. For each data pool, correlations were calculated between the *observed* (STATGRAD) and the *predicted* criterion scores, the latter of which were generated by applying a subgroup's regression weights to its own data (i.e., PREDY11 and PREDY22). The actual *R*s for Subgroups 1 and 2 are .6316 ($r_{\text{STATGRAD x PREDY11}}$) and .6515 ($r_{\text{STATGRAD x PREDY22}}$), respectively. The corresponding R^2s for these two subgroups are .3989 (R_1^2) versus .4245 (R_2^2). Although researchers may choose merely to compare the squared multiple correlations between subgroups, Thompson (1994) suggests that even this task is made problematic by differences in sample size and sampling error. Notwithstanding this caveat, the two sets of *R*s and R^2s for Subgroups 1 and 2 are very similar.

Table 5

Correlation Coefficients from Invariance Analysis

	STATGRAD	PREDY11	PREDY21
PREDY11	.6316[a]		
PREDY12	.5430[a]	.8602[a]	
PREDY21	.5575[b]	-	
PREDY22	.6515[b]	-	.8562[b]

Note. STATGRAD = Final Grade in Statistics Course (*Y*), PREDY11 = Predicted Course Performance using data (X_i) and regression coefficients from the derivation sample (Subgroup 1), PREDY12 = Predicted Course Performance using data (X_i) in the derivation sample (Subgroup 1) and regression coefficients obtained in the validation sample (Subgroup 2), PREDY21= Predicted Course Performance using data (X_i) in the validation sample (Subgroup 2) and regression coefficients obtained in the derivation sample (Subgroup 1), and PREDY22 = Predicted Course Performance using data (X_i) and regression coefficients from the validation sample (Subgroup 2).
[a] $N = 175$. [b] $N = 169$.

Table 5 also provides "shrunken" Rs for the weights obtained in Subgroups 1 and 2, sometimes termed cross-validity coefficients ($r_{yy'}$): .5575 ($r_{\text{STATGRAD x PREDY21}}$) and .5430 ($r_{\text{STATGRAD x PREDY12}}$), respectively. The first of these two correlations ($r_{\text{STATGRAD x PREDY21}}$) was calculated using the data (i.e., observed scores for the regressor and criterion variables) in Subgroup 2 and the weights from Subgroup 1 (*Figure 3, No. 11a*). The latter coefficient ($r_{\text{STATGRAD x PREDY12}}$) was based on the actual scores in Subgroup 1 and the weights from Subgroup 2 (*Figure 3, No. 11b*). The squared cross-validity coefficients are .3108 ($.5575^2$) and .2948 ($.5430^2$). Therefore, the actual Rs and the "shrunken" Rs are quite comparable across the two subgroups.[19]

Invariance analysis represents an empirical process of cross-validation, which also involves the calculation of two invariance coefficients: $r_{\text{PREDY11 x PREDY12}}$ and $r_{\text{PREDY21 x PREDY22}}$. By way of illustration, the first of these two invariance coefficients ($r_{\text{PREDY11 x PREDY12}}$) is a correlation between two (separate) predicted measures of final grade outcome using data in Subgroup 1: PREDY11, which is the product of applying the weights from Subgroup 1, and PREDY12, which is obtained with weights derived in Subgroup 2. A parallel procedure is used to calculate $r_{\text{PREDY21 x PREDY22}}$, that is, by separately applying the weights from Subgroups 1 and 2 to the data in Subgroup 2, and then correlating the two predicted criterion variables (PREDY21 and PREDY22).

---

[19] Refer to Appendix D for a discussion of alternative indictors of shrinkage, that is, squared cross-validity coefficients versus formula-based estimates of ρ_c^2.

The invariance coefficients for weights from Subgroups 1 and 2 are shown in Table 5, that is, .8562 and .8602, respectively. According to Thompson (1994), the invariance coefficient for weights derived from Subgroup 1 ($r_{\text{PREDY21 x PREDY22}}$) *reflects*[20] the degree of shrinkage between the actual R^2 ($.6316^2 = .3989$) for Subgroup 1 and the "shrunken" R^2 ($.5575^2 = .3108$). Further, as both of the invariance coefficients approach one, "more confidence can be vested in the generalizability of the results" (Thompson, 1994, p. 163). In summary, cross-validation and invariance analysis indicate that the results of multiple regression presented in Table 2 are stable across the two subgroups.

External Cross-Validation: Phase-Two Data

Additional questionnaire responses were collected (i.e., phase-two data) to further evaluate the stability of the predictive regression equations derived from the two subgroups of the original (phase-one) data set. The second phase or external replication sample, labelled Subgroup 3, consists of 77 cases. Two predicted measures of statistics course performance were generated by applying the regression coefficients calculated in Subgroups 1 and 2 (of the phase-one data) to the newly collected sample or Subgroup 3: PREDY31 and PREDY32, respectively.[21] For each student in this new sample, there are two predicted values and one actual score for final course performance.

---

[20] Perhaps a better choice of language here might be "indicates."

[21] PREDY31 is the predicted score for statistics course performance generated by applying the regression coefficients from Subgroup 1 to the new sample ($n = 77$), whereas PREDY32 is the predicted criterion variable using the weights from Subgroup 2.

The external cross-validation procedure (*Figure 3, No. 12*) yielded three Pearson product-moment coefficients among the predicted and the observed scores for statistics course performance. When the weights from Subgroups 1 and 2 are applied to phase-two data (*Figure 3, No. 12a & 12b*), the resulting correlations are .5900 ($r_{\text{STATGRAD x PREDY31}}$) and .5328 ($r_{\text{STATGRAD x PREDY32}}$), respectively. These values are highly comparable to the "shrunken" *R*s reported in Table 5, when the regression coefficients derived in Subgroups 1 and 2 are used to predict final grades in the opposing partitions of the phase-one data set: .5575 ($r_{\text{STATGRAD x PREDY21}}$) and .5430 ($r_{\text{STATGRAD x PREDY12}}$). The two squared cross-validity coefficients for the external replication sample ($n = 77$) are .3481 ($.5900^2$) and .2839 ($.5328^2$). These shrunken R^2s compare favorably to the results of internal cross-validation reported earlier. The correlation between the two predicted criterion variables created by separately applying the regression weights from Subgroups 1 and 2 to the new sample ($r_{\text{PREDY31 x PREDY32}}$) is .8420, which further supports the generalizability of the results.

Supplementary Analysis: Person-Related Variables

Studies that investigate the correlates or predictors of statistics performance commonly examine the effects of affective, cognitive, and socio-demographic variables, but seldom all three. Although the consideration of person-related factors is not central to the aims of this thesis, some findings pertaining to the relationship between final grade outcomes and four background variables are provided below:[22]

- highest level of school mathematics (HIGHMATH),
- recency of formal education (LASTSCHL),
- number of years since taking a mathematics course (LASTMATH), and
- gender (GENDER).

The results of analysis of variance procedures comparing achievement levels across categories of these four covariates are presented in Appendix E. Some additional discussion of the relationship between "person" variables and statistics performance is included in Chapter VI, Summary and Conclusions. It should also be noted that the following summary of results is illustrative in nature, as these matters require a more thorough examination than is permitted here.

Highest level of school mathematics. Significant differences in average course grades are evident for two of the four person-related factors noted above: LASTSCHL [$F\,(3, 431) = 5.481, p = .001$] and HIGHMATH [$F\,(3, 410) = 5.968$, $p = .001$]. With regard to the latter variable, achievement increases with exposure to more advanced mathematics curricula ($r = .18, p < .001$). The mean criterion scores corresponding to each of the four categories of HIGHMATH are 4.41, 4.51, 5.14, and 6.49, respectively.

Recency of formal education and of last mathematics course. In this study, statistics performance varies significantly according to the sequencing (or interruption) of students' participation in formal education. The largest proportion

[22] These findings are based on the total sample ($n = 435$), that is, a combination of phase-one and phase-two data.

of the total sample comprises "traditional" students ($M = 5.35$). Of the three categories of non-sequential students, the mean criterion score is lowest for those who had not attended school for *one to two years* ($M = 4.08$). The average final grade among individuals who had interrupted their studies for *more than two but less than five years* ($M = 5.12$) is comparable to that of the "traditional" students, whereas those who had been out of school for *five or more years* exhibit the highest mean achievement score ($M = 6.63$). On a related topic, statistics performance outcomes vary across categories of LASTMATH (i.e., the number of years since last mathematics class) but not significantly (see Appendix E).

Gender. The average statistics grade for female students is higher than that of their male counterparts (i.e., 5.39 versus 5.12), but this difference is not statistically significant [F (1, 433) = 0.9873, $p = .321$]. Indeed, of the 15 regressors used to predict statistics performance in this study, only five exhibit statistically significant differences (ps < .05) by gender (with means for female and male students in parentheses): DEEPMIPS (Ms = 32.34 versus 29.03), SASEVAL (Ms = 35.48 versus 32.31), ATSCOURS (Ms = 14.27 versus 15.24), SURFMIPS (Ms = 24.98 versus 23.65), SASLEARN (Ms = 14.64 versus 13.59), in descending order of the size of F. Further, the complete set of 15 variables accounts for 40.08% of the criterion variance among female students, as compared to 46.02% in the male subsample, that is, R^2s of .4008 ($.6331^2$) and .4602 ($.6784^2$), respectively (see Appendix F).

Commonality analysis was used to examine the relative contributions of the 15 regressor variables to the prediction of female versus male students' course performance. The total data set ($n = 435$) was partitioned into female and male subsamples. Two corresponding sets of 15 regression solutions were then calculated (for both subsamples) in which each variable was entered in the final step of a separate equation. This procedure revealed some noteworthy findings. First, EXPGRADE is the single strongest predictor of the criterion variable in both groups of students. Second, five of the six subscales derived from the Statistics Anxiety Scale and the Attitudes Toward Statistics scale make statistically significant ($ps < .05$) unique contributions to the prediction of female respondents' final grades: ATSFIELD, SASCLASS, ATSLIFE, SASLEARN, and ATSCOURS, in descending order. Only one of the six affective subscales, SASEVAL, adds significantly to the R^2 for the male subsample. Third, of the eight cognitive–learning variables, four show significant unique contributions to the prediction of male students' course performance: INTRFERE, FOCUSATT, DEEPMIPS, and PROCMIPS, in descending order. By comparison, three of these predictors (i.e., PROCMIPS, SURFMIPS, and FOCUSATT) are responsible for significant increments in the R^2 for female students.

CHAPTER V

ANALYSIS OF THE INTERVIEW DATA

This chapter deals with the results of personal semi-structured interviews conducted with nine statistics students. The interview component of this study provided opportunities to probe students' perceptions of the involvement (and interaction) of affective and cognitive factors in their statistics learning. This information was beneficial to both the interpretation of the questionnaire-based data and the consideration of further research (see Chapter VI). Indeed, many of the students' comments during the interviews underscore theoretical issues and concepts raised in earlier sections of this thesis.

Students were guided through a series of discussion points within a flexible and open interview process. Although the content, sequence, and wording of the questions varied somewhat, the interviewing protocol was organized around seven main themes of inquiry:

1. Reasons for taking an introductory statistics course.
2. Exposure to statistics content in school mathematics curricula.
3. Descriptions of the types of learning or study strategies applied or adapted to statistics content.
4. Characterizations of feelings, emotions, attitudes, or anxieties experienced while learning or taking tests in a statistics course.
5. The Statistics Affect Grid (see Appendix G).
6. The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988).
7. Perceptions of statistics self-confidence, including contributing factors and changes over the duration of the course.

With the exception of students' responses to the PANAS, the interviews were largely unstructured in nature. Each of the seven topic areas was introduced briefly, prior to asking any questions, in order to develop a suitable context for discussion. The basic format and flow of the interviews were established by a series of open-ended questions and supplementary prompts (see Appendix H for exemplars). This method of inquiry was selected because it supplied a clear frame of reference in which to conduct the interviews, without overly restricting the content or manner of students' answers. In this regard, participants were encouraged to comment freely when responding to thematic questions. The interviews varied in length from 25 to 58 minutes, with an average of approximately 44 minutes.

Although the majority of the interviewees were "traditional" students, three of the nine participants had not been involved in formal education for at least five years. Indeed, it had been over 10 years since two of the non-sequential students had completed a mathematics-related course. There was also variation in mathematics backgrounds, which ranged from fairly limited to more advanced high-school mathematics. Based on these socio-demographic variables, the interview participants appear reasonably representative of the larger sample. The following sections summarize common themes and observations drawn from discussions with the nine (i.e., five female and four male) students. In all instances, the students' actual names have been replaced by pseudonyms, although gender associations have been maintained.

Reasons for Taking a Statistics Course

Six interviewees indicated that a statistics (or a mathematics) course was required for their intended educational programs (e.g., nursing, business, or social work). Only one student (who had the most limited mathematics background) indicated any degree of apprehensiveness at the prospect of "having" to take statistics, and she indicated that her concern dissipated within the first few classes. The remaining students evidenced a range of affective responses to the requisite statistics course, for example, interest, determination, confidence, and in some instances, affective neutrality.[23] Of those individuals who were not required to take statistics, one thought it might be beneficial to her major and two other (female) students chose the course because of previous success in high-school mathematics.

