1989

The Psychometric Assessment Of Job Satisfaction And Its Relation To Stress In The Workplace

Timothy Douglas Hill

Follow this and additional works at: https://ir.lib.uwo.ca/digitizedtheses

NOTICE

AVIS

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

THE PSYCHOMETRIC ASSESSMENT OF JOB SATISFACTION
AND ITS RELATION TO STRESS IN THE WORKPLACE

by

Timothy D. Hill

Department of Psychology

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
December, 1988

ISBN   0-315-49347-X

Canadä

The Psychometric Assessment of Job Satisfaction
and its Relation to Stress In the Workplace

Abstract

The purpose of this dissertation was threefold: a) to review the current status of the job satisfaction construct, b) to evaluate how well modern scale development guidelines can produce a satisfaction measure that will be reliable and valid, and c) to investigate the relationships among satisfaction, stress, and other organizational outcomes.

The first chapter contains three sections. The first is a historical review. The second focuses on measurement concerns from the review. The third section reviews the shortcomings of the literature and how some of these shortcomings might be overcome by considering job satisfaction from a psychometric perspective.

The second chapter outlines the development of a measure of job satisfaction, the Satisfaction Research Questionnaire. The development strategy, readability, reliability, and construct validity are described. An argument is presented for a classification of satisfaction based on modal profiles. This classification yielded two bipolar modal profiles of scores.

The third chapter presents empirical results from two samples, a cross-Canada study and a student sample. Five content domains were tapped in this study: a) respondent information about themselves, b) response information about their occupations for use in monomethod multitrait comparisons, c) the Satisfaction Research Questionnaire and another measure, (the Job Descriptive Index), d) a measure of the social desirability response bias, and e) a measure of the Type A behavior pattern, as a measure of stress.

The results suggested that the Satisfaction Research Questionnaire would be a viable alternative measure of satisfaction. A robust relationship between satisfaction and stress was evidenced. The differences between this research and historical results were attributed to several factors, including: a) a modern scale construction approach, b) the use of modal profile analysis, and c) a multivariate conceptualization of job satisfaction.

Acknowledgements

It is with sincere gratitude that I wish to thank Dr. Douglas N. Jackson for his guidance throughout my academic career. A receptive student can best learn from the most adept teacher. I wish to also thank Dr. P.A. Vernon and Dr. J.P. Meyer. Both have been very helpful throughout the time that I have known them. Thanks are also extended to those members of the Psychology department that have helped in this effort, and to those who sat on the Senate defense committee. My gratitude is also extended to those employees across Canada who contributed their time to the completion of the survey, the students at the University of Western Ontario who contributed items to the Satisfaction Research Questionnaire, and the others who have completed the questionnaire.

Above all, my deepest thanks go to my wife, Mandy Chin Mien Chooi.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION

One of the most frequently researched aspects of life within organizations has been the affective state of the worker, either individually or as a group. This affective state has become known as "job satisfaction". Large bodies of research literature exist relating job satisfaction to organizational outcomes such as absenteeism, turnover, and worker productivity. (For extensive reviews, see Brayfield & Crockett, 1955; Herzberg, Mausner, Peterson, & Capwell, 1957; Hill, 1985b; Porter & Steers, 1973; Vroom, 1964.) Although the research has typically been correlational, an implicit causal model is usually present. Directional conclusions are often made regarding the effect of satisfaction upon some measure (cf. Organ, 1977).

There has been no strong consensus within this literature about the veridicality of the job satisfaction construct. Opinions have been divided since the publication of the Brayfield and Crockett (1955) and Herzberg et al. (1957) reviews. Brayfield and Crockett questioned the strength of the relationship between satisfaction and organizational outcomes. They felt that the observed median correlation between job satisfaction and organizational outcomes was simply too small to be important. Herzberg et al., however, concluded that the small median correlation was an artifact. They felt that the magnitude of the relationship was obscured by the presence of moderator variables. As a result of these publications, one branch of interest began to close, and another began with the examination of moderator variables.

To date, job satisfaction has continued to enjoy a wide audience. Assumptions about moderator effects have persisted. These assumptions have led to the implication that some groups of satisfied workers would be less likely to leave unexpectedly, less likely to be repeatedly late, or more likely to be more productive workers. The possible organizational benefits from having more of such "satisfied" workers have been intuitively clear, but the empirical findings have not consistently suggested any such enduring relationships.

This chapter provides a historical overview of the development of job satisfaction constructs and theories. The review begins with the period marked by Taylor's (1911) Scientific Management and concludes with a description of the current status of job satisfaction assessment. To complete the literature review, some representative recent meta-analytic reviews are also presented.

Following this review, substantive psychometric issues relevant to the assessment of job satisfaction are presented. Suboptimal scale attributes of popular satisfaction measures are introduced. Psychometric properties that are an impediment to construct valid measurement are discussed. The logical problem of expecting an aggregate measure of job satisfaction to predict any single act criteria (e.g., quitting, or a single rating of performance) will be examined. Finally, a desideratum of modern assessment techniques will be presented. These techniques, in turn, serve to introduce the second chapter, which describes the development of a recent measure of job satisfaction that incorporates modern standards. The third chapter presents the validity data from a Canada-wide study performed with the recent measure. This chapter draws upon the above validity data to focus on the relationships between job satisfaction and several criteria, most notably stress in the workplace. As well, issues in the conceptualization and

measurement of job satisfaction and criterion relationships are presented in support of the argument that some of the poor performance of historical job satisfaction measures has been due to poor psychometric properties.

Scientific Management. The goals of Taylor's Scientific Management were a) to make work related physical activities more efficient (in terms of corporate cost and productivity), and b) to improve the welfare of the workers by increasing their wages and decreasing their level of fatigue.

Taylor (1911) claimed that there were three reasons why his scientific management was needed in America at that time.

> First. To point out, through a series of simple illustrations, the great loss which the whole country is suffering through inefficiency in almost all of our daily acts.
> Second. To try to convince the reader that the remedy for this inefficiency lies in systematic management, rather than in searching for some unusual or extraordinary man.
> Third. To prove that the best management is a true science, resting upon clearly defined laws, rules, and principles, as a foundation. And further to show that the fundamental principles of scientific management are applicable to all kinds of human activities, from our simplest individual acts to the work of our great corporations, which call for the most elaborate cooperation. And briefly, through a series of illustrations, to convince the reader that whenever these principles are correctly applied, results must follow which are truly astounding. (Taylor, 1911, p. 7)

Taylor's now famous brick laying case history clearly reflected this major theme. Taylor's grasp of the scale of the issue of worker productivity was accurate. Taylor felt that Scientific Management could be applied to a diversity of occupational tasks. He did not anticipate detractors or the negative reaction to Scientific Management from America, Britain, and Europe. Scientific Management was, however, well received in Japan. His attempt to install Scientific Management into American, British, and European industries was

frequently perceived as a direct threat to the labor unions and management of the day. In North America Taylor was classified as having simplistic views about worker motivation. Most texts have portrayed him in an unfavorable light.

Taylor made few explicit assumptions about internal states of workers. One of Taylor's arguments was that pay would have the strongest relationship with productivity. A great deal of the controversy about Taylor's work lay in the fact that he argued for paying workers reasonable wages for the work that they could do. Moreover, if those same workers could do better under the auspices of Scientific Management, then they were entitled to appropriate increases in their wages. The assumptions most frequently questioned were that all workers would equally value economic incentives, and whether or not they would be willing to change their work pattern so as to increase their economic return and decrease their fatigue and inefficiency. A typical program of Scientific Management research would entail a) a review of the physical working conditions (a job analysis), and b) an attempt to identify the one most productive system for the production of those goods or services. Correcting worker behavior had three aims, to optimize productivity, to optimize the worker's return, and to minimize worker fatigue.

Figure 1.1 displays a typical flowchart relating selection, training, performance, rewards, and satisfaction in the workplace. An employee enters an organization, and may proceed directly to the job. Since some employees require training before actually performing their tasks, the training box is offset.

- -

Insert Figure 1.1 about here

- -

Figure 1.1 A typical flowchart displaying the relationship between selection, training, performance, and satisfaction.
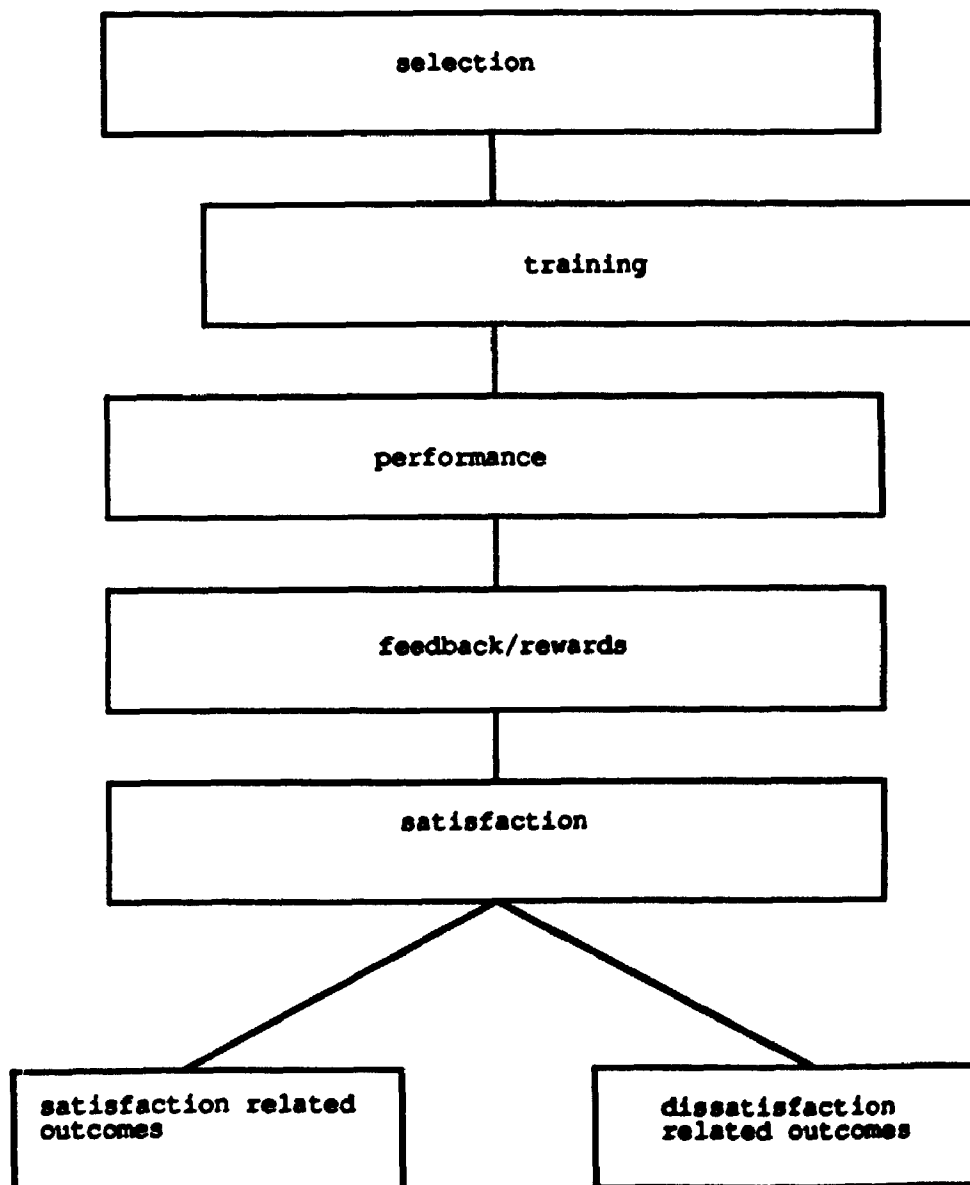
Taylor viewed the selection and training aspects as being critical to performance, and that satisfaction would be a consequence of the cumulative effects of satisfaction with one's level of performance, and the rewards that one receives for one's work. Although Taylor did not use the term "satisfaction", he was attempting to invoke accurate selection and training programs to classify each worker by level of ability. He felt that each worker's quality of life would be improved by accurately placing that worker and providing quality training so that the entire workforce would be "far more prosperous, far happier, and more free from discord and dissension" (Taylor, 1911, p. 29). Controversy surrounded Taylor's campaign for Scientific Management, eventually leading to Taylor being investigated by the Inte state Commerce Commission and the American House of Representatives. Taylor was charged with exploiting workers to reduce organizational costs, and of using more efficient procedures to increase the numbers of the unemployed. Taylor responded that his system of greater efficiency would lead to greater profits for management and workers, and the ability to place workers in situations that would maximally benefit from those specific abilities. These proceedings were cut short by the onset of World War I, and thereafter, Taylor's influence was further diminished with his death in 1915, the first world war, the great depression, and the advent of the human relations movement.

Elton Mayo took part in the final stages of the Scientific Management research at the Hawthorne plants of General Electric. Mayo attempted to dissuade the Hawthorne plant from applying Scientific Management, and it was his influence that gave rise to the "Hawthorne Studies" proper. These studies were probably responsible for the first empirical recognition of worker's attitudes in the workplace.

Hawthorne Studies and the Beginning of the Human Relations Movement. A great deal has been written about the Hawthorne studies, not all of which is relevant to this historical overview. There are, however, four important aspects of the Hawthorne studies that serve to note the separation in thinking from the Scientific Management period. First, the studies presented one of the first cases where researchers explicitly focused on employee's feelings, and how these attitudes could affect their work. Second, the Hawthorne studies were probably responsible for perpetuating certain negative biases against employees, such as the worker being viewed as a tabula rasa and easily manipulated (cf. Bramel & Friend, 1981). Thirdly, some of the methodological strengths and weaknesses of the Hawthorne research designs were subsequently noted (Franke & Kaul, 1978), and therefore made known to other researchers. Fourth, the research introduced the distinction between the way that groups of workers describe their work and the objective reality of the workplace, thereby heralding the recognition of the social dynamics of the workplace. Although Taylor had originally promoted the recognition of workers' feelings, and the scientific analysis of the workplace, history has attributed this to the Hawthorne researchers.

Mayo's (1960) appropriation of some of the accomplishments of Taylor and his colleagues regarding the human side of management was partly responsible for the decline of the Scientific Management movement. The Hawthorne studies focused upon rewards and needs, specifically, how the need for recognition by the workers related to their levels of production. The selection, training, and performance aspects of the workplace were not considered to have any effect upon satisfaction.

Vroom (1964) described the human relations movement as "an attempt to increase productivity by satisfying the needs of employees" (p. 181). The studies of Mayo appear to have been predicated on the belief that job performance was contingent on employee attitudes. The Hawthorne studies set the stage for subsequent researchers who also sought to determine whether or not satisfaction led to performance, or performance to satisfaction.

Hoppock (1935). Hoppock's (1935) large scale studies of overall job satisfaction focused on the differences between individual and group reported levels of satisfaction. He felt that there were variables outside of the workers that affected the levels of job satisfaction. One such extrinsic variable was group membership, i.e., the occupational group or groups to which the worker belonged.

Hoppock posed two general questions. One, how satisfied were the workers in his sample, on an absolute level? Secondly, were some groups of workers more satisfied than other groups of workers, on a relative level?

The recognition of the possibility of individual differences in job satisfaction was just being hinted at. Hoppock found that often the variability in job satisfaction scores *within* an occupation was greater than the variability *between* occupations. Hoppock did not find that all occupational levels were equally satisfied, nor did he overcome the problem of individual variation being greater than group variation in job satisfaction scores. For Hoppock, group membership was thought to have an effect on satisfaction. His expectation was that the within-groups variability in satisfaction would be much less than the variation observed in satisfaction across groups. Hoppock found that the variability in job satisfaction scores within occupational groups was often equal

to, or greater than, the between group variability. Again, the selection, training, and performance concerns have been left absent from this model, as is the feedback/rewards concern. This effect has persisted throughout the literature, and is responsible for the passing of group membership (i.e., belonging to different occupational groups) as a moderator.

Changes In Measurement Practice. The research cited above was primarily concerned with satisfaction as a global construct. Subsequent researchers began to hypothesize that respondents might have different feelings about different aspects (or facets) of their work. A major shift in emphasis seen in the literature following the 1930's was the recognition that global satisfaction could mask specific (facet) satisfactions. Hence, the literature following the above period began progressively to utilize facet satisfaction measures. Yet to collect facet satisfaction data, the researchers required a new tool for the examination of separate dimensions of a global construct.

Likert (1932) and Thurstone (1947) both introduced powerful methods for the collection and analysis of respondent data. Managers and others in business liked the Likert scales because they did not require negatively worded items, and hence these scales became popular. Importantly, one of the first applications of Thurstone's multiple group factor analysis method was in the assessment of job satisfaction. (Wolfle (1940) did not report any factor analysis of workplace attitudes in his pre-1940's literature review.) With the advent of these procedures, several changes were heralded in the examination of satisfaction.

First, researchers were no longer focusing on general satisfaction as a unidimensional construct. What used to be referred to as "moral" or "esprit de corps" were being labelled as the constituent factors produced by Thurstone's

method. While the recognition of the multidimensional nature is important, a virtual moratorium was declared on workplace morale research. It has only been recently that authors have again called on workplace morale as an important aspect of workplace life (cf. McMillan, 1984).

Due to these changes in the assessment and conceptualization of workplace attitudes, the following years of satisfaction research, (roughly from 1950 onward), became the domain of facet satisfaction. For example, the gain in sophistication in conceptualization was accompanied by the wholesale use of ad hoc scales, often single item scales. One potential caveat should be raised about the facet satisfaction literature; it is difficult to determine whether or not the increased sophistication of the analytic tools was adequately realized. The literature may have expanded as rapidly as it did during this period simply because assessment tools were available. Likert scales that did not use negatively keyed items (managers in applied settings did not like negatively keyed items) have been partially responsible for the acquiescence response bias in this literature. Thurstone's factor analytic method was predated by Hotelling's principal components analysis, yet there were no principal components analyses of job satisfaction reported prior to Thurstone's (1947) example in his text. (It is important to note that factor or components analyses would have been prohibitive in terms of personnel cost to compute at that time.) One possible explanation is that the development in rating methodology was partially a result of capitalizing on the available popular instruments.

Schaffer (1953). Schaffer felt that the key to unlocking the mystery of individual and group variation lay in the workers themselves, and not in extrinsic measures, as suggested by Hoppock. Moreover, Schaffer felt that it was

defensible to speak of people being generally satisfied or dissatisfied, as a part of their personality repertoire, or of their psychological "set". This set was thought to affect their satisfaction with work as well. When certain needs of the individual were not met, then tension would ensue, with the amount of that tension being proportional to the press of the unfulfilled need. Schaffer drew heavily on the personality psychology need theorists. While he is given almost universal credit for being the first to invoke formal human needs in the assessment of workplace attitudes, DeMan (1927) predated Schaffer by 26 years.

This was the first large scale empirical research that specifically stated that individual workers had needs, affect, or cognitions that could influence their feelings about their job and subsequently, their behavior in the workplace. Schaffer posited that there were twelve basic needs present in all people, similar to those suggested by Maslow (cf. Maslow, 1943, 1954) and other early need theorists. Schaffer felt that to argue for the equivalence of all needs for all people would be an untenable position. Instead, he sought to identify those needs that were most associated with satisfaction. Schaffer's recognition of individual differences in need fulfillment in the workplace bears close resemblance to several modern need theories of job satisfaction. To discover which needs were most central to job satisfaction, Schaffer asked three types of questions of each respondent. One, the importance of each of the twelve needs for each person. Two, the degree to which each respondent felt that those needs were currently being met or satisfied, and three, a self report measure of overall satisfaction with the job. Schaffer discovered that if a respondent's two most important needs were being met, then that person would rate themselves as being generally satisfied with their job.

Historically, Schaffer's work has been given merit for this discovery of the relationship between one's two most important needs and overall satisfaction. (As above, the selection, training, performance, and feedback/rewards concerns have been left out of the model.) Another laudatory aspect of Schaffer's research was that it was a major turning point in the development of how satisfaction with one's work is conceptualized. The publication of Schaffer's findings and other research of the period led to great hopes for the prediction of workplace performance by satisfaction. Yet there was no explicit statement about whether or not workers could actually do the job, or whether they were correctly selected or trained for their jobs.

As the literature to date has shown, a great deal of progress had been made in the development of the conceptualization of the worker in the workplace, and of the worker's "affective workplace". Most of the research in the 1950's concentrated on the lower half of Figure 1.1, i.e., on defining a simple relationship between satisfaction and a simple measure of an organizational outcome.

As the research literature began to accumulate, literature reviews were published to present an overview of the research findings. The most notable were those of Brayfield and Crockett (1955) and of Herzberg et al. (1957). Each of these reviews was to promote a somewhat different path of future research, although the majority of subsequent research focused on a single aspect of the model first postulated by Taylor.

Brayfield and Crockett (1955). Brayfield and Crockett published the first literature review of the job satisfaction and performance literature, citing research dating from the earliest American investigations relating attitudes and

productivity in an industrial setting (cf. Kornhauser & Sharp, 1932). While the reviewers did suggest that workers' attitudes were related to absenteeism and turnover, they concluded that there was little evidence to support the hypothesis that worker attitudes have any simple or appreciable relationship to performance on the job. They noted in their review the nearly complete lack of statistically significant correlations between satisfaction and performance, and cited this as evidence for support of the null hypothesis of no relationship. Again, selection, training, and feedback/rewards concerns were absent from the literature review. Moreover, the performance measures that were reviewed were simple performance measures, typically viewed as discrete events, and replete with rating errors.

The strict dependence on statistical significance by Brayfield and Crockett had three effects. One, researchers finally had a methodological guide that they might follow to improve the psychometric integrity of their research. Secondly, the Brayfield and Crockett position risked a high Type II error rate, given the fact the studies that were reviewed were all based on small sample sizes because the unit measured was the group of workers, and not the individual workers themselves. This finding led to research focusing on individual workers (e.g., person-environment fit studies) as researchers felt that possible individual differences were being masked behind group averages. Lastly, the review suggested that there was no relationship between satisfaction and performance, a call echoed by many subsequent researchers. These effects were similar to those produced by the next major job satisfaction literature review, that of Herzberg, Mausner, Peterson, and Capwell (1957).

Herzberg, Mausner, Peterson, and Capwell (1957). The literature review of Herzberg et al. covered a similar literature to that of Brayfield and Crockett, but arrived at somewhat different conclusions. Herzberg et al. did not rely on strict statistical significance in their review, but qualified their suggestive findings, as they recognized that any conclusions or generalizations would lack statistical power. They noted that the majority of the satisfaction and performance correlations were in the correct direction, as were the correlations between satisfaction and other work behaviors such as accidents and psychosomatic illnesses. They interpreted a mean correlation whose magnitude was similar to that of Brayfield and Crockett's mean correlation, but deemed that their mean correlation might be interpretable. In short, they argued for moderator or suppressor variables that might have truncated the "true" correlations between satisfaction and organizational outcomes. Once the authors had presented their "satisfaction is real, but inadequately captured" thesis, there was a spate of moderator variables research that sought to better capture the elusive job satisfaction variance. To this end, the subsequent literature flourished with satisfaction studies that used moderator variables, looked for suppressor variables, and used simple criteria.

Vroom (1964). The notorious median correlation between job satisfaction and worker performance of 0.14 was provided by Vroom in his review of 23 studies of the satisfaction performance relationship. Vroom felt that the published literature of his period should have contributed to advanced theories about human motivation, but he did not find this to be so.

> Unfortunately, it does not appear that this potential is being realized. Research in industrial psychology is still largely atheoretical with little use being made of the concepts and models

> which are an integral part of current theories of motivation. There is a marked tendency for industrial psychologists to take their concepts from "everyday" vocabulary. Terms like morale, consideration, participation, fatigue, and vocational interest are seldom given adequate or consistent conceptual definitions. (Vroom, 1964, p. 4)

Again, selection, training, and feedback/rewards were not explicitly a part of Vroom's model of motivation. This median correlation was to become strong evidence for future arguments about whether satisfaction exists at all, and even if it did exist, the lack of relationship with performance made the worth of satisfaction research seem limited. However, Vroom's research was not concerned with factors that constrained this correlation, he simply averaged other peoples' correlations. He did not, for example, consider the variance restricting effects of employee self- and pre-selection, poor performance assessment, or a host of other threats to higher correlations. Nor did he acknowledge that residual variance in the satisfaction-performance relationship might be accounted for by other individual measures. General cognitive ability, for example, would explain a large portion of the variance in performance.

Two Factor Theory. After the publication of the above reviews, and the attendant debates, subsequent research demonstrated that there were certain types of performance that were related to certain employee attitudes. Importantly, it was also recognized that satisfaction and dissatisfaction might be conceptualized as two different phenomena. Herzberg concluded that the two were indeed distinct, that they developed from separate sources, and that they each had unique effects on worker behavior, both in the immediate and long term cases.

Herzberg's Two Factor Theory was based on the results of a study that collected attitudinal information from a sample of engineers and accountants. The workers were asked to list times when they felt particularly good and bad about

their occupations. For each instance, the workers' responses were scored along three continua: a) the situation that led to those good or bad feelings, b) the needs or drives that were related to those situations, and c) how long those good or bad feelings lasted. The results of this study suggested that there were two groups of needs that were related to satisfaction. One, the motivator needs, related to the challenge and nature of the work itself. The second, the "hygiene" needs, encompassed those attributes of the work that related to the psychological and physical environment of the work. Herzberg further suggested that when the motivator needs were not being met that the individual would not be satisfied, but that this lack of need closure would not cause dissatisfaction either. When the hygiene needs were not being met Herzberg hypothesized that the individual would be dissatisfied, and would no longer be so when these needs were met. The empirical research has since shown that there is no clear distinction between what are "satisfiers" and what are "dissatisfiers". This result has been particularly damaging to the simple prospect of maximizing satisfiers and minimizing dissatisfiers.

Lawler (1973). Lawler's (1973) publication provided an expansion of the Porter (1961, 1962) and Porter and Lawler (1968) models of work motivation. The models hypothesized a connection between the actual rewards that one receives for performance and whether the perceived rewards are deemed to be equitable. In his model of satisfaction, Lawler (1971) emphasized the employees' perception of the organization, employee inputs and outcomes, and the inputs and outcomes of others.

Taylor (1911) first argued for the recognition of the fact that employees may not perceive their tasks or rewards in the same manner. Mayo emphasized this possible difference between the employee's real world versus the perceived

world. Additionally, Lawler made the Porter and Lawler model more specific. The respondent was seen as having a perceived amount of reward that he or she feels he or she should receive.

- -

Insert Figure 1.2 about here

- -

Figure 1.2 presents Lawler's model. The perceived amount of reward actually received is shown to be a function of the perceived outcomes of referent others and the actual outcomes received. When the perceived amount that should be received (Box "A" in Figure 1.2) is equal to the perceived amount received (labelled "B"), the employee is hypothesized to have satisfaction with that facet of work. Therefore, Lawler argued that a) A=B is the satisfied state, b) A>B is the state when the worker is dissatisfied, and c) when A<B the worker would report feelings of guilt, inequity, and discomfort.

Lawler's model assumes that the sum of all facet satisfaction can serve to functionally define global job satisfaction, and that excess levels of rewards (outcomes) are always met with dysfunctional reactions. Yet a slight overpayment may not elicit the same affective response as a gross overpayment.

