THE RELIABILITY AND VALIDITY OF IPSATIVE AND NORMATIVE FORMS OF THE HUTCHINS BEHAVIOR INVENTORY

by

Harold William Wheeler Jr.

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Educational Research and Evaluation

APPROVED:

Dennis E. Hinkle, Co-Chairman	David E. Hutchins, Co-Chairman
Lawrence H. Cross	James C. Impara

Michael J. Sporakowski

October, 1986

Blacksburg, Virginia

THE RELIABILITY AND VALIDITY OF IPSATIVE AND NORMATIVE FORMS OF THE HUTCHINS BEHAVIOR INVENTORY

by

Harold William Wheeler Jr.

Committee Co-Chairmen:

Dennis E. Hinkle and David E. Hutchins

Administrative and Educational Services

(ABSTRACT)

The current trend among theorists in counseling and psychotherapy is toward the development of metatheoretical models that can be used to organize, systematically and comprehensively, existing theories and techniques within the discipline. Some models also provide behavior adaptation guidelines for practitioners who wish to adapt their behavior to client behavior patterns.

Hutchins created the metatheoretical TFA System to accomplish the above goals. He also created the Hutchins Behavior Inventory (HBI) to complement the TFA System. The HBI purportedly measures the thinking, feeling, and acting dimensions of behavior upon which the TFA System is built; it thus enables a practitioner to assess the unique, situationally specific, TFA behavior pattern of a client.

At the time of this study, the only evidence concerning the measurement properties of the HBI was for a form that produces ipsative scores (the HBI-I). But ipsative scores possess inherent psychometric properties that cause problems when they are subjected to certain types of statistical analyses. Thus, in this study, a normative form of the HBI (the HBI-N) was designed. The measurement properties of the HBI-I and HBI-N were then investigated and compared. Reliability was investigated using test-retest and internal consistency procedures. Construct-related validity was investigated using four procedures: internal consistency analysis of HBI-N scores; factor analysis of the items comprising the scales of the HBI-N; an analysis of a multitrait-multimethod validity matrix containing scores from the HBI-I, HBI-N, Strong Campbell Interest Inventory (SCII), and Myers-Briggs Type Indicator (MBTI); and a factorial validity analysis of scores from the HBI-N, SCII, and MBTI.

Results indicated that the HBI-I possesses a high degree of reliability. Prior evidence of content-related validity suggested that the three constructs measured by the HBI are the thinking, feeling, and acting dimensions of behavior hypothesized by Hutchins. Some of the construct-related validity results obtained in this study supported this conclusion, while the main body of results

supported the more limited conclusion that the HBI scales measure different, yet to be more clearly identified, constructs. Based on the evidence in this study, the HBI-I seems appropriate for research and clinical use.

DEDICATION

This paper is dedicated to my wife Sharon and to our children, Trilby and William.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the following people without whose help this study would not have been completed.

Appreciation is extended gratefully to Dr. Dennis Hinkle and Dr. David Hutchins who patiently mentored me throughout the completion of every aspect of this study. Their encouragement, support, and understanding were a constant source of inspiration and motivation.

Appreciation is also thankfully extended to the other members of my committee - Dr. Lawrence Cross, Dr. James Impara, and Dr. Michael Sporakowski - who provided invaluable advice and suggestions concerning this study.

Appreciation is extended gratefully, as well, to Dr. Robert Frary whose advice concerning data collection and analysis was indispensable.

Finally, special thanks are extended to Mrs. Dawn Reynolds for her assistance with the word processing package used to prepare this paper.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER ONE

Introduction

The latest trend among theorists and practitioners of counseling and psychotherapy is toward the development and use of metatheoretical models (Ward, 1983). Metatheoretical models are paradigms that provide ways to systematically and comprehensively organize the diverse theories and techniques of counseling and psychotherapy. Organizing theories and techniques according to a particular metatheoretical model draws attention to their unique and common qualities. Practitioners can then select systematically, from among the theories and techniques, whichever ones they ultimately determine are potentially best suited to meeting the specific needs of a particular client in a given situation (Bruce, 1984). This approach, metatheorists believe, enhances the counseling process, making it more efficient and effective.

Some of the metatheoretical models also offer behavior adaptation guidelines. These guidelines can be followed by practitioners who wish to adjust their behavior to the unique behavior patterns of clients. The metatheorists who have proposed such guidelines believe that personal behavior adaptation can strongly influence the effectiveness of practitioners as they attempt to individualize their services to clients.

Hutchins (1979, 1982, 1984) created a metatheoretical model known as the TFA system and designed it to offer assistance to practitioners who wish to achieve the goals outlined above. The TFA system is meant to help practitioners systematically and comprehensively organize theories and techniques on the basis of how they relate to thinking (T), feeling (F), and acting (A) dimensions of behavior. Also, the system offers behavior adaptation guidelines to practitioners wishing to adjust their own behavior to the unique, situationally specific, TFA behavior patterns of their clients. But the TFA system does not stop with behavior adaptation guidelines. The guidelines have been supplemented with an assessment tool, the Hutchins Behavior Inventory (HBI), that purportedly measures the thinking, feeling, and acting dimensions of client behavior.

The HBI has been proposed to enable a practitioner to assess the actual, situationally specific, TFA behavior pattern of a client. This is very important for at least two reasons. First, knowing a client's behavior pattern allows the practitioner to systematically select the combination of theories and techniques (T, F, and/or A oriented) most likely to succeed with that particular client. Secondly, it provides the practitioner with an empirical basis for deciding how to adapt his or her

behavior to the unique, situationally specific, TFA behavior pattern of a client.

Statement of the Problem

A problem arises in that scant empirical evidence exists concerning the important measurement properties of the HBI. Reliability and content-related validity have only been assessed on a preliminary forced-choice form of the HBI which produces ipsative scores (Walker, 1984). Other types of validity for the HBI have not been investigated, and the relationship of the HBI to other assessment instruments is not known.

Rationale for the Study

A major problem has confronted practitioners of counseling and psychotherapy since the inception of the discipline. The problem has involved determining what theories and techniques work best for which individuals, when rendered by which practitioners, under what circumstances (L'Abate, 1981). In spite of repeated research efforts, this question has remained unanswered and the problem unresolved. But the results of such research efforts, although failing to provide clearcut answers, have led to some positive consequences. One has been the realization, among a majority of practitioners, that no one theory or technique offers the best approach to helping all persons in need of counseling and psychotherapeutic services

(Ivey & Simek-Downing, 1980). A second has been the realization that all practitioners do not achieve equal success when using the same theory or technique (Corey, 1986).

Over the past two decades these realizations have stimulated a growing movement among practitioners toward the adoption of a more eclectic approach in their use of theories and techniques (Garfield & Kurtz, 1976; Smith, 1982). From its inception, however, the trend toward eclecticism has been the object of frequent criticism. This criticism has been fostered by a tendency on the part of eclectic practitioners to combine incongruous theories and techniques (Brammer, 1969). Concern over this practice has become pronounced, and increasingly so within the past 10 years.

The growing concern over theories and techniques being combined into a mishmash has stimulated the emergence of another newer movement within counseling and psychotherapy. Members of this new trend are seeking ways to overcome the current prevalent practice of mixing incompatible theories and techniques. One way to do so, they propose, is to systematically and comprehensively organize the diverse theories and techniques of counseling and psychotherapy (Ward, 1983). Doing so, they claim, will allow for the unique and common qualities of the many theories and

techniques to become apparent. The counseling process can then be made more efficient and effective because practitioners will be able to select wisely the theories and techniques best suited to meeting the specific needs of a particular client in a given situation (Bruce, 1984).