Exposure to Statistics Content in High-School Mathematics

When asked about exposure to statistics topics in school mathematics, few students could identify any specific experiences with statistics. One of the non-sequential students did recall some high-school instruction in the use of Venn diagrams, which she now associated with probability concepts. A traditional student indicated that he had dealt superficially with the mean, median, and mode in grade 12 mathematics, although he later remarked that "everything [in the university-level statistics course] was new." Only one student (again, non-sequential) was able to state with any degree of certainty that he had covered

---

[23] As a point of clarification, students' affective responses to being required to take statistics should be distinguished from their reactivity to statistics learning and test-taking processes. These latter two issues are discussed in the section labelled *Affective Responsiveness to Statistics Learning*.

topics in high school that were similar to those extending as far as the first test in introductory statistics. For the most part, these particular students either had not been exposed to statistics content, at least not to any significant extent, or they were unable to make connections between previous and current learning.

Learning and Study Strategies in Introductory Statistics

A major component of the interview dealt with the types of learning and study strategies that students applied explicitly to their statistics course. In many instances, problem solving emerged as a central theme of discussion, which included considerations such as pattern recognition, templates or frames, heuristic devices, practice exercises, homework assignments, and metacognitive and meta-affective processes.

Several of the interviewees remarked on the importance of practicing a variety of problems in preparation for tests and examinations which, for some, included seeking out additional textbooks and exercises. This strategy was considered particularly significant by students who indicated a preference for independent or mastery learning. For example, three of the interview participants routinely sought out supplementary exercises in order to learn how to select formulas and to solve problems, but perhaps more important, to improve their understanding of the content. Beth stated that "I found a need to do a lot of questions to make myself familiar with…working with that kind of material." Similarly, Wanda indicated that she felt a strong need to go over many different types of problems until it "becomes part of me." She went on to say that "if I have more examples…if I'm exposed to more variety, then…I can see it better, you know, I'm more sure of what I'm doing." Wanda deemed repetition an essential strategy for understanding and remembering the material over an extended period of time. Robert similarly indicated that it was not enough for him to know formulas, he had to be able to apply them.

Almost all of the interview participants remarked on the importance of the assignments in learning course material and preparing for evaluations. On a related topic, two of the nine students mentioned that they had employed some form of error analysis to improve their problem-solving skills. For example, Kathy reviewed the errors that she made in homework assignments and practice exercises as part of her preparation for tests. John used a similar technique while studying for tests and examinations. He adapted this strategy when he noticed that he was prone to making certain types of errors in practice problems. John also commented on the role of heuristics in solving non-routine problems, for example, looking at the situation logically or looking for similarities to known problem types.

A number of students raised issues bearing on the distinction between procedural and conceptual learning, or to use Skemp's (1978) terminology, *instrumental* and *relational* understanding. One of the interviewees identified a proclivity to implement procedural or reproductive strategies for learning and taking tests in introductory statistics. Helen stated categorically that "It's all step-by-step…It's very clear." She further remarked that she prepared for statistics tests by organizing lecture materials into blocks (i.e., headings and subheadings), to which she attached key words or cues to trigger solution procedures. This is consistent with Davis' (1984) notion of constructing and using "frames" to solve recognizable problems. When Helen encountered a situation in which she could not make an immediate connection between a given problem and a known procedure (i.e., a formula and an accompanying algorithm), she would "carry" the problem across every available frame until she eventually made a selection. In this sense, Helen used a modified form of "templation" (MacPherson & Rousseau, 1996) for solving non-routine problems.

My conversation with Thomas, who was considering a major in mathematics, highlighted aspects of instrumental and relational understanding. Thomas' comments demonstrate, as Biggs (1985) would attest, that it is not advisable to characterize students as "types" of learners, because they may effectively synthesize very distinctive strategies. Such individuals are sometimes referred to as "versatile" learners due to their capacity to select or adapt specialized tactics to suit variable task demands.

In the initial stages of the interview, Thomas likened his approach to following a recipe, which is characteristic of procedural or reproductive learning. However, Thomas seemed to contradict his earlier statement by saying that "I never memorize. I try not to. [I] just understand the logic or the flow, the order, so you can, you know, if you have to on an exam or a mid-term, you can just kind of recreate." Thomas added that, when he answered test questions, he did not "scroll" through all of the statistical procedures on his formula sheet in order to identify a match; rather, he let them "kind of find each other." Indeed, he described this problem-solution process as "intuitive" in nature, but he also admitted to using analytical procedures or heuristics (e.g., breaking things down into their simplest form). Later in the interview, Thomas indicated a strong interest in understanding the background of formulas (e.g., mathematical derivations or knowing why a procedure is used) and, as such, he was appreciative of instructors who focused on what he termed the "fundamentals." "You can understand where a person's going, you know, with the question or you can understand where they're coming from and where they're going to." Thomas' comments exemplify a blend of procedural and conceptual learning, as well as intuitive and analytical problem solving.

In contrast to Thomas' experiences, Doug recounted some notable difficulties with statistical problem solving, particularly on tests. Doug indicated that he would consistently get lost or confused in tests because of missing information, steps, or variables. His "frustration" over not being able to navigate a solution path sometimes led him to consult with friends and tutors. "I'd just ask…how to do it and they'd show me." Although he felt he understood the problem-at-hand after receiving such assistance, he remarked that "I hardly ever saw it on my own." In a similar fashion, Kathy also spoke to friends or attended laboratory sessions when she was unable to answer questions on the homework assignments. In addition to providing practical assistance, however, these consultations seemed to facilitate Kathy's motivation to learn the material and to complete the exercises. Kathy stated that "I have to know what the process is and why before I really kind of get into it."

Affective Responsiveness to Statistics Learning

In addition to discussing cognitive–learning strategies, the interviews included a series of questions concerning students' attitudes, anxieties, and emotions while learning statistics content and taking tests or examinations. In general, the individuals who agreed to be interviewed revealed quite variable affective reactions to statistics learning, in terms of both the apparent direction and the magnitude of their responses.

I will begin with those students who seemed to express low affective intensity when they were questioned about learning and evaluative components of the statistics course. For example, Beth was neutrally disposed toward introductory statistics at the outset and later developed a positive attitude toward the course. Beth had originally planned to take linear algebra but decided instead to enroll in a statistics course. After speaking to a number of students who had taken introductory statistics, she came to the conclusion that "probably people have different…have brains for different things and some people can handle it and some people can't….I was wondering if my brain would be able to handle it." So, Beth began the course with an open mind and became progressively more positive and more confident over the duration of the course. Michelle was also affectively neutral with regard to introductory statistics but for slightly different reasons. Michelle described herself as a very organized person and "pretty confident," because she seemed to "pick up on it fairly easily." As a result, Michelle felt quite positive about the course and her abilities to learn statistics content.

In contrast, a number of students experienced some degree of anxiety when taking statistics tests or examinations. Wanda expressed a classic textbook description of anxiety-induced "blocking" on tests: "I think, you know, there is this mental blockage…you read the question and there is…one variable missing or it's not that obvious as other questions…and everything freezes…You cannot retrieve your information or your pattern." On a related topic, Kathy had experienced fairly pronounced mathematics test anxiety in grade 12, which subsequently led her to monitor and to make efforts to control her emotional responses in the statistics course. Kathy indicated that she began to experience panic attacks during mathematics examinations because of demands to provide more detailed answers. She suggested that these experiences had contributed to adverse emotional reactions in subsequent evaluative contexts. Further, Kathy felt susceptible to similar types of affective responses (i.e., negative) in the statistics course, for example, when she began an assignment and realized she did not understand the questions.

Kathy described a situation in which she could not remember some of the intermediary steps needed to solve a test question and, because of this, she "started panicking." Kathy's awareness of the deleterious effects of anxiety provoked her to devise meta-affective strategies aimed at preventing such feelings from escalating. One such method involved sensing when she was about to panic and then going back over everything (in the question or problem) very slowly in order to regain control.

The severity of Kathy's emotional reactions was not unique among the interviewed students. For example, Doug also experienced fairly intense negative responses to unsuccessful problem-solving episodes. He described the associated affective states as "frustration" and "a little bit of anger." Further, Doug felt he was somewhat more anxious than friends who were taking statistics, but he emphasized that his feelings of frustration and anger occurred when he was unable to solve a problem because of a missing detail. His self-described aversive experiences seemed to be interconnected with his dependence on others (e.g., friends or the instructor) to show him where he had made mistakes.

Although Thomas admitted feeling some tension when he encountered non-routine problems, he did not characterize such reactions as negative. He indicated that he did not feel anything when he was dealing with recognizable problems, because he only had to check the question and write down the answer. However, Thomas remarked that, in the course of solving non-routine problems, he would sometimes feel tense because he was worrying about his grade or because he felt he should know the material. In this sense, challenging problems evoked some degree of anxiety or tension in Thomas, which he characterized as more positive than negative and "4 out of 10" with regard to magnitude.[24] For Thomas, tension could be a good thing, that is, "it leans toward the positive." He also noted that when he began to see the answer, his feelings of positive affect seemed to intensify: "you start pumping...you feel good." He evinced a kind of exhilaration

---

24 Thomas characterized his self-described moderate responsivity to non-standard problems as "4 out of 10," where "10" referred to the most intense form of emotional reaction.

at being able to solve difficult problems. Later in the interview, Thomas added that these types of affective experiences were more apt to occur outside of tests and examinations.

With regard to changes in affective states over the duration of the course, two non-sequential students indicated that they were fairly anxious at the beginning of the statistics course, largely because they did not know what to expect. For example, Robert expressed a highly positive disposition toward the course, but he found the emphasis on formulas somewhat intimidating at the outset. His concerns seemed to diminish when he determined that "it's just a matter of applying the right formula and putting the numbers in the right spots." Helen experienced a similar shift in affect from the beginning to the end of the statistics course; feelings of nervousness were replaced by enjoyment and self-confidence. When Helen encountered difficulties during tests, she would invoke a meta-affective strategy to control her emotions, cognition, and behavior. In order to prevent what she called "meltdowns," Helen would tell herself to "just relax and access." By her own admission, anxiety caused Helen to feel confused or "emotionally jumbled," and it was at these times that she felt she must intervene.

Statistics Affect: Some Illustrative Events and Experiences

In addition to the more general discussion of attitudes, anxieties, and emotions while learning statistics, the interviewees were asked to identify events that best exemplified four categories of affective response: (1) high-intensity-positive, (2) low-intensity-positive, (3) high-intensity-negative, (4) low-intensity-negative. The object of this exercise was to have students appraise a range of experiences (e.g., solving problems or taking tests) in terms of two dimensions of affect: *direction* and *intensity*. Students were assisted in this task by a pictorial aid, the Statistics Affect Grid (see Appendix G), which represented direction and intensity as separate but intersecting axes or affective continua.

Interviewees were asked to provide exemplars for each of the four quadrants of the Statistics Affect Grid corresponding to the general forms of statistics affect noted above. With regard to *high-intensity-positive*, for example, students were prompted to recount one or more of their most positive experiences in the course. This line of questioning was repeated for the remaining three categories of statistics affect. The Statistics Affect Grid was not used to quantify affective states but rather to facilitate the interviewees' understanding of the task. Table 6 summarizes the types of events that students described as they reflected on their affective or emotional reactions to statistics learning and course demands.

Performance on tests and examinations figured prominently in the minds of seven of the nine interviewees when asked about their most positive experiences in the statistics course. Two students indicated that achieving an understanding of the course material elicited *high-intensity-positive* affect, but this learning issue was just as commonly associated with the *low-intensity-positive* quadrant of the Statistics Affect Grid. Indeed, positive affect (i.e., both low- and high-intensity) was often linked to an admixture of three factors, that is, students' evaluations of their understanding, problem-solving skills, and test scores.

Sources of *high-intensity-negative* affect were somewhat more diverse, ranging anywhere from forgetting to hand in an assignment to receiving a poor

Table 6

Summary of Students' Affective Responses to Statistics Learning or Evaluative Demands

Student	*High-Intensity-Positive*	*Low-Intensity-Positive*	*High-Intensity-Negative*	*Low-Intensity-Negative*
B	combination of understanding and good grades: "knew what I was talking about going into the exam"	working along in the course, doing the exercises, and getting answers correct	sitting in class waiting to actually work on the material (was frustrating)	just sitting in the class and getting bored
D	"understanding a question," that is, feeling "happy" and "satisfied"	"doing a question on my own"	not seeing connections in a multi-stage problem (feeling frustrated)	not knowing how to do a homework question
H	"doing well on a mid-term" or preparing for tests	no response	missing information about how or when to use a formula	forgetting some small detail
J	taking the final exam	understanding concepts: "feels pretty good"	preparing for the final exam	disappointment with results on assignments
K	"getting back a good mark on a test"	"when you've done your homework and you think you've done good"	"frustration on my homeworks," that is, not understanding what is being asked	"missing info in the [lecture] notes"
M	"getting back a good mark on an exam"	"understanding things in class"	"going into an exam…sitting there waiting for the exam"	making a mistake on the homework
R	"aced a test"	no response	not understanding some concepts	no response
T	solving an unfamiliar problem (but not necessarily on a test)	no response	forgetting to hand in an assignment or missing a final examination	not going to class or doing any assignments and then going in to write a mid-term test: "there is no intensity"
W	"finding the right answer" on tests and in the homework	"confidence of teacher in students"	"poor test result," which led to reduced self-confidence	interruptions in class

Note. The letters in the *Student* column of this table are abbreviations for the pseudonyms used throughout Chapter V, that is, (B)eth, (D)oug, (H)elen, (J)ohn, (K)athy, (M)ichelle, (R)obert, (T)homas, and (W)anda.

result on a test or an examination. Some common threads include a perceived lack of understanding of course content, encountering difficulties while solving problems or completing assignments, as well as waiting to take a test or to learn new material (i.e., anticipation). Some of these same factors also appeared to contribute to *low-intensity-negative* affect, but with slightly different connotations. Less intense negative responses were more commonly associated with the homework assignments, for example, not knowing how to answer a question, missing a detail, or making minor errors.