- -

Insert Figure 1.3 about here

- -

Figure 1.3 serves to show that Lawler's model is still one that focuses on only a small part of the total work experience. (Again, selection and training are not specifically included in the model.) Moreover, while ther has been some empirical support for the premise that people's perceptions are influenced by

Figure 1.2 Lawler's (1973) model of facet satisfaction.
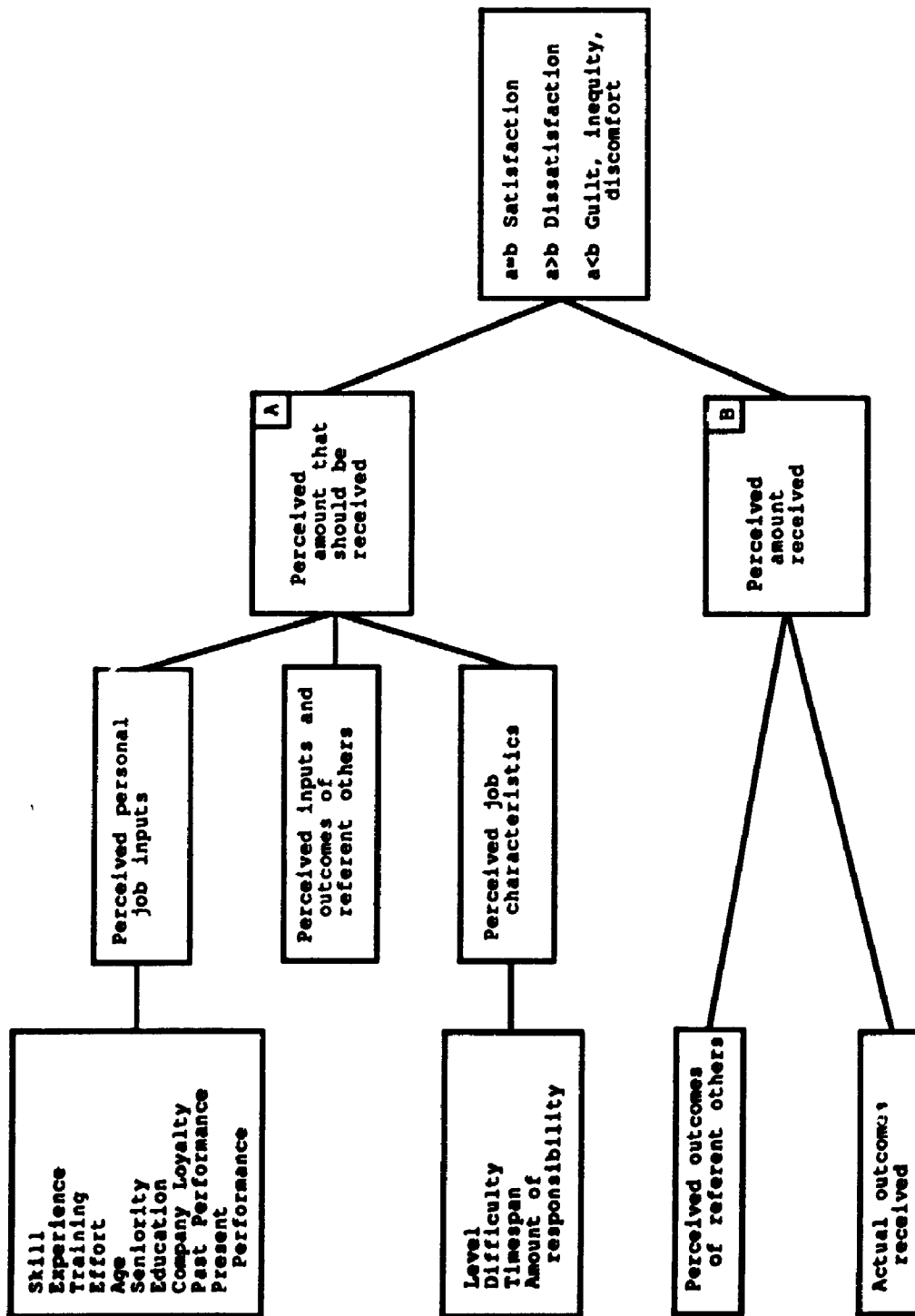
Figure 1.3 Lawler's (1973) model of satisfaction, relating perception and simple performance measures.

simple performance measures, viewed as discrete events

perceived amount of reward expected and perceived amount of reward received

satisfaction

satisfaction related outcomes

dissatisfaction related outcomes
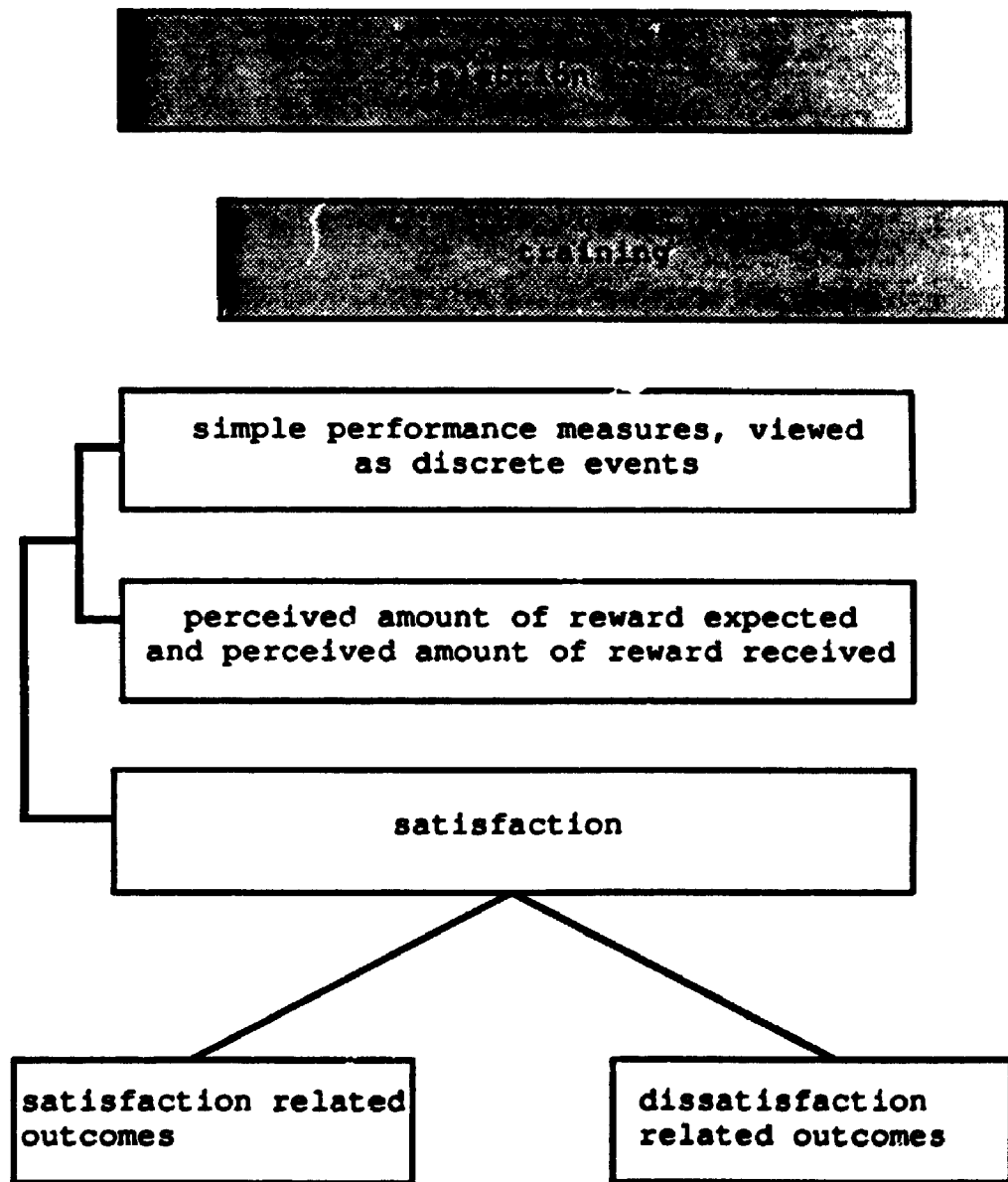
those around them, Lawler's model of satisfaction is somewhat limited in its applicability. Employers would probably wish to improve their employees' levels of satisfaction, but the employer cannot provide an endless supply of happy and equitable Others.

Locke (1976). Locke (1976) defined job satisfaction as perceiving the job one does as "fulfilling or allowing the fulfillment of one's important job values, providing these values are compatible with one's needs" (p. 1342). Values are the subjective ratings of what a respondent wants, desires, or values in the workplace, either consciously or unconsciously. Locke defined needs in a manner similar to that used by earlier need theorists. He has suggested that the relative importance of each satisfaction facet be recognized in any model of satisfaction. Specifically, advocates of Locke's model suggest that a) satisfaction is not adequately described by simply summing facet scores, and b) that a better way to consider overall job satisfaction is to incorporate the relative importance of each satisfaction facet. The weighting by importance scheme is used to elicit the range in satisfaction scores that a simple summed total scores obviates. This was felt to be important because Locke hypothesized that the importance of each facet served to affect the range of scores, not the absolute presence or absence of satisfaction. A respondent may be satisfied with a facet, but not rate that facet as important. Locke's use of weighting by importance was an attempt to answer the same question that Hoppock faced -- the presence of within group variation in satisfaction scores. In Locke's model, a respondent who rates a facet as very important is thought to be affected by even slight variations in that facet, whereas another respondent who does not rate that facet as important is unaffected. Consider the case of satisfaction with temperature. A low importance

respondent has less variation in this facet score than does the high importance respondent. Locke and others have suggested a variety of weighting procedures for determining an overall job satisfaction (OJS) score.

$$OJS=(JSF1)^{W1}+(JSF2)^{W2}+(JSF3)^{W3}+...+(JSFn)^{Wn}$$

Where: JSF_i = Job Satisfaction Facet score for the ith facet, and W_i = the importance rating of the ith facet. Again, the weighting (power function) is used to increase the range of respondent scores (i.e., overall job satisfaction).

Using a seven point scale, the possible overall job satisfaction score for a five facet scale ranges from 5 to 4,117,715. (Or, the range would be from 5 to 49, using multiplicative, not exponential weights.) While this is clearly an increase in the range of the overall job satisfaction score, the meaning of any such new score awaits empirical scaling.

Although Locke's theory was notable for reviving the acknowledgement of worker perceptions, (an idea that _was_ previously advocated by others), the technical criticisms of the weighting by importance scheme are provided below and describe how the model has fallen from popular favor. Locke's theory has been shown to be moderately applicable in the workplace, particularly when a group of workers share a common need. Yet it is frequently difficult to determine what all workers' values and needs are. Locke has not defined the mechanism by which a need becomes important, loses its importance, or affects satisfaction as that need changes. If needs are viewed as traits, then a worker

high on a given trait (e.g., Autonomy) will continue to work in a manner that promotes an autonomous work style. This employee's satisfaction can be improved by providing more autonomous aspects to the job. If needs are viewed as more transient personal attributes, then manipulating the workplace to meet a need might lead to perpetual job redesign, to meet whatever need has not yet been satisfied.

Landy's (1978) Opponent Process Theory. Landy's (1978) application of Solomon and Corbit's (1974) opponent process theory also addressed the within groups satisfaction variation problem. Landy recognized that often a job will be less interesting after several years than that job was after only a few weeks. Rather than simply label this decline in satisfaction "boredom", Landy suggested that there are mechanisms that help individuals maintain their emotional equilibrium. A job is seen as producing either a positive or a negative emotional reaction. If the level of emotion should exceed some value, then an opponent process acts to bring the level of emotion back to some basal rate. Once the level of affect has been returned to the basal rate, the opponent process is no longer required. Theoretically, the repeated use of the opponent process makes the opponent process stronger over time.

Given this simple mechanism, it is possible to explain how an individual can remain in a single job for a long period of time, and become less satisfied -- even though the job itself has not changed. This would explain the greater levels of satisfaction reported at the onset of peoples' tenure in a job than later in their tenure. Although this theory might offer an explanation of boredom, it does not adequately deal with those cases wherein people have a relatively constant level of satisfaction, or increasing levels of satisfaction.

Landy has likened satisfaction to a physiological state, and this reduces the ability of an employer to change something in the workplace so as to improve satisfaction. It would be difficult, at best, to alter the physiological state of an entire workforce or to constantly provide novel workplace features. This inability to test the basic premise of Landy's description of satisfaction renders the theory interesting, but not practicable.

Meta-Analytic Job Satisfaction Literature Reviews. Researchers have recognized that predictor and criterion unreliability, small sample sizes, range restriction, and other factors have a cumulative effect on the size of the sample correlation coefficient. Meta-analytic procedures have been applied to account for some, if not all, of these sources of attenuation on the correlation between satisfaction and performance. The meta-analytic literature reviews still do not take into account any more of the basic satisfaction model than did most of the earlier reviews. (By definition this is the case, since any meta-analysis must always be a quantitative literature review.) With this method in hand, researchers have pursued the relationship between satisfaction and organizational outcomes, often with the hope that the procedure will illuminate any previously undiscovered correlations that are significant in the population but not necessarily in each sample.

The meta-analyses of Iaffaldano and Muchinsky (1985) and Podsakoff and Williams (1986) both computed the expected (population) value of the job satisfaction relationship with job performance. The resultant correlation was 0.17, seeming to confirm the earlier findings of Vroom (1964). Iaffaldano and Muchinsky (1985) concluded their review by presenting an argument which stated that only a) the manifest importance of job satisfaction and performance,

together with b) their (implicitly assumed) relationship, have led to the continuation of satisfaction publications. Their viewpoint is representative of researchers that have failed to find illuminating (mean or corrected) correlations between satisfaction and simple performance measures.

Petty, McGee, and Cavendar (1984) also examined the individual overall job satisfaction and individual job performance relationship. As part of their meta-analyses, they reanalyzed the 15 studies cited by Vroom (1964) and found the same mean correlation of 0.14. The original reliabilities and sample and population standard deviations were not available to the authors. They were still able to show that 40% of the variance between the correlations cited by Vroom was attributable to sampling error. Other artifacts, such as error of measurement and range restriction, accounted for the remaining 60% of the variance.

In their own meta-analysis of a different set of 20 satisfaction performance studies, an attenuation-corrected correlation of 0.31 was reported, with about three-quarters (77%) of the variance across studies due to a combination of sampling error and error of measurement.

Hackett and Guion's (1985) meta-analysis of satisfaction and employee absenteeism concluded that there was a negligible relationship with satisfaction and absenteeism. The authors noted that the distribution of absence data in any sample tends to be highly skewed and quite truncated, thereby constricting any subsequent correlation. Importantly, these effects will occur in most relationships between satisfaction and organizational outcomes because of the nature of the satisfaction outcomes, e.g., absenteeism, lateness, turnover. All are infrequent events with decidedly noncontinuous, nonnormal sample distributions. Moreover,

the data available for most meta-analyses are not the raw item data, but the aggregate data, and these data are not readily amenable to meta-analytic corrections.

Loher, Noe, Moeller, and Fitzgerald (1985), however, reported a somewhat different meta-analysis. Their review included publications that reported correlations between satisfaction and job enrichment, with growth need strength as a moderator. They report a mean correlation between satisfaction and the job characteristics index of .39, with specific correlations ranging from .32 (satisfaction and task identity) to .46 (satisfaction and autonomy). These authors provide an important proviso in that "the results from the moderated studies warn that simply enriching a job will not necessarily hold the same amount of benefit for everyone" (Loher et al., 1985, p. 287). These authors then demonstrate how respondent differences affect the correlations with satisfaction. For high growth need strength (GNS) individuals, the correlation between satisfaction and job characteristics was .68, for low GNS subjects, the correlation was .38. (This notion that the between respondent variation has contributed to the ambiguous job satisfaction literature is discussed below, as a measurement concern.)

The last area that will be covered in the literature review of job satisfaction is the empirical literature that relates job satisfaction to stress and stress-related outcomes. Here, job dissatisfaction is viewed as a stressor with the possible consequences of stress being psychological, behavioral, and psycho-physiological.

Satisfaction and Stress. The interest that the medical profession has had in stress and its physiological and psychological consequences has been mirrored by a similar interest by those in behavioral medicine and lately, I/O psychology.

These latter two areas have generated impressive findings relating low job satisfaction (a stressor) to decreased longevity, dysfunctional symptomatology, boredom, interpersonal conflict, and psychosomatic disorders.

Palmore's (1969, pp. 242-243) research showed that job satisfaction and life satisfaction were the best two predictors of overall longevity. Both were better predictors than the level of physical functioning or the amount of tobacco use (controlling for age). Not only are there direct organizational outcomes related to job satisfaction, but there are individual health/attitude outcomes that should be assessed. Indeed, Palmore (1969a, 1969b) and Palmore and Jeffers (1971) have even argued that life insurance companies should adjust their (actuarially-based) life expectancy tables to include work satisfaction.

Palmore's life expectancy research was the first to apply multivariate methods to the analysis of the determinants of longevity, and the first to include both social and physical variables (Rose & Bell, 1971). In addition to stress being dysfunctional for an individual, stress will also be detrimental to the organization that has such afflicted workers. For example, Beehr (1976) and Beehr, Walsh, and Taber (1976) and Gupta and Beehr (1979) have shown that job stress is related to outcomes valued by the workers (e.g., mental and physical health), as well as by the organization (e.g., such withdrawal behaviors as absenteeism and turnover). Work stressors have even been related to wife abuse severity and frequency (Barling & Rosenbaum, 1986).

Theories relating job satisfaction to stress in the workplace have either looked at a general measure of stress, or have used the Type A versus Type B classification of stress-related behavior patterns. Some researchers (cf. Kahn & Quinn, 1970; Kahn, Wolfe, Quinn, Snoek, & Rosenthal, 1964) have defined stress

as a demand from any facet of working life that has extreme or noxious characteristics. Stress is said to be aversive and to specifically lead to withdrawal behaviors. Alternately, within the Type A/Type B personalogical classification of stress, the Type A person is the individual who is considered to be perpetually racing against the clock, feels a lack of control over his or her life events, and who does not have an efficient coping strategy for dealing with time pressure and conflicts. The Type B person is commonly portrayed as the individual who proceeds at a slower pace, but gets the work done, as well as, if not better, than the Type A counterpart. Friedman (1969) provided a vague definition of the behavior pattern of a typical Type A individual. The pattern was referred to as:

> a chronic struggle to obtain an unlimited number of poorly-defined things from (the) environment in the shortest period of time, and if necessary, against the opposing efforts of other things or persons in the same environment (Friedman, 1969, p. 269).

The development of coronary heart disease has been linked to this behavior pattern. The disease can be hastened in an individual by the excessive levels of sympathetic nervous system arousal and/or adrenalcortical activity (Ross & Glomset, 1976). Coronary heart disease (e.g., lesions and/or hardening of the arteries, angina pectoris, and myocardial infarctions) can be evidenced in an individual without presenting overt symptoms of coronary heart disease. That is, the degree of tissue pathology can be much greater than normal before an individual actually displays angina pectoris symptoms (e.g., severe chest pains) or experiences a heart attack (Glass, 1977; Rapaport, 1980). Thus, the Type A behavior pattern can lead to unseen risk in the form of tissue pathology, and the Type A "person" has been shown to be the individual who is dissatisfied with his or her work.

What have the satisfaction and stress data shown?

A great deal of the literature that has dealt with stress in the workplace has focused on the effects of stress that are detrimental to worker health. Typically, subjective stress has been thought to relate to dysfunctional emotional and physical states. The Type A behavior pattern has often been described as a primary individual characteristic. The literature relating stress to performance has overwhelmingly found that the presence of individual stress can lead to heightened levels of aggression and hostility, and a decline of altruistic and prosocial behaviors. Anxiety, hostility, and depression have been shown to be some of the prime affective consequences of individual stress. The Type A behavior pattern (i.e., the stress-related pattern of behaviors or behavioral style) has also been shown to be related to these mental and physical debilitations. For example, Olson and Tetrick (1987) have shown how the degree of stress a worker feels is inversely related to their levels of job and life satisfaction. Motowidlo, Packard, & Manning (1986) have shown that occupational stress is related to decreased levels of interpersonal and cognitive/motivational aspects of job performance. These authors suggested that stressful events in the workplace can lead to depression and decrements in job performance. Specifically, the Type A pattern correlated with subjective stress, anxiety, hostility, and depression. Tetrick & LaRocco (1987) found both general and facet job satisfaction to correlate negatively with psychological well-being (i.e., anxiety and depression), and positively with understanding of workplace events, the predictability of those events, and the feeling of control over one's work environment.

What improvements are required in the research reviewed in this literature?

The relationship between job (dis)satisfaction and stress has been researched, in part, because of the inability of traditional physiological risk factors successfully to predict coronary heart disease. While most of these findings have been retrospective in nature (cf. Sales & House, 1974), prospective studies (cf. Goldband, 1980; Medalie, Snyder, Groen, Neufeld, Goldbourt, & Riss, 1973) have related satisfaction with angina pectoris. Indeed, Motowidlo et al. (1986) have recently shown that there is a robust relationship between occupational stress and job performance decrements. For the empirical investigation of the relationship between satisfaction and stress to come to fruition, reliable and valid multidimensional measures of both constructs are required. Stress has been typically assessed with reference to the Type A personality, clinical interviews, or suboptimal paper and pencil devices. Few, if any, of these stress measures reflect the multidimensional nature of stress or of the Type A behavior pattern. Similarly, conceptually narrow measures of satisfaction have been invoked in this literature. Therefore, modern measures reflecting the multidimensional nature of stress and of satisfaction should be applied in any future research applications. Yet more than simply being multidimensional, these measures must maintain high levels of construct validity.

Summary of the Empirical Literature

As has been suggested above, the number of publications within the job satisfaction domain has not imbued any particular quality to this research area. Traditional and even meta-analytic reviews have not yielded any consensual finding about the nature of job satisfaction, at least as a consistent predictor of salient organizational outcomes.

Cook, Hepworth, Wall, and Warr (1981) voice their belief that

investigators have not in general learned from their endeavors as successfully as they might. Specifically, there has often been insufficient attention paid to issues of conceptualization and measurement, so that cumulative development of the area has been retarded" (p. 1).

Any informal review of the literature will reveal the fact that the most popular measures of satisfaction have been the single item scales developed on an ad hoc basis. This is one reason for the Cook et al. (1981) criticism. Another reason being that scales have often been used because of their ease of access, rather than their psychometric integrity. It may not be surprising, therefore, that literature reviews have produced such disappointing mean correlations. Moreover, as Cook et al. (1981) suggest, measurement concerns have been overlooked by the vast majority of authors in the satisfaction domain.

In defense of Taylor. It has been noted above that some job satisfaction theorists have criticized Taylor for his "simplistic view of worker motivation". These criticisms were largely responsible for Taylor being ignored in Britain until after the first World War, and even thereafter he was universally met with hostility from labor and management. The reactions towards Taylor in Europe and the United States were of a similar nature. Some authors (cf. Hill, 1988; Ivancevich, Szilagyi, & Wallace, 1977; and Locke, 1982), have noted that Taylor's view was quite revolutionary, both in its breadth of research and influence. Taylor answered the charges of simplistic viewpoints in his own time. First, he claimed that the problem of inefficiency was not the problem of the worker, but of management. Secondly, he reported that workers felt they would be punished if they worked too quickly, which led to the third point that workers tend to work at less than their full capacity. Fourth, what management should be doing

is looking for the most efficient workers, and training them for further efficiency. Lastly, Taylor argued that the workers' output should be tied to their pay system, either by early incentive or some piece-rate system. To these ends, Taylor pioneered piece-rate pay systems, goal-setting procedures (with explicitly defined, difficult but attainable goals), the use of worker performance feedback, and selection, training, and production systems that would all have met modern definitions of validity and utility. Locke, Latham, and Erez (1988) have shown that this turn-of-the-century model of goal setting is as, or more, effective than other goal setting models, then or now.

Modern theorists have decried Taylor's so-called "ignorance towards the workers". These same critics have pointed to examples like Japan, whose industries appear to have succeeded enormously without Taylor-like principles. These data form the basis for popular arguments leveled against Taylorism. The "hard skills" of the Americans versus the "soft skills" of the Japanese are seen as partial explanations for differences in these industries. These "soft skills" have been typically attributed to the influence of Mayo and the Hawthorne studies. The irony in this contention is that it can be shown that Japan has succeeded, in part, because of Taylorism, not because of Mayo. Indeed, Cole (1979) was wrong, Taylor's ideas were not first implemented in the American railway and textile industries. McMillan (1984) points out that:

> In fact, as early as 1908, Junihoko Iwatare of Nippon Electric, Japan's first joint venture, studied Scientific Management at Western Electric and subsequently introduced it to his company. During 1913-1914, Daigoro Yasukawa visited Western Electric in the U.S. and introduced the bonus system at Yasukawa Electric Equipment in 1915. By 1919, the government subsidized the development of the Kyocho Kai (Labor Management Cooperation Society), which had a section on Scientific Management, the Industrial Efficiency Institute. (p. 203).

Taylor's work was popular and influential. His book was translated into Japanese shortly after its American publication, and was known as The secret of saving lost motion. Quickly, nearly two million copies of the translated work were purchased in Japan. Company owners and directors personally bought copies in orders of ten to twenty thousand, or more, and distributed them to their managers. The President of Mitsubishi Goshi Kaisha acquired and then distributed to his workers twenty thousand copies. The President of Kawaska Shipbuilding Yard gave out some fifty thousand copies. The legacy of Taylor can now be seen in the Japanese production strategies. Their kanban inventory management, reduction of waste, quality circles, robotics, and automated factories have become the standards for modern industrial efficiency.

While it is difficult to find controlled experiments relating American, European, and Japanese industries to empirically test the "Japanese" method, McMillan (1984, pp. 274-275) cites research that compares an American, a British, and two Japanese color television assembly plants. Absenteeism in the Japanese plants was half that of the American and British plants, labor productivity was as great, or greater, in the Japanese plants as in the others, satisfaction was greatest in the Japanese plants, and the quality of the products was best in the Japanese plants. (See also Wong, 1973.)

It was the combination of the Japanese focus on production management and Taylor's emphasis that the engineer play a more central role in management studies, that created the environment in which Scientific Management was welcomed in Japan.

What is known about job satisfaction?

Several summary statements can be made about what is known about job satisfaction.

A) Brayfield and Crockett (1955), Vroom (1964) and others were correct in their conclusions that there was indeed no simple relationship between satisfaction and performance. The belief in a simple relationship has been maintained, in part, because it leads to a simple rule for good performance: Satisfy workers and they work well. The job satisfaction and performance literature might also be called the "realm of the small correlation coefficient", and as such, has not provided strong empirical support for this simple prescription for better performance.

B) Job enlargement, job enrichment, worker participation in production, worker autonomy, and group and individual incentives, to name a few topics of research, have had meaningful correlations with job satisfaction. Most of these were the same areas argued for by Taylor, and as argued above, the validity of Taylor's arguments was improperly condemned. The correlation of .68 reported by Loher et al. in their meta-analysis has provided eloquent testimony for these relationships, as does the modern idea that performance appraisal requires a (scientific) job analysis.

C) There has been some recognition of the reasons for the "small correlation coefficients". These reasons have often been sought by meta-analytic methods. For example, Hunter, Schmidt, and Jackson (1982) have provided an outline of the methodological and sampling artifacts that can be accounted for with a meta-analysis. When researchers have applied this procedure, they have discovered that not all population correlations with satisfaction are trivial.

D) There has been some awareness of the role of proper scale construction methodology, but this has not been widespread. The measures that were developed from specific theories of job satisfaction all have several suboptimal scale properties. The most popular satisfaction measures are still single item scales. Specific scale recommendations that have been made (cf. Cook, Hepworth, Warr, & Wall, 1981; Guion, 1981) have not been acted on.

E) Very few job satisfaction researchers have demonstrated an awareness of relevant psychometric issues. For example, the impact of response biases, the problem of aggregated predictors and single act criterion, the lack of construct valid measurement, the problem of extremely skewed and truncated sample data, and so on have been traditionally neglected concerns in the satisfaction literature.

F) There has been some recognition of between person variance, and the possible effects this may have on the total sample correlation coefficients. However, moderator analyses have been improperly employed to investigate relevant hypotheses. High-low, dichotomous, or trichotomous data splits have been used to generate subgroups for moderator variable research, but this methodology has been inappropriately applied.

G) Lastly, the concept of job satisfaction is at risk of being deleted from the study of Industrial and Organizational Psychology. Introductory and senior level I/O Psychology textbooks are now being published without job satisfaction chapters. The hunt for large correlations based on simple relationships has decidedly failed. The absence of such relationships should not undermine the reality of the relationship between satisfaction and other measures, for example, stress, as suggested above.

To assay some of these assessment issues, nine measurement concerns appropriate to the assessment of job satisfaction are presented below. These measurement concerns are: a) the use of weighted scores, b) the use of difference scores, c) the difficulties with moderator variables studies, d) the problems with performance measures, e) the presence of validity-threatening social desirability responding, f) the problem of acquiescence response bias, g) the logical and psychometric problem of aggregate measures linked with discrete, low frequency measures, h) the relationship between reliability, construct validity, and aggregation, and lastly, i) the importance of recognizing between-respondent variation in scale scores. This chapter concludes with a section that describes what is currently required in the assessment of job satisfaction.

Measurement Concerns

There have been several recurring themes in the above literature review. One idea involved the use of difference scores, another the use of weighted scores. Both procedures sought to improve the capture of true score variance by creating an observed score that was supposed to more accurately reflect the respondent's true level of affect. Difference and weighted scores are discussed from a psychometric perspective. Moderator variables have been hypothesized since the recognition of individual differences in satisfaction, but the history of these effects has been inconsistent. Possible reasons for this inconsistency are discussed within the context of testing moderator effects. Response biases, such as desirability and acquiescence, are discussed. Various interpretations of desirability responding are reviewed, as the job satisfaction literature has only rarely recognized this insidious response bias. The impact and control of both biases are discussed. Almost all of the job satisfaction literature has attempted to

correlate job satisfaction with some specific, albeit organizationally relevant, criterion. The folly of expecting an aggregate predictor to correlate highly with a discrete criterion is discussed. Lastly, the measurement concerns relevant to the assessment of between respondent variation are presented. Each of these measurement concerns is now presented below.

Weighted Scores. Locke's value theory, as originally postulated and later revised by himself and others, called for the recognition of the fact that an employee may not value all facets of work equally, or may not think all facets of work are equally important. (This paradigm was a partial descendant of the Schaffer (1953) model, except that Schaffer ranked satisfaction facets by their importance, and did not weight them.) To reflect this (implicit) differential weighting, it was suggested that each facet satisfaction score should be weighted by either the importance of that facet, or by how much the employee valued that work facet. This weighting procedure has been soundly criticized in the past, but its use continues. Typically, these criticisms point out that the predictive efficacy of any weighting procedure is minimal when compared to unit weights.

Proponents of the "weighting by importance" scheme suggested that "if importance is a meaningful dimension, then the response to each item should be weighted by the importance of the item to the employee" (Glennon, Owens, Smith, & Albright, 1960, p. 519). Implied herein was the assumption that the unweighted and weighted sums of importance items should not correlate, thus providing independent sources of variance that could cumulatively account for more variance in a prediction formula.

Dachler and Hulin (1969), Decker (1955), Ewen (1967), Mikes and Hulin (1968), and Schaffer (1953) have all shown that the weighting by importance scheme does not significantly add to prediction in any way. Indeed, Mikes and

Hulin (1968) have demonstrated that unweighted values may predict an outcome more accurately. Moreover, criticisms of selecting some nonunity weights by which to maximize a prediction are well documented within the regression literature (cf. Cohen & Cohen, 1975; Draper & Smith, 1981; Pedhazur, 1982). The weighting-by-importance procedure may make intuitive sense, but empirical examination of the results does not warrant the application of this procedure. Moreover, there is a very high correlation between ratings of satisfaction and weighted by importance satisfaction ratings (Zenisek & Rowney, 1980). Both correlate with general satisfaction in the high 0.90's. Therefore, the incremental validity of posing a question in discrepancy score format would be low -- little new (i.e., non-redundant) information would be added (cf. Bereiter, 1963; Cronbach & Furby, 1970; Lord, 1963; Wall & Payne, 1973; Werts & Linn, 1970; Zenisek & Rowney, 1980).

Wall and Payne (1973) suggested a possible method of salvaging the rationale. They suggested that the respondent be explicitly told to consider how much of a particular facet there currently is on the job, and how much they would like there to be. The respondent would then be told to rate their satisfaction with this information in mind. This procedure tries to gather reliable data by providing an explicitly stated problem set, and asks the question in a standardized manner. Yet such instruction sets would increase the reaction or completion time for each item. This increased (mean) response time may decrease the reliability in the test-retest situation by introducing noncontent related changes of direction in item responding (cf. Fekken & Jackson, 1984; Windle, 1955). Moreover, respondents may already use such an evaluative heuristic when

responding to satisfaction questionnaires. In summary, the arithmetic processes have been thought to be similar to the cognitive processes, but this hypothesis has not received empirical support.

Difference Scores. Another procedure for calculating satisfaction scores has been to compute the difference between what the employee actually feels and what they think they would feel in an ideal situation. This difference is supposed to yield a measure of how satisfied employees are, relative to how satisfied they think they could be. The appeal of such difference scores lies in their self-referential norming. A score is provided that is supposed to reflect the relative satisfaction that each employee feels. This procedure seeks to avoid the problem of the temporal narrowness of most satisfaction measures, and to measure satisfaction in general. The two sets of questions are thought to guide the respondent's cognitive processing to a deeper level of processing, one wherein the respondent pays greater attention to each item. The intuitive appeal of this proposal is that somehow "better" or less biased information will be obtained (cf. Locke, 1969; Porter, 1961).