A number of rubrics have been suggested for this latest trend, such as: creative synthesis (Brammer & Shostrom, 1982), metatheory (Ivey, 1980; Ivey & Authier, 1978; Ward, 1983), general theory (Ivey & Simek-Downing, 1980), meta-modeling (Ivey & Matthews, 1984), technical eclecticism (Lazarus, 1984), and systematic eclecticism (Prochaska & DiClemente, 1984; Smith, 1982). Among these labels, metatheory appears to be the term employed most frequently as a descriptor of this trend. Some theorists who have worked to develop metatheoretical models include Barclay (1983, 1984), Bruce (1984), Egan (1982), Frey and Raming (1979), Hutchins (1979, 1982, 1984), Ivey and Matthews (1984), L'Abate (1981), Prochaska and DiClemente (1984), Thorne (1967), and Ward (1983).

The movement among theorists toward the development of metatheoretical models of counseling and psychotherapy is well under way. Stage one involved the delineation of different models. However, for a given model to progress it must undergo a second stage of development. This second stage involves linking empirical indicants to the concepts

and constructs set forth in the model (Zeller & Carmines, 1980). To accomplish this linking, kinds of measurement appropriate for the purposes of the metatheory must be selected and applied to the metatheoretical model in order to build, test, and if necessary, rebuild the metatheory.

Recently, Hutchins (1984) indicated he had entered this next vital stage in the development of his metatheoretical TFA system. Efforts by Hutchins (1984) and Walker (1984) have led to the development of an instrument, the Hutchins Behavior Inventory (HBI), that is designed to measure the TFA behavior dimensions of his metatheoretical system. The measurement of behavior produced by the HBI is situationally specific. Evidence reported thus far indicates that the form of the HBI investigated by Walker (1984) possesses content-related validity as well as a high degree of reliability.

This initial form of the HBI incorporates a forced-choice item format that produces scores of an ipsative nature. These ipsative scores correspond to one of the three types of measurement - ipsative, normative, and interactive (Cattell, 1944). A problem arises in that ipsative scores are known to have some inherent psychometric properties that cause problems when the scores are subjected to certain kinds of statistical analyses. Guilford (1952, 1954) cautioned that while ipsative scores could be used to

correlate persons, they should not be used to correlate variables. As a result of his warning, many studies were undertaken to investigate the properties of ipsative scores produced by forced-choice instruments. Certain of these studies sought to compare the properties of ipsative scores, produced by a forced-choice form of an instrument, with the properties of normative scores produced by a specially created single stimulus form of the same instrument. Hicks (1970) reviewed many of the conflicting findings produced by such studies and recommended that the development of instruments producing normative scores be substituted, whenever possible, for instruments designed to produce ipsative scores. Anastasi (1976), in concluding her review of the ipsative/normative issue, stated that "the controversy regarding the use of ipsative scores is far from resolved" (p. 530).

Objectives of the Study

Presently only one study has investigated the reliability and content-related validity of the HBI. In that study Walker (1984) examined the measurement properties of the ipsative form of the HBI (referred to in this study as the HBI-I). Since the completion of the Walker study, a single stimulus form of the HBI (referred to in this study as the HBI-N), which produces normative scores, has been devised from the forced-choice form investigated by Walker.

No studies have been conducted to investigate the reliability and validity of the HBI-N.

Therefore, this study had two objectives:

1. Investigate and compare the reliability of scores produced by the forced-choice ipsative (HBI-I) and single stimulus normative (HBI-N) forms of the HBI.
2. Investigate and compare the validity of scores produced by the forced-choice ipsative (HBI-I) and single stimulus normative (HBI-N) forms of the HBI.

CHAPTER TWO

Review of Related Literature

The review of related literature is divided into four sections. Section one contains an overview of the history of eclecticism and metatheory. Some of the better known metatheoretical models are presented in section two. Hutchins' metatheoretical model is outlined in section three. Then, in the fourth section, the role of measurement in developing and testing Hutchins' metatheory via the HBI is explained.

A Brief History of Eclecticism and Metatheory

The roots of counseling and psychotherapy extend far beyond the days of Freud and psychoanalysis. Indeed, one effort to identify the important historical figures who have contributed to the development of counseling and psychotherapy began in 400 B.C. with Hippocrates (Kottler & Brown, 1985). However, it was not until after the conception of Freud's school of psychoanalysis that a proliferation of schools of counseling and psychotherapy began that has continued directly into modern times (Patterson, 1986). Egan (1986) commented upon this staggering increase in the number of helping models and approaches, suggesting that anyone willing to conduct a review of the many sources that keep abreast of the new approaches to psychotherapy will "soon discover a

bewildering number of schools, systems, approaches and techniques, all of which are proposed with equal seriousness and all of which claim success" (p. 7). A 1975 survey of the many therapeutic approaches practiced in the United States was carried out by the National Institute of Mental Health. The survey identified over 130 different approaches in existence at that time (Brammer & Shostrom, 1982). More recently, Prochaska and DiClemente (1984) cited a Time magazine article which reported that there are now 200 different therapies.

Eclecticism

It is not surprising to discover that counselors and therapists have long been perplexed over the apparent contradictions and distinctions among the approaches to counseling and psychotherapy. Garfield (1980) asked the one question that has been posed by many modern counselors and therapists when he queried, "how does one choose from this vast diversity of offerings? How does one know which form of psychotherapy is best?" (p. 31). One popular answer had been to join the emerging eclectic movement within counseling and psychotherapy. But Eclectic theory developed along the same lines as personality and learning theory, where bits and pieces of many different theories and methods were drawn together into a hodgepodge full of "contradictory assumptions and incompatible techniques"

(Brammer, 1969, p. 192). Thus the Eclectic movement quickly developed a negative image.

Because, in the initial stages of eclecticism, adherents practiced an unguided use of theories and techniques, they were often no more effective than the practitioners of the single theory approach (Smith, 1982). Nevertheless, the movement to abandon the single theory approach in favor of some form of eclecticism continued to grow larger and stronger. There have been at least two major issues that have fueled this transition toward eclecticism.

The first point of debate has been over whether or not there is a single best theory of counseling and psychotherapy. Two decades ago Frederick Thorne, a noted clinical psychologist, stated that there are "no schools of psychotherapy that are in themselves alone complete systems providing complete answers for all clinical problems. There is no universal panacea, curing all possible problems" (Thorne, 1967, p. 354). More recently, Downing (1975) noted that "all the theories of counseling and psychotherapy have something to offer the counselor" (p. 3). Smith and Glass (1977), in their study entitled "Meta-Analysis of Psychotherapy Outcome Studies," found evidence that has been regarded by some as a source of support for this view. They reported that

> despite volumes devoted to the theoretical differences among schools of psychotherapy, the results of research demonstrate negligible differences in the effects produced by different therapy types. Unconditional judgments of superiority of one type or another of pscyhotherapy, and all that these claims imply about treatment and training policy, are unjustified. (p. 760)

While some researchers strongly disagreed with this conclusion, the Smith and Glass findings were subsequently corroborated by Landman and Dawes (1982).

Ivey and Simek-Downing (1982) also discarded the idea of a single best theory. They stated that the "evidence is mounting that the professional of the future will require more than a single set of methodological and theoretical answers to meet the needs of an increasingly diverse clientele" (p. 1). Hutchins (1984) stated that "no single theory to date seems adequate for all people" (p. 572). Moreover, Corey (1986) suggested that "valuable dimensions of human behavior can be overlooked if the counselor is restricted to a single theory (p. 2). Finally, Egan (1986) summarized the positive attitude many theorists now hold toward the use of different theories when he said of his own problem-solving model that "it is open to being

corroborated, complemented, and challenged by any other approach, model or school of helping . . . . The needs of clients--not the egos of model builders--must remain central to the helping process" (p. 10).

The second issue that has helped to bring about the transition to eclecticism has been over practitioner flexibility of response. This second issue has received a great deal of attention. Egan (1975) and Cormier and Cormier (1979) stressed the view that the most effective counselors are the ones who can incorporate the widest variety of techniques and can manifest the greatest array of responses. Furthermore, Brammer and Shostrom (1982) cited studies which found that experienced counselors meet "changing therapeutic demands with flexibility of response" (p. 38).