Positive and Negative Affect Schedule (PANAS)

As part of the interview-based discussions of statistics affect, students responded to and commented on the PANAS. As was mentioned in Chapter III, Method, the PANAS includes 20 mood descriptors that subjects rate in terms of *extent* or *frequency* of occurrence. Watson et al. (1988) indicate that this set of adjectives can be separated into two 10-item scales dealing with positive and negative affect (see note in Table 7). The PANAS was employed in this study as an alternative instrument for gathering information about participants' affective responsiveness while preparing for statistics tests.

Students were asked first to reflect on the two-day period just prior to a mid-term test and then to indicate the extent to which they had experienced the affective states implied by the 20 adjectives comprising the PANAS. Eight of the nine interviewees rated each descriptor on a 4-point scale: 1 (*definitely do not feel*), 2 (*do not feel*), 3 (*slightly feel*), and 4 (*definitely feel*). One student (Thomas) indicated that, other than feeling slightly "nervous" just prior to an examination, none of the other adjectives really applied to him. Table 7 below shows participants' individual ratings of the 20 PANAS adjectives and the two summative scales (i.e., POSAFECT and NEGAFECT).

Table 7
Students' Ratings of PANAS Adjectives

Descriptor	B	D	H	J	K	M	R	T	W	*M*	*SD*
Enthusiastic	4	1	4	3	1	2	3	-	4	2.75	1.28
Interested	4	2	4	2	2	2	4	-	4	3.00	1.07
Determined	4	4	4	4	4	4	4	-	4	4.00	0.00
Excited	3	3	3	3	1	3	4	-	4	3.00	0.93
Inspired	4	1	3	3	4	3	3	-	3	3.00	0.93
Alert	3	3	3	4	3	4	4	-	4	3.50	0.54
Active	4	1	3	1	3	3	3	-	4	2.75	1.17
Strong	3	1	3	4	3	3	3	-	3	2.88	0.84
Proud	4	2	4	3	3	3	3	-	4	3.25	0.71
Attentive	4	1	4	3	3	3	4	-	4	3.25	1.04
Scared	2	4	1	2	3	3	2	-	1	2.25	1.04
Afraid	2	3	1	1	3	3	1	-	2	2.00	0.93
Upset	2	3	1	2	2	2	2	-	1	1.88	0.64
Distressed	1	2	1	3	4	4	3	-	1	2.38	1.30
Jittery	2	3	1	2	3	2	3	-	1	2.13	0.84
Nervous	2	4	1	2	3	2	4	3	3	2.63	1.06
Ashamed	1	2	1	1	1	1	1	-	1	1.13	0.35
Guilty	1	3	1	2	3	1	1	-	1	1.63	0.92
Irritable	3	2	1	2	4	3	2	-	3	2.50	0.93
Hostile	1	3	1	1	3	2	1	-	1	1.63	0.92
POSAFECT	37	19	35	30	27	30	35	-	38	31.38	6.30
NEGAFECT	17	29	10	18	29	23	20	-	15	20.13	6.64

Note. The numbers in this table represent the actual rating scores that individual students provided for each of the 20 adjectives comprising the PANAS. POSAFECT = Enthusiastic + Interested + Determined + Excited + Inspired + Alert + Active + Strong + Proud + Attentive. NEGAFECT = Scared + Afraid + Upset + Distressed + Jittery + Nervous + Ashamed + Guilty + Irritable + Hostile.

Individual scores for the POSAFECT and the NEGAFECT scales seem consistent with other information extracted from the interview process. For example, Doug indicated that he felt frustrated, and at times angry, when he encountered difficulties solving problems. In general conversation, he evidenced the least positive and the most negative characterizations of learning in the statistics course, and this pattern is borne out in the results of the PANAS. Doug had the lowest overall score on the positive affect component of this instrument and one of the two highest scores on the negative dimension (i.e., 19 and 29, respectively).

Kathy's remarks in the interview suggested that while she felt positive about her ability to learn statistics, she also experienced some apprehensiveness about completing homework assignments and taking tests or examinations. Her scores on the two affect scales are very similar in numerical value, that is, 27 for POSAFECT and 29 for NEGAFECT. Further, several of the interview participants (e.g., Beth, Helen, and Robert) expressed reasonably positive dispositions toward the course and their abilities to learn statistics content. Again, these students' scores on the two subscales of the PANAS resemble response patterns suggested by other components of the interview.

Although the PANAS was not designed explicitly for use in the fields of mathematics or statistics education research, it may prove useful as an alternative method of inquiring into students' mood states while engaged in statistics learning or evaluative tasks. It appears that the PANAS provides indications of students' affective dispositions toward preparing for statistics tests. However, use of the PANAS in this study was exploratory, and additional research is advisable.

Statistics Self-Confidence

The final aspect of the interview dealt with students' feelings of self-confidence while learning and taking tests in their statistics course. Self-confidence appeared to play a major role in defining the nature of students' affective and cognitive experiences. Robert's comments perhaps best illustrate the enigmatic linkage between confidence and anxiety. He recalled feeling anxious at the beginning of the course and yet confident in his ability to handle the mathematical aspects of the content. Although his feelings of anxiety never totally dissipated, his confidence grew in response to successful test performance and positive encouragement from the instructor. Not unlike Robert, Wanda began the statistics course feeling fairly confident but anxious. However, her confidence level plunged when she received a poor test score. Wanda indicated that it was the instructor's confidence in her ability that enabled her to successfully complete the course.

Helen began the statistics course feeling unsure of what to expect and somewhat anxious. Over the span of the course, her self-confidence escalated to the point where she remarked that "It's my only course where I know I'm going to do well…I'm pretty proud." This transformation seemed to be the product of several interacting factors: being able to follow the lecture material, completing exercises and assignments, constructing templates (and related algorithms), as well as performing well on tests. Helen's confidence stemmed partially from her own deliberate actions to establish control over the content. Thomas' comments further illustrate this point. Thomas suggested that, prior to studying for a test, he was no better able to answer questions or solve statistics problems than anyone else in the course. However, he was very confident in his ability to learn the material in preparation for evaluation: "After mid-terms or exams, I know what's going on."

In speaking about the role of affective issues in learning statistics, Kathy estimated that her feelings or emotions accounted for about 90% of how well she performed in a course. She further remarked that her confidence was almost exclusively a function of how well she thought she knew the material. "My knowledge is a very integral part in how confident I feel in my knowing it." In essence, interest in the course content was a key variable motivating her to work on assignments, which in turn enhanced both her knowledge and her feelings of self-confidence. Doug's confidence was also tied to his success in the homework assignments. However, he expressed mixed feelings about his knowledge and his self-confidence when he said "I knew it and I didn't know it."

In slight contrast, Beth was not apprehensive at the outset of the course, although she did express some initial uncertainty because of previous difficulties with high-school mathematics. Her confidence level increased in conjunction with learning the material on an everyday basis, completing assignments, and performing well on tests. Michelle also indicated that she lacked confidence in high-school mathematics, although she seemed quite confident about the statistics course. Later in the interview, Michelle suggested that she had experienced some mathematics anxiety in high school, primarily because she did not know what to expect or how to prepare for the provincial examinations. Her sense of self-confidence in learning statistics seemed to derive from organized study, careful preparation for the tests, and familiarity with algorithmic problem-solving procedures.

In summary, the interviews provided first-hand accounts of students' affective and cognitive experiences while learning and taking tests in introductory statistics courses. Some general observations are warranted here. First, the fact that statistics was a required course for some of the students did not seem to have any significant or lasting impact on their anxiety levels. A number of individuals did experience varying degrees of apprehensiveness about statistics tests and assignments, but the notion of "having" to take statistics did not seem to overly concern any of the interview participants. Second, it has been suggested that aversive experiences in school mathematics may have long-term negative effects on students' future mathematics learning. Several of the interviewees mentioned that they were somewhat nervous at the beginning of the statistics course because of previous difficulties or affective discomfort in high-school mathematics. Perhaps what is most noteworthy here is that these same students went on to develop positive attitudes toward the statistics course and feelings of self-confidence in their own abilities to learn the content.

Third, a number of the students emphasized the importance of developing and rehearsing pattern-recognition skills, problem frames or templates, heuristic strategies, and metacognitive procedures. Although the interviewees did vary in terms of how systematically they acquired these skills and strategies, it was clear that the homework assignments and, in some instances, additional exercises played a significant role in students' learning and self-confidence. Finally, discussions with the interview participants reinforced the importance of investigating the interrelatedness of affective and cognitive factors in accounting for differences in learning experiences and performance outcomes. The value of this style of research was reinforced by students' remarks about their confidence levels in the course and how they were affected not only by their earlier school experiences but also by their own efforts and those of interested instructors. The information collected in the interviews punctuates the intimate linkage and constant interaction among students' perceptions of and attitudes toward the statistics course, willingness to develop problem-solving strategies, motivation to learn, performance outcomes, and self-esteem.

CHAPTER VI

SUMMARY AND CONCLUSIONS

As an emerging field of academic and pedagogic interest, statistics education has taken direction from mathematics education theory and research. Inquiry into mathematics learning and performance outcomes has, at different times, examined cognitive and affective factors, but seldom simultaneously. The primary goal of this study was to explore the individual and combined effects of affective and cognitive–learning variables on statistics course performance. The research questions, as outlined in Chapter II, Theoretical Considerations, were:

1. Is statistics course performance related to self-reported attitudes toward statistics, as defined in terms of students' perceptions of their current statistics course or the usefulness of statistics?
2. Is the level of statistics anxiety, that is, students' self-reported responsiveness to learning and evaluative processes, related to final grade outcome?
3. Are deep-conceptual versus surface-procedural learning strategies (or learning preferences) differentially related to achievement scores?
4. Is the level of self-reported use of metacognitive problem-solving strategies associated with statistics course performance?
5. Are reported levels of cognitive interference, distractibility, or cognitive concern (i.e., worry) related to final grades?
6. In what ways do measures of statistics anxiety, attitudes toward statistics, learning strategies, metacognitive problem-solving procedures, and cognitive interference interact with statistics course performance?

Overview of Research Design and Methods of Analysis

This study used both quantitative and qualitative methods to collect and analyze data pertaining to the above objectives. The quantitative or questionnaire-based data were assembled in two stages: (1) an existing data set (n = 358) derived from "in-class" surveys administered between 1994 and 1996 and (2) a new sample (n = 77) collected in 1999. The second phase of the research design included semi-structured interviews with nine students who had recently completed an introductory statistics course (see Chapter V). A total of 435 Brandon University students responded to a series of background questions and three multi-item instruments: the Statistics Anxiety Scale (SAS), the Attitudes Toward Statistics (ATS) scale, and the Mathematics Information Processing Scale (MIPS).

Internal (i.e., data-splitting) and external cross-validation procedures were used to analyze the quantitative data. The available sample was randomly partitioned into two subgroups for the purposes of data reduction and cross-validation: Subgroup 1 (the derivation sample) and Subgroup 2 (the validation sample). Stepwise and backward methods of multiple regression were used to analyze the data in Subgroup 1, with the aim of identifying the most efficient group of variables predicting statistics course performance. The subsets of scale items selected from each of the three study instruments were independently submitted to principal components analysis to assist in the specification of the final set of predictors (see Table 1):

- three SAS subscales (SASEVAL, SASLEARN, and SASCLASS)
- three ATS subscales (ATSCOURS, ATSFIELD, and ATSLIFE)
- a measure of students' expected final course grade (EXPGRADE)
- four MIPS learning-strategy subscales (SURFMIPS, DEEPMIPS, ACHVMIPS, and PROCMIPS)
- metacognitive problem solving, Item 54b of the MIPS (METAPROB)
- three MIPS cognitive-attentional subscales (DISTRACT, INTRFERE, and FOCUSATT)

This set of 15 regressor variables was included in the calculation of separate predictive equations for each of the two subgroups of the original data, based on double cross-validation and invariance analysis (see Table 5). The weights derived from Subgroups 1 and 2 were also applied to a new sample, Subgroup 3 (n = 77), as an external replication of the regression results. Commonality analysis was used to explore the relative contributions of the selected variables (both individually and in groups) to the prediction of statistics course performance. For this latter purpose, the 15 regressors were categorized as either affective or cognitive–learning factors. The six subscales extracted from the Statistics Anxiety Scale (SAS) and the Attitudes Toward Statistics (ATS) scale, along with a global measure of statistics self-confidence (EXPGRADE), were treated as affective variables. The scale items selected from the MIPS were used to generate eight cognitive or learning-related variables.

Summary of Findings

Quantitative data. It should again be mentioned that the main purpose of this study was to investigate linkages among affective, cognitive–learning, and performance variables. From an analytical standpoint, this process began with the examination of the existing questionnaire-based data set (n = 358). These data were first partitioned into two subgroups in preparation for simple and double cross-validation procedures of analysis. Subgroup 1 was used to identify and combine scale items for inclusion in the final predictive equation, whereas Subgroup 2 was retained to validate the regression results.

In Subgroup 1, the 15 predictors (i.e., seven affective and eight cognitive–learning variables) account for 39.89% (i.e., actual $R^2 = .6316^2 = .3989$) of the variance in students' final grades (see Tables 2 and 5). When the weights corresponding to this regression solution were applied to the data in Subgroup 2, the resultant shrunken $R^2 = .5575^2 = .3108$. Although the actual and shrunken R^2s associated with simple cross-validation affirmed the overall explanatory value of the regressor variables, additional analyses were carried out to evaluate the stability of the results.