What has been truly surprising is the lack of recognition of the relevant psychometric literature dealing with difference scores. The two problems with a difference score are that a) it tends to ignore the reliabilities of the two original variables, and b) the degree of relationship between these variables. The two original variables must be highly reliable and minimally correlated in order to produce a reliable difference score. The standard formula for the reliability of a difference score is:

$$\text{Reliability}_{\text{difference}} = \text{Reliability}_1 + \text{Reliability}_2 - 2r_{1,2} / (2 - 2r_{1,2})$$

Where: Reliability $_{\text{difference}}$ = the computed reliability of a difference score, Reliability$_1$ is the reliability for the first measure, Reliability$_2$ is the reliability of the second measure, and $r_{1,2}$ is the correlation between the two measures.

For example, consider two hypothetical variables, X and Y, with reliabilities of 0.90 and 0.80, respectively. If the correlation between these two variables was 0.80, then the reliability of the difference score would be only 0.25. As the correlation between any two variables increases, the amount of reliable variance remaining in the difference score decreases, with an attendant loss of reliability. This effect is exacerbated when the two original variables have modest reliabilities. Since the correlation between any satisfaction measure and a version of itself will be nonzero by definition, then often the difference measure is of dubious worth. Statistical regression effects within the sets of scores will also decrease the validity of difference scores, as will any criterion contamination caused by presenting people with the same items.

Moderator Effects. Researchers have often felt that a stronger relationship between satisfaction and some organizational outcome should have been evidenced in their research. When a large correlation is not produced, some researchers have argued that some moderator is present, reducing the absolute size of this relationship. A typical methodology for assessing moderator effects has been to show that various subgroups have separate correlations. Others will perform an analysis of variance on dichotomized data (e.g., "low" satisfaction and "high" satisfaction groups). This ANOVA approach restricts the range of

variability within the total sample, introduces incorrect classification by ignoring the standard error of measurement about the cut score, and serves to inflate the Type II error rate. The correct procedure for analyzing moderator effects is an orthogonal stepwise regression, or a "moderated" multiple regression.

Hunter, Schmidt, and Hunter (1979) criticized moderator variables and relationships research because it often lacks theoretical propositions as to why the degree of the relationship should vary with Y. In the absence of well founded empirical theory the strongest interpretation of the "effect" of Y is simply in terms of differing degrees of extraneous error introduced for differing values of the moderator (cf. Schmidt & Hunter, 1978). Analysis of variance is often the data analysis tool de rigeur, but information lost via simple group divisions is irretrievable. ANOVA is therefore not as sensitive as regression in detecting moderated relationships.

Consider the predictors X and Y, with the criterion Z. Variables X and Y are correlated. It is hypothesized that Y is influencing the amount of variance in Z predicted by X. In any regression application involving such a set of relationships, the reliability of all measures must be taken into account. In the presence of unreliability, more prudent predictions and more stringent theorizing about the sizes of expected effects are required since the size of confidence intervals are wider, and hence less accurate estimates are being generated.

The correlation between X and Z, across some values of Y, is the degree or magnitude of the relationship. Most moderator-type satisfaction research considers only the degree of the relationship.

In order to determine whether or not some third variable affects the degree of the relationship can be discovered by examining the form of the relationship. The appropriate analysis is a regression analysis that contains a

multiplicative interaction term. In this procedure, the predictors are entered into a hierarchical stepwise (i.e., orthogonal) regression, followed by the interaction term. Each regression weight is tested for a significant increase in prediction from the previous entry. In this manner, the contribution to prediction that is unique to the moderator -- independent of the other effects -- is denoted. The degree of the relationship is not interpretable whereas the form of the relationship is. (For a further discussion of moderator variables, see Arnold, 1982.) This procedure has been rarely applied to the investigation of moderator effects.

Performance Measures: A Legacy of Inconsistent Measures. An important measurement concern is the nature of the criterion score that satisfaction is to predict. Typically, satisfaction has been correlated with some organizational concern. This section is aimed at describing the traditionally poor quality of performance measures, and relating this poor quality to a detrimental effect on the job satisfaction and performance correlation.

Guion (1965) presented three methods (or groups of methods) by which information about the performance of a worker can be gathered. These three groups were a) Objective performance data, b) Personnel data, and c) Judgmental data. Each group is presented and criticized below.

Objective Performance Data. This term refers to the "hard" data that can be gathered from the workplace on the basis of frequency counts --such as number of units produced, or amount of sales in a given period. An attribute of these assessments that is rarely noted in the literature is that these measures are not consistent. Authors repeatedly report correlations between satisfaction and productivity. These correlations are assumed to be meaningful, even though the

mean reliability for these performance measures is only .45 (Muchinsky, 1987). Moreover, while these data are less prone to halo, leniency, or central tendency errors, they are fraught with criterion contamination and criterion deficiency.

Personnel Data. Often researchers have presented hypotheses about satisfaction leading to reduced turnover, absenteeism, or some other criterion that can be culled from company's personnel files. While there has been no doubt that these criteria are organizationally relevant, they are all difficult to assess reliably. Typically, assessments are made of single event criteria that are already low frequency behaviors, and frequently there is little or no consensus about definitions. Muchinsky (1977) has criticized the typical approaches to absenteeism that depict it as a unidimensional construct. Similar arguments have been forwarded by others for turnover, "job hopping", accidents, grievances, lateness, and other measures. Blum and Naylor (1968) and Latham and Pursell (1975), for example, have found that record keeping activities associated with the above measures is often poor. "Sloppy record keeping is common" (Muchinsky, 1987, p. 309). Not only is the validity of personnel data poor due to inefficient record keeping practices, but the mean reliability is only .40, with a range of reliabilities that begins at .00 and extends only to .74.

Judgmental Data. Supervisory, self, or peer assessments, are the most popular format for collecting information about employee performance. Guion (1965) found that 81% of the studies in a 5 year period used judgmental data as a criterion for job success. Lent, Aarbach, and Levin (1971) reviewed some 1,500 criterion measures and found that 60% of them were judgmental. There is no doubt as to the popularity of judgmental ratings as a means of assessing employee performance. All of these types of ratings fall prey to leniency, halo, and central

tendency rating errors. All of these sources of error variance contribute to greater restriction of range, and its attendant truncating of any subsequent correlations. The finding that interrater reliability is typically good does not specifically address the impact of these three error variance sources -- it may simply mean that raters all make similar errors. In addition, the validity of the judgmental ratings depends in part on several other factors, such as the quality of the original job analysis, the training of the raters, and so on. The consensus is that the farther away the actual rating items are from concrete and observable workplace behaviors, the lower the validity of the performance assessment. To date, only the Behavioral Observation Scales (BOS) methodology solidly relates the items to be rated with the actual workplace behaviors, by making the items to be rated actual critical incidents from the workplace.

The measurement concerns presented above have used the traditional concepts of reliability and validity to evaluate scale merit. The next section presents a somewhat less traditional threat to reliability and validity, namely social desirability responding. Specifically, desirability responding detracts from the construct validity of satisfaction in a manner unlike the previous concerns.

Desirability as a Threat to Construct Validity. Social desirability as a response bias impacts on the validity and generalizability of all satisfaction research. All self report inventories are subject to some faking and malingering, and most have some items that are recognizable by respondents as being more desirable than others. There is the possibility that respondents do recognize these items and "fake good" so as to present a favorable depiction of themselves. It is not necessarily true that all respondents purposely fake good.

Edwards (1957) described desirability as the tendency for a respondent to "put up a good front", with no proviso that the response involves deliberate deception.

Silverman and Shulman (1970) reported that the presence of desirability responding may by elicited by the respondent's perception of what the experimenter expects, the desire to protect one's image, and the wish to please or frustrate the experimenter. Each of these sources of Desirability responding may take place in the workplace. (See also Arnold & Feldman, 1981; Bateman & Organ, 1983; Hill, 1986b; James, 1973; Orpen, 1974; Wall, 1972.)

The issue of Desirability responding has been well researched in the personality domain. In that empirical literature, the impact of Desirability is often assessed by examining the first unrotated principal component. This dimension is the unweighted linear composite that best expresses the variance in the data set. In personality assessment, there is no general or "g" factor, as there is in the assessment of human intelligence. In the intelligence literature, there is strong evidence for a "g"-intelligence factor, plus a series of related group factors. Additionally, in the assessment of job satisfaction there can be both a "g"-satisfaction and several group (or facet) factors. In the satisfaction domain, researchers tend to use a summed total score -- suggesting the "g"-satisfaction. This "g"-satisfaction could cast some doubt on interpreting the first unrotated principal component as Desirability.

An intermediate solution would be to a) extract the first unrotated principal component from the data matrix of job satisfaction responses, b) refactor the residualized matrix, and c) determine whether any group satisfaction dimensions remain. This procedure would, de facto, assess the impact of the first

unrotated principal component -- regardless of its definition. If meaningful residual factors were found, then the fidelity of the scale would be assured. If, however, there was no systematic variance in the residualized data matrix, then the scale at hand would have failed to both a) preserve its advertised structure, or b) refute the claim that Desirability responding constituted a proportion of the subject's response variance. (One way to evaluate such a claim would be to correlate the factor scores from the first unrotated principal component with the Desirability measure.)

There are three views as to the "correct" interpretation of desirability. One states that desirability deserves a content oriented interpretation. The second argues that Desirability may be construct related, but that the content is not relevant to the scale in question. And the third states that desirability detracts from any discussion of the respondent's true score estimate. Each view has received some support. All are critically presented here to reflect these continuing differences as to the interpretation of desirability.

- - - - - - - - - - - - - - - - - - - -

Insert Figures 1.4, 1.5, and 1.6 about here

- - - - - - - - - - - - - - - - - - - -

Desirability as Content Relevant to the Test. This view dictates that desirability does not play a role at all in determining the observed scores of the respondents. Figure 1.4 presents a diagrammatic model of this relationship. Desirability is viewed as another predictor and is conceptually different from the actual predictors at hand. There is no conceptual overlap indicated amongst the other "predictors", but there invariably is in reality. In a regression formula, the total amount of variance accounted for would be artifactually high, as

Figure 1.4 Desirability as content relevant to the test, but that does not interfere with the predictor-criterion relationship.

Criterion

Predictor   Desirability
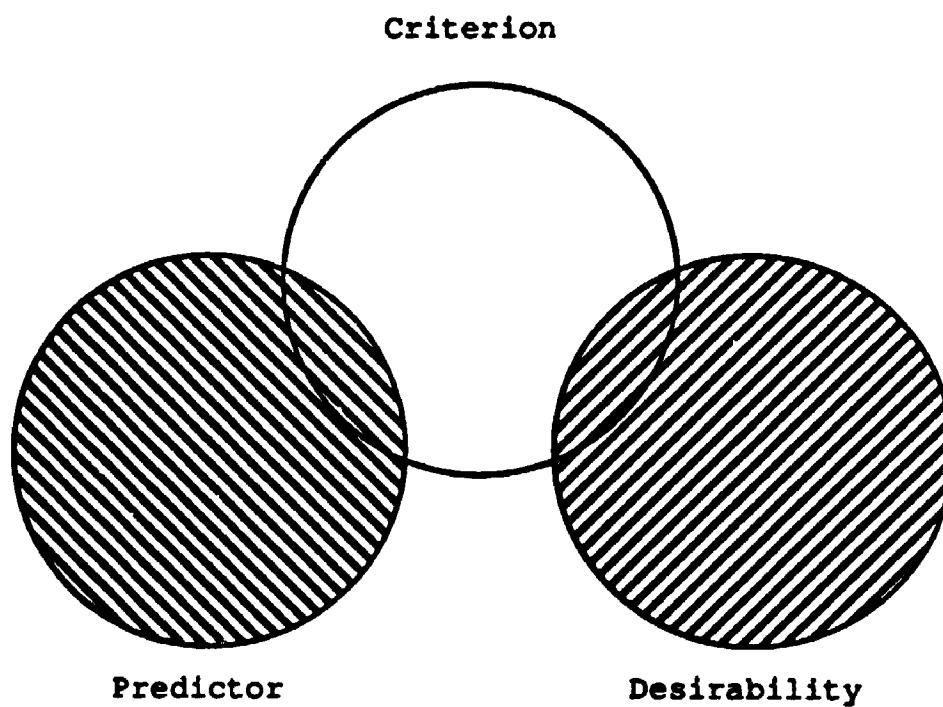
Figure 1.5 Desirability as important content, but different to that of the test.

Criterion

Predictor
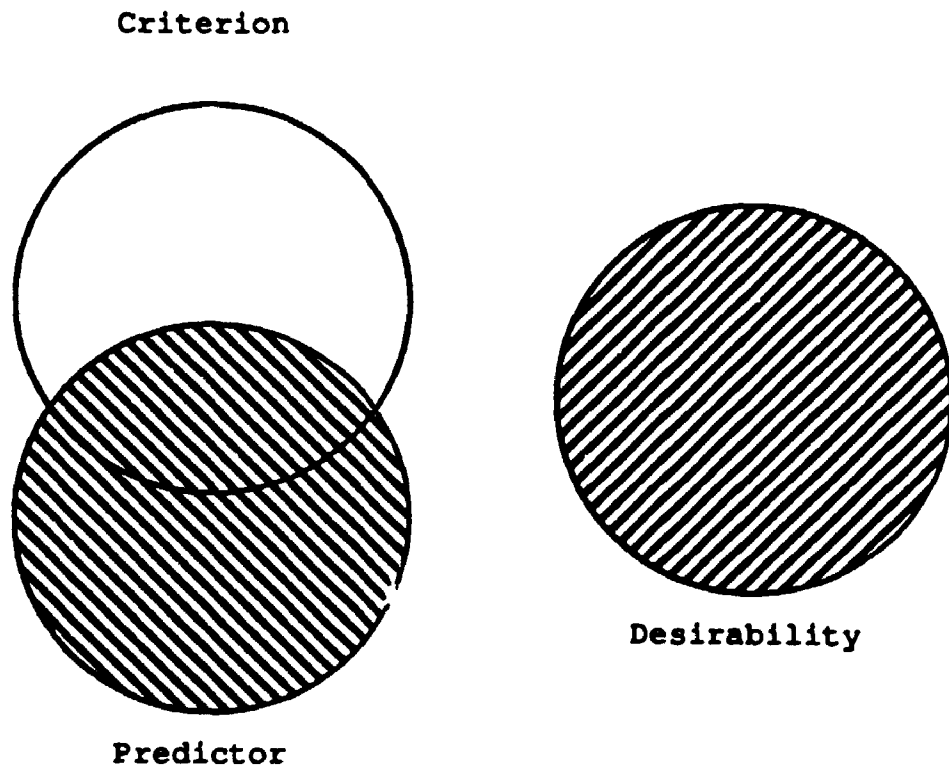
Desirability

Figure 1.6 Desirability as a response bias that artifactually inflates the amount of variance accounted for.
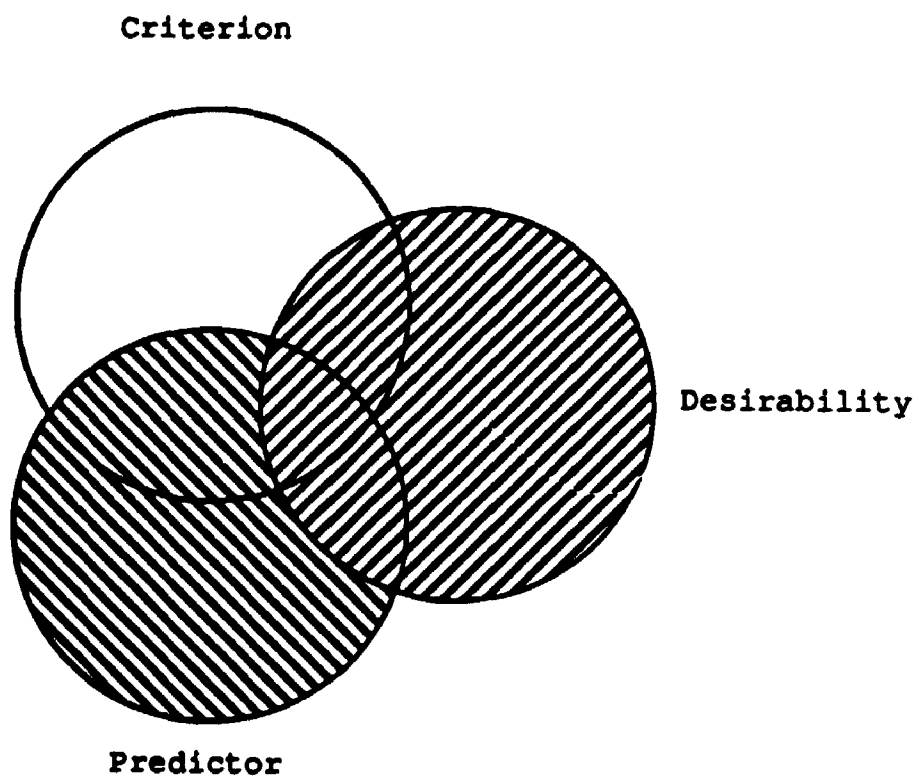
nonsatisfaction (predictor) and nonperformance (criterion) variance if also being accounted for. Research representative of this perspective often operationalizes desirability as a self presentation of worker variable, or as some moderator variable. The criticisms of "moderator variables" research outlined above are relevant here.

Desirability as Content of Worth, But Different. The proponents of this view claim that desirability responding reveals important personalogical information about the respondents, and that desirability information can be interpreted in its own right, albeit not as a component of job satisfaction. This view does not claim that desirability responding is a threat to the construct validity of the scale at hand. Primarily, desirability is viewed as part of a different research hypothesis, and as such, the conceptual description (cf. Figure 1.5) of the predictor-criterion relationship does not include desirability. The pervasiveness of the response bias is not acknowledged at all. For example, the Job Descriptive Index (Smith, Kendall, & Hulin, 1969) considers desirability to be important for personality theorists, but not job satisfaction researchers.

Desirability as a Response Bias. This last view states that desirability is a threat to the construct validity of the scale at hand, and steps must be effected at the earliest stages of item writing so as to control this response bias. Factor analytic studies of the structure of personality scales have repeatedly revealed that a small number of non-content factors have been able to account for large portions of the scale's score variance. Typically two dimensions are found, acquiescence and desirability responding (Edwards, 1963; Jackson, 1960; Jackson, 1967, 1971, 1984; Jackson & Messick, 1961, 1962a, 1962b; Messick & Jackson, 1961; Wiggins, 1962). A great deal of any scale's subscale variance will become

redundant since the scales can be represented by either a single, or a few factors. Moreover, these pervasive factors inflate the between-subscale correlations, thereby threatening the convergent and discriminant validity of the scale in question.

Figure 1.6 shows how the between subscale variance is artifactually increased, graphically shown by the overlap of the variance accounted for. Two subscales that assess diverse attitudes, but that correlate highly, will produce attenuated correlations. This is due to the presence of non-content desirability response variance. Since this third perspective defines desirability as an impediment to the validity and reliability of a predictor criterion relationship, the two previous perspectives are subsumed. It does not matter what one calls desirability responding, but it does matter what one does with a pervasive factor that artifactually inflates between-item and between-subscale correlations. The first view does not impact on the convergent and discriminant validity of the scale at hand to the same degree as the second. The Acquiescence response bias was introduced as another major responses bias. More is presented on this threat below.

Acquiescence Response Bias. Acquiescence is the tendency to answer "yes" or "true" in any assessment situation. This response bias encompasses those who habitually agree with positively keyed items and those who disagree with negatively keyed items (i.e., an item in which a "no" or "false" is keyed positively). A simple solution exists such that the threat of acquiescence can be effectively canceled from any scale. Balancing the number of true-keyed and false-keyed items will overcome this threat to construct validity, as has been

effected in the Personality Research Form (Jackson, 1967, 1984). A cursory inspection of any popular satisfaction scale now in use will reveal the near total use of true-keyed items.

The Problem of Aggregate Predictors and Single Act Criteria. Job satisfaction is typically conceptualized as a general attitude towards an act -- the act of being employed, or simply, one's work. Ajzen and Fishbein (1977) and Fishbein and Ajzen (1974) and other attitude theorists have argued that any attitude behavior correlation will be manifested only when those two variables have been similarly aggregated.

- -

Insert Figures 1.7 and 1.8 about here

- -

Figures 1.7 and 1.8 represent an ideal depiction of this "level of aggregation" concept. The venn diagrams in Figure 1.7 should be familiar to attitude and personality theorists. Figure 1.7a shows a poor mapping of one item onto a given attitude. Figure 1.7b presents the model of choice. Many terms of an attitude scale can map out the range of content. Each item is adding its contribution towards incremental validity by having low correlations with other items and variance sources and a high correlation with construct/content. Each item will a priori have a low reliability because of the highly specific variance that each variable accounts for. The scale can have a very high reliability because variance is additive. The relationship between reliability and the number of items is well known. Figure 1.8a is another venn diagram illustration, now presenting the relationship between an aggregate predictor and a single-act criterion. As shown, the area of overlap in this figure is small. The predictive validity of the

1

1.0

1.1

1.25   1.4   1.6

4.5   2.8   2.5
5.0   3.2   2.2
    3.6
    40   2.0
           1.8
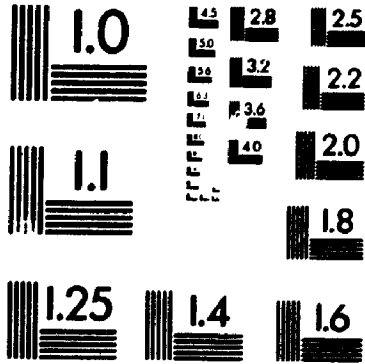
Figure 1.8 Second illustration of the "power of aggregation" concept.

A

Attitude/construct (i.e.,
aggregated predictor).
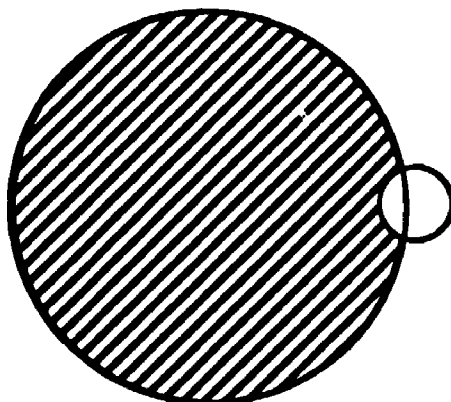
Single-act behaviour

Above, a case with a very small multiple correlation
between the single act behaviour and the construct.
This case will inevitably result in poor prediction and
definition of the construct/attitude. Below, a case
wherein the construct/attitude is more fully defined by
the many behaviours hypothesized to reflect the
aggregated criterion.

B

Attitude/construct (i.e.,
aggregated predictor).

Aggregated
Criterion

1

1.0
1.1
1.25

4.5
5.0
5.6
6.3

2.8
3.2
3.6
4.0

2.5
2.2
2.0
1.8

1.4
1.6

MicroD
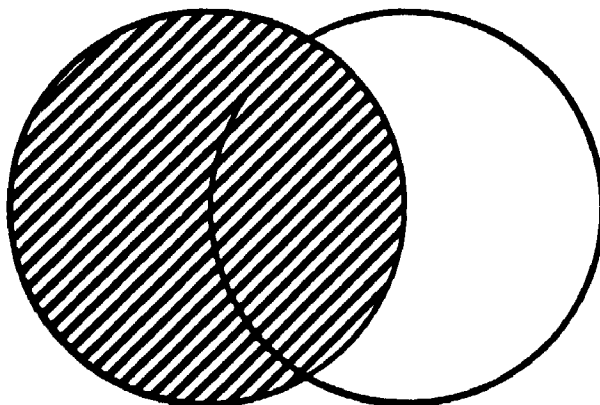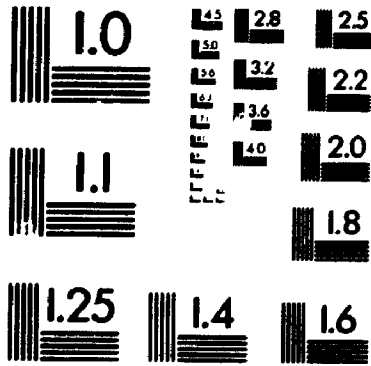
uniformly to disappointment. To summarize, the logical problem of using non-aggregated measures is endemic within psychology and must lead to extremely conservative empirical results. Certain areas of concentration recognize the merit of reliable, construct valid, and aggregated measures, but this recognition is not widespread, and certainly not within the job satisfaction literature.

Between Respondent Variation in Satisfaction Scores. Previous satisfaction theories have placed undue emphasis on simple correlations between satisfaction scale scores and particular organizational outcomes. As has been shown above, most theories have focussed on the correlation between a summed satisfaction scale score and some outcome. Once each person's score is calculated, the respondents are frequently classified as "high satisfaction", or "low satisfaction". This procedure ignores both the recognized loss of statistical power associated with dichotomous data splits, and the within-respondent between-subscale variation in scale scores. Rather than invoke a similar methodology, another approach to respondent classification has been sought in the current research. The similarity of a respondent to a target respondent is an empirical question. Simply, how are Person A's scores similar to those of Person B's? When one person is a respondent, and the other some meaningful average type, then the two sets of scores can be related meaningfully.

Jackson and Williams (1975), Jackson, Holden, Locklin, and Marks (1984), Skinner (1977), and Skinner, Jackson, and Hoffman (1975) have developed a procedure to uncover the basic structure that underlies profile similarity. Their program of research relied on existing psychometric technology, but applied it in a novel manner. This procedure involves an analysis similar to transpose factor analysis (Q factor analysis) with a univocal varimax rotation.

Modal profile analysis involves a singular value decomposition of a matrix of scores by persons. The matrix is standardized and centered in such a way that the factors obtained are identical to those that might be obtained had the intercorrelations of person's profiles been subjected to principal components analysis. In this manner, the similarity of all respondents can be assessed. These principal components reflect different profile types or "modal profiles" of respondents, and each factor loading represents the similarity between that respective respondent and a modal profile. A principal components analysis of the correlation matrix (based on the transposed data matrix) is followed by a univocal varimax rotation. The prior standardization eliminates the variables' differences in means and standard deviations and insures the variables are weighted equally in terms of variances. Few, if any, previous studies of satisfaction have attempted to classify respondents by their profile of scores.

The univocal varimax rotation seeks to ensure that each respondent has a loading on one and only one "average person" or "profile" factor. It begins with a traditional varimax rotation, (and retains the properties of a varimax rotation, e.g., orthogonal dimensions). Respondents without a substantial component loading (after rotation) are classified as not being similar to any of the modal profiles. This classificatory procedure is especially robust when the modal profiles have been developed independently of the group to be classified. In the current research, it is proposed that the modal profiles of the cross-Canada sample be generated a priori using an independent sample. This has the advantage of requiring modal profiles to be generalizable across broad segments of the population. Moreover, by using an independent sample classification scheme, respondents in the new sample can be readily classified by correlating

their profile scores with the modal profile scores derived from the previous profiles. Each person will then have a correlation with each modal profile, in a manner directly analogous to a variable having several loadings across several factors in traditional factor analysis. In the same way that variables can then be grouped into factor scales (items all loading on the same factor), respondents can be gathered into discrete groups, each group reflecting a particular modal profile. The respondents in the sample are sample values from the population of all respondents belonging to that profile.

The use of modal profiles is warranted by their: a) ability to address the between-persons factors question, and b) ability to classify respondents by their similarity to modal profiles.

What Should Be Done to Improve Satisfaction Assessment

A new measure of job satisfaction should be available for researchers so that they can collect data secure in the knowledge that their instrument is free from response biases, reflects modern scale development guidelines, (i.e., construct validity), and assesses the independent content domain(s) of interest.

Jackson and Paunonen (1985) clearly note that "all measures thought to represent psychological attributes, such as behavioral observations, frequency counts, and performance and product evaluations, are subject to the requirements of construct validation" (p. 554). The authors cite an "assessment revolution heralded by the introduction of construct validation". Yet this revolution has had very little impact on how job satisfaction has been conceptualized, assessed, or the construct applied. The ground rules for the development of construct valid measures have been laid out three decades ago (cf. Cronbach & Meehl, 1955; Loevinger, 1957). The "modern scale development guidelines" noted above are not

themselves modern, but their application is. Any new measure of job satisfaction must avoid both groundless theorizing and all of the suboptimal measurement concerns listed above. The development of such a construct valid measure is presented in the next chapter.

CHAPTER TWO: DEVELOPMENT OF THE SATISFACTION
RESEARCH QUESTIONNAIRE

A logical consequence of the research in job satisfaction would be the development of job satisfaction scales. This scale development has indeed paralleled the job satisfaction research, and has to date fostered a large number of satisfaction scales. Over 7,000 satisfaction publications have appeared to date. Some authors have suggested that the bulk of satisfaction research has developed as a partial function of the availability of published scales. The technical quality of these satisfaction measures has not been typically high. In addition, a great deal of satisfaction research has implemented single item scales that were developed ad hoc for specific applied situations.

An Outline for the Development of A New Measure.

As described in the previous chapter, many existing satisfaction questionnaires have failed to incorporate sufficient protection from response biases. The development strategy of the Satisfaction Research Questionnaire was designed to reflect superior scale construction guidelines and be impervious to response biases. The remainder of this chapter presents these construction guidelines, how they were implemented, and what the scale development program has yielded.

Development Strategy for the Satisfaction Research Questionnaire

The merit of any psychological scale rests upon its ability to demonstrate construct validity. Loevinger (1957) made clear that scale construction should begin with explicit definitions of the constructs of interest. This suggestion has become nearly synonymous with the application of construct-oriented scale

construction methods. The definitions serve as the basis of and guide for item writing and scale development. Most of the satisfaction measures cited in the literature have been developed empirically, with little or no attention paid to the substantive content of the scales. One rationale behind the construction of previous scales was the inclusion of items that investigators felt a priori to reflect feelings about the workplace. This rationale can be compared with the construct development practices that demand tractable relationships between items and their scales.

The focus of the creation of the Satisfaction Research Questionnaire items was on the use of objective behaviors associated with the components of job satisfaction that were reported in the literature. Even though "the job itself is embedded in a rich social setting which affects how people characterize and feel about their work" (Salancik & Pfeffer, 1977, p. 446), the job can still be adequately described by a smaller set of descriptors. "Studs" Terkel (1974) was correct in suggesting that many different dimensions (or aspects of the workplace) affect employee attitudes. Nonetheless, when respondents are asked to produce a list of satisfying and dissatisfying things about their jobs, the list is relatively short, and there are communalities in responses across people. Therefore, the argument of Salancik and Pfeffer (1977) that the use of facet satisfaction questionnaires generates response artifacts (via their consistency and priming effects) might be considered moot.