Hutchins (1984) cited several studies of counselor effectiveness conducted by Fiedler (1950a,b), and Auerbach & Johnson (1977), in which the researchers concluded that higher quality counseling relationships were achieved by experienced counselors. Hutchins claimed that "evidence suggests that these experienced counselors are more likely to adapt their personal style of relating to clients than are their less experienced counterparts" (p. 572). Hutchins speculated that counselor adaptation tends to be intuitive and that it increases as a result of experience. Brammer

and Shostrom (1982), however, challenged the idea that counselor adaptation and experience are causally related. They cited a number of replication studies that take issue with Fiedler's claims regarding the affect of counselor experience upon a counselor's ability to adapt. The more rigorous replication studies did not establish a causal link between counselor experience and ability to adapt.

Thus, many experts agree that flexibility of response is an important characteristic for practitioners to possess. However, there is a great deal of disagreement over the nature of the relationship between experience and practitioner behavior adaptation or flexibility of response. Some, like Hutchins, speculate that the relationship is a causal one. Others, like Brammer and Shostrom, adopt a more cautious position; while they agree that a relationship may exist, they hold that no causality has been established.

Metatheory

Although eclecticism within counseling and psychotherapy continues to grow, a great deal of caution has been expressed about using theories and techniques in an unguided fashion. Thus, there is a need for models that can serve to guide practitioners in their use of theories and techniques. Prochaska and DiClemente (1984) addressed the urgency of this need for new models to guide the use of theories and techniques when they warned that "unless

increasing differentiation is matched by more effective forms of integration, crisis rather than creativity will be the result" (p. 1). Numerous theorists in the field of counseling and psychotherapy have recently been investing their time, energy, and creativity into the task of developing models that systematically and comprehensively organize the diverse theories and techniques of counseling and psychotherapy. Ward (1983) spoke of this new movement when he noted that theorists are no longer emphasizing the development of new theories; rather, they are working to develop the kind of models (alluded to above) that can serve "to guide the conceptualization and application of theories and techniques" (p. 154). Smith (1982) also discussed this movement. In his review of current literature in counseling and psychotherapy he reported the existence of "a trend in the direction of creative synthesis, masterful integration, and systematic eclecticism" (p. 802).

Patterson (1973) foresaw the emergence of this new trend toward a more systematic eclecticism. He predicted that any loosely organized eclectic position possessing underlying consistency among the phenomena incorporated into the position would eventually be developed, by way of research and experience, into a formal theory. Theorists involved in the new trend to develop models that are systematic and comprehensive have, in essence, been

searching for just such underlying consistencies among the theories and techniques of counseling and psychotherapy.

While there is a natural inclination to try and use the name eclecticism to label this new movement (i.e., Palmer, 1980), Smith (1982) has noted that "there seems to be a dissatisfaction with the term eclecticism per se" (p. 808). Thus, it is not surprising to find that new rubrics are being applied to this movement. Some currently being used are: creative synthesis (Brammer & Shostrom, 1982), general theory (Ivey & Simek-Downing, 1980), meta-modeling (Ivey & Matthews, 1984), metatheory (Ivey, 1980; Ivey & Authier, 1978; Ward, 1983), systematic eclecticism (Egan, 1982; Prochaska & DiClemente, 1984; Smith, 1982), and technical eclecticism (Lazarus, 1984). Of these rubrics, metatheory appears to be used most frequently.

Metatheoretical Models

The list of metatheoretical models being developed by contemporary theorists is growing rapidily. Each theorist proposes a different metatheoretical framework for organizing theories and techniques. A discussion of some of the better known metatheories follows.

Microcounseling

Ivey and Authier (1978) proposed a metatheoretical model called microcounseling. Microcounseling provides a conceptual framework that allows for a systematic

examination of counseling theories in order to identify commonalities as well as differences among them. It involves the use of microskills which "are communication skill units of the interview that will help you develop a more intentional and rounded ability to interact with a client" (p. 4). A major goal of this approach is to help the counselor or therapist identify alternatives that can be used with different types of clients.

The Problem-Solving Model

Egan (1982) developed a problem-solving metatheoretical model. He stated that the approach offers the type of conceptual framework necessary for integrative or systematic eclecticism. He claimed that there is increasing awareness that problem-solving approaches have real value. The model focuses upon identifying and organizing the skills clients need "both to collaborate with helpers in the helping relationship and to face problems in living their day-to-day lives" (Egan, 1982 p. 9). Egan noted that these skills may be drawn from any of the current theories of counseling and psychotherapy.

The E-R-A Model

L'Abate (1981) developed the E-R-A metatheoretical model. In this model an emphsis is placed upon the development of competencies within relationships. Three basic aspects - effects (E), reasons (R), and actions (A) -

make up the model. L'Abate explained that in counseling "E is the structural input, R is the processing throughput, and A is the outcome output" (p. 263). It is hypothesized in this theory that there is an awareness factor which provides feedback for each of the E-R-A components and for the context of each relationship.

Multimodal Therapy

Lazarus (1984) developed a metatheoretical approach known as multimodal therapy. His model involves unlimited eclecticism with regard to techniques; but little or no eclecticism with regard to counseling theories. In this model an assumption is made that "there is no unitary way to approach peoples' problems" (p. 496). Therapeutic flexibility is recommended as the most effective counseling posture. Clients are believed to respond most favorably to this type of relationship and interactive style.

Seven aspects or modalities of the human temperament and personality are identified in multimodal therapy. They are: behaviors, affective processes, sensations, images, cognitions, interpersonal relationships, and drugs/biological functions (BASIC I.D.).

In multimodal therapy the counselor starts by bridging to the client. Lazarus (1984) explained that "bridging refers to a procedure in which the therapist deliberately tunes into the client's preferred modality before branching

off into other dimensions that seem likely to be more productive" (p. 493). Once bridging is accomplished, tracking becomes important. Lazarus (1984) explained that "tracking refers to a careful examination of the `firing order' of the different modalities" p. 493). The counselor can identify the firing order of the client in a particular situation by using bridging and tracking techniques. The counselor can then intervene by helping the person learn to restructure their firing pattern so that it is more suitable to the situation in which they are experiencing problems. Reliable assessment of the BASIC I.D. is critically important in order to make full use of this metatheoretical approach.

The Dynamic Developmental Continuum of Counseling Goals

Bruce (1984) developed a metatheoretical model that centers around matching counseling process models, goals, and intervention techniques to a client's everdeveloping need system. Bruce claimed that the model, which he named the dynamic developmental continuum of counseling goals, synthesizes counseling approaches in a way that is theoretically and philosophically consistent. The model offers a differential approach to counseling practice and relies on an "assessment of client readiness for a particular modality" (p. 259). While Bruce emphasized the need for assessment, he also recognized "the limitations of

empirical science to explain the full range of human behavior or to validate the full range of counseling practice" (p. 259).

The Transtheoretical Approach

Prochaska and DiClemente (1984) developed a metatheoretical model known as the transtheoretical approach. They created the model to "serve as a systematic guide for eclectic therapists seeking to help clients with some of the most common yet complex clinical problems" (p. 2).

The model uses the concept of change as a foundation for integrating theories and techniques. It is proposed in the model that client problems exist at any of five increasingly complex levels. To solve a problem, at any level, the client must move through five stages of change. To move progressively through the stages of change requires the application of 10 different processes of change, each of which is incorporated during a specific stage of change, or during a transition from one stage of change to another. The model attempts to account for how people change by overcoming problems with therapeutic assistance, as well as how people change by resolving problems on their own in their normal environment.

A major strength of the transtheoretical approach is that instruments (in the form of questionnaires) have been

developed to assess the core concepts of the model - levels of problems, stages of change, and processes of change. Furthermore, research evidence supporting the model has been cited by Prochaska and DiClemente (1984).