The predictive efficiency of the 15 regressors was further examined by means of double (internal) cross-validation. This procedure required the calculation of two additional solutions. Statistics course outcomes were first regressed on all of the predictors using the data in Subgroup 2, and the corresponding weights were then applied to the data in Subgroup 1. The actual and shrunken R^2s are .4245 ($.6515^2$) and .2948 ($.5430^2$), respectively. As a final evaluation of these findings, the weights generated in Subgroups 1 and 2 were used (separately) to predict statistics performance in new data ($n = 77$): .3481 ($.5900^2$) and .2839 ($.5328^2$). The shrunken R^2s associated with this external validation procedure compare quite favorably to the results of (internal) double cross-validation.

In addition to investigating the combined effects of all 15 regressor variables, this study explored the *usefulness* (Darlington, 1968) of affective versus cognitive–learning variables in predicting statistics course performance. Commonality analysis indicated that six predictors make significant ($ps < .05$)

unique contributions to the R^2 (i.e., after partialing out the effects of all other variables): EXPGRADE, FOCUSATT, METAPROB, INTRFERE, SURFMIPS, and PROCMIPS.[25] These variables are listed in descending order of uniqueness or relative contributions to the total explained variance. Of these six predictors, only one, EXPGRADE, is associated with the affective domain, whereas the remainder are classified as cognitive–learning factors. When treated as a single block, the eight cognitive–learning variables increment the R^2 by .1223 after entering all of the affective variables (including EXPGRADE) in the first step of the regression equation (see Tables 3 and 4). Therefore, a number of the cognitive and learning-related variables extracted from the Mathematics Information Processing Scale (MIPS) significantly impact final course grades.

The results of commonality analysis suggest that the affective variables, with the exception of EXPGRADE, account for somewhat less of the variance in students' final course grades than do the cognitive–learning factors. This was borne out by an examination of the relative contributions (to the explanation of the criterion) attributable to the affective variables, both before and after removing the effects of the cognitive–learning predictors. For example, when the order of entry begins with (1) the complete block of eight cognitive–learning variables, followed by (2) the combined effects of the three anxiety and three attitude subscales, and (3) EXPGRADE in the final step, the increment in $R^2 = .2207 + .0501 + .1281$, in order of contribution.

---

[25] These are the results of commonality analysis when each of the predictors is entered in the final step of 15 separate regression solutions.

Of the seven affective regressors, EXPGRADE appears to be the single most powerful predictor of the criterion variable, either alone or in conjunction with other affective or cognitive factors. Consequently, when EXPGRADE is grouped with the statistics anxiety and attitude variables, and entered in the final step of regression analysis (i.e., following entry of the eight cognitive–learning variables), the complete block of affective variables accounts for an additional .1782 of the variance in course performance.

Overall, then, the 15 predictors account for approximately 40% of the total criterion variance in Subgroups 1 and 2 (i.e., actual R^2s of .3989 and .4245, respectively).[26] Second, the weights derived from each of the two subgroups yield comparable shrunken R^2s or squared cross-validity coefficients when they are applied to opposing partitions of the original data (see Table 5): .3108 and .2948. Third, similar findings are noted when the regression coefficients calculated in Subgroups 1 and 2 are used to predict statistics course performance in the new sample ($n = 77$): .3481 and .2839.

Qualitative data. The examination of the interview data provided additional insights into the involvement and the interaction of affective and cognitive factors in statistics learning and performance outcomes. What follows is a brief review of several main points of interest stemming from the interviews.

The information extracted from the interview component of this study illuminated a range of affective responses that students linked fairly directly to

---

[26] These actual R^2s are obtained when regression analysis is allowed to optimize on chance in each of the respective data sets.

predisposing factors such as mathematics background, early experiences learning mathematics, and problem-solving activities. At the most fundamental level, perceptions of weak (entry-level) mathematics skills or poor past performance led some students to question their abilities to learn and demonstrate knowledge of statistics content. In this sense, antecedent learning experiences may shape students' beliefs about self (as a learner), problem solving, and the discipline as a whole. Although attitude structures are generally thought both to condition and to be conditioned by new learning, it is unclear to what extent positive experiences can counteract well-established negative dispositions. This topic warrants additional research, most notably with regard to the evolution of statistical beliefs and attitudes, as well as their interaction with anxiety and performance.

The analysis of interview data also drew attention to the potential impact of affective factors on statistical problem solving. Students identified a range of learning and evaluative situations (or events) with which they associated distinctive affective responses. It is noteworthy that some of the major affective issues currently under study by statistics (and mathematics) education researchers emerged in discussions with introductory statistics students: self-deprecatory beliefs and attitudes, fear of evaluation, self-confidence, meta-affective processes, satisfaction, aesthetic appreciation, and excitement. Allen (1997) similarly states that "Emotion may have either an organizing and focusing effect, or a disruptive and distracting effect" (p. 32). Also, McLeod's (1988, 1989a) conceptualization of emotional states in mathematical problem solving is particularly relevant to the analysis of statistics affect, that is, *magnitude*, *direction*, *duration*, *level of awareness*, and *level of control*.

Although the anxiety and attitude scales used in this study include references to problem solving, the interviews supplied valuable contextual details about students' responses to solving problems. One of the main reasons why empirical studies of statistics (or mathematics) affect and performance have yielded weak or discrepant findings may be that conventional instruments do not address the highly changeable nature of affective states under variable learning or evaluative conditions. Therefore, qualitative, quantitative, and mixed-mode research on statistics affect, in general, and more specifically with regard to problem-solving processes, is important to the advancement of statistics education theory and practice.

A second major theme of (qualitative) inquiry addressed students' awareness and use of specialized strategies for learning statistics content. Many of the interview discussions concerned problem-solving competencies, for example, constructing problem frames, developing pattern-recognition skills, practicing a variety of problem types, instantiating formulas, analyzing errors, and applying heuristics. All of these issues have received extensive attention in mathematics education literature and, based on the accounts provided in the interviews, there is a parallel need to examine statistical problem solving. What remains in question, and should perhaps be the topic of further study, is how or to what extent mathematical problem-solving theory and research apply to statistics learning.

Over the course of the interviews, it became apparent that students differed noticeably in terms of their strategic orientations to statistics course material. Some students placed emphasis on identifying and routinizing algorithms with direct attachments to recognizable problem types, whereas others expressed greater concern about their understanding of the subject matter. In this sense, the students' remarks reinforced the theoretical and practical value of Skemp's (1976) differentiation between *instrumental* and *relational* understanding, particularly with regard to conceptual versus procedural conceptions of learning statistics content.

One final matter of significance to this study involves the linkage between affective and cognitive components of learning. Allen (1997) has suggested that emotion "has a direct influence on cognition, heuristic reasoning, metacognition, and belief systems in problem solving" (pp. 45-46). This interplay between affective and cognitive–learning variables is manifest in both the results of regression analysis and the students' accounts of learning and problem solving in introductory statistics. With regard to the interviews, for example, Helen's edict "just relax and access" aptly conveys the intimate connection between affective states and memory (i.e., storage and retrieval) processes.

In summary, the interviews elicited personalized insights into the nature of students' affective responses to statistics learning, problem-solving, and evaluative contexts. Many of the interviewees' comments shed light on the variability of their affective states (i.e., direction and intensity) in conjunction with differential task demands. It was valuable to hear individuals express *in their* *own language* how they had elected to learn, study, remember, and apply course content. Many of the students remarked on the affective (e.g., satisfaction or frustration) and cognitive (e.g., algorithms and heuristics) fabric of statistical problem solving. This topic of discussion, perhaps more than any other, illustrated the complex interaction among affective and cognitive–learning factors. Hence, the results of analyzing both the questionnaire- and interview-based data reinforced the importance of investigating theoretical and empirical linkages among affective, (meta)cognitive, and statistics performance variables.

Suggestions for Further Research

The final section of the thesis identifies some possible directions for future research on statistics achievement, specifically with regard to person-related, affective, and cognitive factors. Only brief mention is made of background variables, because the primary goal of this study was to explore the individual and combined effects of affective and cognitive–learning variables on grade outcomes.

Person-related factors. Studies that investigate the relationships between "person" factors and statistics achievement have yielded inconclusive results and indeterminate grounds for further inquiry. However, there is some indication in the literature that future research on person-related correlates or predictors of statistics course performance should more carefully examine the mediating effects of affective and cognitive variables. Increasingly, researchers are beginning to analyze the involvement of interacting variables (e.g., attitudes, self-confidence, and content skills) in an effort to understand and explain the impacts of person-related factors.

First, a number of studies have reported positive associations between statistics achievement and mathematics-related variables such as basic arithmetic skills (Gourgey, 1984), years of high-school and college mathematics (Elmore & Vasu, 1986), high-school average for all mathematics courses (Goldstein & High, 1992), mathematics ability (Elmore et al., 1993), and highest level of school mathematics (this study, see Chapter IV). Although these findings support a widely held view that statistics performance is linked to mathematics background variables, prospective research could explore and analyze the essential bases of this relationship. Very little attention has been focused on clarifying whether (and, if so, how) prerequisite knowledge, reasoning, and computational skills affect statistics learning. This line of inquiry could also address the transferability of mathematics learning and problem solving to statistics curricula.

Second, a number of studies have compared the statistics achievement scores of traditional versus non-sequential (or adult) students. Although this line of inquiry has revealed inconsistent findings, it has provided some direction for future research. Sagaria (1989) reports that non-traditional students performed at a significantly higher level than traditional students. The author suggests the need to "understand" the attitudinal, motivational, experiential, and information-processing characteristics of different student populations. The results of this thesis are consistent with Sagaria's study (see Chapter IV). In contrast, Gourgey (1984) finds a negative correlation between age and achievement outcomes, but she too remarks on the role of elemental affective and cognitive factors: negative attitudes, low mathematics self-concept, mathematics anxiety, misconceptions about mathematics (i.e., beliefs), and weak arithmetic skills. Both studies illustrate the importance of examining underlying variables in explaining performance differentials between sequential and non-traditional students.

Third, there is some indication in the literature that statistics performance may be inversely related to the number of years since taking a mathematics course (e.g., Gourgey, 1984). Harvey et al. (1985) report a negative correlation between this particular mathematics background variable and scores on a preliminary statistics test, but they also find a positive association with final examination performance. The findings of this thesis (see Chapter IV) show non-significant performance differences across categories of a corresponding variable (i.e., LASTMATH). Again, it may be useful to carry out analyses of the affective, cognitive, and background factors that may be interacting to produce inconsistent findings.

Finally, gender difference in statistics achievement is one of the most evocative person-related topics of investigation. Gender-based research has reported three types of findings: (1) lower grades for female students as compared to their male counterparts (Feinberg & Halperin, 1978), (2) non-significant performance differences (Bradley & Wygant, 1998; Buck, 1985, 1987; Goldstein & High, 1992; Harvey et al., 1985; Mogull, 1989; Woehlke & Leitner, 1980; and this study, see Chapter IV), or (3) statistically significant results favoring females (Brooks, 1987; Elmore & Vasu, 1986; Schram, 1996). The extent of these discrepancies suggests the involvement of other factors. For example, Schram

(1996)[27] recommends that further research examine the interactions of gender with affective variables (e.g., anxiety and attitudes). On a related point, Elmore and Vasu (1986) remark that "attitudes toward feminism help to explain why, despite their weaker mathematical backgrounds, women were able to perform so well [in statistics]" (p. 221). Given the multitude of factors that may affect statistics achievement, it is important to conduct research that examines the manner in which contributing variables (i.e., affective, cognitive, and skill-based) combine to explain outcomes across diverse populations.

Affective factors. As was noted in Chapter I, Introduction, inquiry into the linkages between affective variables and mathematics learning dates back to the middle of the twentieth century (e.g., Dreger & Aiken, 1957). Academic interest in mathematics affect increased steadily throughout the 1970s and burgeoned in the 1980s, particularly with regard to anxiety and attitude constructs. The discussion of affective issues expanded somewhat in the late 1980s with the consideration of emotions and their potential impact on mathematical problem solving (Mandler, 1989a, 1989b), but has since declined.

Notwithstanding waning interest in mathematics attitudes and anxieties, attention to affective issues in statistics learning and problem solving has intensified. Concern about students' attitudes toward statistics has evolved from a general understanding that affective variables, including beliefs, dispositions, and anxiety levels, may impact learning experiences. This increased focus on statistics

---

[27] Schram's (1996) article is based on a meta-analysis of 13 studies reporting relationships between gender and applied statistics (i.e., methods courses in education, psychology, and business).

affect is evidenced in the proliferation of survey instruments designed to measure attitudes toward statistics (see Schau, Stevens, Dauphinee, & Vecchio, 1995). Statistics anxiety scales, by comparison, are far fewer in number and dimensionally less sophisticated. However, future efforts to refine or construct instruments for use in the affective domain should be more firmly based on conceptual or theoretical foundations.

Research on statistics performance commonly finds negative, though weak, correlations with self-reported measures of statistics anxiety (e.g., Zeidner, 1991). In contrast, positive attitudes toward statistics are typically associated with higher achievement levels (see Harvey et al., 1985; Zimmer & Fuller, 1996). This general pattern of relationships is also reported in the present study, although commonality analysis suggests that anxiety and attitude variables may be more closely related to female as opposed to male students' final course grades (see Chapter IV and Appendix F). However, the results of regression analysis indicate that subscales derived from the Statistics Anxiety Scale and the Attitudes Toward Statistics scale are less efficient predictors of course performance than cognitive or learning-related variables.