- - - - - - - - - - - - - - - - - - - -

Insert Table 2.1 about here

- - - - - - - - - - - - - - - - - - - -

Table 2.1 presents a critical comparison between several scale attributes

Table 2.1 Comparison of satisfaction scales and their properties.

| Scale Property | Multi-scale Satisfaction Scales | | | | |
	SRQ[1]	JDI	Q&S	NSQ	MSQ
year published	1985	1969	1965[1]	1962	1967
n of items before item analysis	1670	>100	N/A	N/A	80
n of items after item analysis	80	72	33	13	100
n of scales	5	5	6	5	20
n of items per scale	16	9 or 18	3 to 11	4 to 11	5 each
equal length scales	yes	no	no	no	yes
orthogonally developed scales	yes	no	no	no	no
balanced scales	yes	no	no	no	no
standardized norms	yes	no	no	no	no
endorsement proportions (difficulties) provided	yes	no	no	no	no
item-factor analyses available	yes	no	no	no	no
number of responses per item	2	3	4	7	5
use of difference scores	no	no	no	yes	no

Table 2.1 Comparison of satisfaction scales and their properties, continued.

Multi-scale Satisfaction Scales

Scale Property	SRQ[1]	JDI	Q&S	NSQ	MSQ
readable across levels	yes	yes	no	no	yes
acceptable internal consistencies	yes	yes	yes	yes	yes
construct development practice	yes	no	no	no	no
response bias controlled	yes	no	no	no	no
developed using explicit theory	yes	no	no	no	yes
overall or facet satisfaction	both	both	both	both	both

Notes.

[1] SRQ - Satisfaction Research Questionnaire (Hill, 1985c), JDI - Job Descriptive Index (Smith, Kendall, & Hulin, 1969), Q&S - Quinn and Staines (1969), NSQ - Need Satisfaction Questionnaire (Porter, 1961,1962), MSQ - Minnesota Satisfaction Questionnaire (Weiss, Dawis, England, & Lofquist, 1967).

[2] The facet-specofoc satisfaction questionnaire were based on surveys completed in 1969 and 1973. The 1977 survey did not represent any attempt at the generation of new items.

for four popular scales, and the scale attributes of the Satisfaction Research Questionnaire. Each scale attribute is briefly described below, with a statement of how each scale attribute will contribute to or detract from a scale's construct validity.

Year Published. Inspection of these row entries readily illustrates an interesting historical point. As most of the other scales have made extensive use of previously published satisfaction items, the year publication is a liberal estimate of the age of each of the original item pools. A reason for including the date of publication is that the workplace is not a static phenomenon. Items that are two or three decades old may no longer reflect the modern workplace. The four extant scales do not reflect any recent attempt at item development.

Number of Items Before Item Analysis. In order for a scale developer to ensure that the final version of a scale retains only those items that maximally benefit that scale, a large initial item pool is required. As is the case in some of the scales in Table 2.1, some scale developers began with the assumption that the items they initially generated were reliable and related to their construct. Having a large initial item pool affords the scale developer the chance of deleting large numbers of poor items while retaining a scale of sufficient length to generate an acceptable reliability.

Number of Items After Item Analysis. A comparison of the number of items after item analysis indicates that only one satisfaction measure has a high ratio of initial items to final items. Simply put, this evidence suggests that the recent measure has probably been successful at deleting suboptimal items. A scale that finished with more items than it began with clearly has had no opportunity to remove such suboptimal items.

Number of Subscales, Items Per Subscale, and Equal Length Subscales.
The number of subscales that are contained in a scale reflects the bandwidth-fidelity dilemma. A researcher can either collect extremely reliable information about a narrowly defined content area, or collect somewhat less reliable information from many diverse content areas. Moreover, the final use of the scale must guide the number of constructs to be assessed. If a scale is to be used for collecting individual data, and not group data or aggregate data, then reliability must not be sacrificed. The number of subscales in the Satisfaction Research Questionnaire was based on these considerations. Three of the other scales have a similar number of subscales, with the exception of the Minnesota Satisfaction Scale, with twenty scales. Each Satisfaction Research Questionnaire subscale has sixteen items, as opposed to the earlier scales that have a varying number of items per subscale. The presence of unequal length subscales puts the scale user at risk of unequally representing the constructs at hand. Additionally, since subscale reliability is partially determined by the length of the subscale, such facet measures will often have different reliabilities.

Orthogonally Developed Subscales. The degree to which the subscales of a scale were developed to be independent measures of their respective constructs bears on several issues. Orthogonal scales preserve the researcher's ability to speak of pure constructs, and these "clearer" measures permit more accurate presentations of profiles of subscale values. Subscales should have minimal correlations with other subscales. At the item level, items that are minimally correlated with other items have a better chance of correlating with their own total score. The construct validity of the scale would be attenuated if the subscales had artifactually high intercorrelations. In addition to the subscales

losing their ability to map out the appropriate content domains, inflated between-subscale correlations force the loss of any differential validity that the subscales may have possessed. Orthogonally developed scales aid in the conceptualization of independent facets in a multi-facet model. Moreover, subsequent analyses, (e.g., regression using principal components scores) are further aided by the presence of orthogonally developed subscales.

Balanced Subscales. Do the scale's subscales have equal numbers of positively-keyed and negatively-keyed items, and are there equal numbers of positively and negatively phrased items? If not, then the respondent scores may be partly due to the acquiescence response bias, where respondents simply respond in the keyed direction most of the time. Since in the past large portions of scale respondents have claimed to be very satisfied with their work, the charge of acquiescence response bias probably deserves to be investigated, or at least recognized within this literature. Hill (1985b) has shown that a large portion of the response variance of one of the major satisfaction scales in Table 2.1 can be traced to the acquiescence and social desirability response biases. The lack of orthogonal subscales and the absence of balanced subscales can lead to such response styles that distort the true score in such self-report situations.

Standardized Norms. So that the end-user of the scale information can make an accurate interpretation of scale scores, scale score norms must be provided. In the absence of such norms, interpretations could become meaningless. The absence of large-scale normative sample information would be analogous to administering an intelligence scale without having estimates of the population mean and standard deviation. The norms of a satisfaction measure must also reflect the current diversity of the workplace. The norming procedure

should therefore have been a recent endeavor, or at least one that continuously updates the normative database. Of all of the scales in Table 2.1, only one can lay claim to a recent cross-country data collection and a continuously updated normative database.

Endorsement Proportions Provided. The endorsement proportion of an item directly indicates the frequency with which an item is endorsed in the keyed direction. An item that has an endorsement proportion of 0.50 has the best chance of discriminating between respondents. On the other hand, an item that has an endorsement proportion of .99 would be practically useless since everyone would endorse that item. As described below, the average endorsement proportion of a scale should be about 0.50, reflecting an ability to maximally discriminate among respondents. The item endorsement proportion should be viewed as a relevant part of the item analysis procedure. Only the recent satisfaction measure has incorporated this attribute into its scale development. In the absence of incorporating this information into the scale development, the scale author(s) should provide endorsement proportions to demonstrate the items' abilities to discriminate among respondents. None of the four major extant scales has provided such information.

Item-Factor Analyses Available. An empirical demonstration that the scale actually measures its advertised constructs can be afforded by subjecting the items to an item-factor analysis. This step is crucial when the scale developer has claimed that the scale was developed along factor-analytic guidelines. For example, Quinn and Staines (1977) have claimed that their scale was developed along such lines. The authors, however, have not provided any further reference to these analyses and users of the scale have not had access to the empirical information that would validate the authors' claim.

Number of Responses. Another major decision that a scale constructor must face is the number of response categories to allow per item. Traditionally, scales have used a two category (agree-disagree, true-false) response format for ease of administration and scoring. The other popular format is a Likert-type response format. Hill (1985a) has shown that item response formats that use an even number of response categories (e.g., two and six response categories) are useful in maintaining the factor analytic structure of that scale. That is, a scale administered with two response options will most likely have the same structure as the same scale administered with six response options. This does not hold true with an odd number of response options. A seven response format that contains a neutral or "uncertain" middlemost option most probably will not replicate the original scale structure. In the odd number of options case, the response determinants behind the variance in the "neutral" responses are rarely the same as behind the variance of other responses. Few of the scales in Table 2.1 make use of a small, even number of response options.

Use of Difference Scores. One theory of job satisfaction assessment dictates that a respondent should indicate the current level of satisfaction they perceive with some facet of their work, and then rate their level of satisfaction with the "ideal" facet. The difference between the two scores is thought to indicate the actual level of satisfaction that the respondent feels. It is noteworthy that only one of the scales in Table 2.1 makes use of such a difference score. The psychometric problems associated with difference scores have been outlined above, but the use of such measures continues.

Readable Across Levels of Ability. To answer the scale, respondents must be able to read the items. Therefore, the use of simple and relevant items is of paramount importance, and again, this is an item attribute that should be present

at the earliest level of item creation. When items have been created to be generalizeable across occupations and levels of reading ability, the resultant measure can be applied to a diversity of work environments. Both the Quinn and Staines measure and the Need Satisfaction Questionnaire have been criticized as being only appropriate for white collar workers.

Internal Consistency. Another cornerstone of scale development has been to ensure that the scale and its subscales are maximally reliable. All of the scales in Table 2.1 can lay claim to good reliabilities. However, subscale reliability is a necessary, but not sufficient, condition to guarantee superior measures. The validity of any given measure can be low, even if the measure has an acceptable internal consistency.

Construct Development Practice. A review of the literature of the four extant satisfaction measures has indicated that none of these measures has invoked a construct centered scale development practice. The majority of the items on these measures were culled from previous inventories and adjective checklists.

Response Bias Controlled. Central to the validity of any measure is the demonstration of freedom from response biases. Such response biases will inflate the between-subscale correlations, decrease the validity of the measures, and inflate the error score component of any resultant observed score. In the four extant measures, there is no mention of controlling for the effects of either social desirability responding or acquiescence response bias, whereas in the recent measure, response bias control was a central aspect of the scale development.

Developed Using Explicit Theory. Test development must begin at the level of the definition of the relevant constructs and then proceed through a series of planned empirical and rational item and scale development stages (cf.

Jackson, 1971; Loevinger, 1957). Others have argued cogently for the importance of the construct in psychological scale development, and it was this approach to scale construction that was applied to the first scale in Table 2.1. None of the other four scales in Table 2.1 was specifically designed from the beginning to foster construct, convergent, and discriminant validity while minimizing the impact of response biases. Of the four extant scales, only the Minnesota Satisfaction Questionnaire had an explicit theory of scale development.

Overall or Facet Satisfaction. Both applied and academic audiences have assessed satisfaction as a single continuum and as a series of discrete facets. All scales in Table 2.1 consist of five or more subscales, and all scale authors recommend using these facet scores for diagnostic, descriptive, and prescriptive purposes. However, the four extant satisfaction measures typically also recommend summing these correlated subscale scores to produce an aggregate total score for the same purposes as the facet or subscale scores. As presented above, there may be little utility in expecting a general measure, comprised of related facets, to possess any differential validity. The description of the contents of Table 2.1 has served to describe some of the properties of the major satisfaction inventories. The following section provides a prescription for the development of a recent measure of satisfaction, a prescription that was followed in the development of the Satisfaction Research Questionnaire.

How to construct a "modern" scale

The above description of fundamental scale attributes should indicate the importance of recogniz⁻ g as many threats to construct validity as possible. This section describes the scale development program invoked for the construction of the Satisfaction Research Questionnaire.

Insert Table 2.2 about here

Literature Reviews for Constructs of Interest. One aim of this research was to investigate the degree to which the application of modern scale construction principles could produce a job satisfaction scale that showed appropriate psychometric properties. In short, the premise was that the accurate assessment of popular constructs would better serve to advance the field than would the development of a new job satisfaction "theory" or some novel scale developed with suboptimal construction strategies. To this end, the job satisfaction literature was extensively reviewed. It was found that both job satisfaction theorists and applied researchers supported a small number of job satisfaction constructs -- usually portrayed as satisfaction facets. These constructs described what the respondent might be aware of, or interact with, in the work environment. Such constructs were: the co-workers that one has, the supervisors one works under, the people who are below one's level in the organization, the type and amount of benefits that the organization offers, the status one has in the organization, the amount of pay one receives for the work one does, whether or not one feels equipped for work, the availability, or perceived probability, of a promotion, the perceived equity of organizational rewards, the physical working conditions, and other dimensions of the workplace.

Definition of Constructs. Jackson (1971) stated that an aspect of the development of a new scale is the substantive component of validity, of which construct validation is an integral part. Twelve item writers (students enrolled in an undergraduate psychometrics course) reviewed the satisfaction subscales of

Table 2.2 Rules used to develop the Satisfaction Research Questionnaire.

Step 1.   Derive accurate definitions of relevant constructs and scales.

Step 2.   Review constructs until judges reach consensus as to definitions.

Step 3.   Write items to reflect both positive and negative behavioral exemplars of each construct. Avoid ambiguous or extreme items. Write items in the first person singular, with equal numbers of positively and negatively keyed items.

Step 4.   Review all items for their content, relationship to scale, and grammar.

Step 5.   Select a subset of items (1,000) that can be completed by subjects in a two hour testing session. (The first 1,000 items from the 1,670 item pool were selected.)

Step 6.   Have subjects with employment experience complete the questionnaire and the Personality Research Form Desirability subscale. An eight item measure of general cognitive ability was administered for screening purposes.

Step 7.   Delete all incomplete respondent data, or data that suggest random responding, or that fall below the criterion value of the ability measure. Number of valid cases is 119. Begin item analysis.

Step 8.   Create two 500 by 500 variable by variable matrices for item analsysis. Subject matrices to principal components analyses.

Step 9.   Retain those items with loadings >= 0.30 on their one construct (factor) only. Rank items on the basis of their loadings (i.e., content saturation).

Step 10.   Rank items by their item endorsement proportions (difficulties) and reject those items with extreme values.

Step 11.   Rank items by their Differential Reliability Index values, and reject poor items.

Step 12.   Review scales for equal numbers of positively and negatively loading items to ensure balanced scales.

Step 13.   Review items again for construct relevance.

Step 14.   Final editorial review. Average scale reliability was .87, reliability of total scale was .93.

Step 15.   Generate modal profiles. Begin cross-Canada mail-out.

previous instruments to determine which facets had been of primary interest to researchers. It was found that the previous research had focussed on the following facets of the workplace: a) the challenge that the work provides, b) the people that one works with, c) the feeling of comfort one experiences in the workplace, d) the people one works for or under, e) the remuneration one receives for work, f) having the tools, information, or training for the job, and g) the supervisors that one works under. The original definitions of these subscales (constructs) were reviewed by the item writers, and revised until a consensus was reached about the construct definitions. In all, six constructs survived this review process. They were: a) Challenge, b) Comfort, c) Co-Workers, d) Pay, e) Resource Adequacy, and f) Supervisors.

A high score on the Challenge subscale would indicate a worker who felt that his or her job provided them with sufficient challenge, provided oppor nities to display their workplace abilities, and felt that their job had continued to capture their interest. Someone who had a high score on the Comfort subscale would be a worker that reported feeling free from worry and anxiety while at work, and was thus able to work more efficiently. Such a person would report that they enjoyed being in the workplace itself. An individual with a high score on the Co-Workers subscale would report having a good working relationship with others in the workplace, possibly having several co-workers as social friends. A high Pay subscale score would indicate respondents that felt they were receiving an appropriate level of remuneration and workplace status for the work that they do. A high Resource Adequacy score would indicate that the worker felt prepared to do their job, in terms of information, equipment, and

training. Lastly, a person with a high Supervisor score would indicate satisfaction with those that are responsible for the supervision of the job tasks performed by the respondent.

Item Writing. Each item writer was instructed to generate as many items as possible. Each item was designed to: a) reflect its own scale content (i.e., construct), b) be phrased in the first person singular, c) be a concrete behavioral exemplar, d) avoid ambiguous or "wordy" styles, e) use simple vocabulary, f) use generic (i.e., generalizable) terms for equipment, machines, etc., g) avoid the use of conjunctions that would confound the item construct relationship (e.g., "I enjoy the challenge of my job and my pay"), and h) avoid content that almost everyone would endorse or not endorse (e.g., "I like to get paid for my work" and "I would like to be fired immediately", respectively). Items were also written to reflect the "dissatisfied" pole of each satisfaction continuum to map out more fully relevant variance. Also, items were positively and negatively worded so that the acquiescence response bias might be controlled for by having equal numbers of "agree" and "disagree" items in the final version of the scale. Complete items were then subjected to editorial review and poor items (i.e., poor grammar, structure, or duplicates) were either revised or deleted. In total, 1,670 items were eventually completed.

Data Collection. Employed subjects and subjects with a recent work history were tested in groups of 5 to 16. Each subject was allotted 2 hours to complete the items, and extra time was provided for those subjects who required it. Some subjects participated as an introductory psychology course requirement, and some participated on a voluntary basis. Subjects could easily complete the first 1,000 of the 1,670 items. In all, 135 respondents completed the questionnaire

consisting of the satisfaction items, the Personality Research Form E Social desirability scale, and an 8 item general cognitive ability scale. Of these respondents, 119 complete cases were retained. Other cases were deleted because of missing data in the principal measures, or did not pass the general cognitive ability scale.

Commencement of Item Analyses. After the completion of the item pool and the data collection, a planned series of item analyses was begun. This sequential plan was designed to select items that maximally reflected their construct and minimally represented any response bias. To this end, items were reviewed for their content saturation (item-component loadings), ability to maximize reliability (item endorsement proportions), and their freedom from the desirability response bias (Differential Reliability Index values).

Principal Components Analysis with Varimax Rotation. A series of exploratory principal components analyses was undertaken to assess the dimensionality of the item pool. (Due to computer hardware constraints, it was not possible to factor analyze the full 1,000 by 1,000 item correlation matrix. Instead, an odd item 500 item correlation matrix and an even item numbered correlation matrix were generated. The discussion below will treat the analyses as a single 1,000 by 1,000 principal components analysis. All factors from all analyses were compared with their odd- and even-item counterparts, and thus the split halves were replicated.) Solutions ranged from a 1 factor general solution to a 10 factor solution to determine whether there was any empirical evidence of factors/subscales collapsing into a smaller number of similar dimensions, or bifurcating to produce a larger number of discrete dimensions. It was found that a 5 factor solution best reflected the original subscale/construct definitions, even though 6 dimensions were originally hypothesized.

These factors were: a) Challenge, b) Comfort, c) Co-Workers, (but only reflecting those co-workers at the same or lower level in the employment hierarchy), d) Pay, which was interpreted as a fairness of pay, recognition, and stature in the workplace, and e) Resource Adequacy, wherein supervisors were seen more as providers of resource material (e.g., training, information, and materials) than co-workers. These analyses provided strong empirical support for the original subscale definitions. While the sample size of 119 might be considered too small for such an item-factor analysis, there is empirical evidence to the contrary. Guadagnoli and Velicer (1988) have shown that the critical issue in determining the stability of principal component loadings for a given sample size is their salience. Items that are written and edited to have high saturations on the constructs for which they were keyed in general will have large component loadings. The traditional "rules" regarding the ratio of number of subjects to number of variables appear to be based on a misconception. Using the Guadagnoli and Velicer (1988) formula for sample size, the sample size was found to be sufficient.

Ranking and Deletion of Items by Their Component Loadings. Each item in the five factor solution was reviewed for its item-component loading with its own factor/subscale and all other factors. Those items that uniquely reflected their own content were retained. Items with multiple loadings were deleted. The cut-off value in the item retention decision was an item component loading of .40 or greater. Items were gathered into their own subscales on the basis of this examination of item homogeneity.

Ranking and Deletion of Items by Their Endorsement Proportions. Each item's mean endorsement proportion was calculated. Items that elicit responses in the keyed direction from all respondents do not discriminate between

respondents. That is, any item that has zero variance would not yield any differential information and would not contribute to the reliability of the subscale. Items with extreme endorsement proportions were therefore deleted. Importantly, there were very few items with extreme endorsement proportions. A secondary concern frequently raised during discussions of deletion of items on the basis of endorsement proportions is that some zero variance items might yield important information in that they signal a unanimous response in the sample. For example, advocates of this position might argue that it would be important to know that the entire sample was displeased with their rate of pay. This information about the consensus would be valuable when considering an application of the completed instrument, but is secondary when the instrument is being constructed.

Ranking and Deletion of Items by Their Freedom From Desirability. Jackson (1984) presented an item index designed to indicate the degree to which an item's content-related variance relates to its desirability-related variance. Jackson (1984) has asserted that:

> The requirement that items correlate highest with their proper scale, and particularly, the requirement that this relationship not be substantially a function of desirability bias, tends to produce finished scales with relatively high reliability and discriminant measurement properties. Scales constructed in this way tend not to be spuriously correlated due to desirability. In the absence of such spurious correlations, optimal conditions exist for discriminating among individuals on the basis of their different scale scores, and for making inferences regarding test patterns. (p. 31.)

The Differential Reliability Index (DRI) is the square root of the squared item-scale correlation minus the squared item-desirability correlation. As the above excerpt shows, attention to item development and item testing yields benefits in terms of overall quality of the finished instrument, as well as the

quality of decisions made regarding the similarities of patterns of respondent scores. Therefore, for each item, the Differential Reliability Index value was computed, items were ranked on their index values, and poor items were deleted.

Final Review of Items for Construct Relevance. After all of the above item analyses were completed, the items were reviewed yet another time. This final review was primarily to guarantee the construct relevance of each item, but also included checks for: a) grammar and spelling, b) item construction, c) equal numbers of positively and negatively keyed items, d) reliability of the subscales, and e) the readability of the subscales, which is described below. After this final review, the Satisfaction Research Questionnaire consisted of 5 subscales, consisting of 16 items each, with each subscale having 8 true-keyed and 8 false-keyed items. (Sample items are contained in Appendix A.)

Readability Analyses. The Satisfaction Research Questionnaire was subjected to a readability analysis using a commercially available software package. This procedure was employed so as to generate an empirical assessment of the reading level required to understand and respond to the Satisfaction Research Questionnaire. The program provided a summary containing several indices which are described below.

A) Readability Index. The Readability Index ranges from a low of 1.0 (first grade level) to a maximum level of 50.0 (unreadable). A value of 12 or greater would indicate that the material is suited for college educated readers, as well as suggesting that the material is written in a complex manner and is difficult to understand. The Readability Index value for the Satisfaction Research Questionnaire was 4.42, indicating that the scale respondents would require a fourth grade level of education or better.

B) **Strength Index**. The Strength Index measures the strength of delivery of the document's message, or more simply, the frequency of use of unnecessary qualifiers, uncommon words, or complex sentence structure. A desirable aspect of a scale intended for broad application would be a good Strength Index value. A Strength Index value of 0.00 would indicate a weak wordy writing style, whereas reports and technical papers should have a value of 0.50 or greater. Manuals and business letters should have a recommended value of 0.80 or greater. The Satisfaction Research Questionnaire Strength Index value was 0.65, indicating an absence of wordy items, or items that are vague. This value places the Satisfaction Research Questionnaire Strength Index value between an average report and an average business letter. an acceptable value for those who are in the workplace.

C) **Descriptive Index**. This index measures the use of adjectives and adverbs in a given piece of text. The presence of these modifiers alters the meaning of nouns and verbs, and a too frequent implementation of these modifiers detracts from the clarity of the text. The range of values for this index starts from 0.10 (the writing is terse and choppy) to 0.90 (the writing is too descriptive). The Satisfaction Research Questionnaire Descriptive Index value was 0.65, indicating that the use of adjectives and adverbs is within the normal range.

D) **Jargon Index**. Since jargon is a direct block to communication, and a threat to the ability of respondents from diverse occupational backgrounds to understand scale items, the Satisfaction Research Questionnaire was subjected to this last readability test as well. A value of 0.00 for this index would indicate the absence of jargon, or, at most, the presence of jargon that is within acceptable levels. A value of 0.50 or greater would suggest that people outside of a

particular area of subject matter concentration would have difficulty understanding the content. The Satisfaction Research Questionnaire value for this index was 0.00, strongly suggesting that workers from diverse backgrounds would have little, if any, difficulty in comprehending the meaning of the items.

Reliability of the Satisfaction Research Questionnaire Subscales.

In the derivation sample (n=119) the Satisfaction Research Questionnaire total score reliability was .93. The reliabilities for each Satisfaction Research Questionnaire subscale were: Challenge, .85; Comfort, .81; Co-Worker, .96; Pay, .86, and Resource Adequacy, .87. The mean alpha reliability for the five Satisfaction Research Questionnaire subscales was .87. Overall satisfaction with one's work can be viewed as a reliable construct that is itself comprised of several minimally correlated and highly reliable dimensions. The ideal mean item endorsement proportion for a scale is 0.50, this being the value that maximizes a scale's reliability (Cronbach & Warrington, 1952). The mean item endorsement proportion for the Satisfaction Research Questionnaire was 0.52.

- -

Insert Table 2.3 about here

- -

Structural Fidelity of the Satisfaction Research Questionnaire

One of the aims of the development of the satisfaction measure was the development of independent measures of the facets of satisfaction. To determine empirically whether this goal of subscale independence was met, the 80 by 80 matrix of Satisfaction Research Questionnaire item correlations was subjected to a principal components analysis. Table 2.3 presents the factor pattern correlations from a principal components analysis of the Satisfaction Research Questionnaire

Table 2.3 Factor pattern correlations from a principal components analysis of the Satisfaction Research Questionnaire subscales. (Five dimensions extracted with direct oblimin rotation, delta = 0.0.) Below, the subscale correlations from the normative (n=119) sample.

Factor Pattern Correlations

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1.00	-.11	+.15	-.21	+.15
Factor 2	-.11	1.00	-.15	+.24	-.01
Factor 3	+.15	-.15	1.00	-.09	+.08
Factor 4	-.21	+.24	-.09	1.00	-.13
Factor 5	+.15	-.01	+.09	-.13	1.00

Scale Correlations

	Challenge	Comfort	Co-Workers	Pay	Resource Adequacy
Challenge	1.00				
Comfort	.12	1.00			
Co-Workers	.27	.24	1.00		
Pay	.35	.26	.23	1.00	
Resource Adequacy	.28	.33	.22	.28	1.00

scales. Five components were extracted, and rotated with a direct oblimin rotation, delta = 0.0, to permit a non-orthogonal final solution. The factor pattern intercorrelations were low, all less than an absolute value of 0.25. The average of the off-diagonal elements of this matrix was approximately -0.01. The lower half of this same table presents the between-subscale correlations for the same sample. Again, intercorrelations were low. The normative intercorrelations for the Job Descriptive Index, however, were as high as .52 (Smith et al., 1969, 77-78).

Freedom from Desirability Responding in the Satisfaction Research Questionnaire

Several average values were computed to ascertain whether or not the Satisfaction Research Questionnaire scales and items were relatively free from desirability responding. These analyses were based on the derivation sample. The mean correlation between the five Satisfaction Research Questionnaire subscales and the Personality Research Form desirability scale was +0.11, and the mean correlation between an item and the Personality Research Form desirability scale was +0.06. Jointly, these two mean correlations provide support for the hypothesis that desirability responding has been controlled in the Satisfaction Research Questionnaire.

- -

Insert Table 2.4 about here

- -

Absence of desirability responding is a necessary, but not sufficient, condition for the acceptance of construct validity. Each item must have its highest correlation with its own scale, and each item must have a low portion of its response variance associated with desirability responding. Table 2.4 provides

Table 2.4 Satisfaction Research Questionnaire mean item information.

0.11	Mean correlation between the Satisfaction Research Questionnaire subscales and the Personality Research Form Desirability scale.
0.55	Mean correlation between a Satisfaction Research Questionnaire item and its own scale (i.e., content saturation).
0.06	Mean correlation between a Satisfaction Research Questionnaire item and the Personality Research Form Desirability scale.
0.55	Mean Differential Reliability Index Value for the Satisfaction Research Questionnaire items.
0.52	Mean item endorsement proportion for the Satisfaction Research Questionnaire items.
0.33	Mean between-scale correlation for the Satisfaction Research Questionnaire scales.

the mean item information from the Satisfaction Research Questionnaire. The mean correlation between a Satisfaction Research Questionnaire item and its own scale was 0.55, and all items had their highest correlation with their own scale. The mean Differential Reliability Index value for the Satisfaction Research Questionnaire items was 0.55, strongly suggesting that, on average, the content saturation of the Satisfaction Research Questionnaire items was not heavily influenced by desirability responding.

Modal Profiles of Satisfaction Research Questionnaire Respondents

One approach to the classification of respondents by their satisfaction scores has been to divide the respondent group(s) into those respondents with above average and below average scores.

- -

Insert Table 2.5 about here

- -

Table 2.5 shows that each of the Satisfaction Research Questionnaire subscales is reliable, and has a relatively narrow confidence interval about its respective mean. Indeed, the standard errors about the mean subscale scores are less than one standard deviation. Ordinarily, researchers would dichotomize these subscale scores in an attempt to classify respondents as high or low on that subscale. These robust subscale characteristics might be interpreted as evidence for such a dichotomization of scores in that the high reliabilities and narrow standard errors would reduce the misclassification and loss of information that accompany dichotomization. However, one aspect of this research was the empirical discovery of average or "modal" patterns of satisfaction in respondents, and not the traditional, narrow, classification of the past. (Even with acceptable

Table 2.5 Scale means, standard deviations, reliabilities, standard errors of measurement, and the 95% and 99% confidence intervals for the Satisfaction Research Questionnaire (SRQ) subscales.

Scale	Mean	Sd	r^{xx}	SE^M	95%CI	99%CI
Challenge	9.09	2.21	.85	0.85	1.76	2.20
Comfort	7.89	1.78	.81	0.78	1.52	2.00
Co-Worker	7.37	1.97	.96	0.40	0.77	1.02
Pay	8.01	2.04	.86	0.76	1.49	1.97
Resource Adequacy	7.82	1.87	.87	0.68	1.32	1.75
Total Score	40.18	5.10	.93	0.14	2.65	3.48

reliabilities and standard errors, dichotomization would still force a loss of information.)

Using the modal profile analysis procedure described in the first chapter, two bipolar modal profiles were extracted. Various numbers of modal profiles were extracted in an exploratory manner. One solution was to first assign all respondents to a general profile, and then repeat the modal profile analysis for each pole of this bipolar construct. Other solutions examined the series of (bipolar) profiles that emerged when various numbers of modal profiles were extracted. It was decided that two bipolar modal profiles would be retained. This solution met the demands of parsimony, and was consistent with a variety of numbers of factors rules. The number of respondents that could be classified, the classification efficiency, the total percent of variance explained by the modal profiles, and the percent of variance explained for subjects that met the acceptance (epsilon entrance) criterion were other factors that guided the number of profiles decision. With a large number of profiles, all subjects can be classified, but the aim of dimension reduction would be lost. The two bipolar modal profile solution yielded a classification efficiency of 79.8 percent, indicating that nearly 80 percent of the sample could be accurately classified.