The Four I Model

Barclay (1984) has been developing a metatheoretical counseling framework for over 25 years. He labeled his approach the Four I model and described it as an assessment paradigm (Barclay, 1983). This model begins with an evaluation that includes the identification of empirical, behavioral and psychometric characteristics of a person. These characteristics are then integrated into a multi-method synthesis that matches specific skill deficits identified by the assessment to alternative inferred treatments. Finally, the treatments are implemented and evaluated. The theory was developed for use in elementary schools that adopt a preventive paradigm. Barclay (1984) equated the approach, which incorporates a computerized multi-trait and multi-method assessment technology called the Barclay Classroom Assessment System (BCAS), to the aptitude-treatment interaction research approach discussed by Cronbach and Snow (1977). The researchers who have adopted the aptitude-treatment interaction approach have been criticized by Cronbach and Snow for being less productive than desired. However, Barclay suggested his

approach, with its computerized assessment technology, may be able to overcome the two major problems Cronbach and Snow cited. These two problems were: a failure to use a broad enough set of personological data when defining aptitude; and a failure in developing long-lasting and well defined treatments. The system is designed as a team approach and utilizes teachers, counselors, school psychologists, and other school personnel.

Hutchins' TFA System

Hutchins (1979, 1982, 1984) developed a unique metatheoretical model called the TFA system. He designed the model to help answer the question that has been posed frequently by leading authorities within the field of counseling and psychotherapy - What counseling and psychotherapeutic theories and techniques work best, for which clients, when used by which practitioners, under what circumstances (Garfield, 1980; Hutchins, 1984; Ivey & Authier, 1978; Ivey & Simek-Downing, 1980; L'Abate, 1981; Lazarus, 1984)?

The goal of the TFA system is to help clients resolve their problems by helping them change their behavior. Client behavior, according to Hutchins (1979), includes three related dimensions - thinking, feeling, and acting. Therefore, the practitioner's task is to carefully select intervention strategies that "are specifically designed to

affect the client's thoughts, feelings, or actions" (p. 529). The TFA system possesses two major features that help make it possible for a practitioner to achieve the goal of client problem resolution through behavior change.

The first feature of the TFA system is its ability to serve as a guide for systematically and comprehensively organizing the various theories and techniques of counseling and psychotherapy. The theories and techniques are organized according to whichever dimension of behavior - thinking, feeling, or acting - they primarily address. Ward (1983) stated that this strategy for organizing theories and techniques is basically sound. Ellis (1982) also agreed with this strategy but stressed that when the major schools of counseling and psychotherapy are placed into such categories, the principal schools actually "significantly overlap in their goals, processes, and intervention methods" (p. 7). Ellis emphasized that this overlap among the schools is of greater importance today than it was in the past.

The second feature of Hutchins' TFA system is the set of guidelines it offers for practitioners to use when they desire to adapt their behavior to the unique, situationally specific behavior patterns of their clients. Much has already been said about the issue of practitioner behavior adaptation. Therefore, it merely remains to be noted that

these guidelines are a particularly strong feature of the TFA system.

Hutchins (1984) summed up the two major features of his TFA system when he stated,

> In my opinion evidence indicates that the "art" of paychotherapy and counseling comes mainly from experience, by means of which counselors intuitively synthesize elements of theory and techniques and adapt their personal relationship to the uniqueness of each client. (p. 575)

Like Barclay (1984), Bruce (1984), Prochaska and DiClemente (1984), and Lazarus (1984), Hutchins believes that assessment is one key to the success of a systematic approach to counseling and psychotherapy. In addition to encouraging practitioners to assess client TFA behavior patterns using observational techniques, Hutchins (1984) also reported that "extensive testing is underway on an instrument to classify individual TFA orientations" (p. 575). This heavy emphasis that is being placed upon assessment by various metatheorists draws attention to the role measurement plays in the development and testing of metatheories.

The Role of Measurement in Developing and Testing Metatheories

Within the behavioral sciences there has been widespread recognition and appreciation of the many ways in which measurement is put to use throughout the process of theory development and testing.

Translating Concepts and Constructs into Variables and Indicants

Blalock (1970) noted that the introduction of measurement into the theory development process enables us "to clarify our theoretical thinking and to suggest new variables that should be considered" (p. 88-9). Maranell (1974) pointed out that one important application of measurement occurs when scales are developed; such scales are the result of "the empirical process of translating theoretical ideas and concepts into variables" (p. xi). Zeller and Carmines (1980) discussed the complexity of the process of developing measuring devices such as scales. They noted that "it is the very vagueness, complexity, and suggestiveness of concepts that allow them to be empirically referenced with varying degrees of success at different times and places" (p. 3). They stressed that empirical indicants representing concepts need to be as specific and exact in design as theoretical formulations and research settings will permit. Yet they cautioned that indicants can

never be expected to represent completely the meaning of theoretical concepts because the small set of empirical indicants chosen is drawn from a domain containing an almost infinite number of indicants representing the concept.

Within the metatheoretical movement there has been a growing effort on the part of theorists to translate their metatheoretical ideas and concepts into measurable variables via scales, inventories, tests, or other measuring devices. Ward (1983), who developed a metatheoretical approach in the area of group counseling, discussed the need for this emphasis. He stated that models which can more accurately and usefully assess the way clients function in the counseling/therapeutic process are needed in order to ensure that the metatheoretical models used to guide theory selection result in increased counseling effectiveness. Pedazur (1982) summed up the need for new and better assessment tools by stating, "the goal of bridging the gap between theory and observed behavior by constructing highly valid and reliable measures deserves greater attention, sophistication, and expertise" (p. 232).

Eysenck, Wakefield, and Friedman (1983) cautioned that the psychological assessment systems needed for the future must move toward dimensional systems of measurement and away from the categorical ones like the DSM-III system (1980). These new dimensional assessment systems will have to be

based on a scale of measurement other than the nominal scale that only categorizes. Further, the systems will need to be based on some form of untransformed (interactive) or transformed (ipsative or normative) evaluation of raw measurement defined by Cattell (1944).

Testing the TFA System via the Development and Investigation of the Hutchins Behavior Inventory (HBI)

Hutchins (1979, 1982, 1984) attempted to classify counseling theories and techniques and then link corresponding theories and techniques using his TFA model. Further, he recommended the model be used to assess counselor and client behavior patterns. Finally, he proposed guiding deliberate counselor behavior adaptation to each client through the application of the TFA model. Hutchins (1984) stated that these three elements "combine to promote a good working relationship between counselor and client" (p. 575). While the TFA system can be used "to categorize and synthesize major patterns of behavior," it is not intended "to 'pigeon hole' people" (p. 573). Indeed, the Hutchins Behavior Inventory (HBI) was created and designed to produce dimensional measures of each of the three related behavior constructs - thinking, feeling, and acting - that comprise Hutchins' metatheoretical model.

Walker (1984) reported that the HBI possesses both content-related validity and a high degree of reliability.

However, the revised form of the HBI that Walker studied incorporated a forced-choice item format that produced scores of an ipsative nature. Guilford (1952) defined such ipsative scores by contrasting them with normative scores, which are better known, and better understood. According to Guilford, normative scores show the extent to which one particular trait is possessed by different people. Ipsative scores show the extent to which a number of traits are possessed by one particular individual. Thus, inter-individual differences in a trait can be compared using normative scores; intra-individual differences, on the other hand, can be compared using ipsative scores.

Guilford (1952, 1954) cautioned that ipsative scores for variables should not be subjected to statistical procedures that involve correlation techniques. Many studies were subsequently undertaken to investigate the actual properties and appropriate uses of ipsative scores produced by forced-choice measuring instruments (Broverman, 1961, 1962; Clemans, 1966; Davis & Chissom, 1981; Gleser, 1972; Radcliffe, 1963). Additional studies were carried out that compared the properties and performance of ipsative and normative scores produced, respectively, by forced-choice and single stimulus forms of the same instrument (Bartlett, Quay, & Wrightsman, 1960; Heilbrun, 1963; Karr, 1958; Knapp, 1964; Merritt & Marshall, 1984; Scott, 1968; Wright, 1961).