It is apparent from the findings of this and other related studies that conventionally used anxiety and attitude instruments make relatively small contributions to the explanation of statistics achievement. The continued and meaningful use of questionnaire-based methods of investigating statistics affect will depend at least partially on researchers' willingness to address a number of substantive issues. First, there is a persistent lack of conceptual and operational clarity in the examination of statistics anxiety and attitude constructs. For example, scales designed to measure attitudes toward statistics commonly include references to anxiety responses such as emotional discomfort, feelings of nervousness, panic responses, or fear.

Second, more attention should be focused on the dimensional structure of multi-item instruments while they are being constructed. More often than not, developers and researchers use factor-analytic procedures to identify and label empirical dimensions rather than theoretical or conceptual definitions. Third, students are seldom involved in the process of scale construction (e.g., wording or phrasing), which may account for interpretive problems and artificial or misleading data. Finally, research on affect seems "to be a collection of generally unrelated clumps of studies on issues like motivation, attitude, and causal attributions" (McLeod, 1992, p. 590).

Although previous studies of statistics affect have tended to use traditional quantitative methods of inquiry and analysis, there is a growing recognition that qualitative techniques can add significantly to the understanding of statistics learning and problem solving. For example, the interview component of this study provided insight into the interplay among students' beliefs, self-attributions, motivational dispositions, and achievement scores. The investigation of affective issues could profit from a combination of both quantitative and qualitative methods of data collection and analysis. This will require both the refinement of qualitative and quantitative methods and the development of relevant theoretical frameworks.

Cognitive factors. There is an extensive body of literature on information-processing models and the potential negative effects of anxiety on cognition (see Chapter II), but there has been comparatively little research on the interrelatedness of cognitive-attentional factors with statistics performance. Cognitive variables have often been limited to measures of mathematics background and arithmetic skills. Preferred study procedures, information storage and retrieval strategies, pattern-recognition skills, and other information-processing factors have rarely been addressed in statistics education research.

Recent theoretical and empirical analyses of the involvement of general study orientations with affective, cognitive, and behavioral dimensions of learning show promise. Biggs (1987), for example, developed the Study Process Questionnaire to identify characteristic differences in student learning approaches: *deep*, *surface*, and *achieving*. However, scales or inventories of this type are often composed of Likert-format items phrased in *content-independent* language and, as such, they may not capture the motivational or cognitive features of learning (*content-specific*) statistics material.

Some researchers have suggested that students' learning motives and strategies are interwoven with attitudinal variables and anxiety proneness. Biggs (1985), for instance, associates the potential for negative affect more so with surface approaches, while deep orientations offer greater possibilities for "affectively satisfying" (p. 187) learning experiences. Marton and Säljö (1984) likewise indicate that intrinsic motivation and low anxiety are associated with a deep approach, whereas extrinsic motivation, perception of threat, and anxiety are linked to a surface approach. Furthermore, Mills (1991) has discussed the involvement of affective processes, state of mind, self-esteem, and enjoyment of learning with metacognitive awareness. He states that "A more intrinsic, self-regulated state of motivation derives from the higher, metacognitive self-as-agent" (Mills, 1991, p. 67). These authors' views represent an important new prospect for research on statistics learning, that is, the integration and combined consideration of cognitive and affective variables in performance studies.

The inclusion of the Mathematics Information Processing Scale (MIPS) in this thesis permitted the investigation of several types of cognitive–learning variables: statistics learning strategies, metacognitive problem solving, and cognitive-attentional deployment during tests or examinations. It is noteworthy that five predictors extracted from the MIPS make significant unique contributions to the prediction of statistics course performance: attentional focus, metacognitive problem solving, cognitive interference, surface-disintegrated learning, and procedural learning. These results affirm the theoretical and empirical value of distinguishing between *relational* versus *instrumental* understanding (Skemp, 1978, 1986) and *deep* versus *surface* study strategies (Biggs, 1985). Further research is needed, both qualitative and quantitative, that explores the nature of students' preferred or habituated learning strategies and the potential linkages to affective states and performance outcomes.

In conclusion, the results of regression and commonality analysis indicate that conventional measures of statistics anxiety and attitudes are weakly related to final course outcomes and, as such, there may be some merit in designing more refined, theoretically guided indicators of affective variables. The study findings also suggest the benefits, both explanatory and predictive, of additional inquiry into differential (i.e., habitual or preferential) modes of learning statistics and cognitive-attentional responses to evaluative conditions. Ware and Chastain (1989) have recommended closer examination of the "variety of skills, attitudes, and motivational forces" (p. 10) that may contribute to introductory statistics performance. Much greater attention could be focused on the interactions among anxieties, attitudes, learning strategies, cognitive-attentional variables, and performance outcomes.

One of the underlying goals of this thesis concerned the furtherance of multiple method (i.e., integrative) research on affective and cognitive components of statistics performance. The research questions, data sources (i.e., quantitative and qualitative), and analytical procedures identified in the present study were selected expressly to investigate the involvement of affective and cognitive–learning factors in statistics course performance. Indeed, the examination of a broader spectrum of cognitive factors derived from learning approach theory, information-processing models of memory, mathematical metacognition, and cognitive-attentional theory represents a novel approach to the investigation of statistics learning. It is hoped that the design characteristics, analytical techniques, and findings of this thesis will lead to additional inquiries into the underlying affective and cognitive factors contributing to statistics learning and performance outcomes.

REFERENCES

Aiken, L. R., (1974). Two scales of attitude toward mathematics. *Journal for Research in Mathematics Education*, *5*, 67-71.

Allen, B. D. (1997). Emotion and its influences on mathematical problem solving. *Dissertation Abstracts International*, *58*, 12A. (University Microfilms No. AAG9818816)

Ang, R. P. (1998). Use of the double cross-validation and bootstrap methods to estimate replicability of results of multiple regression. *Perceptual and Motor Skills*, *86*, 1143-1152.

Barnard, G. A. (1974). Discussion of Professor Stone's paper "Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society, Series B*, *36*, 133-135.

Bendig, A. W., & Hughes, J. B., III. (1954). Student attitude and achievement in a course in introductory statistics. *Journal of Educational Psychology*, *45*, 268-276.

Benjamin, M., McKeachie, W. J., Lin, Y.-G., & Holinger, D. P. (1981). Test anxiety: Deficits in information processing. *Journal of Educational Psychology*, *73*, 816-824.

Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education*, *57*, 247- 261.

Bessant, K. C. (1997). The development and validation of scores on the Mathematics Information Processing Scale (MIPS). *Educational and Psychological Measurement*, *57*, 841-857.

Biggs, J. B. (1985). The role of metalearning in study processes. *British Journal of Educational Psychology*, *55*, 185-212.

Biggs, J. B. (1987). *Study Process Questionnaire: Manual*. Melbourne: Australian Council for Educational Research.

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, *63*, 3-19.

Birenbaum, M., & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational statistics. *Educational Research*, *36*, 93-98.

Blalock, H. M., Jr. (1987). Some general goals in teaching statistics. *Teaching Sociology*, *15*, 164-172.

Bradley, D. R., & Wygant, C. R. (1998). Male and female differences in anxiety about statistics are not reflected in performance. *Psychological Reports*, *82*, 245-246.

Briars, D. J. (1983). An information-processing analysis of mathematical ability. In R. F. Dillon & R. R. Schmeck (Eds.), *Individual differences in cognition* (Vol. 1, pp. 181-204). New York: Academic Press.

Brooks, C. I. (1987). Superiority of women in statistics achievement. *Teaching of Psychology*, *14*, 45.

Brush, L. R. (1978). A validation study of the Mathematics Anxiety Rating Scale (MARS). *Educational and Psychological Measurement, 38*, 485-490.

Buck, J. L. (1985). A failure to find gender differences in statistics achievement. *Teaching of Psychology*, *12*, 100.

Buck, J. L. (1987). More on superiority of women in statistics achievement: A reply to Brooks. *Teaching of Psychology*, *14*, 45-46.

Carter, D. S. (1979). Comparison of different shrinkage formulas in estimating population multiple correlation coefficients. *Educational and Psychological Measurement*, *39*, 261-266.

Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education*, *61*, 287-292.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, *69*, 161-182.

Davis, R. B. (1984). *Learning mathematics: The cognitive science approach to mathematics education*. Norwood, NJ: Ablex Publishing.

Dreger, R. M., & Aiken, L. R., Jr. (1957). The identification of number anxiety in a college population. *Journal of Educational Psychology, 48*, 344-351.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183-201.

Elmore, P. B., Lewis, E. L., & Bay, M. L. G. (1993). *Statistics achievement: A function of attitudes and related experiences*. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction Service No. ED 360 324)

Elmore, P. B., & Vasu, E. S. (1986). A model of statistics achievement using spatial ability, feminist attitudes and mathematics-related variables as predictors. *Educational and Psychological Measurement*, *46*, 215-222.

Entwistle, N. (1988). Motivational factors in students' approaches to learning. In R. R.. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 21-51). New York: Plenum Press.

Feinberg, L. B., & Halperin, S. (1978). Affective and cognitive correlates of course performance in introductory statistics. *Journal of Experimental Education*, *46*, 11-18.

Fennema, E., & Sherman, J. A. (1986). *Fennema-Sherman Mathematics Attitudes Scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males*. Madison, WI: Wisconsin Center for Education Research.

Gal, I. & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Toward an assessment framework. *Journal of Statistics Education* [On-line serial], *2*(2). Available E-mail: archive@jse.stat.ncsu.edu Message: Send jse/v2n2/gal

Garofalo, J., & Lester, F. K., Jr. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education, 16*, 163-176.

Goldstein, J., & High, R. V. (1992). *Identifying cognitive and affective variables as they relate to the successful completion of business statistics*. Paper presented at the Adelphi University Colloquium, Garden City, NY. (ERIC Document Reproduction Service No. ED 350 921)

Gough, Sister M. F. (1954). Mathemaphobia: Causes and treatments. *The Clearing House*, *28*, 290-294.

Gourgey, A. F. (1984). *The relationship of misconceptions about math and mathematical self-concept to math anxiety and statistics performance.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 254 417)

Greeno, J. G. (1980). Analysis of understanding in problem solving. In R. H. Kluwe & H. Spada (Eds.), *Developmental models of thinking* (pp. 199-212). New York, NY: Academic Press, Inc.

Hald, A. (1990). *A history of probability and statistics and their applications before 1750.* New York: John Wiley & Sons, Inc.

Hart, L. E. (1989). Describing the affective domain: Saying what we mean. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 37-45). New York: Springer-Verlag.

Harvey, A. L., Plake, B. S., & Wise, S. L. (1985). *The validity of six beliefs about factors related to statistics achievement.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED 262 965)

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*, 33-46.

Herzberg, P. A. (1969). The parameters of cross-validation. *Psychometrika*, Monograph No. 16, *34*, 1-70.

Huberty, C. J., & Mourad, S. A. (1980). Estimation in multiple correlation/ prediction. *Educational and Psychological Measurement*, *40*, 101-112.

Hudak, M. A., & Anderson, D. E. (1990). Formal operations and learning style predict success in statistics and computer science courses. *Teaching of Psychology*, *17*, 231-234.

Institute for Personality and Ability Testing. (1957). *Handbook for the IPAT Anxiety Scale Questionnaire (Self Analysis Form)*. Champaign, IL: The Institute.

Kenney, P. A., & Silver, E. A. (1997). *Results from the sixth mathematics assessment of the National Assessment of Educational Progress*. Reston, VA: NCTM.

Kirkland, K., & Hollandsworth, J. G., Jr. (1980). Effective test taking: Skills-acquisition versus anxiety-reduction techniques. *Journal of Consulting and Clinical Psychology*, *48*, 431-439.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance-testing system. *Psychological Bulletin*, *85*, 410-416.

Korchin, S. J. (1964). Anxiety and cognition. In C. Scheerer (Ed.), *Cognition: Theory, research, promise* (pp. 58-78). New York: Harper & Row.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: The classification of educational goals. Handbook II: Affective domain.* New York: David McKay.

Kromrey, J. D., & Hines, C. V. (1995). Use of empirical estimates of shrinkage in multiple regression: A caution. *Educational and Psychological Measurement*, *55*, 901-925.

Kulm, G. (1980). Research on mathematics attitude. In R. J. Shumway (Ed.), *Research in mathematics education* (pp. 356-387). Reston, VA: National Council of Teachers of Mathematics.

Kurtz, A. K. (1948). A research test of the Rorschach test. *Personnel Psychology*, *1*, 41-53.

Lester, F. K., Garofalo, J., & Kroll, D. L. (1989). Self-confidence, interest, beliefs, and metacognition: Key influences on problem-solving behavior. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 75-88). New York: Springer-Verlag.

Lester, D., & Hand, S. (1989). Locus of control and mathematics anxiety as predictors of performance in a psychological statistics course. *Perceptual and Motor Skills*, *69*, 504.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*, 975-978.

MacPherson, E. D. (1966). *Some correlates of anxiety in learning programmed mathematics*. Unpublished doctoral dissertation, Washington State University.

MacPherson, E. D., & Rousseau, L. A. (1996). *A nomological network for elementary school mathematics*. (7th ed.). Manuscript submitted for publication, University of Manitoba, Winnipeg, Manitoba, Canada.

Mandler, G. (1989a). Affect and learning: Causes and consequences of emotional interactions. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 3-19). New York: Springer-Verlag.

Mandler, G. (1989b). Affect and learning: Reflections and prospects. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 237-244). New York: Springer-Verlag.

Manitoba Education and Training. (1995). *K-4 mathematics: Manitoba curriculum framework of outcomes and grade 3 standards*. Winnipeg, MB: The Crown in Right of Manitoba.