The first modal profile is notable for its extremely elevated Pay score, (nearly two standard deviations above the mean).

- -

Insert Figures 2.1 through 2.4 about here

- -

As Figure 2.1 shows, the modal profiles utilize standard scores with a mean of 50 and a standard deviation of 10. The first modal profile can also be

Figure 2.1 Modal Profile One. (Low Challenge and High Pay profile.)

Challenge

Comfort

Co-Workers

Pay

Resource
Adequacy

```
30      40      50      60      70
```

SRQ scale scores (T-scores)
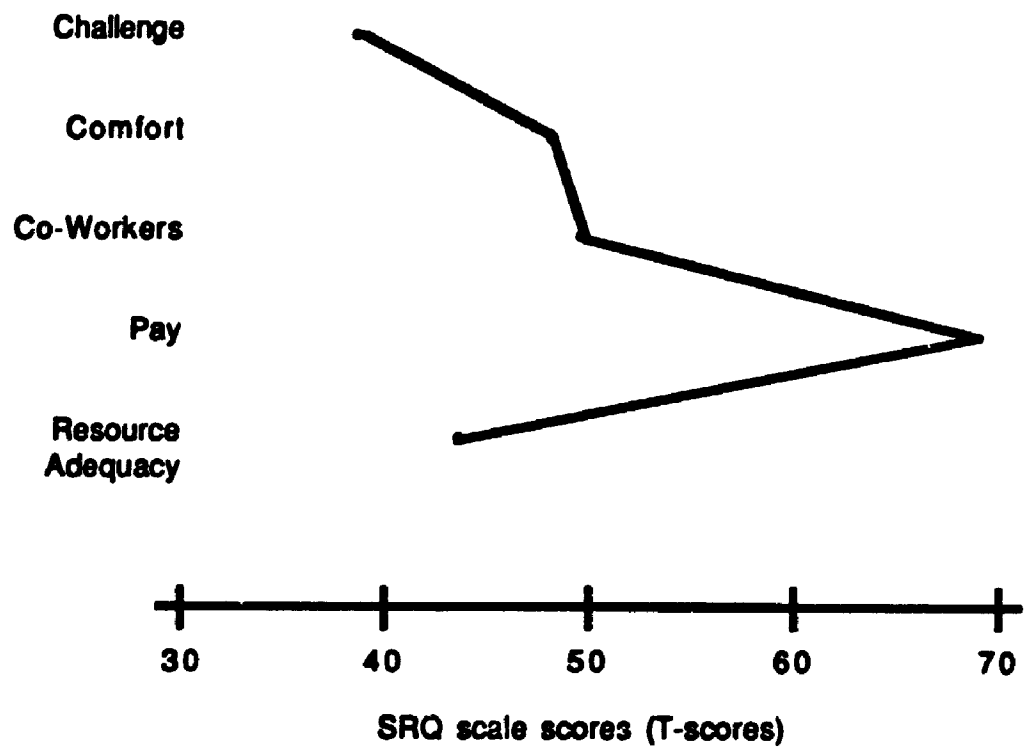
Figure 2.2 Modal Profile Two. (High Challenge and Low Pay profile.)

Challenge

Comfort

Co-Workers

Pay

Resource
Adequacy

30   40   50   60   70

SRQ scale scores (T-scores)
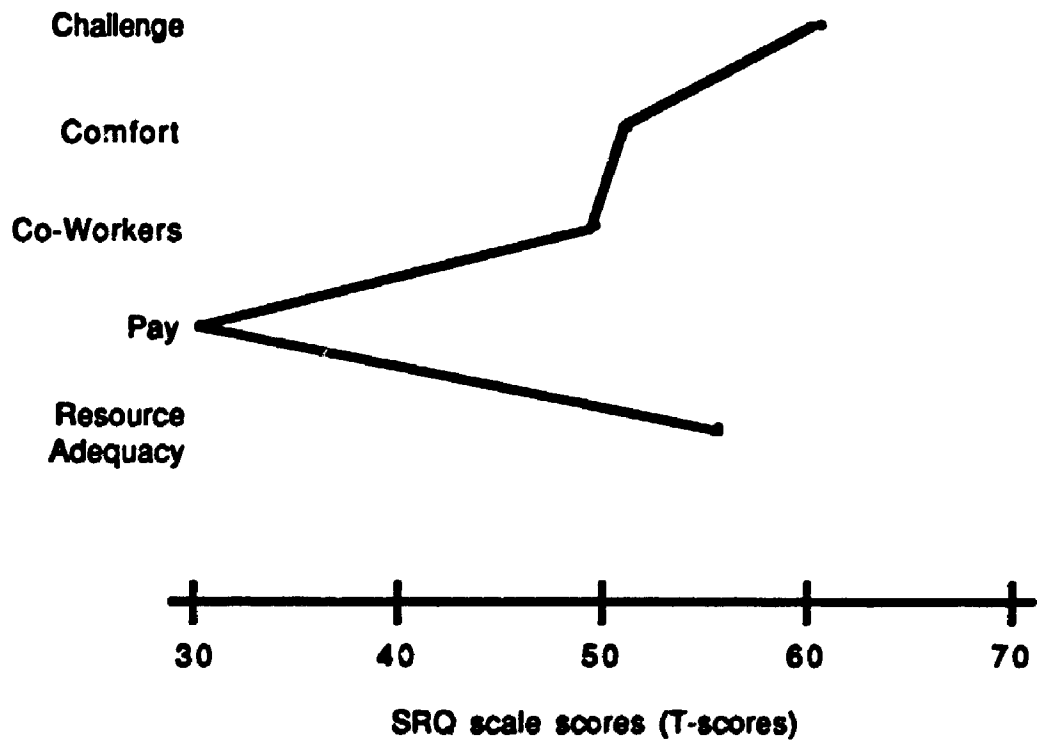
(aggregate) predictor is low as the criterion is a) not similarly aggregated, b) a (time or event) specific behavior, and c) not reliable. Figure 1.8b presents the theoretical increase in both predictive validity and utility when the criterion is a) similarly aggregated, b) a general disposition or intention to behave, and c) reliable. Psychometrically, ᴛhe similarity of level of sophistication of measurement is well known. Unfortunately, few Industrial and Organizational psychologists are aware of this. (But cf. Fisher (1979) for an example of the "pro-aggregate criterion" viewpoint from the satisfaction domain. See also, Rushton, Brainerd, & Pressley (1983) for a developmental example of the merit of aggregation, and Jackson & Paunonen (1985) for an argument for the use of aggregation in the personality domain).

Reliability, Construct Validity, and Aggregation. Each of these three scale attributes is dependent on the other two. Reliability determines the upper bound for validity, and aggregation increases reliability. If there are several heterogeneous constructs being treated as a single unity by using a single total score, then aggregation will reduce reliability (Cronbach & Meehl, 1955; Loevinger, 1957). The predictive validity of the intellectual ability construct is well documented. The assessment of intellectual ability provides an illustrative example of the aggregation concept. Green (1978) calculated that the mean reliability of an item from a typical measure of intellectual performance (the SAT) was only .18. The (aggregated) scale reliability was .95. Muchinsky (1977) reported that the reliability of absence measures that were typically single event measures ranged from .00 to .74, with an average of .40. Tryon (1973) noted that even in the intelligence and scholastic aptitude domains, expecting any discrete item to bear a high relationship to any predictor (or criterion) would lead
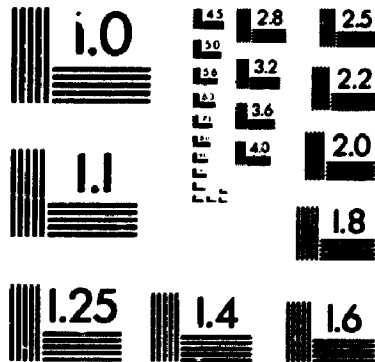
Figure 2.3 Modal Profile Three. (Low tertiary coronary heart disease symptoms profile.)

Challenge

Comfort

Co-Workers

Pay

Resource
Adequacy

SRQ scale scores (T-scores)

30 40 50 60 70
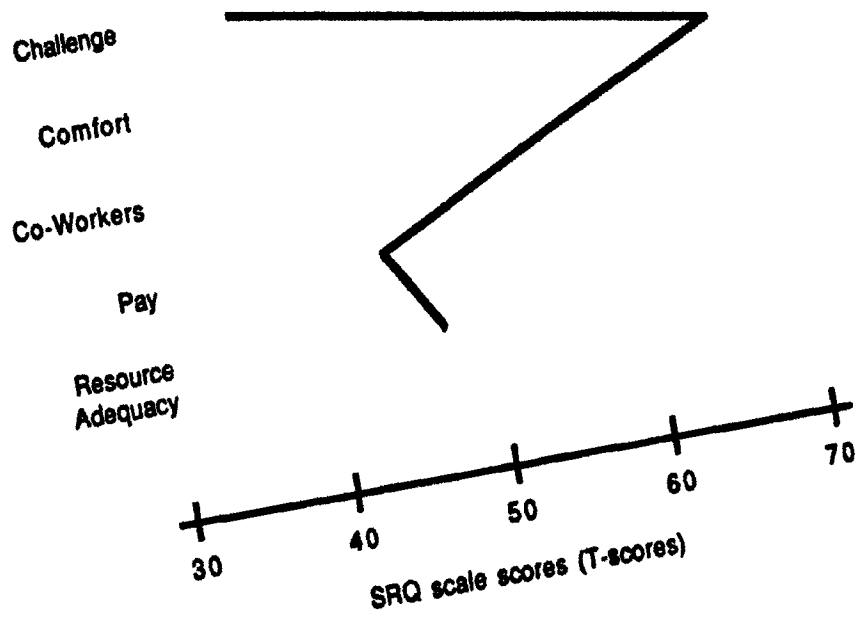
Figure 2.4 Modal Profile Four. (High tertiary coronary heart disease symptoms profile.)

Challenge

Comfort

Co-Workers

Pay

Resource
Adequacy

30   40   50   60   70

SRQ scale scores (T-scores)
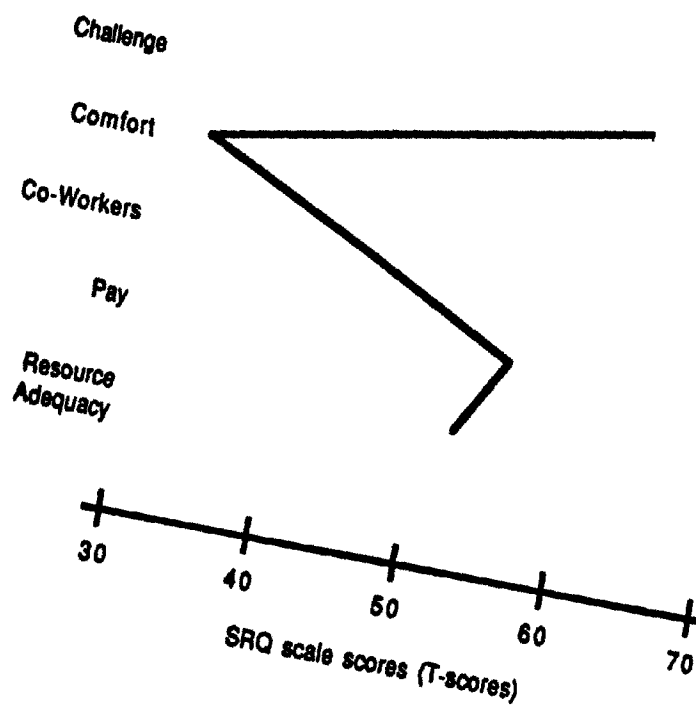
recognized by its near-average Comfort and Co-Workers scores. The Challenge and Resource Adequacy scores are somewhat depressed. This profile can be interpreted as a "good pay, but lackluster job" profile. Twenty-two percent of the respondents fell into this profile.

The second modal profile (Figure 2.2) is the opposite pole of the previous profile. Now, the Pay scale is nearly 2 standard deviations below the mean, while the Challenge and Resource Adequacy scales are above the mean. The Comfort and Co-Workers scales are still about average. This profile accounted for some 20 percent of the respondents. This profile can be interpreted as a "has a challenging job that does not pay well" profile.

The third profile (Figure 2.3) has a markedly depressed Challenge score and a notably elevated Comfort score. The Co-Worker, Pay, and Resource Adequacy scales are average. The combination of the low Challenge score and the high Comfort scale suggest a respondent that very much enjoys working in a routine environment. Moreover, since there was a strong inverse relationship between the Satisfaction Research Questionnaire Comfort scale and the Scale A measure (described in greater detail below), this profile has been called the "Low stress symptoms risk" (Dysfunctional Subjective Health Symptoms) profile. Some 22 percent of the respondents fell into this profile.

The fourth modal profile (Figure 2.4) has been labelled the "High stress symptoms risk" (Dysfunctional Subjective Health Symptoms) profile, as it is the opposite of Profile Three. The extremely depressed Comfort scale value is readily seen. The elevated Challenge scale, in combination with the depressed Comfort scale suggest a worker that is continually striving for goal attainment (i.e., meet

new challenges), but finds the workplace an aversive environment that does not afford him or her any respite. Roughly 15 percent of the respondents were similar to this modal profile.

With the extraction of two (bipolar) modal profiles, roughly 20 percent of the respondents were unclassifiable (with epsilon = 0.50), yet the classification efficiency was 79.83, with over half of the variance in scores (60.05 percent of the total variance) explained by the profiles. Finally, the percent of variance explained for the respondents meeting the epsilon entrance criterion was 71.45. (Epsilon is the "entrance onto the factor" criterion, and can be thought of as the minimal respondent-component loading required for classification. A value of 0.50 yields a conservative classification and reduces the likelihood of incorrectly classifying a respondent by chance.)

The quantitative extraction of these four profiles has empirically demonstrated that all patterns of respondent satisfaction scores are not similar. The folly of expecting a single total score to serve as a complete classificatory scheme should be clear. For example, the correlation between satisfaction and pay has been found in some studies, and not found in others. This finding may have occurred because not all respondents belong to a profile with a large amount of Pay scale variance. Similarly, job enrichment (i.e., increasing the challenge of the job) would not be effective for all employees. The degree to which the respondents within each profile group differ will be examined in the next chapter.

CHAPTER THREE: EVALUATING THE NEW SATISFACTION MEASURE

The third chapter presents the results from data gathered from a cross-Canada sample of employed Canadians. Respondents in this sample completed a questionnaire comprised of several scales including the Satisfaction Research Questionnaire. The purpose of this questionnaire was to assess the validity of the Satisfaction Research Questionnaire. This chapter examines the validity of the recent measure.

In the first section, the design of the validity study is described and the measures are presented. The respondents in this study are then described.

The empirical examination of the scale's validity follows Loevinger's (1957) three components of construct validity. The substantive component of validity began with the description of the reliability of the scale in the previous chapter. Other reliability information is introduced here, as well as additional empirical data relating to the stability and parallel forms of the recent satisfaction measure. As a part of the structural component of validity, relationships between the recent satisfaction measure and organizationally relevant measures are presented. The external component of validity focuses on the relationships among the three principal measures, between the principal measures and the secondary measures across the modal profiles, and among the secondary measures across the modal profiles. As the previous chapter contended that desirability responding is a threat to the construct validity of satisfaction measures, a section describing the empirical relationships between desirability

and the two satisfaction measures is provided. The next section of this chapter briefly addresses sex differences in the pattern of satisfaction in the Canadian workplace.

This third chapter also serves as an introduction to the final chapter, where a proposal for a new model of satisfaction is presented. This model incorporates the importance of recognizing individual differences in satisfaction, the principles of selection, training, performance management, and feedback, and individual differences in general affect and personality.

Overview of the Validity Study

The validity study involved the distribution of a collection of measures to a sample of Canadian workplaces, ranging from the Maritime provinces to Vancouver, British Columbia. The aim of this mail-out cross-Canada study was to determine whether or not employed Canadians had any profiles of satisfaction, and whether or not the Satisfaction Research Questionnaire was an improvement over existing measures.

The Respondents. A preliminary list of organizations was created using the companies listed in The 100 best companies to work for in Canada (Innes, Perry, & Lyon, 1986). From this publication, a mailing list of current addresses of executive officers was generated. A letter of introduction was mailed to each executive officer responsible for personnel or human resources. This letter briefly described the research project, and asked the company to do two things on behalf of the project. The first was to agree to participate in the project. The second was to arrange for the random distribution of sealed questionnaire envelopes throughout the company hierarchy. Each sealed envelope contained a letter of introduction, the questionnaire, a computer-readable response sheet, and

a stamped, self-addressed return envelope so that the respondent could return the response sheet directly to the investigator. Each candidate company received a sample questionnaire envelope containing these materials. Each company also received an "intention to participate" return sheet, whereby they could indicate their a) agreement to participate and a request for testing material, b) agreement in principle, but with a request for more information, c) refusal to participate, or d) the name of a more appropriate company officer. A copy of the letter of introduction, the "intention to participate" sheet and the "additional information about the study" letter are all included in Appendix B. Of these one hundred companies, twenty-three decided to participate and were sent testing material for distribution.

The Measures. To this end, the recently developed Satisfaction Research Questionnaire was included in a battery of work attitude measures. There were three primary measures. They were the Satisfaction Research Questionnaire (Hill, 1985c), the Job Descriptive Index (Smith et al., 1969), and the Survey of Work Styles (Jackson & Mavrogiannis, 1987a). Six secondary measures were selected to evaluate the convergent and discriminant validity of the satisfaction measure. The Job Descriptive Index is the most popular facet measure of satisfaction and was included for monotrait mo¬ method comparisons with the Satisfaction Research Questionnaire. The Survey of Work Styles, a recent construct-valid, six facet multidimensional measure of the Type A behavior pattern, was included as a measure of stress. To assess the impact of desirability responding upon both satisfaction measures, the desirability subscale from the Personality Research Form (Jackson, 1984) was included. Additionally, a small number of secondary scales were included. The Scale A from the Survey of Work Styles, a measure of

absence from work, a role ambiguity scale, a good self maintenance scale, a scale to measure the presence of subjective health dysfunctional symptoms, and a subjective measure of workplace performance were also included. Each of these measures is described in greater detail in the next section.

The Job Descriptive Index.

The Job Descriptive Index has proven to be the most popular facet measure of satisfaction. Only single item "home-made" scales have been more popular. The scale has five subscales, Work on Present Job (18 items), Pay (9 items), Opportunities for promotion (9 items), supervision (18 items), and co-workers (18 items).

Development Rationale for the Job Descriptive Index. Smith et al., (1969) have defined job satisfaction as the feelings that a worker has about his/her job, and they endorsed the use of facet subscales as they feel that there are different feelings corresponding to different aspects of any job. The Job Descriptive Index is unusual in its use of a trichotomous scoring and response format. The authors recommended that an endorsement in the keyed (i.e., satisfied) direction should be scored a "3", a response in the opposite to keyed direction should be scored a "0". The authors then suggest that a neutral or "?" response be scored "1", as they felt that this response would be more indicative of dissatisfaction than satisfaction.

Initially, the Job Descriptive Index authors provided a set of respondents with an adjective checklist, and asked each respondent to describe their job by indicating "yes", "?", or "no" for each adjective. With this scheme, the authors had no way of determining whether or not an item was a satisfier or a dissatisfier, or even whether the item was being interpreted in a consistent manner. For

example, the Job Descriptive Index item, "hot", may have been endorsed to reflect the fact that the workplace was overly warm, or that the workplace was an exciting and dynamic environment.

Secondly, the authors used a manner of triadic scoring, wherein each respondent used the same adjectives to rate their present job, the job that they would most like to have, and the job that they would least like to have. Smith et al., (1969) scored an item as a satisfier if that item was scored in the keyed direction for both the current and best jobs. An item was a dissatisfier if it was endorsed in that manner on both the present and worst jobs. If an item was endorsed in the same way on the present, best, and worst jobs, then that item was scored as neutral because it did not distinguish the best from the worst job.

As a third method of developing the Job Descriptive Index response key, the authors applied a direct scoring scheme. In a pilot study, if an item was endorsed more frequently for the best job, then that item was keyed in the positive direction. If an item was endorsed most frequently for the worst job, that item was scored in the negative direction. When the authors divided the response data into high-low splits (on the basis of the total score), the dissatisfied respondents used the "?" response more than did the satisfied respondents. The authors concluded that the "?" response is more indicative of dissatisfaction than satisfaction, and therefore awarded those responses with a weight closer to "0" than to "3".

Origin of the Job Descriptive Ir ~x Items. Smith et al., (1969) originally culled the Job Descriptive Index items from other job satisfaction inventories, and from lists of adjectives and phrases that they felt would apply to various aspects of a typical job (Smith et al., 1969, p. 71). To determine whether or not

their scale would generate sufficient response variance, "preliminary JDI scales were administered to 17 janitors, 25 secretaries, and 16 cafeteria workers at Cornell university" (Smith et al., 1969, p. 71).

Item Analyses of the Job Descriptive Index. Smith et al., (1969) administered their preliminary scale to an "accidental sample" of 317 Cornell university students, and discarded all items that did not differentiate between the respondents' best and worst jobs. The authors then supplemented the diminished item pool by asking 81 randomly selected employees of a New York state farmers' co-operative to give their reasons for calling a certain job their best job, and certain jobs their worst jobs. The authors conducted this last opinion poll to "ensure that no important items indicative of satisfaction or dissatisfaction have been omitted during the original item-selection procedure" (Smith et al., 1969, p. 72). Smith et al., (1969) then gathered Job Descriptive Index data from 163 men and 73 women that were randomly chosen from three companies. These respondents were divided into high and low satisfaction groups (median split). The authors' only assurance of equal representation of positively and negatively keyed items was their examination of the proportional differences between the high and low groups in the use of true-keyed and false-keyed items.

The authors presented a mean item validity of .48 for their new scale, and a mean item intercorrelation of .30. The mean validity refers to the average Job Descriptive Index subscale correlation with the Faces (Kunin, 1955) scale of satisfaction. They also presented a listing of the ranges of item intercorrelations per scale (Smith et al., 1969, p. 73) that shows the presence of items that correlate negatively with their own scale. The range of item correlations for the Work subscale was -.16 to .63. For the Pay subscale, the range was -.08 to .58,

for the Promotions subscale, the range was .18 to .76, for the Supervision subscale, -.16 to .78, and lastly, for the Co-Workers scale, the range was -.10 to .66. These interitem correlations suggest that the scales' items were not entirely homogeneous, even after item analyses in the derivation sample. The authors reported subscale internal consistencies that were modest. The subscale reliability was .73 for the Work scale, .67 for the Pay scale, .75 for the Promotions scale, .77 for the Supervision scale, and their Co-workers subscale had a reliability of .78. These reliabilities were computed using data from the derivation sample. The authors reported that "by discarding one or two more items in each area, somewhat more reliable scales could have been developed, but it was decided to retain all items ... in order that all scales would have either 9 or 18 items" (Smith et al., 1969, p. 74).

The two sets of item analyses, the opinion poll, and the absence of a rationale to link items to their scale in a substantive manner all had a cumulative effect that can be seen in the relatively poor reliabilities of the Job Descriptive Index subscales. When the authors gathered test-retest data on their scale, they found fairly high test-retest correlations. A three year test-retest administration using 45 farmers from the above farmer's co-operative yielded test-retest correlations that ranged from .45 to .75 for the Job Descriptive Index subscales.

Since the Job Descriptive Index subscales are presented as separate subscales, and not mixed into a omnibus format, the authors performed a Latin square design analysis of variance to determine whether or not the order in which the subscales were presented affected the scale scores. They found no such order effect.

The Job Descriptive Index authors claimed to develop unique subscales that were as orthogonal as possible. The mean between-subscale correlations that the authors report, however, indicate substantial overlap. For male respondents, the mean between-subscale intercorrelation was .37, (ranging from .29 to .42). For female respondents, the mean value was .32, (ranging from .19 to .52). The authors themselves point out that "(I)t is evident that nearly all the scale intercorrelations are quite high" (Smith et al., 1969, p. 78). Rather than collapse some subscales, or refrain from asserting that the subscales are orthogonal, the authors argued that the subscales still measured different areas of job satisfaction. They claimed that some job aspects would be more important for some people than others, and that different job aspects would be related to personal differences in background. Moreover, the authors finally asserted that their subscales were conceptually different without necessarily being psychometrically different. The authors claim that the Job Descriptive Index "is directed toward specific areas of satisfaction rather than global or general satisfaction" (p. 69) and

> It is evident that nearly all the scale co·relations are quite high. We have shown previously, however, that the JDI scales measure discriminably different areas. (p. 78.)

Little evidence has been provided by the Job Descriptive Index authors to suggest that scale development procedures to enhance construct validity were ever applied to their scale. There has never been any formal definitions of the Job Descriptive Index subscales or the content that they are supposed to assess. Items were culled from other inventories and were not written to reflect any specific theoretical content domain that the subscales were supposed to represent.

Moreover, when a scale administrator completes an administration of the Job Descriptive Index, there are no norms available by which the administrator can gather relative information about his/her respondent sample. When the Job Descriptive Index is used, the scale user must refer to the few published reports that make explicit their Job Descriptive Index subscale scores. Additionally, the item responses are often difficult to interpret as the items are just single words or terse phrases, not behavioral exemplars or attitudinal items about jobs in general. The repeated use of some items ("Bad", "Boring", "Lazy", and "Intelligent" are all used twice) further complicates the interpretation of subscale scores by changing the subjective weighting of the items. Some items have unclear associations with their own subscale. Like the previous "hot" example, responses to the co-worker subscale item "fast" may mean that one's peers are bright or swift. Not all subscales have the same length. The Pay and Promotions subscales are half as long as the other Job Descriptive Index subscales. The scale authors recommended doubling the means for these two scales, but did not empirically determine whether doubling the scales' totals was equivalent to equating those two nine item scales with their eighteen item counterparts. The scale authors attempted to control (indirectly) for the acquiescence response bias by generating roughly equal numbers of true and false keyed items, but did not actually generate equal numbers or evaluate response bias empirically.

Item Format of the Job Descriptive Index: Defense of Dichotomous Format. There were five considerations that were reviewed in deciding whether to use the traditional trichotomous Job Descriptive Index response format (yes, ?, no), or a dichotomous format.

The first consideration was that the use of a neutral category might prompt or cue various response sets in the respondent data. There is an extant psychometric literature that has dealt exclusively with response formats. This literature suggests that the decision to endorse an item with a "?" response may be attributable to certain other sources of response variance than responses to item content (cf. Hill, 1985a). While respondents might actually be neutr 1 or uncertain, low social desirability of the item, evasiveness, avoidance, indecision, or other response styles might also be responsible. Rejecting the trichotomous format in favor of the dichotomous format resolves this issue.

A second issue is clerical in nature, in that the computer-readable answer sheet made available for this research provided a dichotomous response format.

A third argument against the traditional scoring was that the atypical weighting scheme used by Smith et al., lacks theoretical justification and appears arbitrary. A disagree is given a "0", a "?" response is scored as a "1", and an agree response is given a "3". The authors do not provide any additional information to substantiate their contentions about the scoring, other than to declare that they felt a "?" response was actually a response in the dissatisfied direction.

A fourth contention was the hypothesis that, given the chance, respondents would not use the "?" option that frequently in any case. Most of the Job Descriptive Index items are single words or short phrases and are fairly obvious. The authors themselves suggested that no more than a low level of reading ability is required to understand their scale. In a sample of working undergraduate students, few items were endorsed with the "?" response more than 10% of the time. Clearly, the "?" option was not used that frequently.

The fifth contention against the use of the trichotomous format was a hypothesis related to the last point. It was hypothesized that the two response formats for this scale would yield highly similar scale scores. The same subjects were given both the dichotomous and trichotomous formats. (There were no significant order or presentation effects, and the order of presentation was randomly determined.) When this hypothesis was tested, the dichotomous and trichotomous scale score pairs all correlated 0.82 or greater. The two formats of the Work scale yielded highly similar scores ($r=.88$), as did the Pay, Promotions, Supervision, People, and Total scores, (.89, .82, .89, .84, and .79, respectively). Clearly, the use of the dichotomous response format is defensible from the perspective of substantive keying of items to scale, the elimination of potential response styles, and the fact that the two formats yield nearly identical results in any event.

The Survey of Work Styles

The Survey of Work Styles was developed using the construct validity approach to scale construction in a manner similar to the development of the Satisfaction Research Questionnaire. The Survey of Work Styles has six subscales, each reflecting a content domain related to the Type A Behavior Pattern. Jenkins, Zyzanski, & Rosenman (1978) defined those with the Type A Behavior Pattern as possessing a sense of time urgency, easily aroused anger and hostility, and extreme levels of competitiveness, impatience, and striving for achievement. The subscales of the Survey of Work Styles reflect these contents areas. They are Impatience, Anger, Work Involvement, Time Urgency, Job Dissatisfaction, and Competitiveness. The Type A Behavior Pattern is associated with vigorous verbal and psychomotor mannerisms, a sense of time urgency, easily aroused anger and

hostility, and extremes of competitiveness, impatience, and achievement striving. In the past, these Type A attributes and their relationship with coronary heart disease have received wide attention. However, prior to the development of the Survey of Work Styles, there were few assessment tools available. The primary measures were the Structured Interview, the Jenkins Activity Schedule, and the Framington Type A Scale. The Structured Interview has been recognized as the superior device of the three, but is expensive and time consuming to administer and score, and only yields a general, global measure thereby possessing little power to differentiate among respondents. Moreover, a history of poor inter-rater reliability has plagued the device, along with a host of rater biases that have detracted further from the validity of this measure. The Jenkins Activity Schedule and the Framington Type A Scale are both paper and pencil measures. Both were designed to overcome the shortcomings of the Structured Interview, but reviewers nd empirical data have indicated that both of these scales are suboptimal in many regards. These measures were developed in the absence of a recognition of the threat of response biases, have relatively poor classification rates, and have failed to assess the multidimensional nature of the Type A Behavior Pattern. The Survey of Work Styles was developed to overcome the problems previously encountered with other paper and pencil questionnaires. In an early study using a sample of business managers, the Survey of Work Styles yielded reliabilities superior to the previous measures (administered to the same sample), and surpassed the classification rates of the other paper and pencil measures to a statistically significant degree. An additional feature of the Survey of Work Styles was its job Dissatisfaction subscale, which permits a direct comparison with the recent satisfaction measure.

The Survey of Work Styles "Scale A". The authors of the Survey of Work Styles identified those items that had the highest loadings on the unrotated first principal component of the Survey of Work Styles. These items, when aggregated into a single scale, were labelled the Scale A scale. The test authors found that this scale was predictive of the Structured Interview Type A and non-Type A classification. The authors completed a classification study to determine which of the Type A measures had the highest classification rate, when compared with the Structured Interview classification. The Survey of Work Styles Scale A scale had the highest classification rate. The Survey of Work Styles Scale A measure can therefore be viewed as a viable measure of the Type A Behavior Pattern.