Furthermore, experts such as Radcliffe (1963), Clemans (1966), and Hicks (1970) examined extensively the properties of ipsative scores; they also wrestled with questions concerning the circumstances under which ipsative and normative scores are most useful. Hicks (1970) concluded, after reviewing the mathematical and empirical properties of ipsative and nonipsative measures, that

> although nonipsative (normative or forced-choice normative) measuring instruments can be highly effective in most assessment situations, purely ipsative instruments, on the other hand, possess such extensive psychometric limitations that use of such instruments is not recommended. (p. 167)

On the other hand, Clemans (1966) pointed out that there are times when the forced-choice technique that produces ipsative scores is the only way to get a valid measure of a variable. This is due to problems that can exist in creating instruments with independent items that may be subject to numerous kinds of response bias, especially if the items represent socially undesirable choices (Cronbach, 1946, 1950; Rorer, 1965; Travers, 1951). Clemans (1966), at the same time, cautioned that ipsative measurement be used only when normative measurement failed to produce equally valid measures of behavior.

In order to further investigate the TFA system it seemed useful to study the properties of the HBI scores which serve to measure the thinking, feeling, and acting constructs of the metatheory. The two most general properties of any instrument are validity and reliability (Standards for Educational and Psychological Testing, 1985). Reliability is primarily a performance issue focusing upon the extent to which an instrument produces consistent scores that are relatively free of random errors. Validity is mainly a theoretical issue focusing upon the extent to which an instrument produces scores that achieve the purpose for which they are intended. To accomplish this second goal the scores must be relatively free of non-random error (Carmines & Zeller, 1979). Impara and Stoker (1985) have noted that "although validity is the prime consideration, a good test must also be reliable. . . . A test with no reliability cannot be valid for any purpose" (p. 65).

There are numerous types of reliability coefficients. Parallel forms, test-retest, and internal consistency are the most familiar. Because each type accounts for different sources of random error, procedures followed in deriving any particular type of coefficient should be thoroughly described (Standards for Educational and Psychological Testing, 1985).

There are also many types of validity. Those commonly recognized are content-related validity, criterion-related validity, and construct-related validity (Standards for Educational and Psychological Testing, 1985). Different types of investigations are required to gather information aimed at establishing each of these. The most frequent investigations are: correlational, internal consistency, group differences, changes in performance, and investigations of the test taking process itself (Cronbach & Meehl, 1955; Helmstadter, 1964). Both the reliability and validity of the scores produced by the current forms of the HBI need to be investigated in order to determine the adequacy of the instrument.

Summary

Counseling and psychotherapy have moved beyond unidimensional approaches and mishmash eclectisim, arriving at the current attempts to develop systematic, comprehensive, metatheoretical approaches to helping people. While numerous metatheoretical models have been proposed, Hutchins' TFA system is one of the few that offers a measurement tool, the Hutchins Behavior Inventory (HBI), that can be used in developing and testing the metatheory. It seemed imperative that the properties of the scores produced by the HBI be investigated in order to determine the HBI's usefulness as a measure of the TFA constructs.

CHAPTER THREE

Research Design

The purpose of this study was to investigate the measurement properties of the HBI. There were two objectives that pertained to this purpose. The first objective was to investigate and compare the reliability of scores produced by the ipsative (HBI-I) and normative (HBI-N) forms of the HBI. The second objective was to investigate and compare the validity of scores produced by the two forms. To fulfill these objectives, a research design was developed and is presented in this chapter. Three aspects of the design are discussed: the instrumentation; subjects and procedures; and research questions and analyses appropriate for each question.

Instrumentation

Four different inventories were incorporated into this study. The inventories used were the HBI-I, the HBI-N, the Strong Campbell Interest Inventory (SCII), and the Myers-Briggs Type Indicator (MBTI).

The Hutchins Behavior Inventory (HBI)

The HBI-I. The HBI was designed by Hutchins to measure behavior. According to Hutchins (1986),

> Members of the helping professions use the term behavior somewhat differently from the way it is used in other contexts. As it is used during the

helping process, behavior includes a person's thoughts, feelings, and actions. (p. 2)

Having already built his TFA system around this definition of behavior, Hutchins designed the ipsative form of the HBI - the HBI-I - to produce dimensional measures of these thinking (T), feeling (F), and acting (A) aspects of behavior. Definitions of the T, F, and A dimensions of behavior proposed by Hutchins are summarized in Appendix A. Hutchins (1986) noted that practitioners need to be cognizant of how these three aspects of behavior interact. He proposed that altering the mixture of these three elements of behavior "can trigger a snowballing effect with tremendous positive or negative results in behavior" (p.2).

Walker's study of the revised HBI-I. In 1984 Walker identified the 15 words used to make up the items of the HBI-I (see Table 1). Five words represented thinking behavior, five represented feeling behavior, and five represented acting behavior. These T, F, and A words were paired with the words in all other groups to create the 75 pairs of forced-choice items found in the HBI-I. Next, she developed a method for scoring the inventory that involved both raw and intensity scores. The HBI-I was then administered to four groups of private liberal arts college students. The reliability of the instrument was investigated by analyzing the TFA frequency scores.

Table 1

The Fifteen Words Used in the Various Forms of the HBI to Create Items for the TFA Behavior Scales

Behavior Scale	Words
Thinking	
	Logical
	Contemplative
	Curious
	Rational
	Analytical
Feeling	
	Sensitive
	Compassionate
	Emotional
	Caring
	Concerned
Acting	
	Initiating
	Decisive
	Spontaneous
	Assertive
	Doing

Walker determined the internal consistency of the TFA frequency scores by computing a Cronbach coefficient alpha for each group's ipsative T, F, and A test scores. Alphas were also computed for each group's ipsative T, F, and A retest scores. The reported alpha coefficients ranged from .78 to .98, with only four of the 12 alphas below .90.

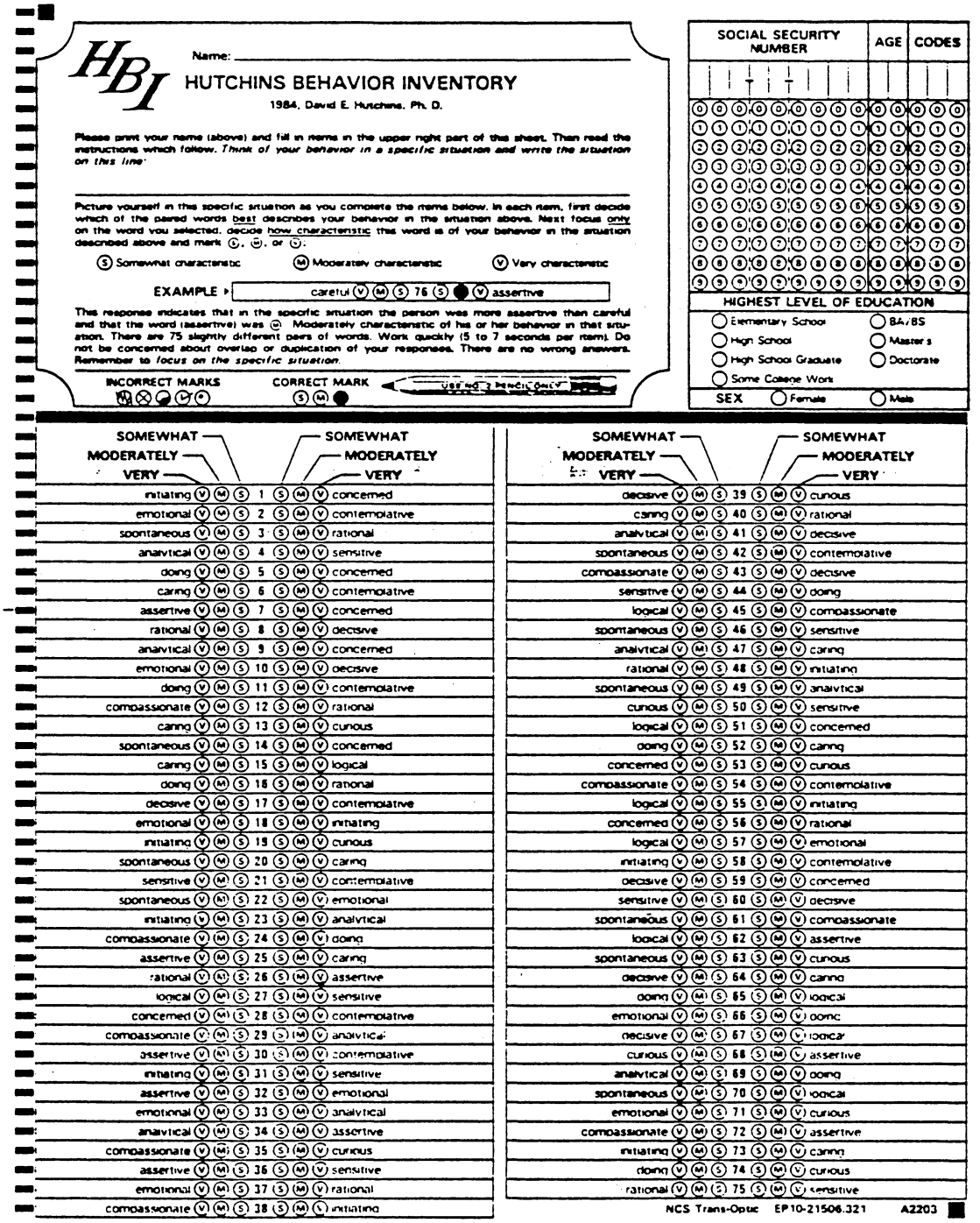
Walker also computed 15-minute and 16-day test-retest reliability coefficients for T, F, and A frequency scores. Two groups participated in the 15 minute test-retest procedure. The six Pearson product-moment correlation coefficients computed (two for each behavior component - T, F, and A) ranged from .84 to .93. Two other groups participated in the 16-day test-retest procedure. The six Pearson product-moment correlation coefficients computed (again, two for each behavior component - T, F, and A) ranged from .71 to .88.