Manitoba Education and Training. (1996). *Grades 5 to 8 mathematics: Manitoba curriculum framework of outcomes and grade 6 standards*. Winnipeg, MB: The Crown in Right of Manitoba.

Manitoba Education and Training. (1997). *Senior 1 mathematics: Manitoba curriculum framework of outcomes and senior 1 standards*. Winnipeg, MB: The Crown in Right of Manitoba.

Marton, F., & Säljö, R. (1984). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning* (pp. 36-55). Edinburgh: Scottish Academic Press.

May, F. B. (1969). An improved taxonomical instrument for attitude measurement. *College Student Survey*, *2*, 33-38.

McLeod, D. B. (1988). Affective issues in mathematical problem solving: Some theoretical considerations. *Journal for Research in Mathematics Education*, *19*, 134-141.

McLeod, D. B. (1989a). The role of affect in mathematical problem solving. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 20-36). New York: Springer-Verlag.

McLeod, D. B. (1989b). Beliefs, attitudes, and emotions: New views of affect in mathematics education. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 245-258). New York: Springer-Verlag.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). Toronto: Maxwell Macmillan Canada.

Mill, J. S. (1872). *A system of logic: Ratiocinative and inductive.* (8th ed., Vol. II). London: Longmans, Green, Reader, and Dyer.

Mills, R. C. (1991). A new understanding of self: The role of affect, state of mind, self-understanding, and intrinsic motivation. *Journal of Experimental Education*, *60*, 67-81.

Mogull, R. G. (1989). Comparative gender performance in business statistics. *Educational Research Quarterly*, *13*, 2-10.

Mosier, C. I. (1951). Problems and designs of cross-validation. *Educational and Psychological Measurement*, *11*, 5-11.

Mossholder, K. W., Kemery, E. R., Harris, S. G., Armenakis, A. A., & McGrath, R. (1994). Confounding constructs and levels of constructs in affectivity measurement: An empirical investigation. *Educational and psychological Measurement*, *54*, 336-349.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression: A second course in statistics*. Reading, MA: Addison-Wesly Publishing Co.

Mueller, J. H. (1980). Test anxiety and the encoding and retrieval of information. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 63-86). Hillsdale, NJ: Erlbaum.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

Naveh-Benjamin, M. (1991). A comparison of training programs intended for different types of test-anxious students: Further support for an information-processing model. *Journal of Educational Psychology*, *83*, 134-139.

Naveh-Benjamin, M., McKeachie, W. J., & Lin, Y.-G. (1987). Two types of test-anxious students: Support for an information processing model. *Journal of Educational Psychology*, *79*, 131-136.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction*. (2nd ed.). Orlando, FL: Harcourt Brace College, Publishers.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. (3rd ed.). Orlando, FL: Harcourt Brace College Publishers.

Perney, J., & Ravid, R. (1990). *The relationship between attitudes toward statistics, math self-concept, test anxiety and graduate students' achievement in an introductory statistics course*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA. (ERIC Document Reproduction Service No. ED 318 607)

Picard, R. R., & Berk, K. N. (1990). Data splitting. *The American Statistician*, *44*, 140-147.

Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, *79*, 575-583.

Quetelet, L. A. J. (1969). *A treatise on man and the development of his faculties*. (A facsimile reproduction of the English translation of 1842 with introduction by S. Diamond.). Gainsville, FL: Scholars' Facsimiles & Reprints. (Original work published in 1835)

Reece, C. C., & Todd, R. F. (1989). *Math anxiety, attainment of statistical concepts, and expressed preference for a formal-deductive cognitive style among beginning students of research.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED 318 608)

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology, 19*, 551-554.

Richardson, F. C., & Woolfolk, R. L. (1980). Mathematics anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 271-288). Hillsdale, NJ: Erlbaum.

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, *40*, 235-238.

Roberts, D. M., & Reese, C. M. (1987). A comparison of two scales measuring attitudes towards statistics. *Educational and Psychological Measurement*, *47*, 759-764.

Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change.* San Francisco: Jossey-Bass.

Rounds, J. B., Jr., & Hendel, D. D. (1980). Measurement and dimensionality of mathematics anxiety. *Journal of Counseling Psychology*, *27*, 138-149.

Rozenboom, W. W. (1978). Estimation of cross-validated multiple correlation: A clarification. *Psychological Bulletin*, *85*, 1348-1351.

Russell, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology*, *37*, 345-356.

Sagaria, S. D. (1989). *Teaching traditional and non-traditional age individuals: How should methods, expectations, and standards differ?* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 308 363)

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, *46*, 929-938.

Schau, C., Stevens, J., Dauphinee, T. L., & Vecchio, A. D. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, *55*, 868-875.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition. In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189-215). Hillsdale, NJ: Erlbaum.

Schram, C. M. (1996). A meta-analysis of gender differences in applied statistics achievement. *Journal of Educational and Behavioral Statistics*, *21*, 55-70.

Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Education*, *61*, 293-316.

Shulte, A. P., & Smart, J. R. (1981). Preface. In *Teaching Statistics and Probability* (pp. ix-x). 1981 Yearbook. Reston, VA: National Council of Teachers of Mathematics.

Silver, E. A. (1987). Foundations of cognitive theory and research for mathematics problem-solving instruction. In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 33-60). Hillsdale, NJ: Erlbaum.

Skemp, R. R. (1978). Relational and instrumental understanding. *The Arithmetic Teacher*, *26*, 9-15.

Skemp, R. R. (1986). *The psychology of learning mathematics*. (2nd ed.). New York: Penguin Books.

Speth, C., & Brown, R. (1988). Study approaches, processes and strategies: Are three perspectives better than one? *British Journal of Educational Psychology*, *58*, 247-257.

SPSS, Inc. (1993). *SPSS base system syntax reference guide: Release 6.0*. Chicago, IL: SPSS Inc.

Stevens, J. (1986). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge, MA: Belknap Press.

Suinn, R. M., Edie, C. A., Nicoletti, J., & Spinelli, P. R. (1972). The MARS, a measure of mathematics anxiety: Psychometric data. *Journal of Clinical Psychology*, *28*, 373-375.

Thompson, B. (Ed.). (1993). Statistical Significance Testing in Contemporary Practice: Some Proposed Alternatives with Comments from Journal Editors [Special issue]. *Journal of Experimental Education*, *61*(4).

Thompson, B. (1994). The pivotal role of replication in psychological research: Empirically evaluating the replicability of sample results. *Journal of Personality*, *62*, 157-176.

Tobias, S. (1979). Anxiety research in educational psychology. *Journal of Educational Psychology*, *71*, 573-582.

Tobias, S. (1985). Test anxiety: Interference, defective skills, and cognitive capacity. *Educational Psychologist*, *20*, 135-142.

Uhl, N., & Eisenberg, T. (1970). Predicting shrinkage in the multiple correlation coefficient. *Educational and Psychological Measurement*, *30*, 487-489.

Wachtel, P. L. (1967). Conceptions of broad and narrow attention. *Psychological Bulletin*, *68*, 417-429.

Ware, M. E., & Chastain, J. D. (1989). *Person variables contributing to success in introductory statistics*. Paper presented at the Annual Meeting of the Southwestern Psychological Association, Houston, TX. (ERIC Document Reproduction Service No. ED 309 927)

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.

Wertheimer, M. (1959). *Productive thinking*. New York: Harper & Brothers.

Wherry, R. J., Sr. (1975). Underprediction from overfitting: 45 years of shrinkage. *Personnel Psychology, 28*, 1-18.

Wine, J. D. (1980). Cognitive-attentional theory of test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 349-385). Hillsdale, NJ: Erlbaum.

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement, 45*, 401-405.

Woehlke, P. L. (1991). *An examination of the factor structure of Wise's "Attitude Toward Statistics" scale*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED 337 500)

Woehlke, P. L., & Leitner, D. W. (1980). Gender differences in performance variables related to achievement in graduate-level educational statistics. *Psychological Reports, 47*, 1119-1125.

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology, 61*, 319-328.

Zimmer, J. C., & Fuller, D. K. (1996). *Factors affecting undergraduate performance in statistics: A review of literature*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Tuscaloosa, AL. (ERIC Document Reproduction Service No. ED 406 424)

Appendix A

Fennema-Sherman Mathematics Attitudes Scales: Instruments Designed to Measure Attitudes Toward the Learning of Mathematics by Females and Males

Appendix A

Fennema-Sherman Mathematics Attitudes Scales: Instruments Designed to Measure Attitudes Toward the Learning of Mathematics by Females and Males

Fennema and Sherman (1986) developed "nine, domain specific, Likert-type scales measuring important attitudes related to mathematics learning" (p. i); of these, two are presented below:

Confidence in Learning Mathematics Scale

1. Generally I have felt secure about attempting mathematics.
2. I am sure I could do advanced work in mathematics.
3. I am sure that I can learn mathematics.
4. I think I could handle more difficult mathematics.
5. I can get good grades in mathematics.
6. I have a lot of self-confidence when it comes to math.
7. I'm no good in math.
8. I don't think I could do advanced mathematics.
9. I'm not the type to do well in math.
10. For some reason even though I study, math seems unusually hard for me.
11. Most subjects I can handle O.K., but I have a knack for flubbing up math.
12. Math has been my worst subject.

Effectance Motivation in Mathematics Scale

1. I like math puzzles.
2. Mathematics is enjoyable and stimulating to me.
3. When a math problem arises that I can't immediately solve, I stick with it until I have the solution.
4. Once I start trying to work on a math puzzle, I find it hard to stop.
5. When a question is left unanswered in math class, I continue to think about it afterward.
6. I am challenged by math problems I can't understand immediately.
7. Figuring out mathematical problems does not appeal to me.
8. The challenge of math problems does not appeal to me.
9. Math puzzles are boring.
10. I don't understand how some people can spend so much time on math and seem to enjoy it.
11. I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself.
12. I do as little work in math as possible.

Appendix B

Principal Components Analysis, with Varimax Rotation, of Predictor Variables Selected from the SAS, the ATS, and the MIPS

Appendix B

Principal Components Analysis, with Varimax Rotation, of Predictor Variables Selected from the SAS, the ATS, and the MIPS

Table B1

Rotated Factor Coefficients for Selected SAS Items

Factor Labels/Scale Items	I	II	III
Factor I: Statistics Evaluation Anxiety (SASEVAL)			
11. Thinking about an upcoming test one hour before.	.872	.147	.095
22. Taking a final examination in a statistics course.	.871	.050	-.090
31. Thinking about an upcoming test five minutes before.	.844	.019	.137
4. Studying for a statistics test.	.688	.356	.146
16. Being given a "pop" quiz in a statistics course.	.656	.142	-.508
42. Waiting to have a statistics test returned.	.625	.385	.075
33. Receiving your final statistics grade in the mail.	.613	.254	-.062
13. Not knowing the formula needed to solve a problem.	.593	.083	-.481
15. Being given a homework assignment of many difficult problems.	.552	.258	-.459
32. Being asked to explain how you arrived at a solution.	.469	.343	-.060
Factor II: Statistics Learning Anxiety (SASLEARN)			
23. Reading about a statistics formula.	.176	.797	.039
25. Looking through the pages of a statistics text.	.146	.706	-.009
8. Starting a new chapter in a statistics book.	-.048	.675	.256
35. Adding up the results of a survey or poll.	.139	.655	.294
41. Being told how to interpret probability statements.	.219	.650	.050
17. Listening to another student explain a statistics formula.	.202	.640	-.042
3. Opening a statistics book and seeing a page of problems.	.348	.586	-.118
19. Getting ready to study for a statistics test.[a]	.426	.533	.098
18. Solving a problem such as: If x=12, and y=4, then the ratio of x to y is equal to ___?	.131	.522	.288
Factor III: Statistics Course Anxiety (SASCLASS)			
29. Walking to a statistics class.	.111	.475	.748
27. Walking into a statistics class.	.093	.548	.700

Note. $N = 178$. Permission for this scale from R. M. Suinn. Any reproduction is prohibited without written permission from: R. M. Suinn, Department of Psychology, Colorado State University, Ft. Collins, CO 80523. The wording of some SAS items has been abridged.

[a] Item 19 of the SAS was included in the additive scale SASEVAL, as part of the process of maximizing the R^2 in the derivation sample.

Table B2

Rotated Factor Coefficients for Selected ATS Items

Factor Labels/Scale Items	I	II	III
Factor I: Attitude Toward the Course (ATSCOURS)			
2. The thought of being enrolled in a statistics course makes me nervous.	.896	-.031	.000
15. I get upset at the thought of enrolling in another statistics course.	.872	.095	.137
18. I feel intimidated when I have to deal with mathematical formulas.	.816	-.027	-.012
7. I see being enrolled in a statistics course as a very unpleasant experience.	.807	.100	.125
Factor II: Attitude Toward the Field (ATSFIELD)			
21. My statistical training will help me better understand the research being done in my field of study.	-.042	.766	.257
6. I have difficulty seeing how statistics relates to my field of study.	.053	.755	.200
1. I feel that statistics will be useful to me in my profession.	.058	.694	.359
13. Statistics is a worthwhile part of my professional training.	.177	.690	.232
3. A good researcher must have training in statistics.	-.004	.617	-.241
22. One becomes a more effective "consumer" of research findings if one has some training in statistics.	-.024	.572	.226
Factor III: Attitude Toward Statistics in Everyday Life (ATSLIFE)			
24. Statistical thinking can play a useful role in everyday life.	.084	.141	.761
26. I feel that statistics should be required early in one's professional training	-.084	.252	.670
5. Most people would benefit from taking a statistics course.	.255	.198	.649

Note. N = 179. Permission from S. L. Wise, Department of Educational Psychology, University of Nebraska, Lincoln, NE 68588-0345.