The Absence Scale. Absence from work may be the result of dissatisfaction with any single satisfaction facet, or any combination of facets. This broadband relationship has been responsible for the fact that the average correlation between satisfaction and absence has typically been greater than mean satisfaction and performance correlations. People absent themselves from their workplaces for a variety of reasons. Several authors have noted that both the definition and reliability of absence measures have been equivocal. Typical measures of absenteeism are simple frequency counts, either from personnel files or from self-reports. Mikalachki and Gandz (1979) have argued that a conceptual problem arises when trying to use a general measure of absenteeism. They argue that a general measure is inappropriate because it obscures the different reasons that a person might be absent. They argue that it is important to know whether a person missed several days of work a year for work-related or not work-related reasons. They further propose that this related-to-work dimension be subdivided so that any researcher that wishes to include an absenteeism measure actually

include several diverse absence measures. They suggest three absenteeism indices. An Inactivity index, to measure the number of scheduled days or time lost to absenteeism. A Participation index, to reflect the degree to which absenteeism is widespread among a group. Lastly, they suggest an Incidence Rate index, to measure the frequency with which episodes of absenteeism occur in a workforce. Their suggestion has probably been the most ambitious to date. Unfortunately, there has been very little empirical research that would either support or refute the validity of Mikalachki and Gandz's (1979) claim.

The current research reviewed the absenteeism and job satisfaction literature to determine what were the most frequent or important reasons that a person might be absent. There was no one universally accepted scale, rather there were twelve reasons for absence, ranging from work-related accidents to domestic accidents, union-related absences, various types of illnesses, and so on. Table 3.1 provides a complete listing of all absence measures and their intercorrelations. The literature yielded certified absences (certified medical illness, certified work-related accident absences, certified domestic accident absences) contractual absences (jury duty absences, bereavement absences, union activities absences, and the ubiquitous "other" category), and uncertified noncontractual absences, (such as disciplinary suspensions, personal or family absences, uncertified medical absence, unreported absences, and suspension from work).

- -

Insert Table 3.1 about here

- -

Using the cross-Canada sample (n=388), the mean correlation among these

Table 3.1 Correlations among single item absence measures. Based on cross-Canada sample , n=388. (Mean correlation of .47.)

	0[1]	1	2	3	4	5	6	7	8	9	10	11	12
1[2]	.61	-											
2	.69	.53	-										
3	.51	.33	.28	-									
4	.44	.33	.44	.32	-								
5	.71	.47	.53	.32	.37	-							
6	.75	.58	.68	.41	.36	.70	-						
7	.64	.52	.42	.31	.30	.54	.60	-					
8	.63	.32	.46	.28	.28	.45	.51	.43	-				
9	.76	.51	.58	.36	.35	.70	.71	.64	.62	-			
10	.80	.57	.62	.44	.41	.75	.80	.65	.58	.89	-		
11	.63	.52	.51	.31	.26	.49	.57	.47	.48	.55	.63	-	
12	.53	.42	.29	.35	.18	.38	.47	.38	.40	.39	.48	.48	-
	0	1	2	3	4	5	6	7	8	9	10	11	12

[1] Column "0" entries are the item loadings on the first unrotated principal component derived from the above correlation matrix.

[2] Absence measures are the number of missed days of work in a calender year. The items are: 1) the number of unreported absences, 2) absences due to suspension, 3) reported illnesses, 4) unreported illnesses, 5) work-related accidents, 6) domestic accidents, 7) jury duty leave, 8) bereavement absences, 9) unreported absences, 10) disciplinary suspensions, 11) personal excuse absences, and 12) "other" reasons. The reliability for the 12 item aggregate scale was .91.

twelve items was .47, ranging from .18 to .89. On the basis of the respondent data, it can be argued that the proposition for a multidimensional representation for absenteeism is weak. All absence measures have substantial correlations with one another, much in the same manner as items on the same scale. To test this analogy empirically, the absence measures were treated as items on the same scale, and were subjected to reliability analyses. The absence "scale" in this research is this aggregate. The null hypothesis was that the internal consistency of the aggregate measures would be low, as each question should reflect a substantively different source of true score variance. The actual reliability of 0.91 permitted rejection of this null hypothesis, and provided evidence for a single measure of absenteeism -- or alternately, for the use of multiple measures of absenteeism which are then expressed as a total scale score. After the reliability analyses of the new absence measure, it was subjected to a series of principal components analyses. Various numbers-of-factors solutions were examined. (The numbers of factors rules included a) the Scree test; b) Eigenvalues greater than one; c) Velicer's (1976) MAP test; and d) split-half replications.) Inspection of these solutions confirmed the creation of a unidimensional absence scale. A weak two factor solution was suggested, but did not replicate itself, suggesting poor factor reliability.

Job Ambiguity Scale. Caplan and Jones (1975) investigated the relationship between job ambiguity, work load, and the Type A behavior pattern. Their study incorporated a three item job ambiguity scale of acceptable reliability. Caplan (1971) presented a four-item measure, as well. The original seven items were phrased in the form of questions. For example, "Are your work objectives defined well enough for you?". These items were rewritten so that they

were direct statements that a respondent may or may not make about their job. For example, "My work objectives are defineu well enough for me". The rephrased items from these job ambiguity scales were incorporated into the overall questionnaire. Caplan found that high levels of ambiguity were significantly related to higher levels of work and personal stress.

Good Self Maintenance Scale. Exercise and the practice of other health-promoting self maintenance behaviors have been linked to the decrement of physiological stress (cf. Bruning & Frew, 1987). These personal behaviors have not been in the forefront of traditional stress research, but have been appearing more frequently in the recent literature. Most of the literature to date has focused on self-report perceptions, while recent research has investigated the validity of frequency counts of good self maintenance behaviors. Recent research has suggested that individual abilities to cope with stressors may be linked to coping abilities, with good self maintenance behaviors providing one such set of coping abilities. This literature suggests that good self maintenance behaviors can be viewed as either coping abilities, or as mechanisms that promote individual resistance to stressors. The current research reviewed the stress and satisfaction literature and identified those self maintenance behaviors that were related to lower levels of stress. Such behaviors included having a regular exercise program, the amount of exercise, rating of the quality of the diet, whether the individual smoked heavily, whether the individual was overweight, and whether a physician would say that the individual was in good health (cf. Parkes, 1987). These behaviors were incorporated into the mailout questionnaire as items that a respondent could endorse in a yes or no manner. For example, a respondent would reply yes or no to the following item: "My physician would say that I was in good health".

Subjective Health Dysfunctional Symptoms Scale. Several authors have noted that a constellation of symptoms seems to predict coronary heart disease. The earlier presence of these symptoms was noted in subjects that had coronary heart disease. These symptoms are termed subjective health dysfunctions symptoms because they are serious maladies, but they do precede and predict coronary heart disease. Such symptoms include having a cold or the flu more frequently than others, having repeated headaches, muscle soreness, lower back pain, or feeling irritable for no apparent reason. Several of these symptoms were phrased as statements that a respondent could agree that they had or disagree that they had. These items constituted the subjective health dysfunctional symptoms scale.

Subjective Performance Scale. One problematic aspect of job satisfaction research has been the inability to gather "hard" and valid performance ratings or measures. In an attempt to gather this information, each questionnaire respondent answered questions asking whether they felt they were as productive as the people they worked with, whether there were co-workers who could do their job better than themselves, whether they were as productive as they could be, and whether or not they were one of the most productive people at work. Fisher (1979) suggested that "more specific attitude measures are better predictors of single-act criterion" (p. 349). The above items were designed to be specific, concrete behavioral exemplars of the attitude "performance". Since this was a subjective measure of worker productivity, the resultant scale was labelled the "subjective" performance scale.

Description of the Cross-Canada Sample

Mean values for background measures. There were nearly equal numbers of males and females in the cross-Canada sample. There were 151 males, 150 females. Eighty-seven respondents did not report their sex. The mean age of the respondents was 33 years. A third of the sample was 30 years of age or less (31%), 18% were between the ages of 31 and 40, 7% were between the ages of 41 and 50, 6% were 50 or over. Some 38% of the respondents did not report their age. Thirty-five percent of the sample was married, 12% were divorced, and 7% were widowed. Some 58.2% of the respondents did not report their marital status. The majority of the respondents (51%) had no children, 13.1% had three to four children, and the remaining respondents (35.8%) did not report their number of children. Over half of the sample was paid by salary (53.4%). A small portion received incentive pay (4.1%). Twenty-six percent reported their type of remuneration as "other", and 15.7% did not report at all. Of those respondents who reported their salary range, 17% had a salary of less than $15,000, 72.7% fell into the $21-30,000 range, and 10.3% were in the $41-50,000 annual salary range. Some 10.1% had completed only grade school, 65.4% had obtained a college diploma or university degree after high school, and 24.5 had a graduate degree. No one had received only a high school diploma, possibly reflecting a relatively bimodal distribution of selection criteria for these companies. A third of the sample (32.5%) reported receiving some additional job-related education.

Those reporting their tenure of employment indicated that 70.9% had worked for a year or less with their present employer. Some 29.1% had a tenure of between six and ten years. This reflected the length of time that they had spent in their current position. The majority (87.9%) had been in their current position for less than a year, while 12.1% had a tenure of six to ten years. The

short-lived tenures were also reflected in the tenure with previous employers. The majority (79.2%) of respondents held their last job for a year or less, while only 20.8 had six to ten years with their previous employer. When asked why they left their previous position, 14.7% said "not enough pay", 3.9% left because of poor resources, 6.6% quit because of a former supervisor, and 74.8% left for "other" reasons.

When asked whether they expected any promotion in their jobs, 59.8% said they did. Some 89.2% also said that they were expecting that promotion within the next calender year. Only 10.8% said that they were expecting a promotion within the next month. When asked a standard "Have you ever thought of quitting your job?" question, 62.9% said they had. Over a third (36.9%) had looked for alternate employment, and 51.6% of those that had looked, did so in the week before they completed the survey. The proportion of the sample falling into each modal profile group was somewhat different from the original normative sample, with 5.7% of the sample falling into the first modal profile, 13.1% being classified as belonging to the second profile, 25.7% in the third profile, and 29.6% belonging to the high risk profile. About a quarter of the respondents were unclassifiable using the criteria described in the second chapter.

Validity of the Satisfaction Research Questionnaire

Loevinger (1957) suggested that any validation procedure be divided into three components: a) substantive, b) structural, and c) external components. The substantive component of validity refers to the degree to which the items of a test reflect at a conceptual level their own construct. The structural component makes reference to the hypothetical structures that the investigator has developed prior to test development. The last component, the external component, refers to

how well the hypothetical structures relate to actual content-similar measures outside of the test. The following sections present the results from the cross-Canada sample along the components outlined by Loevinger. The substantive component of validity was addressed by the item, subscale, and scale reliability and validity evidence provided in the second chapter. The following section presents the results and discussion relating to the structural fidelity of the Satisfaction Research Questionnaire in the cross-Canada sample. Various factor analytic results are presented, as are the results depicting desirability responding as a threat to the structural fidelity of the test. The third component, Loevinger's external component, presents the relationships between the Satisfaction Research Questionnaire, construct-similar, and construct dissimilar tests so as to provide monomethod multitrait validation (cf. Campbell & Fiske, 1959) of the recent satisfaction measure.

Structural Component of Validity

This first section describing the results of the validity analyses begins with presentation of the reliability values for each of the three principal measures, followed by additional reliability information about the Satisfaction Research Questionnaire, and then reliability information about the six secondary measures. The data referred to are from the cross-Canada survey.

Job Descriptive Index Subscale Reliabilities. The five subscales for the Job Descriptive Index are a) Work, b) Pay, c) Opportunity for Promotions, d) Supervisors, e) People at Work, and a total Job Descriptive Index score. The reliabilities for the above Job Descriptive Index subscales in the cross-Canada sample were: .78, .64, .46, .85, .83, respectively for the five subscales, with a

mean internal consistency of .71. (These values were greater than the trichotomous scale reliabilities reported by Smith et al., (1969)). The reliability for the total score was .90.

Satisfaction Research Questionnaire Subscale Reliabilities. The five subscales of the Satisfaction Research Questionnaire are a) Challenge, b) Comfort, c) Co-Workers, d) Pay, and e) Resource Adequacy, plus a total score. The reliabilities for these subscales in the same sample were: .81, .84, .81, .81, and .84, respectively, with a mean reliability of .82. The reliability for the total scale was .91.

Stability of the Satisfaction Research Questionnaire Scores: I. To determine empirically the stability of the Satisfaction Research Questionnaire subscales, four sets of reliabilities were compared. These four sets were: a) the original sample reliabilities, b) the cross-Canada sample reliabilities, c) and two additional administrations of the test to one sample (i.e., test-retest).

- - - - - - - - - - - - - - - - - - - -

Insert Table 3.2 about here

- - - - - - - - - - - - - - - - - - - -

As can be seen from Table 3.2, the Satisfaction Research Questionnaire scales are reliable across diverse samples. The reliabilities for the original sample (n=119), the cross-Canada sample (n=388), and the student sample (n=183) at two times are included. The last column presents the test-retest reliabilities, using the student sample. A three month interval separated the two student administrations.

Stability of the Satisfaction Research Questionnaire Scales: II. Students in an undergraduate psychology (Industrial and Organizational Psychology) class who were currently employed, or had employment during the previous summer,

Table 3.2 Satisfaction Research Questionnaire subscale and total reliability estimates across four samples.

	Sample				
	Original	Cross-Canada	Student$_1$	Student$_2$	Test-Retest (3 months)
Subscale					
Challenge	.85	.81	.88	.88	.82
Comfort	.81	.84	.76	.81	.69
Co-Workers	.96	.81	.85	.85	.80
Pay	.87	.81	.80	.82	.69
Resource Adequacy	.86	.84	.85	.85	.56
Total	.93	.91	.91	.92	.75

were asked to complete the Satisfaction Research Questionnaire. Three months later, they were asked to use the same job as before, and completed the Satisfaction Research Questionnaire a second time. (In fact, they were not informed that they were completing the same scale twice, but they were explicitly told to use the same job as before and label that job on their answer forms.) By correlating the scale scores from time 1 with those of time 2, the test retest reliabilities were determined. The test retest reliability for the Challenge scale was .82, .69 for the Comfort scale, .80 for the Co-Worker scale, .69 for the Pay scale, .56 for the Resource Adequacy, and .75 for the total scale score. These results show that the Satisfaction Research Questionnaire has yielded moderately stable scale scores.

Survey of Work Styles Subscale Reliabilities. The six subscales of the Survey of Work Styles are a) Impatience, b) Anger, c) Time Urgency, d) Work Involvement, e) Job Dissatisfaction, and f) Competitiveness. The reliabilities for these subscales in the cross-Canada sample, using a dichotomous response format were: .76, .67, .77, .78, .85, and .37, respectively, with a mean reliability of .76. The reliability for the total score was .86. Mavrogiannis, Jackson, and Howard (1987, p. 8) reported the reliabilities as: Impatience .82, Anger .71, Work Involvement .82, Time Urgency .81, Job Dissatisfaction .82, and Competitiveness .80. Their reliability for the total score was .90. These reliabilities were from data gathered using the authors' five-point response scale, not the two-point response format of the mail-out.

Reliabilities For Secondary Measures. In addition to the principal measures employed in this research, there were six additional measures employed in the cross-Canada sample. The Scale A subscale, based on a set of the Survey

of Work Styles items, yielded a reliability of .73. An absence from work scale, with twelve items, had a reliability of .92. The seven item job ambiguity scale had a reliability of .73. The scale measuring good self maintenance behaviors also had seven items, and yielded a reliability of .69. Lastly, the four item measure of subjective worker performance had a reliability of .54. As these measures had unequal lengths, the Spearman Brown formula was used to determine the reliabilities of these measures might have been if they were all sixteen item scales. For iluustrative and comparison purposes, the Spearman Brown corrected reliabilities were: a) the absence scale, from .92 to .94, b) the job ambiguity scale, from .73 to .86, c) the good self maintenance scale, from .53 to .72, d) the subjective health dysfunctional symptoms scale, from .69 to .73, and lastly, e) the subjective performance measure, from .54 to .82. These corrected reliabilities provide empirical support for the unidimensional nature of each of these measures.

Intraindividual Reliability Analyses. Two parallel forms of the Satisfaction Research Questionnaire were developed using the cros.-Canada data. There were three main reasons for this development of a variation of the original scale. One, many research applications require shorter, parallel forms. Two, parallel forms are convenient for test-retest purposes. Three, a correlation between the two arrays of parallel scores for a single individual represents the stability of that individual's pattern of responding, (i.e., the intraindividual reliability). The intraindividual reliability coefficient can be used as a "confidence" measure by which the stability of the respondent score profile can be assessed. A low value would indicate a less purposeful pattern of responding, and therefore a less reliable profile. The creation of psychometrically equivalent

forms, however, is not a straightforward task. Gulliksen (1950) dedicated an entire chapter to the development and validation of equivalent forms of the same test.

Gulliksen (1950) suggested that items could be paired by their biserial correlations and their population endorsement proportions. Pairs are then separated into two parallel forms. Item-pairing in the current research was effected with an additional constraint, that the scales of the final version of the parallel forms had to have equal numbers of positively and negatively loading items. The item pairs were separated, and the two arrays of Satisfaction Research Questionnaire item scores were correlated. This correlation was corrected by the Spearman-Brown formula to achieve the correlation between original length forms of the test. This corrected intraindividual reliability coefficient is reported on the computer-generated booklet returned to the subject, and is used to indicate the robustness of the respondent's test profile.

Classification of the Cross-Canada Sample. In order to classify the respondents in the cross-Canada sample, each respondent's five Satisfaction Research Questionnaire subscale scores were correlated with the subscale values generated for the normative modal profiles. In this manner, all respondents were classified in a manner that: a) used an external classification criterion rather than a within sample criterion, b) assessed profile shape similarity (vis a vis the between profile correlations), and c) also assessed profile scatter and elevation (vis a vis the original modal profile analysis that incorporated scatter, shape, and similarity). All four modal profiles replicated successfully.

In the following sections, the dimensionality of all three of the principal measures in this study was examined.

Factor Analytic Studies of the Three Principal Measures. To assess the dimensionality of the three main measures of this study, each measure was subjected to a series of planned (item) principal components analyses.

Replication of factor structures across split halves. To assess the factor reliability of the two satisfaction scales, the respondent data were separated into odd and even case number files. Although the substantial sample size of this study should have ensured stable factor loadings, a comparison of the factor loadings across the odd-even case numbers files was effected to demonstrate factor reliability. In each data set, five solutions were computed for each satisfaction scale. Not all solutions are presented below. The full rank (i.e., five subscales/factors), two factor, and one factor solutions are presented because of the hypothesized existence of this many factors in the two scales.

Five Factor Solutions. When the odd-even five factor solutions were compared for both satisfaction measures, the respective factors (now scales, due to the use of a full rank solution) replicated nearly perfectly. All scale pairs were therefore replicated. This solution is nearly trivial, however, since a) both tests were designed to assess five subscales, and b) the full rank solution is not as important as the reduced rank solutions.

Two Factor Solutions. Analyses during the modal profile analyses indicated that the Satisfaction Research Questionnaire (SRQ) yielded a satisfactory two factor solution. The first factor was marked by large loadings from the Challenge (.89), Co-Worker (.53), and Pay (.78) subscales. The second dimension was indicated by high loadings from the Comfort (.90) and Resource Adequacy (.77) subscales. The first factor was interpreted as a "Social Entrenchment in the Workplace" factor. The respondent item responses on this

factor indicated that workers felt they were well established in the social workplace, confident of their status, and able to seek greater challenge because of this entrenchment. The second factor was interpreted as a "Feels Provided For" factor. The pattern of item responses to this factor suggested a respondent that felt comfortable in his or her workplace, could concentrate on doing a task well, and was provided with the tools, information, and support from above to complete that task. It was hypothesized that this structure would replicate across the odd-even split halves.

The SRQ two factor solution showed a robust replication across the split halves. The respective dimensions were strongly related across samples, and minimally correlated with their alternate factors. The correlations between the two factor pairs across samples were .97 and .89. These results were replicated with an orthogonal Procrustes rotation and Harman's (1976) congruence coefficient.

The Job Descriptive Index authors have suggested that their scale measures general satisfaction or just five facets, and therefore, no cross-replicated two factor solution was hypothesized.

The Job Descriptive Index data showed little evidence of a replicated 2 factor solution. The correlations between the 2 factor pairs across samples were .47 and .60. This lack of replication was also evidenced in the orthogonal Procrustes rotation and the Harman congruence coefficient.

One Factor Solutions. It was hypothesized that the Satisfaction Research Questionnaire (SRQ) two factor structure would not yield a robust one factor solution. Even though the Job Descriptive Index (JDI) authors supported a "g-satisfaction" interpretation, it was hypothesized that the one factor replication of

the JDI would be poor. The idiosyncratic effects of the hypothesized response biases would operate to reduce factor reliability. Both hypotheses were supported by the data. The two factor SRQ solution replicated better than the one factor JDI solution which yielded a between sample correlation of only .49 using orthogonal factors (again, this finding was replicated with Harman's coefficient of congruence).

External Component of Validity

In the previous sections the substantive and structural components of the validity of the Satisfaction Research Questionnaire have been assessed. In this section the external component of the recent satisfaction measure's validity is examined. First, the Pearson product moment correlations between the Satisfaction Research Questionnaire, the Job Descriptive Index, and the Survey of Work Styles are reviewed. Secondly, the canonical correlations between the three principal measures are examined to determine what relationships exist between these measures on an overall level. Lastly, the relationships between the Satisfaction Research Questionnaire and each of the six secondary measures are reviewed at the general (i.e., total sample) and the profile level.

Convergent Validation of the New Satisfaction Measure: Correlations. It was hypothesized that each Satisfaction Research Questionnaire subscale would have its largest correlation with that single Job Descriptive Index scale that most closely reflected the item content of that subscale or construct. It was further hypothesized that the absolute magnitude of these correlations would be constrained by the presence of response biases in the Job Descriptive Index subscales.

- - - - - - - - - - - - - - - - - - -

Insert Table 3.3 about here

- - - - - - - - - - - - - - - - - - -

As can be seen in Table 3.3, each Satisfaction Research Questionnaire subscale did indeed possess its highest correlation with its content similar Job Descriptive Index subscale, and the appropriate content similar single items. The Satisfaction Research Questionnaire Challenge subscale and the Job Descriptive Index Work subscale correlation represents one such construct similar pair. The correlation (r = .70, p < .001) is clearly the largest column entry, signalling convergent validation for that construct. (The fact that only some 50% of the variance between the two subscales was accounted for is hypothesized to be due to the effects of the response biases associated with the Job Descriptive Index subscale.) The same pattern can be observed for all other construct similar pairs. The one exception to this rule was the Job Descriptive Index Promotions subscale which was substantively related to both the Satisfaction Research Questionnaire Pay and Resource Adequacy subscales, and again, the respective correlations of .30 and .32 are the largest column entries.

Convergent Validation of the New Satisfaction Measure: Canonical Correlations. Since the three main measures in this study are conceptually related, the actual degree of overlap was computed. The canonical correlation was chosen for this task, since it presents the correlation between the linear composites of each set of scores, thereby expressing the relationship between the first general component of each test, i.e., each test in its canonical form.

Job Descriptive Index and Satisfaction Research Questionnaire. The canonical correlation between the Job Descriptive Index and the Satisfaction Research Questionnaire was 0.82, accounting for some 67.2% of the common

Table 3.3 Convergent validation of the Satisfaction Research Questionnaire subscales. Correlations with Job Descriptive Index (JDI) subscales and single item measures. Underlined values indicate the location of the hypothesized largest content-similar correlations.

Satisfaction Research Questionnaire Subscales

JDI Subscales	Chall	Comf	Co-Wo	Pay	Res	Total
Work	.70	.34	.31	.53	.43	.67
Promotions	.27	.22	.19	.30	.32	.39
People	.45	.18	.50	.36	.34	.51
Pay	.35	.24	.14	.74	.29	.51
Supervision	.37	.44	.25	.39	.50	.58
Total	.64	.40	.42	.64	.55	.77

- -

Single Items						
Challenging	.48	-.03	.07	.35	.14	.26
Can relax at work	.08	.54	.18	.18	.28	.40
Pleased with other people	.26	.23	.42	.18	.28	.40
Pleased with pay	.21	.17	.07	.62	.17	.36
Have resources	.05	.35	.17	.12	.45	.35
Happy with job	.41	.26	.22	.36	.34	.45
Would tell a friend	.29	.34	.25	.35	.36	.46

variance. Redundancy analyses (cf. Tabachnik & Fidell, 1983, p.157) indicated that 31.6% of the variance in the Satisfaction Research Questionnaire was redundant with the Job Descriptive Index.

Satisfaction Research Questionnaire and Survey of Work Styles. The canonical correlation between the Satisfaction Research Questionnaire and the Survey of Work Styles was also 0.82 (accounting for 67.2% of the common variance) with the Survey of Work Styles Job Dissatisfaction scale included, and 0.65 without that same scale (accounting for 42.0% of the common variance). Redundancy analyses indicated that the Satisfaction Research Questionnaire had more in common with (i.e., greater redundancy) the full Survey of Work Styles measure, than with the same measure less the Job Dissatisfaction subscale.

Job Descriptive Index and Survey of Work Styles. The canonical correlation between the Job Descriptive Index and the Survey of Work Styles was .84 (accounting for 71.3% of the common variance), with the Survey of Work Styles Job Dissatisfaction scale included, and 0.37 with that scale omitted, accounting for 13.7% of the common variance. The Job Descriptive Index had little redundant variance with the stress measure when the Job Dissatisfaction subscale was not considered.

Inspection of the canonical variates indicated that the majority of variance in the Job Descriptive Index and Satisfaction Research Questionnaire canonical correlation was due to the Job Descriptive Index Work and Pay scales, and the Satisfaction Research Questionnaire Challenge and Pay scales.

The Job Descriptive Index and Survey of Work Styles canonical correlation variance was mostly due to the Survey of Work Styles Anger and Job Descriptive Index Work on Present Job scale contributions, in the analysis that did not include the Survey of Work Styles job dissatisfaction scale.

Examining the Satisfaction Research Questionnaire and Survey of Work Styles canonical relationships led to the importance of the Satisfaction Research Questionnaire Comfort scale and the Survey of Work Styles Time Urgency and Work Involvement scales, when the Survey of Work Styles Job Dissatisfaction scale was not included. When this scale was included, the Satisfaction Research Questionnaire Comfort scale and Survey of Work Styles Job Dissatisfaction subscales best represented the contributions of each respective test. This suggests that the Satisfaction Research Questionnaire Comfort scale would serve as a good predictor of the Survey of Work Styles total score, or the Scale A measure based upon the Survey of Work Styles items. This possible relationship is examined in greater depth when the relationships between the Satisfaction Research Questionnaire and the secondary measures are presented in the next section. This presentation begins with the Scale A measure.

The subsequent sections describe the relationships between the Satisfaction Research Questionnaire and the secondary measures across modal profiles. The Pearson product moment correlations are presented for each section. Also presented for each subsequent section are the correlations corrected for attenuation due to unreliability in both measures. Additionally, the correlations corrected for restriction of variance in one variable are presented. For demonstrative purposes, the correlations corrected for both unreliability and variance attenuation in one measure are also presented. The division of the sample into four modal profile groups was intended to reflect four modal profile populations. However, some may view this classification as an arbitrary division of the data, similar to dichotomizing scores. The second set of corrected correlations address this possible issue by correcting for range restriction. Once

identified with the modal profile analysis program, these unique profile groups can be likened to samples from unique populations. It is valid to examine the within profile correlations to determine whether there are differential patterns of correlations between profiles. Since, however, these correlations contain a measure whose variance will be restricted (i.e., the Satisfaction Research Questionnaire subscales), some correction for this variance attenuation in one variable must be entertained. McNemar (1969, p. 162) provides the formula for the correction of a correlation when the variance of one of the variables is attenuated.

- -

Insert Table 3.4 about here

- -

Satisfaction Research Questionnaire Subscales and the Scale A Score.

Jackson and Mavrogiannis (1987a) have identified those Survey of Work Style items that best predict coronary heart disease. These items, when aggregated, form the Scale A measure. It was found that all of the Satisfaction Research Questionnaire scales had significant correlations with this Scale A measure in the total sample, substantiating the hypothesis that lower levels of satisfaction are related to greater probability of being positively diagnosed as "Type A", (i.e., possessing the coronary prone behavior pattern). In the entire sample, the Satisfaction Research Questionnaire Challenge, Comfort, Co-Worker, Pay, and Resource Adequacy scales had correlations of -0.11 (p < .05), -.51 (p < .001), -.20 (p < .05), -.26 (p < .001), and -.32 (p < .001), with the Survey of Work Styles Scale A scale, respectively. The Satisfaction Research Questionnaire total scale also had a substantial correlation with the Survey of Work Styles Scale

Table 3.4. Correlations between the Satisfaction Research Questionnaire subscales and the Type A measure, across modal profiles.

Profile

	Correlations			
Subscale	1	2	3	4
1 (n=22)				
Challenge	-.50*	-.65	-.42	-.55
Comfort	-.34	-.43	-.36	-.46
Co-Workers	-.52*	-.68	-.56	-.72
Pay	-.50*	-.65	-.71	-.89
Resource Adequacy	-.11	-.14	-.15	-.19
2 (n=51)				
Challenge	-.09	-.12	-.09	-.11
Comfort	-.24*	-.31	-.46	-.57
Co-Workers	-.03	-.04	-.04	-.05
Pay	-.24*	-.31	-.24	-.31
Resource Adequacy	-.23	-.29	-.30	-.38
3 (n=100)				
Challenge	-.28*	-.36	-.24	-.31
Comfort	-.40*	-.51	-.62	-.77
Co-Workers	-.10	-.13	-.09	-.12
Pay	-.39*	-.51	-.42	-.55
Resource Adequacy	-.17	-.22	-.23	-.29
4 (n=115)				
Challenge	-.01	-.01	-.02	-.02
Comfort	-.43*	-.55	-.63	-.78
Co-Workers	-.18*	-.23	-.27	-.35
Pay	-.18*	-.23	-.38	-.49
Resource Adequacy	-.37*	-.47	-.43	- 54

Note. Correlation 1 is the uncorrected correlation. Correlation 2 is the unreliability corrected correlation. Correlation 3 is the variance attenuation corrected correction. Correlation 4 is the original correlation corrected for unreliability and variance attenutation. * indicates a significant correlation.