To conclude this discussion of Walker's study of the HBI, three things should be noted by way of summary. First, the procedures followed to investigate Hutchins' original HBI were carefully designed and carried out in order to ensure that the new form would possess content-related validity. Secondly, computing the test-retest reliability coefficients for the ipsative frequency scores produced by the forced-choice form seemed reasonable; however, computing coefficient alphas on the ipsative test and retest frequency

scores was questionable. Finally, Walker's study did not investigate the criterion-related nor construct-related validity of the HBI.

Hutchins' adjustments on the HBI-I. Shortly after Walker completed her study of the HBI-I, Hutchins added demographic items to the instrument and additional instructions for completing the inventory. These instructions for the HBI-I can be seen in Figure 1. The instructions provided subjects with guidance in selecting a specific situation to focus on while responding to the 75 items. An example was included that demonstrated how subjects were to complete the items of the inventory along with a reminder for subjects to continue focusing on the specific situation while completing all 75 items. In addition, a set of intensities were provided for subjects to use when rating how characteristic a selected word was of their behavior in the specific situation. This last change was the most significant because subjects were now being asked to report how characteristic a word was of their behavior in the specific situation rather than how frequently the behavior occurred in the situation.

Scoring the HBI-I. Three methods were used in this study to score the HBI-I. Each method produced a different score. The first method of scoring was used to produce ipsative T, F, and A choice scores. These ipsative choice



HBI

Name: _____

HUTCHINS BEHAVIOR INVENTORY

1984, David E. Hutchins, Ph. D.

Please print your name (above) and fill in items in the upper right part of this sheet. Then read the instructions which follow. *Think of your behavior in a specific situation and write the situation on this line:*

_____

Picture yourself in this specific situation as you complete the items below. In each item, first decide which of the paired words best describes your behavior in the situation above. Next focus only on the word you selected, decide how characteristic this word is of your behavior in the situation described above and mark (S), (M), or (V).

(S) Somewhat characteristic (M) Moderately characteristic (V) Very characteristic

EXAMPLE ▸ careful (V) (M) (S) 76 (S) ● (V) assertive

This response indicates that in the specific situation the person was more assertive than careful and that the word (assertive) was (M) Moderately characteristic of his or her behavior in that situation. There are 75 slightly different pairs of words. Work quickly (5 to 7 seconds per item). Do not be concerned about overlap or duplication of your responses. There are no wrong answers. Remember to *focus on the specific situation.*

INCORRECT MARKS ⊗ ◒ ◔ ◉

CORRECT MARK (S) (M) ● USE NO. 2 PENCIL ONLY

SOCIAL SECURITY NUMBER	AGE	CODES

HIGHEST LEVEL OF EDUCATION

○ Elementary School ○ BA/BS
○ High School ○ Masters
○ High School Graduate ○ Doctorate
○ Some College Work

SEX ○ Female ○ Male

SOMEWHAT MODERATELY VERY		SOMEWHAT MODERATELY VERY
initiating (V) (M) (S)	1	(S) (M) (V) concerned
emotional (V) (M) (S)	2	(S) (M) (V) contemplative
spontaneous (V) (M) (S)	3	(S) (M) (V) rational
analytical (V) (M) (S)	4	(S) (M) (V) sensitive
doing (V) (M) (S)	5	(S) (M) (V) concerned
caring (V) (M) (S)	6	(S) (M) (V) contemplative
assertive (V) (M) (S)	7	(S) (M) (V) concerned
rational (V) (M) (S)	8	(S) (M) (V) decisive
analytical (V) (M) (S)	9	(S) (M) (V) concerned
emotional (V) (M) (S)	10	(S) (M) (V) decisive
doing (V) (M) (S)	11	(S) (M) (V) contemplative
compassionate (V) (M) (S)	12	(S) (M) (V) rational
caring (V) (M) (S)	13	(S) (M) (V) curious
spontaneous (V) (M) (S)	14	(S) (M) (V) concerned
caring (V) (M) (S)	15	(S) (M) (V) logical
doing (V) (M) (S)	16	(S) (M) (V) rational
decisive (V) (M) (S)	17	(S) (M) (V) contemplative
emotional (V) (M) (S)	18	(S) (M) (V) initiating
initiating (V) (M) (S)	19	(S) (M) (V) curious
spontaneous (V) (M) (S)	20	(S) (M) (V) caring
sensitive (V) (M) (S)	21	(S) (M) (V) contemplative
spontaneous (V) (M) (S)	22	(S) (M) (V) emotional
initiating (V) (M) (S)	23	(S) (M) (V) analytical
compassionate (V) (M) (S)	24	(S) (M) (V) doing
assertive (V) (M) (S)	25	(S) (M) (V) caring
rational (V) (M) (S)	26	(S) (M) (V) assertive
logical (V) (M) (S)	27	(S) (M) (V) sensitive
concerned (V) (M) (S)	28	(S) (M) (V) contemplative
compassionate (V) (M) (S)	29	(S) (M) (V) analytical
assertive (V) (M) (S)	30	(S) (M) (V) contemplative
initiating (V) (M) (S)	31	(S) (M) (V) sensitive
assertive (V) (M) (S)	32	(S) (M) (V) emotional
emotional (V) (M) (S)	33	(S) (M) (V) analytical
analytical (V) (M) (S)	34	(S) (M) (V) assertive
compassionate (V) (M) (S)	35	(S) (M) (V) curious
assertive (V) (M) (S)	36	(S) (M) (V) sensitive
emotional (V) (M) (S)	37	(S) (M) (V) rational
compassionate (V) (M) (S)	38	(S) (M) (V) initiating