Table B3

Rotated Factor Coefficients for Combined Analysis of SAS and ATS Items

Scale Items	I	II	III	IV	V
SAS22	.842	.006	-.139	-.025	.091
SAS11	.815	.196	-.089	-.066	.140
SAS31	.774	.129	.034	-.061	.264
SAS16	.743	-.169	-.311	.041	-.079
SAS13	.693	-.160	-.123	.146	.023
SAS15	.659	-.028	-.262	.166	-.205
SAS33	.650	.206	-.041	-.012	-.037
SAS42	.643	.402	-.017	-.072	.048
SAS4	.640	.382	-.201	-.010	-.058
SAS32	.505	.279	-.088	.107	.023
SAS27	-.006	.849	.091	-.106	.133
SAS29	-.015	.818	.113	-.093	.137
SAS35	.128	.700	-.181	.036	-.200
SAS8	-.037	.693	-.132	-.036	.055
SAS23	.212	.659	-.385	.075	-.153
SAS18	.122	.603	-.070	.103	-.246
SAS41	.264	.554	-.240	-.171	.021
SAS25	.173	.552	-.372	.002	.195
SAS19	.444	.500	-.151	-.026	-.070
SAS17	.250	.494	-.305	.178	-.070
SAS3	.377	.395	-.391	.058	.141
ATS15	-.099	-.095	.837	.241	-.043
ATS2	-.263	-.245	.785	.006	.013
ATS7	-.112	-.154	.761	.153	.086
ATS18	-.283	-.195	.720	.013	-.001
ATS1	.045	-.017	.049	.807	.183
ATS6	.095	-.076	.014	.770	.188
ATS13	-.047	-.044	.118	.764	.155
ATS21	.083	.016	-.063	.609	.515
ATS5	-.013	.067	.332	.509	.138
ATS22	-.003	-.078	-.092	.288	.689
ATS24	.049	.060	.247	.218	.550
ATS26	.004	.258	.041	.299	.537
ATS3	.069	-.158	-.054	.100	.525

Note. $N = 178$.

Table B4
Rotated Factor Coefficients for Selected MIPS Items: Statistics Learning Strategies

Factor Labels/Scale Items	I	II	III	IV
Factor I: Surface-Disintegrated Learning (SURFMIPS)				
34. I feel "dumb" in mathematics or statistics...	.763	.056	.039	.103
26. I do not grasp mathematics easily.	.758	.136	-.005	.082
35. I am unsure what test questions mean or what they are asking me to do.	.719	.062	-.025	.055
13. I find myself going over and over problems without making much progress.	.606	-.033	-.060	.118
36. I find it difficult to go through all of the steps that are needed to solve problems.	.603	-.157	-.101	.133
31. I rely heavily on my instructors to show me how to perform statistical procedures.	.559	.056	.115	-.192
22. I prefer that instructors explain examples step by step.	.484	.165	.308	-.161
23. I am unable to understand material just presented, before the instructor moves on.	.447	.046	-.068	.228
33. It is difficult for me to change the way that I perform calculations or think about problems.	.443	-.169	-.001	.223
52. I would perform better on tests/exams, if I were allowed to write outside the classroom.	.360	-.004	.026	.058
Factor II: Deep-Conceptual Learning (DEEPMIPS)				
3. I create diagrams, pictures and/or flow charts to improve my understanding.	-.001	.657	.030	-.074
43. I keep track of the types of errors I make when I perform various calculations.	.064	.619	.120	-.219
44. I write out the meanings of statistical formulas in my own words.	.151	.614	-.274	-.101
46. I study by clustering text/lecture material into "chunks": e.g., topic, formula...	.095	.611	.200	.094
15. I organize my study according to general subject areas, sub-categories, etc.	.030	.604	.348	.051
47. I rehearse or repeat problems until they become routine.	.018	.589	.320	.127
48. I prepare for tests by looking for associations and relationships between ideas.	-.211	.575	.103	.100
5. I take more time to complete assignments and practice exercises than most students.	.208	.484	.019	.051

Table B4 Continued

Rotated Factor Coefficients for Selected MIPS Items: Statistics Learning Strategies

Factor Labels/Scale Items	I	II	III	IV
Factor II: Deep-Conceptual Learning (DEEPMIPS)				
1. I practice many different types of problems (or examples) as a routine part of study.	-.077	.478	.228	.117
49. I look for key words or phrases which will help me remember how to solve problems.	-.052	.459	.390	.057
8. I do not move on to a new section until I understand the central explanations...	-.147	.398	.371	.079
Factor III: Strategic Learning (ACHVMIPS)				
9. If I learn how to solve a problem on my own, I am less apt to forget it.	.022	.108	.673	.038
11. If I have difficulty solving a problem, I stop and rethink my procedure.	-.015	.224	.625	-.102
12. I do not feel comfortable with my learning until I know how and when to apply a formula.	.218	.130	.599	-.153
2. I memorize formulas or computational steps when I prepare for tests.	-.045	.308	.527	-.070
37. I do most of my studying a few days before the test.	-.060	-.127	.513	.095
28. If I spend a long time studying something, I feel that I should know it for the test.	.198	.113	.510	-.151
24. I like working out problems without assistance from others.	-.155	.223	.470	.125
Factor IV: Procedural Learning (PROCMIPS)				
17. I do not care if I know what a statistics is used for, as long as I can perform calculations.	.089	-.032	.026	.792
16. I do not understand the real object or purpose behind statistical/mathematical procedures.	.337	.226	-.044	.680
25. I am unsure of the central message(s) in lectures and assignments.	.318	.071	-.048	.585

Note. $N = 179$. Copyright 1997 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc. The wording of some MIPS items has been abridged for parsimony of presentation.

Table B5

Rotated Factor Coefficients for Selected MIPS Items: Cognitive-Attentional Focus

Factor Labels/Scale Items	I	II	III
Factor I: Distractibility (DISTRACT)			
54r. I get too agitated to analyze all aspects of the question.	.773	.077	-.183
54f. I have difficulty determining what information in the question or problem is crucial to the solution.	.665	.358	.092
54t. I focus a lot of attention on cues which I later find out are irrelevant to the solution.	.631	.169	-.062
54x. I get caught up examining all of the details in problems.	.584	.252	.055
54l. I have difficulty focusing my attention on the details of each problem or question.	.526	.329	-.190
54b. I either know how to do a problem or I do not.	.497	.084	.052
Factor II: Cognitive Interference (INTRFERE)[a]			
54e. I think about what sort of mark I am going to receive.	.125	.730	-.234
54h. I am preoccupied with what other will think of my performance.	.158	.573	-.315
54i. I spend a lot of time waiting for the answer to come to me.	.478	.566	-.089
54o. I rely on immediate recall and sight recognition to answer questions.	.142	.513	.024
54d. I catch myself thinking about things other than the actual test questions.	.190	.513	-.033
54w. I experience "mental blocks."	.498	.512	-.142
54y. I recall similar practice problems and reproduce steps.	-.423	.428	.374
Factor III: Focused Attention (FOCUSATT)			
54p. I shut out any doubts about my performance and focus on completing questions.	-.076	-.309	.706
54q. I restrict my analysis to what I think is the most important information.	.076	-.171	.675
54a. I examine all aspects of each questions before beginning an answer.	.011	-.011	.662
54k. I can control how excited/aroused I get when I encounter a difficult problem.	-.390	.079	.447

Note. $N = 179$. Copyright 1997 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc. The wording of some MIPS items has been abridged.

[a] Items 54h and 54y of the MIPS were included in the additive scale FOCUSATT based on the examination of alternate factor solutions and the optimization of the predictive equation.

Appendix C

Multiple Regression Analysis of Variables Predicting Statistics Course Performance in the Derivation and Validation Samples

Appendix C

Multiple Regression Analysis of Variables Predicting Statistics Course Performance in the Derivation and Validation Samples

Table C1

Summary of Multiple Regression Analysis (Enter Method) for Variables Predicting Statistics Course Performance in the Derivation Sample

Predictor	*R*	R^2	R^2 Change	*B*	*SE B*	β	*F* Change
1. SASEVAL	.1851	.0343	.0343	-0.033	0.028	-0.098	6.14**
2. SASLEARN	.1953	.0381	.0039	0.052	0.061	0.089	0.69
3. SASCLASS	.2024	.0410	.0029	0.119	0.181	0.055	0.51
4. ATSCOURS	.2401	.0577	.0167	-0.128	0.080	-0.163	3.01*
5. ATSFIELD	.2654	.0704	.0128	0.062	0.070	0.073	2.32
6. ATSLIFE	.2918	.0851	.0147	-0.072	0.138	-0.041	2.70
7. SURFMIPS	.3346	.1120	.0268	-0.091	0.045	-0.203	5.05**
8. DEEPMIPS	.4075	.1660	.0541	0.025	0.032	0.064	10.76***
9. ACHVMIPS	.4091	.1674	.0014	0.007	0.050	0.011	0.27
10. PROCMIPS	.4377	.1916	.0242	0.221	0.112	0.158	4.90**
11. METAPROB	.4567	.2086	.0171	0.502	0.224	0.170	3.51*
12. DISTRACT	.4612	.2127	.0040	-0.042	0.075	-0.055	0.83
13. INTRFERE	.4729	.2236	.0109	-0.147	0.069	-0.188	2.27
14. FOCUSATT	.5204	.2708	.0472	-0.180	0.065	-0.203	10.35***
15. EXPGRADE	.6316	.3989	.1281	0.666	0.114	0.430	33.90***
(CONSTANT)				7.080	2.862		

Note. $N = 175$. SASEVAL = Statistics Evaluation Anxiety, SASLEARN = Statistics Learning Anxiety, SASCLASS = Statistics Course Anxiety, ATSCOURS = Attitude Toward Statistics Course, ATSFIELD = Attitude Toward Field of Statistics, ATSLIFE = Attitude Toward Statistics in Everyday Life, SURFMIPS = Surface-Disintegrated Learning, DEEPMIPS = Deep-Conceptual Learning, ACHVMIPS = Strategic Learning, PROCMIPS = Procedural Learning, METAPROB = Metacognitive Problem Solving, DISTRACT = Distractibility, INTRFERE = Cognitive Interference, FOCUSATT = Focused Attention, and EXPGRADE = Expected Final Course Grade.
* $p < .10$. ** $p < .05$. *** $p < .01$.

Table C2

Summary of Multiple Regression Analysis (Enter Method) for Variables Predicting Statistics Course Performance in the Validation Sample

Predictor	*R*	R^2	R^2 Change	*B*	*SE B*	β	*F* Change
1. SASEVAL	.1866	.0348	.0348	-0.010	0.027	-0.031	6.03**
2. SASLEARN	.1878	.0353	.0005	-0.089	0.051	-0.179	0.08
3. SASCLASS	.2960	.0876	.0523	0.431	0.184	0.204	9.46***
4. ATSCOURS	.3040	.0924	.0048	-0.043	0.069	-0.057	0.87
5. ATSFIELD	.3345	.1119	.0195	0.222	0.064	0.261	3.57*
6. ATSLIFE	.3750	.1406	.0287	-0.400	0.137	-0.220	5.42**
7. SURFMIPS	.4212	.1774	.0368	-0.041	0.048	-0.092	7.21***
8. DEEPMIPS	.4265	.1819	.0045	0.083	0.030	0.215	0.87
9. ACHVMIPS	.4662	.2173	.0354	-0.147	0.055	-0.204	7.20***
10. PROCMIPS	.5029	.2529	.0356	0.271	0.116	0.194	7.52***
11. METAPROB	.5039	.2539	.0010	-0.034	0.208	-0.012	0.21
12. DISTRACT	.5041	.2542	.0003	-0.036	0.066	-0.053	0.05
13. INTRFERE	.5058	.2559	.0017	-0.087	0.062	-0.122	0.36
14. FOCUSATT	.5433	.2951	.0393	-0.168	0.064	-0.183	8.58***
15. EXPGRADE	.6515	.4245	.1293	0.669	0.114	0.455	34.38***
(CONSTANT)				7.210	2.593		

Note. $N = 169$.
* $p < .10$. ** $p < .05$. *** $p < .01$.

Appendix D

Comparisons of Indicators of Shrinkage

Appendix D

Comparisons of Indicators of Shrinkage

Herzberg (1969) has noted that "In applications, the population regression function can never be known and one is more interested in how effective the *sample* regression function is in *other* samples. A measure of this effectiveness is r_c, the sample cross-validity" (p. 4). Although cross-validity coefficients vary across samples, their *average* value will be approximately equal to the *population* cross-validity coefficient (ρ_c). As an alternative to carrying out internal or external cross-validation research to evaluate the predictive efficiency of multiple regression coefficients, a large number of formulas have been developed to estimate the shrinkage in R^2 when cross-validating to another sample.[28] Herzberg's (1969) formula for ρ_c^2, which he attributes to Darlington (1968), has been characterized as yielding "a more severe and realistic estimate of how much prediction power is lost" under cross-validation (Stevens, 1986, p. 80).