A measure (r = -.44, p < .001). This value was not exceeded by the Job Descriptive Index total score correlation with the Scale A measure (-.28, p > .05). The Satisfaction Research Questionnaire Comfort scale had the largest relationship with the Scale A measure, suggesting that being uncomfortable in one's workplace is indeed a correlate of the Scale A score.

The correlations from the modal profiles are presented in Table 3.4. Inspection of the correlations involving the Satisfaction Research Questionnaire modal profiles yielded a differential pattern of relationships between the subscales and the Scale A measure. The first modal profile group had significant correlations with the Scale A measure and the Challenge, Co-Workers, and Pay subscales, as opposed to the second modal profile group that had significant correlations with the Comfort and Resource Adequacy subscales. These findings suggest that low levels of satisfaction are predictive of the Type A Behavior Pattern for the first group, whereas in the second profile, Comfort and Resource Adequacy scores correlate with Scale A responses. This pattern differs somewhat for the third and fourth modal profile groups. The third modal profile group yields significant Scale A scale correlations with the Challenge, Comfort, and Pay subscales. The fourth modal profile group has significant correlations with the Comfort, Co-Workers, Pay, and Resource Adequacy subscales. Again, low satisfaction subscale scores are predictive of high Scale A scores. The presence of four significant Scale A correlations in the fourth modal profile is consistent with the "at risk" label of this profile.

This hypothesis is further validated when one examines the relationship between the Satisfaction Research Questionnaire Comfort scale and the behavioral exemplars of subjective dysfunctional health symptoms (as possible precursors of

physiological stress in the entire sample). Consider, for example, the correlations between the Comfort subscale and the following single items: being overweight (r = -.13 p<.05), always having cold hands and extremities (r = -.10 p<.05), generally feeling stress (r =-.51 p<.001), always feeling tired for no reason (r = -.30 p<.001), would be able to receive a rating of good health (r = +.17 p<.001), having muscle fatigue for no reason (r = -.25 p<.001), feeling irritable for no reason (r = -.17 p<.001), feeling bored for no reason (r = -.16 p<.001), feeling anxious for no specific reason (r = -.27 p<.001), always feeling under pressure (r = -.45 p<.001), lacking any clear expectations about their job (r = -.22 p<.001). All correlations were significant and in the predicted direction.

Several orthogonal stepwise regression analyses were used to predict the secondary measures. Each prediction entered the Satisfaction Research Questionnaire five subscales in one step, and the profile classification in the second step. This was performed to determine the unique contribution of the profile classification. The Scale A measure was the only criterion that yielded a significant (standardized) beta for the unique contribution of the profile classification. This finding suggests that there are significant differences in the Scale A measure across profiles, and that the differences in the other secondary measures across profiles may be due to unique subscale correlations, or relationships with some other variables. As well, the bipolar nature of the modal profiles may tend to "cancel out" specific variance in each other. For example, if Profile one is high on some attribute, and Profile Two is low on that same attribute, then including both Profiles in the same regression might be less informative.

- -

Insert Table 3.5 about here

- -

Satisfaction and Absence. A consistent theme of the present research is that the recognition of the patterns in inter-individual responses will lead to greater precision in classification. Table 3.5 presents the Satisfaction Research Questionnaire scale correlations with the aggregated absence measure. The first column presents the correlations with the absence measure and the Satisfaction Research Questionnaire scale scores for the four modal profiles. Columns two through four present the same correlations, corrected for unreliability, corrected for variance attenuation in one variable, and finally corrected for both unreliability and variance attentuation.

Across the first two modal profiles, a substantial relationship with absence is evidenced. Modal profile one would seem to be the respondent group that contains the satisfaction-absence relationship. The Satisfaction Research Questionnaire Challenge and Comfort scales possess the highest correlations, making substantive sense. For example, people would not be repeatedly absent if they were dissatisfied with their amount of pay. The total Satisfaction Research Questionnaire scale score correlation with the absence measure was -.47 (uncorrected), providing support for the hypothesis that lower satisfaction is related to greater absenteeism. Moreover, the majority of the variance lies in the correlations with Challenge and Comfort. No other profile provides such information.

- -

Insert Table 3.6 about here

- -

Table 3.5. Correlations between the Satisfaction Research Questionnaire subscales with the absence measure, across modal profiles.

Profile

Subscale	Correlations 1	2	3	4
1 (n=22)				
Challenge	-.65*	-.75	-.55	-.65
Comfort	-.65*	-.74	-.68	-.77
Co-Workers	-.33	-.38	-.36	-.41
Pay	-.08	-.09	-.12	-.14
Resource Adequacy	-.08	-.09	-.11	-.12
2 (n=51)				
Challenge	+.18	+.21	+.18	+.20
Comfort	+.12	+.14	+.23	+.26
Co-Workers	+.21	+.24	+.26	+.30
Pay	+.23	+.27	+.23	+.26
Resource Adequacy	+.20	+.23	+.26	+.30
3 (n=100)				
Challenge	-.01	-.01	-.01	-.01
Comfort	-.16	-.18	-.26	-.30
Co-Workers	-.06	-.07	-.05	-.06
Pay	-.06	-.07	-.07	-.08
Resource Adequacy	-.15	-.17	-.20	-.23
4 (n=115)				
Challenge	-.12	-.14	-.18	-.21
Comfort	+.03	+.03	+.05	+.05
Co-Workers	-.08	-.09	-.12	-.14
Pay	-.07	-.08	-.15	-.17
Resource Adequacy	-.10	-.11	-.12	-.13

Note. Correlation 1 is the uncorrected correlation. Correlation 2 is the unreliability corrected correlation. Correlation 3 is the variance attenuation corrected correction. Correlation 4 is the original correlation corrected for unreliability and variance attenutation. * indicates a significant correlation.

Table 3.6. Correlations between the Satisfaction Research Questionnaire subscales and the job ambiguity measure, across modal profiles.

Profile

Subscale	Correlations 1	2	3	4
1 (n=22)				
Challenge	-.35	-.46	-.29	-.38
Comfort	-.33	-.42	-.35	-.44
Co-Workers	-.52*	-.68	-.56	-.72
Pay	-.47*	-.61	-.67	-.85
Resource Adequacy	-.29	-.37	-.39	-.49
2 (n=51)				
Challenge	-.31*	-.40	-.30	-.39
Comfort	-.31*	-.40	-.58	-.72
Co-Workers	-.32*	-.42	-.39	-.50
Pay	-.31*	-.40	-.31	-.40
Resource Adequacy	-.24*	-.31	-.31	-.40
3 (n=100)				
Challenge	-.27*	-.35	-.23	-.30
Comfort	-.20*	-.26	-.32	-.41
Co-Workers	-.21*	-.27	-.19	-.24
Pay	-.28*	-.36	-.31	-.40
Resource Adequacy	-.34*	-.43	-.45	-.56
4 (n=115)				
Challenge	-.08	-.10	-.12	-.16
Comfort	-.42*	-.54	-.61	-.76
Co-Workers	-.20*	-.26	-.30	-.39
Pay	-.23*	-.30	-.48	-.61
Resource Adequacy	-.51*	-.65	-.58	-.73

Note. Correlation 1 is the uncorrected correlation. Correlation 2 is the unreliability corrected correlation. Correlation 3 is the variance attenuation corrected correction. Correlation 4 is the original correlation corrected for unreliability and variance attenutation. * indicates a significant correlation.

Satisfaction and Job Ambiguity. Table 3.6 presents the same series of corrected correlations, only now between satisfaction and job ambiguity. Nearly all satisfaction subscales across profiles yielded significant correlations with the Job Ambiguity measure. This finding suggests that low levels of satisfaction are related to greater levels of job ambiguity. Greater job ambiguity was significantly related to lower Co-Workers and Pay satisfaction in Profile One, all subscales in Profile Two, all subscales in Profile Three, and the Comfort, Co-Workers, Pay, and Resource Adequacy subscales in Profile Four. This pervasive relationship between job ambiguity and satisfaction would suggest that the reduction of job ambiguity should lead to greater levels of satisfaction in all profile groups.

- -

Insert Table 3.7 about here

- -

Satisfaction and Good Self Maintenance. The Good Self Maintenance scale was designed to reflect the frequency of beneficial self maintenance behaviors carried out by the respondents in the cross-Canada sample. In the total sample, greater levels of satisfaction were related to greater scores on this measure for the Comfort, Co-Workers, and Pay satisfaction subscales.

- -

Insert Table 3.8 about here

- -

Satisfaction and Subjective Health Dysfunctional Symptoms. All satisfaction subscales yielded significant correlations with this measure in the total group. In Table 3.8, the Profile One respondents had only the first two

Table 3.7. Correlations between the Satisfaction Research Questionnaire subscales and the good self maintenance measure, across modal profiles.

Profile

Subscale	Correlations 1	2	3	4
1 (n=22)				
Challenge	**			
Comfort				
Co-Workers				
Pay				
Resource Adequacy				
2 (n=51)				
Challenge	+.48*	+.73	+.47	+.72
Comfort	+.16	+.24	+.31	+.46
Co-Workers	+.23	+.35	+.28	+.43
Pay	+.16	+.24	+.16	+.24
Resource Adequacy	+.13	+.19	+.17	+.25
3 (n=100)				
Challenge	+.04	+.06	+.03	+.05
Comfort	+.12	+.18	+.20	+.29
Co-Workers	-.04	-.06	-.04	-.05
Pay	-.06	-.09	-.07	-.10
Resource Adequacy	-.14	-.21	-.19	-.28
4 (n=115)				
Challenge	-.01	-.02	-.02	-.02
Comfort	+.39*	+.58	+.57	+.82
Co-Workers	+.50*	+.76	+.72	+.99
Pay	+.51*	+.78	+.96	+.99
Resource Adequacy	+.42*	+.63	+.48	+.71

Note. Correlation 1 is the uncorrected correlation. Correlation 2 is the unreliability corrected correlation. Correlation 3 is the variance attenuation corrected correction. Correlation 4 is the original correlation corrected for unreliability and variance attenutation. * indicates a significant correlation. ** indicates no correlation due to missing good self maintenance data for profile 1.

Table 3.8. Correlations between the Satisfaction Research Questionnaire subscales and the tertiary coronary heart disease measure, across modal profiles.

Profile

Subscale	Correlations 1	2	3	4
1 (n=22)				
Challenge	-.60*	-.80	-.50	-.70
Comfort	-.58*	-.76	-.61	-.79
Co-Workers	-.11	-.15	-.12	-.16
Pay	-.10	-.13	-.15	-.20
Resource Adequacy	-.17	-.22	-.23	-.30
2 (n=51)				
Challenge	-.02	-.03	-.02	-.03
Comfort	-.17	-.22	-.33	-.43
Co-Workers	+.11	+.15	+.14	+.18
Pay	+.11	+.15	+.11	+.15
Resource Adequacy	-.14	-.18	-.18	-.24
3 (n=100)				
Challenge	-.21*	-.28	-.18	-.24
Comfort	-.36*	-.47	-.57	-.72
Co-Workers	-.21	-.28	-.19	-.25
Pay	-.18	-.24	-.20	-.26
Resource Adequacy	-.06	-.08	-.08	-.11
4 (n=115)				
Challenge	-.16	-.21	-.24	-.32
Comfort	+.52*	+.68	+.74	+.93
Co-Workers	-.35*	-.47	-.52	-.68
Pay	-.13	-.17	-.28	-.37
Resource Adequacy	-.24*	-.32	-.28	-.36

Note. Correlation 1 is the uncorrected correlation. Correlation 2 is the unreliability corrected correlation. Correlation 3 is the variance attenuation corrected correction. Correlation 4 is the original correlation corrected for unreliability and variance attenutation. * indicates a significant correlation.

subscales yielding significant values, whereas none the of the Profile Two correlations were significant. As in the Profile One group, the Profile Three respondents have significant correlations with the first two subscales. In the last profile, the "at risk for coronary heart disease" profile, the Comfort, Co-Workers, and Resource Adequacy subscales yielded significant values.

- -

Insert Table 3.9 about here

- - - - - - - - - - - - - - - - - - - -

Satisfaction and Performance. The results portrayed in Table 3.9 are similar to the previous Table in that one group of respondents has a marked relationship with the measure of interest. As before, there is no discernible relationship between satisfaction and productivity when the correlations based upon the total sample are used. The correlation between the Satisfaction Research Questionnaire total score and the productivity measure for the first profile respondents was .59 (uncorrected). The productivity of these respondents would seem to depend more upon favorable levels of comfort, being satisfied with pay and status within the organization, and having the resources (tools, information, and support from above) to do the job. These three Satisfaction Research Questionnaire subscales all had greater correlations with Performance than the Challenge and Co-Workers subscales.

- -

Insert Table 3.10 about here

- -

Relationships Among the Secondary Measures Across Profiles. Table 3.10 presents the correlations among the six secondary measures for the total sample.

Table 3.9. Correlations between the Satisfaction Research Questionnaire subscales and the subjective performance measure, across modal profiles.

Profile

Subscale	Correlations 1	2	3	4
1 (n=22)				
Challenge	+.47*	+.71	+.39	+.61
Comfort	+.53*	+.79	+.56	+.82
Co-Workers	+.46*	+.70	+.49	+.74
Pay	+.52*	+.79	+.74	+.99
Resource Adequacy	+.51*	+.76	+.66	+.93
2 (n=51)				
Challenge	+.18	+.27	+.18	+.27
Comfort	+.14	+.21	+.27	+.40
Co-Workers	-.07	-.11	-.09	-.13
Pay	+.06	+.09	+.06	+.09
Resource Adequacy	-.09	-.13	-.12	-.17
3 (n=100)				
Challenge	+.16*	+.24	+.14	+.21
Comfort	+.09	+.13	+.15	+.22
Co-Workers	+.04	+.06	+.04	+.05
Pay	+.05	+.08	+.05	+.08
Resource Adequacy	+.06	+.09	+.08	+.12
4 (n=115)				
Challenge	+.31*	+.47	+.46	+.67
Comfort	+.04	+.06	+.06	+.09
Co-Workers	+.13	+.20	+.20	+.30
Pay	-.02	-.03	-.04	-.06
Resource Adequacy	+.12	+.18	+.14	+.21

Note. Correlation 1 is the uncorrected correlation. Correlation 2 is the unreliability corrected correlation. Correlation 3 is the variance attenuation corrected correction. Correlation 4 is the original correlation corrected for unreliability and variance attenutation. * indicates a significant correlation.

Table 3.10 Correlations among the six secondary measures, for the cross-Canada sample (n=388).

Scale A	.73	.07	.18	-.47	.58	.05
Absenteeism	.06	.92	.05	-.30	.45	-.01
Job Ambiguity	.13	.04	.73	.03	.27	-.24
Good Self Maintenance	-.29	-.21	.02	.53	-.55	.36
Dysfunctional Subjective Health Symptoms	.41	.36	.19	-.33	.69	.05
Performance	.03	-.01	-.15	.19	-.03	.54

Note. Boldface diagonal values are the reliabilities for the second measures. Above diagonal elements are the unreliability corrected correlations, corrected values greater than unity are reported as .99. Below diagonal elements are raw correlations. Underlined values are significant at p<.05.

The boldface values in the main diagonal are the reliabilities for each measure, the correlations are contained within the lower triangle, and the correlations corrected for unreliability are the italicized values above the main diagonal. In this section, the relationships among the six secondary measures are reviewed. In the total sample there was a significant positive correlation between the Scale A and the Subjective Health Dysfunctional Symptoms measures. The significant negative correlations between the Scale A and Good Self Maintenance measures and between the Subjective Health Dysfunctional Symptoms and Good Self Maintenance measures provide empirical support for the validity of all three measures. For example, the correlation between the Subjective Health Dysfunctional Symptoms measure and the Scale A measure was +.41. The corrected value for this relationship was +.58. The Subjective Health Dysfunctional Symptoms measure correlated significantly with the Scale A measure (r = .41, r(corrected) = .58), the Absenteeism measure (r = .36, r(corrected) = .45), the Job Ambiguity measure (r = .19, r(corrected) = .27), and the Good Self Maintenance measure (r = -.33, r(corrected) = -.55). All these correlations were in the predicted direction. For example, that good self maintenance habits should predict lower levels of subjective health dysfunctional symptoms.

- - - - - - - - - - - - - - - - - - - -

Insert Table 3.11 about here

- - - - - - - - - - - - - - - - - - - -

Modal Profile One. Table 3.11 also presents the correlations among the six secondary measures in the lower triangle. The main diagonal contains the total sample reliabilities, and the italicized above-diagonal values are the corrected

Table 3.11 Correlations among the six secondary measures, for Profile One (n=22).

Scale A	1.00	-.10	.11	*	-.06	-.60
Absenteeism	-.08	1.00	.09	*	.99	-.41
Job Ambiguity	.14	.07	1.00	*	-.17	-.73
Good Self Maintenance	*	*	*	*	*	*
Dysfunctional Subjective Health Symptoms	-.04	.86	-.12	*	1.00	.99
Performance	-.37	-.29	-.46	*	-.86	1.00

Note. Above diagonal elements are the unreliability corrected correlations, corrected values greater than unity are reported as .99. Below diagonal elements are raw correlations. Underlined values are significant at p<.05. * denotes excessive missing for this variable in this profile group.

correlations. The notable relationships for this group were a) the significant positive correlation between Subjective Health Dysfunctional Symptoms and Absence, b) the significant negative correlation between Performance and the Scale A measure, and 3) the significant negative correlation between Performance and Job Ambiguity.

- -

Insert Table 3.12 about here

- -

Modal Profile Two. As Table 3.12 clearly shows, there were no significant correlations between any pair of the six secondary measures.

- -

Insert Table 3.13 about here

- -

Modal Profile Three. For this group, the Scale A measure correlated significantly and negatively with the Good Self Maintenance measure and significantly and positively with the Subjective Health Dysfunctional Symptoms measure. Table 3.13 shows that absenteeism was predicted by both Job Ambiguity and Subjective Health Dysfunctional Symptoms, while there was a negative relationship between Good Self Maintenance and Subjective Health Dysfunctional Symptoms.

- -

Insert Table 3.14 about here

- -

Modal Profile Four. Table 3.14 shows that the Scale A measure correlated significantly and positively with the Job Ambiguity and Subjective Health

Table 3.12 Correlations among the six secondary measures, for Profile Two (n=51).

Scale A	1.00	-.18	-.01	-.30	.11	.02
Absenteeism	-.15	1.00	-.02	.26	.13	-.09
Job Ambiguity	-.01	-.02	1.00	-.02	-.06	-.27
Good Self Maintenance	-.18	-.18	-.01	1.00	.56	.71
Dysfunctional Subjective Health Symptoms	.08	.10	-.04	.34	1.00	.29
Performance	.01	-.07	-.17	.38	.18	1.00

Note. Above diagonal elements are the unreliability corrected correlations, corrected values greater than unity are reported as .99. Below diagonal elements are raw correlations. Underlined values are significant at p<.05.

Table 3.13 Correlations among the six secondary measures, for Profile Three (n=100).

Scale A	1.00	.16	-.01	-.72	.79	.06
Absenteeism	.13	1.00	.23	-.17	.50	-.01
Job Ambiguity	-.01	.19	1.00	.37	.21	-.21
Good Self Maintenance	-.45	-.12	.23	1.00	-.78	.07
Dysfunctional Subjective Health Symptoms	.56	.40	.15	-.47	1.00	-.26
Performance	.04	-.01	-.13	.04	-.16	1.00

Note. Above diagonal elements are the unreliability corrected correlations, corrected values greater than unity are reported as .99. Below diagonal elements are raw correlations. Underlined values are significant at p<.05.

Table 3.14 Correlations among the six secondary measures, for Profile Four (n=115).

Scale A	1.00	-.06	.45	-.40	.73	.06
Absenteeism	-.06	1.00	.05	-.46	.68	.10
Job Ambiguity	_.37_	.04	1.00	-.27	.27	-.27
Good Self Maintenance	-.29	-.32	-.17	1.00	-.99	-.06
Dysfunctional Subjective Health Symptoms	_.52_	.54	.19	_-.66_	1.00	.28
Performance	.04	.07	-.17	-.03	.17	1.00

Note. Above diagonal elements are the unreliability corrected correlations, corrected values greater than unity are reported as .99. Below diagonal elements are raw correlations. Underlined values are significant at p<.05.

Dysfunctional Symptoms measures in this group. The latter measure also correlated significantly (and positively) with the Absence measure and significantly and negatively with the Good Self Maintenance measure.

To summarize, there was empirical evidence of convergent and discriminant validity of the six secondary measures, and of the utility of the modal profiles. The relationships with the Scale A measure, for example, might be considered unclear by examining the total sample correlations only. Examination of the within-profile correlations with the Scale A measure, however, has shown that the first group's performance was related to the Scale A measure, the second group had no correlations with the Scale A measure, the third group had a robust relationship with poor self maintenance, and finally, the fourth group was notable for the Job Ambiguity and Subjective Health Dysfunctional Symptoms relationships with the Scale A measure. These differential results across profile groups empirically demonstrate some of the poor utility of expecting a single predictor-criterion relationship to remain s ,ble across a variety of individual differences. The data suggest that meaningful discussions of workplace criteria should acknowledge that a) some respondent classificatory procedure (e.g., modal profile analysis) should be applied to classify groups of people, or b) the presence of individual difference variables in the workplace (e.g., personality traits or tvpes) be recognized, or c) ideally, both a) and b).

Satisfaction and Desirability Responding. To test the hypothesis that the variance in the Job Descriptive Index scores attributable to desirability responding was greater than the corresponding score variance in the Satisfaction Research Questionnaire, the unrotated first principal components of both tests

were correlated with the desirability measure. The hypothesis was supported, with the correlation between desirability and the first unrotated principal component of the Job Descriptive Index being 0.35. The corresponding value for the Satisfaction Research Questionnaire was 0.19.

Although the difference between the two correlations was as predicted, the size of these correlations might have been constrained for at least three reasons.

First, the type of desirability assessed by the Personality Research Form desirability scale may not be precisely the same desirability as that underlying the attribution of undesirable job qualities.

Second, the endorsement of items such as "stupid", "lazy", and "intelligent" to describe a job might be different from the bias measured by the Personality Research Form Desirability scale. To address this issue, the Job Descriptive Index items were correlated with the desirability scale score. In the Work on Present Job scale, the items with the highest correlation with desirability were: "fascinating", "satisfying", "challenging", "gives sense of accomplishment", "good", "useful", "boring", and "respected". Those such items on the Present Pay scale were: "income inadequate for normal expenses", "insecure" (negative correlation), and "bad". For the Opportunities for Promotion scale, the relevant items were: "fairly good chance for promotion" (negative correlation), "unfair promotion policy", and "dead-end job". The Supervision on Present Job scale yielded the following items: "asks my advice", "doesn't supervise enough", "bad", "impolite", and "influential". In the People on Your Present Job scale, the items with the highest correlation with the desirability scale score were, "smart", "slow", "unpleasant", "stupid", and "intelligent". It may be that the "desirable" response

mode for respondents is to complain about their jobs in a manner that depicts the jobs as: 1) interesting but not sufficiently challenging, 2) having a low but secure income, 3) having little chance for promotion, 4) supervised by poor supervisors, and 5) having co-workers that are pleasant, but not as capable as the respondent. This pattern of responses is similar to the attribution error that respondents make when asked to rate their own ability -- namely excessive positive leniency errors. Meyer (1980) found that when a sample of engineers were asked to rate their own performance relative to other engineers, most of the respondents felt that they were better than 75% of their workforce. The picture of a prototypical respondent is a person who claims to be a good worker, but who also claims that she or he "could be better if given a break".

Third, endorsement of some items, such as those above, may indicate actual dissatisfaction and therefore not correlate with desirability. Therefore, another approach to the resolution of desirability responding in the Job Descriptive Index at the item level is presented in the next section.

Desirability Responding in the Job Descriptive Index. It has been suggested that the suboptimal test development practices associated with the Job Descriptive Index might leave the instrument more susceptible to desirability or other non-content sources of response variance. To test this hypothesis, several analyses were completed. It was found, for example, that the test has items whose correlations with the desirability scale are greater in magnitude than those items' correlations with their own scale.

- -

Insert Table 3.15 about here

- -

Table 3.15 Correlations among the six secondary measures, for the cross-Canada sample (n=388).

	Correlation between item and its scale	Correlation between item and desirability	DRI value	Item endorsement proportion
Mean	.39	.16	.40	.70
Standard Deviation	.22	.10	.14	.18
Minimum	-.69	-.24	.00	.18
Maximum	.71	.34	.70	.96

In the cross-Canada mail-out sample, the average item-scale correlation for the Job Descriptive Index was .39 (this value was .55 for the Satisfaction Research Questionnaire, as indicated above). The mean item-desirability correlation was .16 (.11 for the Satisfaction Research Questionnaire). The average Item Efficiency Index value for the Job Descriptive Index items was .40 (as opposed to .55 for the Satisfaction Research Questionnaire). Lastly, the average item endorsement proportion for the Job Descriptive Index was .40, as compared to an ideal value of .50, and the value of .52 from the Satisfaction Research Questionnaire.

Another approach to the resolution of desirability responding in the Job Descriptive Index issue can be effected by looking at the pattern of "contaminated" Job Descriptive Index items. It was hypothesized that the Job Descriptive Index items with the highest correlations with desirability would also have high correlations with the unrotated first principal component of the Job Descriptive Index. Since it was also hypothesized that this first component would correlate with desirability, it was also postulated that the pattern of correlations between desirability and the first unrotated principal component would match the pattern of salient Job Descriptive Index item-total correlations. In short, it was predicted that the pattern of item-scale correlations would match the pattern of item-desirability correlations, indicating a presence of this response bias. Items that had such correlations with desirability were deleted from the Satisfaction Research Questionnaire during development stages.

For illustrative purposes, across each Job Descriptive Index subscale, a dichotomous classification was made regarding each item's a) correlation with desirability, b) correlation with the factor scale reflecting the first unrotated

principal component of this test, and c) the item-subscale correlation. The tetrachoric correlations between these indexes were calculated to indicate the similarity of these 3 patterns. In the Job Descriptive Index Work on Present Job subscale, for example, there is an expected correlation between the items and the subscale. However, there was also a relationship between the pattern of item-desirability relationships and both the pattern of item-subscale correlations and the pattern of item-first unrotated principal component loadings. These findings suggested that the pattern of responses that the Job Descriptive Index authors would interpret as meaningful share a fundamental structure with the desirability response bias.

The previous sections have documented the sub-optimal test development practices implemented in the Job Descriptive Index creation. Those test development practices had not been designed to protect the test from this response bias. It has been shown that the test as a whole is confounded by desirability responding, and this last section has illustrated this contamination at the item level. While the original Job Descriptive Index authors had contended that this bias was not present in their test, it has been empirically demonstrated that any hypothesized bona fide item-construct relationships are highly suspect.

Sex Differences in Satisfaction. To determine whether or not any of the Job Descriptive Index or Satisfaction Research Questionnaire items yielded different item endorsement proportions for males and females, the p levels were examined for both tests across sex.

The Satisfaction Research Questionnaire. The endorsement proportions for the Satisfaction Research Questionnaire items were calculated for male (n=151) and female (n=150) respondents in the cross-Canada sample. Almost all items had

a between sex endorsement proportion difference of less than .10. Eight items had an endorsement proportion difference greater than .10, but no item yielded a p-value difference greater than .21. The first of the eight items was answered more frequently by males, and was the "Work gives me time to think about things" items (p-value difference = .21). Females in the sample reported feeling "pretty relaxed in my work area" more often than males (p-value difference = .14). Females in the cross-Canada sample also endorsed the following items more frequently: "Some of my co-workers have become good friends" (p-value difference = .14); "I really think my workplace brings out my best" (p-value difference = .13); "There is always something urgent and pressing at work" (p-value difference = .12); "Basically I do the same thing every day" (p-value difference = .15); "There is always someone to give me directions about resources" (p-value difference = .15), and; "There is always someone to explain information to me" (p-value difference = .20).

Given that only 10% of the Satisfaction Research Questionnaire items had endorsement proportion differences between the sexes of .10 to .21, it was concluded that males and females respond similarly to the items. Moreover, there was little evidence to suggest that there were sex differences in endorsement proportions particular to any single scale. This finding was also evidenced at the mean levels. A possible explanation for some of these observed differences may be that there was a confounding of jobs and sex, and types of jobs, e.g., offices, and others. There may have also been a sex by job type interaction, i.e., more females in offices, more males in warehouses, although since the mailout did not gather specific job title information, these issues cannot be addressed.

As a replication of the above analysis, the Industrial/Organizational Psychology course subjects (University Sample) data were utilized. The male respondent item endorsement proportions were compared with the female respondent item endorsement proportions. This sample contained 135 respondents, consisting of 81 females and 54 males. Again, the two groups were very similar in their pattern of item responses. The only item that yielded a substantial between sex endorsement proportion difference was the item "I think my workplace brings out my best", for which females endorsed the item in the keyed direction more frequently than did the males (p-value difference = .24).

Therefore, in both the Cross-Canada and University samples, males and females were responding to the Satisfaction Research Questionnaire Items in a similar manner. Indeed, upon inspection of the profile groups, there was no single profile that was markedly sex-specific.

The Job Descriptive Index. A similar inspection of item endorsement proportions was effected for the Job Descriptive Index items. Nine of the 72 items yielded between sex p-value differences of .10 or greater, (with no difference exceeding .18). Females endorsed the following Work On Present Job items more frequently: "Pleasant" (p-value difference = .17); "Satisfying" (p-value difference = .13), and; "Good" (p-value difference = .11). Males, however, endorsed the following items on the same subscale more frequently: "Creative" (p-value difference = .18), and; "Routine" (p-value difference = .14). Males answered the Present Pay item "Highly paid" (p-value difference = .10) more frequently than did females, as well as the following "People on Your Present Job" items: "Ambitious" (p-value difference = .10), and; "No privacy" (p-value difference = .10). Females endorsed the Supervision on Present Job item "Tactful"

(p-value difference .15) more frequently than males. Yet, on the whole, the Job Descriptive Index items did not distinguish between males and females. The lack of discrimination between respondent groups by using endorsement proportions reflects the test authors' earlier attempts to eliminate all such effects.

As with the Satisfaction Research Questionnaire, the university sample did not yield any different patterns of responses than the cross-Canada sample for the Job Descriptive Index.

CHAPTER FOUR: SUMMARY AND CONCLUSIONS

The aim of the first chapter of this dissertation was to illustrate the need for a modern measure of job satisfaction. The second chapter outlined a set of ideal requirements for such a modern measure of job satisfaction, and then attempted to create such a modern measure. In the previous chapter, the empirical evidence relating to this recent measure was presented. This final chapter begins with a brief summary of each of the previous chapters. After these summaries, a new model of satisfaction is proposed.