SOMEWHAT MODERATELY VERY		SOMEWHAT MODERATELY VERY
decisive (V) (M) (S)	39	(S) (M) (V) curious
caring (V) (M) (S)	40	(S) (M) (V) rational
analytical (V) (M) (S)	41	(S) (M) (V) decisive
spontaneous (V) (M) (S)	42	(S) (M) (V) contemplative
compassionate (V) (M) (S)	43	(S) (M) (V) decisive
sensitive (V) (M) (S)	44	(S) (M) (V) doing
logical (V) (M) (S)	45	(S) (M) (V) compassionate
spontaneous (V) (M) (S)	46	(S) (M) (V) sensitive
analytical (V) (M) (S)	47	(S) (M) (V) caring
rational (V) (M) (S)	48	(S) (M) (V) initiating
spontaneous (V) (M) (S)	49	(S) (M) (V) analytical
curious (V) (M) (S)	50	(S) (M) (V) sensitive
logical (V) (M) (S)	51	(S) (M) (V) concerned
doing (V) (M) (S)	52	(S) (M) (V) caring
concerned (V) (M) (S)	53	(S) (M) (V) curious
compassionate (V) (M) (S)	54	(S) (M) (V) contemplative
logical (V) (M) (S)	55	(S) (M) (V) initiating
concerned (V) (M) (S)	56	(S) (M) (V) rational
logical (V) (M) (S)	57	(S) (M) (V) emotional
initiating (V) (M) (S)	58	(S) (M) (V) contemplative
decisive (V) (M) (S)	59	(S) (M) (V) concerned
sensitive (V) (M) (S)	60	(S) (M) (V) decisive
spontaneous (V) (M) (S)	61	(S) (M) (V) compassionate
logical (V) (M) (S)	62	(S) (M) (V) assertive
spontaneous (V) (M) (S)	63	(S) (M) (V) curious
decisive (V) (M) (S)	64	(S) (M) (V) caring
doing (V) (M) (S)	65	(S) (M) (V) logical
emotional (V) (M) (S)	66	(S) (M) (V) doing
decisive (V) (M) (S)	67	(S) (M) (V) logical
curious (V) (M) (S)	68	(S) (M) (V) assertive
analytical (V) (M) (S)	69	(S) (M) (V) doing
spontaneous (V) (M) (S)	70	(S) (M) (V) logical
emotional (V) (M) (S)	71	(S) (M) (V) curious
compassionate (V) (M) (S)	72	(S) (M) (V) assertive
initiating (V) (M) (S)	73	(S) (M) (V) caring
doing (V) (M) (S)	74	(S) (M) (V) curious
rational (V) (M) (S)	75	(S) (M) (V) sensitive

NCS Trans-Optic EP10-21506.321 A2203

Figure 1. The Current HBI-I

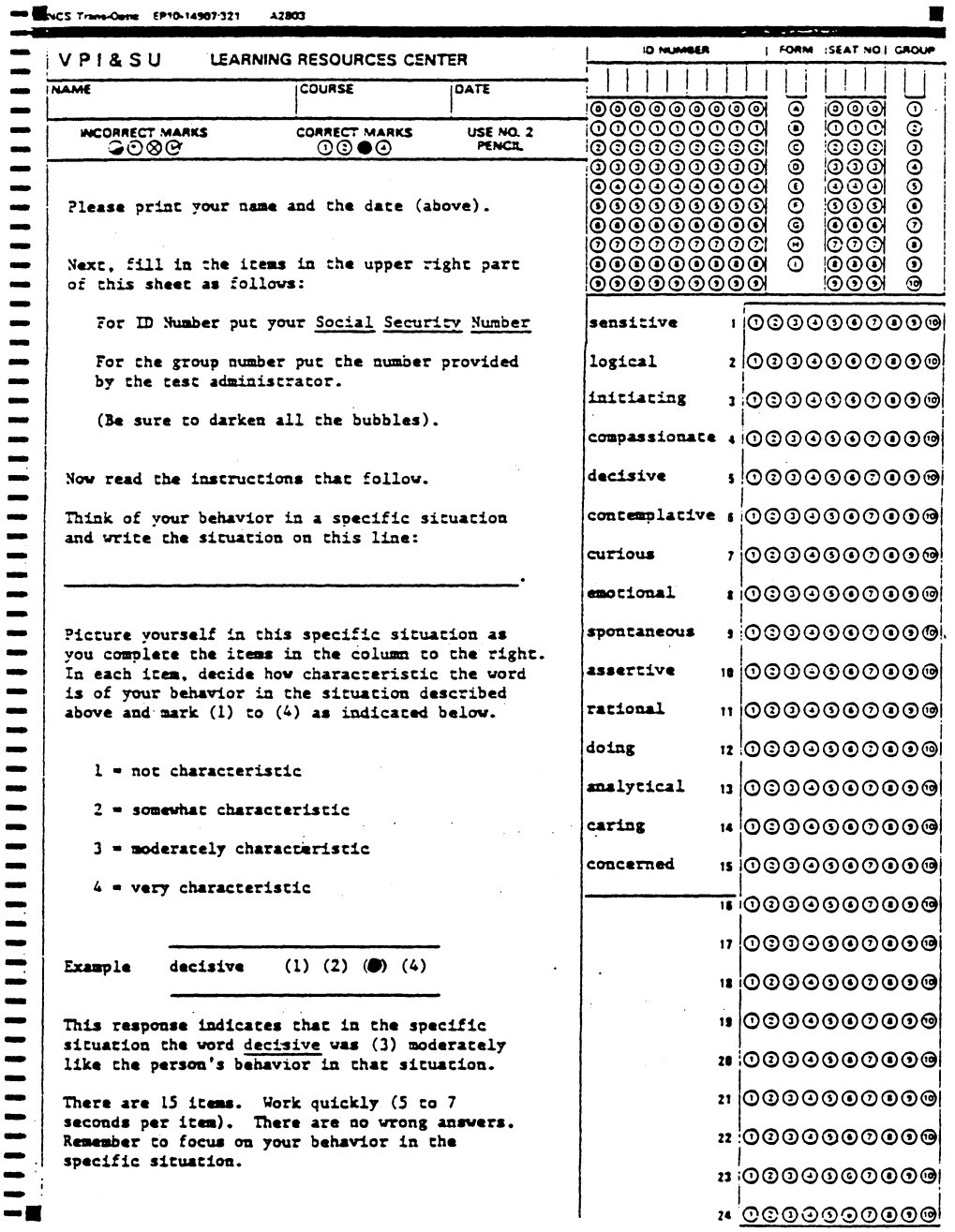
scores corresponded to the ipsative raw scores produced in Walker's study. In this study, T choice scores were derived by simply counting up the total number of T words chosen by a subject. F and A choice scores were derived in the same manner using F words, and then A words, respectively.

The second method of scoring was used to obtain ipsative T, F, and A intensity scores. These intensity scores corresponded to the ipsative frequency scores produced in Walker's study. In this study, subjects selected a word from the pair of words offered in an item. They then indicated how characteristic the word was of their behavior in the specific situation by responding to a likert-type scale consisting of three intensities; s = somewhat (assigned a value of 1), m = moderately (assigned a value of 2), or v = very (assigned a value of 3). The HBI-I T average intensity scores were derived by: 1) summing up the values (intensities) assigned to all T words selected; and 2) dividing by the number of T words selected. The F, and A, average intensity scores were derived in the same manner using F words, and A words, respectively.

The third method of scoring the HBI-I used in this study was designed to produce ipsative TA, AF and FT bipolar scores. These scores indicate where subjects' responses placed them on each of three bipolar behavior continuums: thinking to acting - TA; acting to feeling - AF; and feeling

to thinking - FT. The HBI-I TA bipolar scores for subjects were computed by determining the number of times subjects chose A words over T words whenever T and A words were paired. The HBI-I AF bipolar scores for subjects were computed by determining the number of times subjects chose F words over A words whenever A and F words were paired. Finally, the HBI-I FT bipolar scores were computed by determining the number of times subjects chose T words over F words whenever F and T words were paired. There were 25 TA word pairs, 25 AF word pairs, and 25 FT word pairs on the HBI-I. Therefore, HBI-I TA, AF, and FT bipolar scores could range from 0 to 25.

The HBI-N. The second form of the HBI investigated in this study was the HBI-N (Figure 2). This form was designed to produce independent T, F, and A scales and normative scores for each scale. The HBI-N incorporated the same 15 words used in the HBI-I. Each word was used as a separate item; thus, there were only 15 items on the form. The instructions on the HBI-N directed subjects to select a specific situation to focus on while completing all 15 items of the inventory. The instructions then explained that subjects were to use a likert-type scale to indicate how characteristic each word was of their behavior in the specific situation upon which they were focusing. The likert-type scale consisted of four intensity choices: (1)



NCS Trans-Optic EP10-14907-321 A2803

V P I & S U LEARNING RESOURCES CENTER

NAME | COURSE | DATE

INCORRECT MARKS ⊘ ⊙ ⊗ ⊖ | CORRECT MARKS ① ② ● ④ | USE NO. 2 PENCIL

Please print your name and the date (above).