Table D1 presents the squared coefficients of cross-validity *for the regression weights* generated in Subgroups 1 and 2, as well as two estimates of ρ_c^2 based on calculations suggested by Darlington (1968) and Rozenboom (1978). Formulas proffered by Darlington and Rozenboom represent "estimators for the true squared validity coefficient, ρ_v^2, which is approximately the average, over many samples, of the square of the sample cross-validity coefficient" (Huberty & Mourad, 1980, p. 105). When Darlington's (1968) expression is applied to the

[28] Huberty and Mourad (1980) compare shrinkage formulas estimating ρ_c^2, whereas Carter (1979) and Uhl and Eisenberg (1970) discuss parallel formulas for estimating ρ^2.

actual R^2s for Subgroups 1 and 2 (i.e., .3989 and .4245), the resulting estimates of shrinkage are .2756 and .3016, respectively. Very similar values are generated with Rozenboom's (1978) formula: .2862 and .3124. The estimated values of ρ_c^2 using both formulas compare quite favorably to the "shrunken" R^2s for weights derived from Subgroups 1 and 2, that is, .3108 ($.5575^2$) and .2948 ($.5430^2$).

Table D1

Comparisons of Indicators of Shrinkage: Squared Cross-Validity Coefficients Versus Formula-Based Estimates of ρ_c^2

	SUBGROUP 1	SUBGROUP 2
"Shrunken" R^2 (R_{cv}^2)	.3108	.2948
Darlington's Est (ρ_c^2)[a]	.2756	.3016
Rozenboom's Est (ρ_c^2)[b]	.2862	.3124

[a] $\text{Est}(\rho_c^2) = 1 - \left(\frac{N-1}{N-k-1}\right)\left(\frac{N-2}{N-k-2}\right)\left(\frac{N+1}{N}\right)(1-R^2)$

[b] $\text{Est}(\rho_c^2) = 1 - \left(\frac{N+k}{N-k}\right)(1-R^2)$

Appendix E

Means, Standard Deviations, and F Values for Person-Related Between-Group Differences in Statistics Course Performance

Appendix E

Means, Standard Deviations, and F Values for Person-Related Between-Group Differences in Statistics Course Performance

	n	*Mean*	*SD*	*F*	*p*
I. Total sample	435	5.27	2.87	-	-
II. Highest level of school math[a]				5.97	.001
Math 200 or lower	27	4.41	3.10		
Math 301 or applied math	47	4.51	2.67		
Math 300 or equivalent	279	5.14	2.88		
Advanced math courses	61	6.49	2.50		
III. Years since formal schooling				5.48	.001
Within the past year	321	5.35	2.79		
One to two years	51	4.08	3.08		
More than two, but not five	33	5.12	3.20		
Five or more years	30	6.63	2.14		
IV. Years since last math class				0.57	.633
Within the past year	220	5.30	2.95		
One to two years	99	4.98	2.89		
More than two, but not five	65	5.55	2.72		
Five or more years	51	5.35	2.67		
V. Gender				0.99	.321
Female	247	5.39	2.80		
Male	188	5.12	2.95		

Note. The sample size for comparisons of statistics achievement by highest level of school mathematics differs from the total ($N = 435$) because of missing values.
[a] $N = 414$.

Appendix F

Summary of Multiple Regression Analysis (Enter Method) for Variables Predicting Statistics Course Performance among Female and Male Students

Appendix F

Summary of Multiple Regression Analysis (Enter Method) for Variables Predicting Statistics Course Performance among Female and Male Students

	Female[a]				Male[b]			
Predictor	R	R^2	R^2 Change	F Change	R	R^2	R^2 Change	F Change
1. SASEVAL	.1670	.0279	.0279	6.83***	.2835	.0803	.0803	15.64***
2. SASLEARN	.1750	.0306	.0028	0.67	.2901	.0842	.0038	0.74
3. SASCLASS	.2417	.0584	.0278	6.97***	.3052	.0932	.0090	1.76
4. ATSCOURS	.2512	.0631	.0047	1.18	.3541	.1254	.0322	6.49**
5. ATSFIELD	.2917	.0851	.0219	5.61**	.3647	.1330	.0076	1.54
6. ATSLIFE	.3371	.1137	.0286	7.52***	.3682	.1356	.0026	0.52
7. SURFMIPS	.3974	.1579	.0442	12.19***	.4065	.1653	.0297	6.15**
8. DEEPMIPS	.4019	.1616	.0037	1.01	.4457	.1986	.0334	7.16***
9. ACHVMIPS	.4082	.1667	.0051	1.40	.4480	.2007	.0021	0.44
10. PROCMIPS	.4627	.2141	.0475	13.84***	.4822	.2325	.0318	7.05***
11. METAPROB	.4694	.2203	.0062	1.81	.4966	.2466	.0141	3.15*
12. DISTRACT	.4700	.2209	.0006	0.18	.4998	.2498	.0032	0.71
13. INTRFERE	.4728	.2235	.0026	0.75	.5322	.2833	.0335	7.81***
14. FOCUSATT	.5012	.2512	.0277	8.32***	.5847	.3418	.0585	14.76***
15. EXPGRADE	.6331	.4008	.1496	55.94***	.6784	.4602	.1184	36.18***

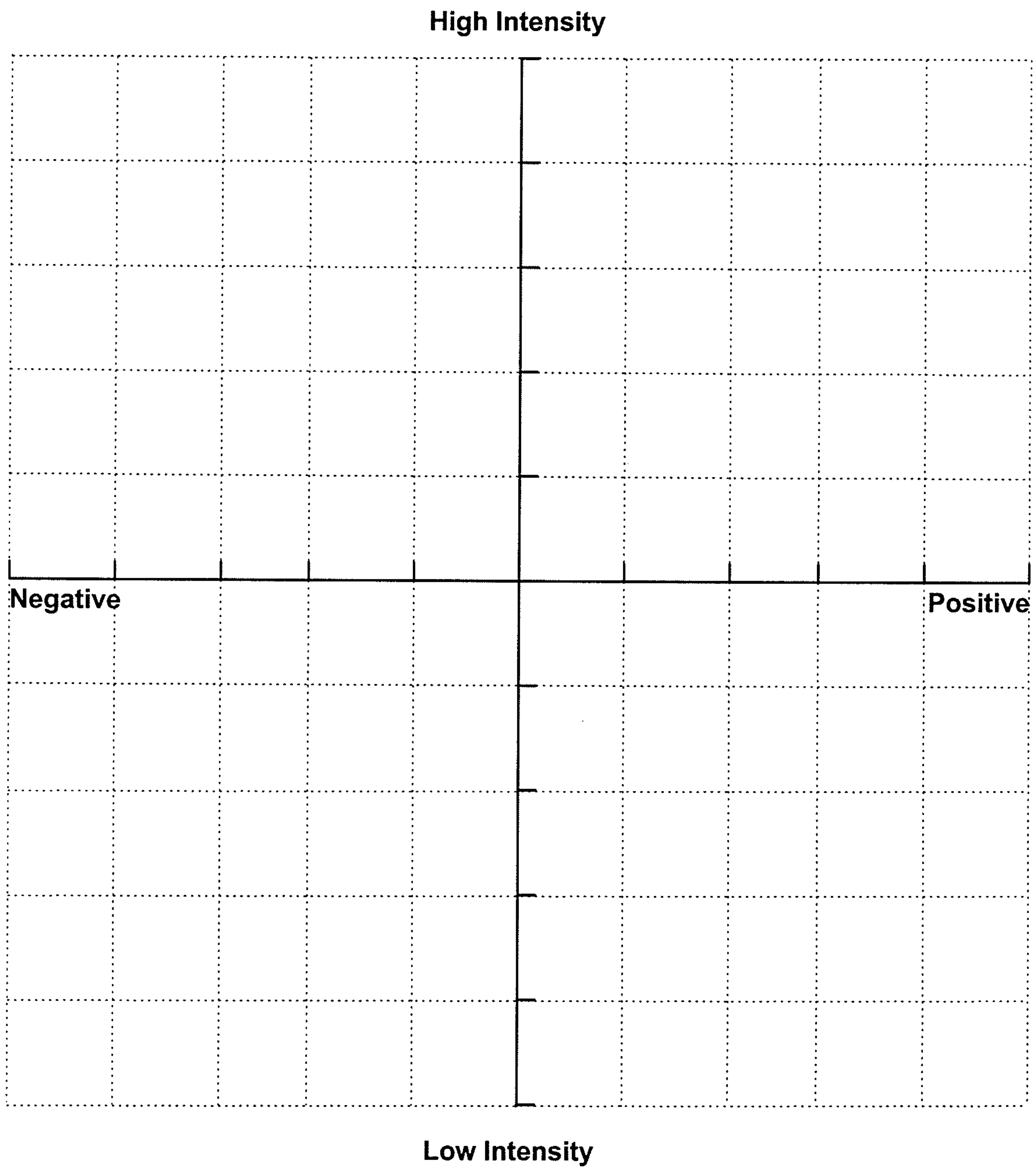
[a]$N = 240$. [b]$N = 181$.
* $p < .10$. ** $p < .05$. *** $p < .01$.

Appendix G

Statistics Affect Grid

Appendix G

Statistics Affect Grid



Appendix H

Exemplary Interview Questions and Prompts

Appendix H

Exemplary Interview Questions and Prompts[29]

1. Why did you take an introductory statistics course?
    - Was the statistics course a degree requirement or an elective?
    - Did it concern you in any way that the statistics course was required?
    - Did the fact that you were required to take statistics affect the way you approached the course?
    - In what way is learning statistics connected to your degree program?

2. Do you recall covering anything in the high-school curriculum that was related to your introductory statistics course?
    - What aspects of the statistics course were most familiar to you?
    - Were you exposed to statistics content in high-school mathematics?

3. Did you use or develop any particular techniques for learning statistics content?
    - How did the homework assignments relate to (or assist) your learning?
    - Did you rehearse your learning in this course (e.g., practice exercises)?
    - What did you do if you encountered a problem you did not recognize?
    - Did your learning in this course differ from that of other subjects?
    - What strategies did you use to solve problems?
    - How did you go about learning the course material?

4. Did your attitudes, feelings, or emotions play a role in your learning or performance in the statistics course?
    - How would you describe your general attitude toward the course?
    - How would you characterize yourself with regard to anxiety about the course content or tests/examinations?
    - How did you feel if you encountered a problem you did not recognize?
    - What would make you react (i.e., become anxious) in the statistics course? Did any aspects of the course make you feel anxious?
    - What was the most emotionally intense aspect of this course?
    - Did your attitudes toward statistics change at all during the course?

---

[29] These queries are paraphrased (as opposed to verbatim) examples of open-ended questions used to introduce and focus the interviews. Thematic questions (e.g., those identified with numbers or letters) were generally asked of all students and, depending on their responses, one or more supplementary prompts (e.g., bulleted exemplars) were then used to extend discussions.

5. Discussion of the Statistics Affect Grid:
    a) *High-Intensity-Positive*—What was the most positive event or experience of the statistics course?
    b) *High-Intensity-Negative*—Can you think of anything of similar intensity (i.e., high) that you would describe as a negative experience? What event in the course made you feel most anxious (or most negative)?
    c) *Low-Intensity-Negative*—Can you think of anything you would describe as negative but lower in intensity, that is, something that made you feel negative but not as negative as (referring back to the answer for high-intensity-negative)?
    d) *Low-Intensity-Positive*—Was there anything in the course that made you feel positive but not as positive as (referring back to the answer for high-intensity-positive)?

6. The Positive and Negative Affect Schedule (PANAS):
    a) First, interviewees were asked to reflect on the two-day period just prior to the mid-term test.
    b) Second, students were instructed to consider and then rate the 20 "mood descriptors" included in the PANAS in terms of how well each described their feelings or emotional states during the specified time frame (see Chapter III for the list of PANAS adjectives).
    c) Third, if students made remarks or comments during this process, I asked for clarification.

7. What role did your self-confidence play in learning statistics?
    - How confident were you in your abilities to learn statistics content?
    - How would you describe your confidence during the course?
    - Did your confidence level change over the duration of the course? If so, how?

Appendix I

Psychometric Properties of the Mathematics Information Processing Scale

Appendix I

Psychometric Properties of the Mathematics Information Processing Scale

The reliability and validity of scores on the Mathematics Information Processing Scale (MIPS) were first discussed by Bessant (1997). The 87 items that comprise the MIPS can be grouped into three distinct sets of elements dealing with (1) statistics learning strategies (52 items), (2) metacognitive problem-solving procedures (9 items), and (3) cognitive-attentional deployment under evaluation (26 items). Bessant (1997) examined the factor structure, reliability, and discriminant validity of scale responses provided by a sample of 340 introductory-level statistics students.

In the 1997 study, all 87 of the MIPS items were submitted to principal axis factoring (oblimin rotation), from which five factors were extracted and labelled as follows: Metacognitive Problem Solving (METAPROB), Surface-Disintegrated Study (SURFMIPS), Deep-Associative Study (DEEPMIPS), Performance Preoccupation (PREOCUPY), and Strategic Study (STRATEGC). The reliability estimates (i.e., Cronbach's alpha coefficients) for scale scores representing these five dimensions are .89, .88, .86, .85, and .72, respectively. The five MIPS factors correlated as expected with measures of mathematics-related anxiety and attitudes toward statistics. For example, METAPROB is weakly or negatively associated with statistics anxiety and moderately correlated with favorable attitudes toward statistics. Similar patterns are noted among the affective correlates of the

DEEPMIPS and STRATEGC scales. In contrast, SURFMIPS is linked to evaluation anxiety and negative affect.

Finally, with regard to the use of the MIPS in this investigation, it should be pointed out that only 49 of the original 87 items were retained for factor and regression analysis. Based on the results of the principal components method, with varimax rotation, the 49 items selected from the MIPS were used to construct eight subscales for inclusion in the final regression equation predicting statistics course performance. Notwithstanding similarities in the factor labels specified in this thesis and the paper by Bessant (1997), differences in analytical procedures and scale compositions preclude direct comparisons between these two studies.