Chapter one portrayed satisfaction as a concept that was once popular, but has since faded from most views of Industrial and Organizational Psychology. The search for simple predictor-criterion correlations of large magnitude has decidedly failed, and this failure has been partially responsible for the decline of the satisfaction construct. Moreover, it was suggested that poor measurement practices were also partially responsible for the decline. Various calls have been made to salvage the worth of assessing workplace attitudes. Earlier theories of organizational behavior were shown to have some merit.

The second chapter built upon the first chapter's recommendations for a better job satisfaction measure, and centered on the premise that a modern approach to test construction could begin to overcome the historical shortcomings of the area. The recent satisfaction measure was introduced, its development guidelines were made explicitly clear, and the procedure for classifying respondents by their patterns of satisfaction was demonstrated empirically.

Chapter three was designed to compare the Satisfaction Research Questionnaire with a popular extant measure. Comparisons were made along several lines. The stability of the recent satisfaction measure was demonstrated

using several reliability schemes. The construct validity of the recent measure was examined by investigating the measure's relationships with the social desirability response bias, criterion correlations with other satisfaction measures, and a measure of the Type A Behavior pattern. In general, the fidelity of both the recent measure and the test development rationale were supported. Clearly, more research is required to fully map out the desirability response bias as it applies to self-report measures of workplace attitudes. The desirability measure applied in this research may have been more appropriate for personality research rather than work attitudes research. The findings of this dissertation, however, do suggest a desirability-like pattern of responses, where endorsement still reflects a positive self-presentation bias. Research should be designed to generate items that would reflect this content. For example, items that foster the "I'm a good worker (regardless of my level of ability)" response bias.

The application of the modal profile classification scheme indicated that the pattern of satisfaction is related to the presence of the Type A Behavior Pattern. The labelling of the fourth Profile as the "at risk for stress" profile was substantiated by the elevated Scale A scores for that profile, as well as elevated values for the job ambiguity and subjective health dysfunctional symptoms measures. The empirical literature corroborates this relationship between stress, the Type A Behavior Pattern, job ambiguity, symptomatology, and satisfaction. The results also indicated that some members of all profile groups could be classified as being Type A. A direct implication of this finding is that intervention need not be aimed exclusively at the Profile Four members. Broadband intervention would still be valid. It was also shown that respondents can be absent for a variety of correlated reasons. This finding led to the

speculation that some workers may only require an excuse in order to absent themselves from work. The historical morale literature suggested that if morale was poor, then workers would find any reason to be absent from work. Future research might investigate the tendency to be absent from work as an individual coping mechanism for dealing with stressors, or as a personalogical attribute (cf. Adler & Golan, 1981; Breaugh, 1981; Cheloha & Farr, 1980; Clegg, 1983; Dittrich & Carrel, 1979). The negative relationship between absence and Challenge and Comfort might suggest that immediate workplace interventions review the complexity of the job and the attributes of the workplace related to low Comfort scores.

The uniform finding that job ambiguity is related to lower satisfaction across all profile groups pointed to job ambiguity as a pervasive aspect of the Canadian workplace. Either people expect unrealistic levels of information about their job, or more probably, employers have not fared well in their attempt to communicate organizational goals and immediate worker goals. Taylor long ago described the declaration of explicit goals as a cornerstone of performance and improved attitude, but many employers have not agreed with this view. To the extent that good self maintenance behaviors are linked to attitudes for some workers, then modern programs aimed at improving employee fitness might be well-founded.

Good self maintenance behaviors are inversely related to subjective health dysfunctional symptoms. The possibility exists that the subjective health dysfunctional symptoms measure is more representative of hypochondriasis or work-induced symptoms than of genuine coronary disease symptomatology. Even if the dysfunctional subjective health symptoms should prove to not predict

subsequent mortality, the current results have suggested that the establishment of such "good self maintenance" programs may be linked to lower absenteeism, lower Scale A values, and reduced job ambiguity. Such possibilities highlight the need to identify worker patterns of satisfaction and stress. A workplace assessment program could identify those workers that have a negative relationship between Scale A and performance. Overall satisfaction -- morale -- to use the historical term, has been used to describe an aspect of job satisfaction. A proposal is suggested below for a modern theory of job satisfaction that incorporates this construct.

A Modern Theory of Job Satisfaction

One aim of this research project has been to explore further the reasons that people differ in their levels of job satisfaction. Historically, the concept has been poorly assessed, and the differences between respondents' pattern of responses have been ignored. Yet even when these aspects have been considered, other reasons for individual differences in satisfaction remain.

Earlier models of satisfaction (e.g., Taylor) proposed that satisfaction in the workplace was a function of several workplace attributes. These attributes included the selection, training, and performance feedback aspects of employed life. More recent theories of worker satisfaction have tended not to incorporate these dimensions of the workplace when considering satisfaction. Rather, popular models have sought to include a variety of moderator variables. These efforts have probably been confounded by the lack of distinction between personality measures as correlates of satisfaction and personality as a predictor of workplace performance.

This dissertation proposes that satisfaction again be viewed in the larger context of the entire workplace -- including its procedural attributes. These attributes should be considered as they determine (via extensive preselection) the final placement of individuals into positions. Without consideration of this preselection effect, large portions of variance related to satisfaction, other attitudes, and organizational outcomes will be constrained. A source of preselection that applicants perform can be seen in their use of lay or implicit personality theory when applying for employment. When economic forces, ability, and training permit, applicants will tend to self-select for those positions that are most congruous with their own personalities. When applicants are constrained to accept those positions that do not match their abilities, training, or personality types, the upper range of their satisfaction scores will be attenuated. For example, consider a computer programmer high on the Personality Research Form traits of Need Order, Understanding, and Endurance. This person wi'l probably be less pleased with the position of gas station attendant than with the position of advanced grade systems analyst.

In short, any consideration of satisfaction should incorporate reliably assessed organizational variables, profiles versus single scale classification procedures, and individual difference variables such as personality and cognitive ability profiles. Previous models of satisfaction and workplace attitudes have not typically acknowledged this wider scope. Earlier publications tend to focus on minute aspects of organizational life, often as a method for maintaining experimental rigor. In order to recognize the breadth of this diversity, future models must incorporate construct-valid multidimensional measures and recognize many situational variables.

Such a model would incorporate the following: a) individual differences in satisfaction, b) the principles of selection, training, performance management, and feedback, and c) individual differences in general affect (i.e., mood), general cognitive ability, and personality. These model elements should be brought together under the auspices of modern scale development procedures and classification schemes similar to those applied in the current research.

REFERENCES

Adler, S., & Golan, J. (1981). Lateness as a withdrawal behavior. **Journal of Applied Psychology, 66,** 544-554.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. **Psychological Bulletin, 84,** 888-918.

Arnold, H.J. (1982). Moderator variables: A clarification of conceptual, analytic and psychometric issues. **Organizational Behavior and Human Performance, 29,** 143-174.

Arnold, H.J., & Feldman, D.C. (1981). Social desirability response bias in self-report choice situations. **Academy of Management Journal,** 24, 377-385.

Barling, J., & Rosenbaum, A. (1986). Work stressors and wife abuse. **Journal of Applied Psychology, 71,** 346-348.

Beehr, T.A. (1976). Perceived situational moderators of the relationship between subjective role ambiguity and role strain. **Journal of Applied Psychology,** 61, 35-40.

Beehr, T.A., Walsh, J.T., & Taber, T.D. (1976). Relation of stress to individually and organizationally valued states. **Journal of Applied Psychology,** 61, 41-47.

Bereiter, C.E. (1963). Some persistent dilemma in the measurement of change. In C.W. Harris (Ed.), **Problems in measuring change.** Madison: University of Wisconsin Press. Pp. 3-20.

Blum, M.L. & Naylor, J.C. (1968). **Industrial psychology: Its theoretical and social foundations.** New York: Harper & Row.

Bramel, D., & Friend, R. (1981). Hawthorne, the myth of the docile worker, and class bias in psychology. **American Psychologist, 36,** 867-878.

Brayfield, A.H., & Crockett, D.W. (1955). Employee attitudes and employee performance. **Psychological Bulletin, 52,** 396-424.

Breaugh, J.A. (1981). Predicting absenteeism from prior absenteeism and work attitudes. **Journal of Applied Psychology, 66,** 555-560.

Bruning, N.S., & Frew, D.R. (1987). Effects of exercise, relaxation, and management skills training on physiological stress indicators: A field experiment. **Journal of Applied Psychology, 72,** 515-521.

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. **Psychological Bulletin, 56,** 81-105.

Caplan, R.D.. (1971). **Organizational stress and individual strain: A social-psychological study of risk factors in coronary heart disease among Administrators, Engineers, and Scientists.** Institute for Social Research: Ann Arbor, MI.

Caplan, R.D., & Jones, K.W. (1975). Effects of workload, role ambiguity, and Type A personality on anxiety, depression, and heart rate. **Journal of Applied Psychology, 60,** 713-719.

Cheloa, R.S., & Farr, J.L. (1890). Absenteeism, job involvement and job satisfaction in an organizational setting. **Journal of Applied Psychology, 65,** 467-473.

Clegg, C.W. (1983). Psychology of employee lateness, absence and turnover: A methodological critique and an empirical study. **Journal of Applied Psychology, 68,** 88-101.

Cohen, J., & Cohen, P. (1975). **Applied multiple regression/correlation analysis for the behavioral sciences.** Hillsdale, NJ: Lawrence-Erlbaum.

Cole, R. (1979). Work, mobility, and participation. A comparative study of American and Japanese industry. Berkley, CA: University of California Press.

Cook, J.D., Hepworth, S.J., Wall, T.D., & Warr, P.B. (1981). The experience of work: A compendium and review of 249 measures and their use. London: Academic Press.

Cronbach, L.J., & Furby, L. (1970). How should we measure change -- or should we? Psychological Bulletin, 74, 68-80.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Cronbach, L.J., & Warrington, W.G. (1952). Efficiency of multiple-choice tests as a function of spread on item difficulties. Psychometrika, 17, 127-147.

Dachler, H.P., & Hulin, C.L. (1969). A reconsideration of the relationship between satisfaction and judged importance of environmental and job characteristics. Organizational Behavior and Human Performance, 4, 252-266.

Decker, R.L. (1955). A study of three specific problems in the measurement and interpretation of employee attributes. Psychological Monographs, 69.

DeMan, H. (1927). Joy in work. New York: Holt.

Dittrich, J.E., & Carrell, M.R. (1979). Organizational equity perceptions, employee job satisfaction, and departmental absence and turnover rates. Organizational Behavior and Human Performance, 24, 29-40.

Draper, N.R., & Smith, H. (1981). Applied regression analysis. (2nd. ed.). New York: John Wiley and Sons.

Edwards, A.L. (1957). The social desirability variable in personality assessment and research. New York: Dryden.

Edwards, A.L. (1963). A factor analysis of experimental social desirability and response set scales. **Journal of Applied Psychology, 47,** 308-316.

Ewen, R.B. (1967). Weighting components of job satisfaction. **Journal of Applied Psychology, 51,** 68-73.

Fekken, G.C., & Jackson, D.N. (June, 1984). Predicting psychological test item response stability. Paper presented at the Annual convention of the Canadian Psychological Association, Ottawa.

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. **Psychological Review, 81,** 59-74.

Fisher, C.D. (1980). On the dubious wisdom of expecting job satisfaction to correlate with performance. **Academy of Management Review, 5,** 607-612.

Franke, R.H., & Kaul, J.D. (1978). The Hawthorne experiments: First statistical interpretation. **American Sociological Review, 43,** 623-643.

Freidman, M. (1969). **Pathogenesis of coronary artery disease.** New York: McGraw-Hill.

Glass, D.C. (1977). **Behavior patter, stress and coronary disease.** Hillsdale, NJ: Erlbaum.

Glennon, J.R., Owens, W.A., Smith, W.J., & Albright, L.E. (1960). New dimension in measuring moral. **Harvard Business Review, 38,** 106-107.

Goldband, S. (1980). Stimulus specificity of physiological response to stress and the Type A coronary-prone behavior pattern. **Journal of Personality and Social Psychology, 39,** 670-679.

Green, L.J. (1978). In defense of measurement. **American Psychologist, 33,** 664-679.

Guadagnoli, E., & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. **Psychological Bulletin, 103,** 265-275.

Guion, R.M. (1965). **Personnel testing**. New York: McGraw-Hill.

Guion, R.M. (1981). Problems of measurement in organizational psychology. In J.A. Sgro (Ed.), **Virginia Tech symposium on applied behavioral sciences.** Vol.1. Lexington, Mass.: Lexington. Pp. 233-255.

Gulliksen, H. (1950). **Theory of mental tests**. New York: Wiley.

Gupta, N., & Beehr, T.A. (1979). Job stress and employee behaviors. **Organizational behavior and Human Performance, 23**, 373-387.

Hackett, R.D., & Guion, R.M. (1985). A reevaluation of the absenteeism-job satisfaction relationship. **Organizational behavior and Human Performance, 35**, 340-381.

Harman, H.H. (1976). **Modern factor analysis**. 3rd ed. rev. Chicago: The University of Chicago Press.

Herzberg, F., Mausner, B., Peterson, R.O., & Capwell, D.F. (1957). **Job attitudes: Review of research and opinion**. Pittsburgh, PA: Psychological Service of Pittsburgh.

Hill, T.D. (1985a). **Item format and structure of the Basic Personality Inventory**. Unpublished Master of Arts Thesis, Department of Psychology, University of Western Ontario.

Hill, T.D. (1985b). **Job satisfaction since 1977**. Department of Psychology Research Bulletin # 635, University of Western Ontario, London, Ontario, Canada.

Hill, T.D. (1985c). **The Satisfaction Research Questionnaire (SRQ)**. Unpublished test.

Hill, T.D. (1988). **Is it finally time to recognize F.W. Taylor?** Department of Psychology Research Bulletin # 668, University of Western Ontario, London, Ontario, Canada.

Hoppock, R. (1935). **Job satisfaction**. New York: Harper.

Hunter, J.E., Schmidt, F.L., & Hunter, R. (1979). Differential validity of employment tests by race: a comprehensive review and analysis. **Psychological Bulletin, 86**, 721-735.

Hunter, J.E., Schmidt, F.L., & Jackson, G.B. (1982). **Meta-Analysis: Cumulating research findings across studies**. Beverly Hills, CA: Sage.

Iaffaldano, M.T., & Muchinksy, P.M. (1985). Job satisfaction and job performance: A meta-analysis. **Psychological Bulletin, 97**, 251-273.

Innes, E., Perry, R.L., & Lyon, J. (1986). **The Financial Post's 100 best companies to work for**. Toronto: Collins.

Ivancevich, J.M., Szilagyi, A.D., & Wallace, Jr., M.J. (1977). **Organizational behavior and performance**. Santa Monica CA: Goodyear.

Jackson, D.N. (1960). Stylistic response determinants in the California Psychological Inventory. **Educational and Psychological Measurement, 10**, 339-346.

Jackson, D.N. (1967). **Personality Research Form Manual**. Port Huron, MI: Research Psychologists Press.

Jackson, D.N. (1971). The dynamics of structured personality tests: 1971. **Psychological Review, 78**, 229-248.

Jackson, D.N. (1984). **Personality Research Form - E Manual**. Port Huron, MI: Research Psychologists Press.

Jackson, D.N., Holden, R.R., Locklin, R.H., & Marks, E. (1984). Taxonomy of vocational interests of academic major areas. **Journal of Educational Measurement, 21**, 261-275.

Jackson, D.N., & Mavrogiannis, A. (1987a). **The Survey of Work Styles**. Research Psychologists Press: London, Ontario, Canada.

Jackson, D.N., & Mavrogiannis, A. (1987b). **An Introduction to the Survey of Work Styles. Research Bulletin # 663, Department of Psychology, University of Western Ontario, Canada.**

Jackson, D.N., & Messick, S.J. (1961). Acquiescence and desirability as response determinants on the MMPI. **Educational and Psychological Measurement, 21, 771-790.**

Jackson, D.N., & Messick, S.J. (1962a). Response styles and the assessment of psychopathology. In S. Messick & J. Ross (Eds.), **Measurement in personality and cognition.** New York: Wiley. Pp. 129-155.

Jackson, D.N., & Messick, S.J. (1962b). Response styles on the MMPI: Comparison of clinical and normal samples. **Journal of Abnormal and Social Psychology, 65, 285-299.**

Jackson, D.N., & Paunonen, S.V. (1985). Construct validity and the predictability of behavior. **Journal of Personality and Social Psychology, 49, 554-570.**

Jackson, D.N., & Williams, D.R. (1975). Occupational classification in terms of interest patterns. **Journal of Vocational Behavior, 6, 269-280.**

Jenkins, C., & Zyzanski, S.J., & Rosenman, R.H. (1979). **Jenkins Activity Survey: Manual.** New York: Psychological Corporation.

Kahn, R.L., & Quinn, R.P. (1970). Role Stress: A framework for analysis. In A. McLean (Ed.), **Occupational mental health.** Chicago: Rand McNally.

Kahn, R.L., Wolfe, D.M., Quinn, R.P., Snoek, J.D., & Rosenthal, R.A. (1964). **Organizational stree: Studies in role conflict and ambiguity.** New York: John Wiley and Sons.

Kornhauser, A.W., & Sharp, A. (1932). Employee attitudes: Suggestions from a study in a factory. **Personnel Journal, 10, 393-401.**

Kunin, T. (1955). The construction of a new type of attitude measure. **Personnel Psychology, 8,** 65-77.

Landy, F.J. (1978). An opponent process theory of job satisfaction. **Journal of Applied Psychology, 63,** 533-547.

Latham, G.P., & Pursell, E.D. (1975). Measuring absenteeism from the opposite side of the coin. **Journal of Applied Psychology, 60,** 369-371.

Lawler, E.E. (1971). **Pay and organizational effectiveness: A psychological view.** New York: McGraw-Hill.

Lawler, E.E. (1973). **Motivation in work organizations.** Monterey, CA: Brooks/Cole.

Lent, R.H., Aarbach, H.H., & Levin, L.S. (1971). Predictors, criteria, and significant results. **Personnel Psychology, 24,** 519-533.

Likert, R.A. (1932). A technique for the measurement of attitudes. **Archives of Psychology, No.** 140.

Locke, E.A. (1969). What is job satisfaction? **Organizational Behavior and Human Performance, 4,** 309-336.

Locke, E.A. (1976). The nature and causes of job satisfaction. In M.D. Dunnette (Ed.), **Handbook of industrial and organizational psychology.** Chicago: Rand McNally.

Locke, E.A. (1982). The ideas of Frederick W. Taylor: An evaluation. **Academy of Management Review, 7,** 7-24.

Locke, E.A., Latham, G.P., & Erez, M. (1988). The determinants of goal commitment. **Academy of Management Review, 13,** 23-39.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. **Psychological Reports, 3,** 635-694.

Loher, B.T., Noe, R.A., Moeller, N.L., & Fitzgerald, M.P. (1985). A meta-analysis of the relation of job characteristics to job satisfaction. Journal of Applied Psychology, 70, 280-289.

Lord, R.M. (1963). Elementary models for measuring change. In C.W. Harris (Ed.), Problems in measuring change. Madison: University of Wisconsin Press.

Maslow, A.H. (1943). A theory of human motivation. Psychological Review, 50, 370-396.

Maslow, A.H. (1954). Motivation and personality. New York: Harper.

Mavrogiannis, A., Jackson, D.N., & Howard, J.H. (1987). Validation of the Survey of Work Styles: A profile measure of the Type A Behavior pattern. Research Bulletin # 661, Department of Psychology, University of Western Ontario, Canada.

Mayo, E. (1960). The human problem of industrial civilization. New York: Viking Press.

McMillan, C.J. (1984). The Japanese industrial system. Berlin: de Gruyter.

McNemar, Q. (1969). Psychological statistics (4th ed.). New York: Wiley.

Medalie, J.H., Snyder, M., Groen, J.J., Neufeld, H., Goldbourt, U., & Riss, E. (1973). Angina pectoris among 10,000 men: Five year incidence and univariate analysis. American Journal of Medicine, 55, 583-594.

Messick, S.J., & Jackson, D.N. (1961). Acquiescence and the factorial interpretation of the MMPI. Psychological Bulletin, 58, 299-304.

Meyer, H.H. (1980). Self-appraisal of job performance. Personnel Psychology, 33, 291-296.

Mikalachki, A., & Gandz, J. (1979). Measuring absenteeism. Relations Industrielles, 34.

Mikes, P. M., & Hulin, C.L. (1968). Use of importance as a weighting component of job satisfaction. **Journal of Applied Psychology, 52,** 394-398.

Motowidlo, S.J., Packard, J.S., & Manning, M.R. (1986). Occupational stress: Its causes and consequences for job performance. **Journal of Applied Psychology, 71,** 618-629.

Muchinsky, P.M. (1977). Employee absenteeism: A review of the literature. **Journal of Vocational Behavior, 10,** 316-340.

Muchinsky, P.M. (1987). **Psychology applied to work: An introduction to industrial and organizational psychology.** 2nd ed. Chicago: Dorsey.

Olson, D.A., & Tetrick, L.E. (April, 1987). **A multivariate analysis of a stress-support-satisfaction model.** Paper presented at the Second Annual Conference of the Society for Industrial and Organizational Psychology, Atlanta.

Organ, D.W. (1977). A reappraisal of the satisfaction causes performance relationship. **Academy of Management Review, 2,** 46-53.

Orpen, C. (1974). Social desirability as a moderator of the relationship between job satisfaction and personal adjustment. **Personnel Psychology, 27,** 103-108.

Palmore, E.B. (1969a). Physical, mental, and social factors in predicting longevity. **Gerontologist, 9,** 103-108.

Palmore, E.B. (1969b). Predicting longevity: A follow-up controlling for age. **Gerontologist, 9,** 247-250.

Palmore, E.B., & Jeffers, F. (Eds.). (1971). **Prediction of life span. Recent findings.** Boston, MA: Health Lexington.

Parkes, K.R. (1987). Relative weight, smoking, and mental health as predictors of sickness and absence from work. **Journal of Applied Psychology, 72,** 275-286.

Pedhazur, E.J. (1982). **Multiple regression in behavioral research.** (2nd. ed.) New York: Holt, Rinehart, and Winston.

Petty, M.M., McGee, G.W., & Cavender, J.W. (1984). A meta-analysis of the relationships between individual job satisfaction and individual performance. **Academy of Management Review, 9,** 712-721.

Podsakoff, P.M., & Williams, L. (1986). The relationship between job performance and job satisfaction. In E.A. Locke (Ed.) **Generalizing from laboratory to field settings.** Lexington, MA: Lexington Books. Pp. 207-253.

Porter, L.W. (1961). A study of perceived need satisfaction in bottom and middle management jobs. **Journal of Applied Psychology, 45,** 1-10.

Porter, L.W. (1962). Job attitudes in management: I. Perceived deficiencies in need fulfillment as a function of job level. **Journal of Applied Psychology, 46,** 375-384.

Porter, L.W., & Lawler, E.E. (1968). **Management attitudes and performance.** Homewood, IL: Dorsey Press.

Porter, L.W., & Steers, R.M. (1973). Organizational, work, and personal factors in employee turnover and absenteeism. **Psychological Bulletin, 80,** 151-176.

Quinn, R.D., & Staines, G.L. (1977). **The 1977 quality of employment survey.** Michigan: Survey Research Center.

Rapaport, E. (1980). Coronary artery disease. In the American Heart Association, **The American Heart Association Heartbook: A guide to prevention and treatment of cardiovascular diseases.** New York: E.P. Dutton. Pp. 176-186.

Rose, C.L., & Bell, B. (1971). **Predicting longevity. Methodology and critique.** Lexington, MA: Heath Lexington.

Ross, R., & Glomset, J.A. (1976). The pathogenesis of artherosclerosis. New **England Journal of Medicine, 295,** 369-377, 671-676.

Rushton, J.P., Brainerd, C.J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. **Psychological Bulletin, 94,** 18-38.

Salancik, G.R., & Pfeffer, J. (1977). An examination of need-satisfaction models of job attitudes. **Administrative Science Quarterly, 22,** 427-456.

Sales, S.M., & House, J.S. (1971). Job dissatisfaction as a possible risk factor in coronary heart disease. **Journal of Chronic Diseases, 23,** 861-873.

Schaffer, R.H. (1953). Job satisfaction as related to need satisfaction in work. **Psychological Monographs, 67.**

Schmidt, F.L., & Hunter, J.E. (1978). Moderator research and the law of small numbers. **Personnel Psychology, 31,** 215-232.

Silverman, I., & Shulman, A.D. (1970). A conceptual model of artifact in attitude change surveys. **Sociometry, 33,** 97-107.

Skinner, H.A. (1977). "The eyes that fix you": A model for classification research. **Canadian Psychological Review, 18,** 142-151.

Skinner, H.A., Jackson, D.N., & Hoffman, H. (1975). Alcoholic personality types: Identification and correlates. **Journal of Abnormal Psychology, 83,** 658-666.

Smith, P.C., Kendall, L.M., & Hulin, C.L. (1969). **The measurement of satisfaction in work and retirement.** Chicago: Rand McNally.

Solomon, R.L., & Corbit, J.D. (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. **Psychological Review, 81,** 119-145.

Tabachnik, B.G., & Fidell, L.S. (1983). Using multivariate statistics. New York: Harper and Row.

Taylor, F.W. (1911). The principles of scientific management. New York: Harper and Brothers Publishers.

Terkel, L. (1974). Working. New York: Pantheon.

Tetrick, L.E., & LaRocco, J.M. (1987). Understanding, prediction, and control as moderators of the relationship between perceived stress, satisfaction, and psychological well-being. Journal of Applied Psychology, 72, 538-543.

Thurstone, L.L. (1947). Multiple Factor Analysis. Chicago: University of Chicago Press.

Tryon, R.C. (1973). Basic unpredictability of individual responses to discrete stimuli presentation. Multivariate Behavioral Research, 8, 275-279.

Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 321-327.

Vroom, V.F. (1964). Work and motivation. New York: Wiley.

Wall, T.D. (1972). Overall job satisfaction in relation to "social desirability", age, length of employment, and social class. British Journal of Clinical Psychology, 11, 79-81.

Wall, T.D., & Payne, R. (1973). Are deficiency scores deficient? Journal of Applied Psychology, 58, 322-326.

Werts, C.E., & Linn, R.L. (1970). A general linear model for studying growth. Psychological Bulletin, 73, 17-22.

Wiggins, J.S. (1962). Strategic, method, and stylistic variance in the MMPI. Psychological Bulletin, 59, 224-242.

Windle, C. (1955). Further studies of test-retest effect on personality questionnaires. Educational and Psychological Measurement, 15, 246-253.

Wolfle, D. (1940). **Factor analysis to 1940.** Chicago, IL: University of Chicago Press.

Wong, M.L. (1973). **Managerial ideologies in industrial society: The construction of the corporate world view.** Ph.D. Dissertation, University of Michigan.

Zenisek, T.J., & Rowney, J.I.A. (1980). A note on job satisfaction measurement. **Canadian Journal of Behavioural Science, 12,** 393-396.

APPENDIX A

Sample Satisfaction Research Questionnaire Items

INSTRUCTIONS

In this booklet you will find a series of general statements that people might use to describe their feelings about various aspects of their work. Read each statement carefully and decide whether or not you would use such a statement about your work. Then indicate your answer on the answer sheet provided. If you agree with a statement, or feel that you would say such a thing about your job, then blacken the T that corresponds to that statment. If you disagree with a statement, or feel that you would not say such a thing about your job, then blacken the F that corresponds to that statement. Answer every statement either true or false, even if you are not completely sure of your answer.

You must remember that some statements are very general, and a given term may apply across several situations. For example, equipment may refer to office machines or to heavy industrial equipment. Remember that the terms in this questionnaire are general.

(c) 1985 Timothy D. Hill

Please turn to the next page and begin.

001. I rarely wan† to work on the harder problems at work.

002. Work gives me time to think about things.

003. I would not think of going out with my co-workers.

004. The money I make does not make me feel disappointed.

005. Maintenance matters are usually overlooked at work.

006. I enjoy working on tasks that require greater intensity.

007. I am burdened with too much work.

008. I am glad to have the co-workers that I do.

009. I do not make enough money.

010. My boss does not make ridiculous decisions.

011. My job requires little mental skill.

012. I feel pretty relaxed in my work area.

013. I do not like associating with my co-workers.

014. My job is an important part of my life.

015. There are too many rules and regulations at work.

APPENDIX B

Copy of Solicitation Material

The University of Western Ontario

Research Unit on Work and Productivity
Room 6436
Social Science Centre
London, Ontario, Canada
N6A 5C2

Mr. R. W. Crofoot
BP Canada Inc.
First Canadian Place 57th Floor
Toronto Ontario M5X 1G8

Friday 27, February, 1987

Dear Mr. Crofoot:

I am a graduate student doing research towards my Ph.D. under the supervision of Douglas N. Jackson Ph.D., Director, Research Unit on Work and Productivity, The University of Western Ontario. The goal of my research is to contribute to the knowledge of the relationships between job attitudes and stress in the workplace. To this end, the Research Unit has devised a research questionnaire assessing both job attitudes and patterns of stress-related behaviours. It is expected that this information could be used to improve the selection of people who are most likely to be satisfied in their work, thus increasing productivity and reducing the incidence of job-related stress.

I would greatly appreciate the opportunity to collect data from BP Canada Inc. employees as I very much wish to make this research reflect current Canadian work environments.

The purpose of this letter is to make two requests of your office. One, to grant permission for this research, which only involves completion of this take-home questionnaire and its return to me by stamped, self-addressed envelope. Secondly, for your office to distribute a small number of questionnaires randomly within your office, or various branch offices. The number of questionnaires would be 100 or 10% of your company's workforce, which ever is less. This second task could easily be completed by clerical staff.

I recognize that time in business is a valuable commodity, and have therefore kept the materials as brief as possible. I have enclosed for your inspection a copy of the letter of introduction and the research questionnaire that each employee would receive.

...2

At the completion of the study I would be pleased to send your firm a summary report that describes the average responses of your company's employees, comparing their data to the Canadian averages. This report should be useful for the detection of both a) patterns of stress and b) patterns of satisfaction in the workplace. Let me assure you that each individual's questionnaire responses will be held in complete confidence. Furthermore, data identifying your organization with survey results will not be released in any form, published, or otherwise.

I thank you in advance for your cooperation, and have enclosed a stamped, self-addressed envelope in which you can return the sheet that indicates your company's willingness to participate in this research.

Sincerely,

Timothy D. Hill B.A., M.A.

/enclosures