Next, fill in the items in the upper right part of this sheet as follows:

For ID Number put your Social Security Number

For the group number put the number provided by the test administrator.

(Be sure to darken all the bubbles).

Now read the instructions that follow.

Think of your behavior in a specific situation and write the situation on this line:

_____.

Picture yourself in this specific situation as you complete the items in the column to the right. In each item, decide how characteristic the word is of your behavior in the situation described above and mark (1) to (4) as indicated below.

1 = not characteristic

2 = somewhat characteristic

3 = moderately characteristic

4 = very characteristic

Example decisive (1) (2) (●) (4)

This response indicates that in the specific situation the word decisive was (3) moderately like the person's behavior in that situation.

There are 15 items. Work quickly (5 to 7 seconds per item). There are no wrong answers. Remember to focus on your behavior in the specific situation.

ID NUMBER | FORM | SEAT NO | GROUP

sensitive	1	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
logical	2	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
initiating	3	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
compassionate	4	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
decisive	5	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
contemplative	6	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
curious	7	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
emotional	8	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
spontaneous	9	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
assertive	10	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
rational	11	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
doing	12	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
analytical	13	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
caring	14	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
concerned	15	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	16	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	17	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	18	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	19	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	20	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	21	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	22	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	23	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
	24	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩

Figure 2. The Current HBI-N

not characteristic, (2) somewhat characteristic, (3) moderately characteristic, and (4) very characteristic.

Scoring the HBI-N. Two methods were used to score the HBI-N. The first method produced normative T, F, and A intensity scores. These intensity scores corresponded to the intensity scores produced by the HBI-I; however, there were some major differences in the scores and how they were derived. First, the HBI-N produced normative scores instead of ipsative scores. Next, the HBI-N used a likert-type scale that offered four intensities, whereas the HBI-I offered only three. This was arranged so that the HBI-N and HBI-I would produce scores as similar as possible. On the HBI-I, a subject might never have preferred a particular behavior word when it was matched with other words. Thus, this non-preferred word would never have been chosen. By offering the "not characteristic" choice on the HBI-N, a subject could exercise this non-preferrence. The final difference between the scores was that the T, F, and A intensity scores produced by the HBI-N could range in value from 0 to 15; compared to a range of only 1.0 to 3.0 for the T, F, and A average intensity scores produced by the HBI-I.

The second method of scoring the HBI-N was used to produce normative bipolar scores. These scores represented an attempt to create bipolar scores that would correspond as closely as possible to the HBI-I bipolar scores. The HBI-N

TA, AF, and FT bipolar scores were produced by subtracting normative T, F, and A intensity scores from one another. The TA normative bipolar scores were produced by subtracting the subjects' T intensity scores from their A intensity scores. The AF normative bipolar scores were produced by subtracting the subjects' A intensity scores from their F intensity scores. Finally, the subjects' FT normative bipolar scores were produced by subtracting the subjects' F intensity scores from their T intensity scores. The HBI-N bipolar scores could range from -15 to +15, while the HBI-I bipolar scores could range from 0 to 25. Table 2 lists all of the different scores analyzed in this study.

The Strong Campbell Interest Inventory (SCII)

The Strong Campbell Interest Inventory is the result of an integration of Holland's personality theory with the Strong Vocational Interest Blank (Campbell, 1974). Since the personality dimensions hypothesized by Holland (1973) will be incorporated into a multitrait-multimethod validity study, that aspect of the SCII will be discussed below.

Holland (1973) stated that there are six personality types--realistic, investigative, artistic, social, enterprising, and conventional. He claimed that each person exemplifies one, or a combination, of these personality types.

Table 2

Inventories and Scores Used in the Study

Inventory	Score	Range
HBI-I	T, F, A Choice Scores	0 to 50
	T, F, A Average Intensity Scores	1 to 3
	TA, AF, FT Bipolar Scores	0 to 25
HBI-N	T, F, A Intensity Scores	0 to 15
	TA, AF, FT Bipolar Scores	-15 to 15
SCII	General Occupational Theme Scores R, I, A, S, E, C	-
MBTI	Continuous Preference Scores E-I, S-N, T-F, J-P	-

Note. The score ranges for the SCII and MBTI can be found by looking in the manuals of the respective instruments.

Holland proposed that each personality type represents a cluster of occupational stereotypes that are intrinsically appealing to an individual dominant in that type. Therefore, six general occupational themes were structured into the SCII to measure the extent to which an individual possesses each of the six personality types. Definitions of each SCII personality type have been provided by Campbell & Hansen (1981) and are summarized in Appendix B.

It was a comparison of verbal descriptions of the SCII personality types and HBI behavior dimensions that led to the decision to include the SCII scales in the multitrait-multimethod validity analysis of the HBI. The six SCII personality type verbal descriptions (appendix B) were compared to the verbal descriptions of the three HBI behavior dimensions (appendix A). Similarities and differences were noted between certain personality types and behavior dimensions. For demonstration purposes, appendix D contains examples of how the similarities were established between the HBI scales and four of the SCII scales. It was the identification of such similarities, as well as differences, that resulted in the selection of the SCII for this study. A procedure for arriving at hypotheses concerning the relationships between the various personality types and behavior dimensions was developed and applied, and will be explained later.

Description of the SCII. The latest version of the SCII consists of 325 items. In the test manual it is explained that a person responds to most items by indicating whether they like, feel indifferent toward, or dislike the item (Campbell & Hansen, 1981). The inventory has seven sections that include: occupations (131 items), school subjects (36 items), activities (51 items), amusements (39 items), types of people (24 items), preference between two activities (30 items), and personal characteristics (14 items).

The SCII must be computer scored. The results are reported in either an interpretive or profile layout. Both report forms cover the same basic information: general occupational themes, basic interest scales, occupational scales, and administrative indices and special scales. The SCII is simple to administer individually or in groups and takes the average adult about 30 minutes to complete. The reading difficulty is at approximately the sixth-grade level for most items.

The reliability and validity of the general occupational themes are reported here since the scores on these themes are the ones being used in this study. Campbell and Hansen (1981) reported that the median test-retest reliability coefficients for the six themes are as follows: .91 for two weeks (range = .85 to .93); .86

for 30 days (range = .84 to .91); and .81 for three years (range = .78 to .87). Intercorrelations between the themes range from +.59 (conventional and enterprising) to -.20 (conventional and artistic).

The validity of the Holland personality typologies measured by the SCII General Occupational Themes has been studied extensively. Scales to measure the personality typologies were first developed for Holland's Vocational Preference Inventory (VPI). Holland (1978) reported in the VPI manual that most construct validity studies done on the VPI scales (well over 100) provided evidence in accord with theoretical expectations.

Holland and Campbell subsequently constructed new scales to measure the six personality types. First, they used items from the Strong Vocational Interest Blank, and more recently they used items from the Strong Campbell Interest Inventory (Campbell & Hansen, 1981). Many correlational, group differences, and other types of construct validity studies of the SCII scales used to measure the Holland personality types have been reported by Holland (1985) and Gottfredson, Holland, and Ogawa (1982). The evidence reported has been primarily positive.

The Myers-Briggs Type Indicator (MBTI)

The MBTI is an instrument that is based upon Jung's theory of type, in which he suggested that variations in the

behavior of people can be attributed to the difference in how persons prefer to use perception and judgment (Myers, 1962). The MBTI measures perceiving and judging as defined by Myers. According to Myers (1962), perception involves "the processes of becoming aware,--of things or people or occurrences or ideas" (p. 1); while judgment involves "the processes of coming-to-conclusions about what has been perceived" (p. 1).

The indicator consists of four different bipolar indices, each of which measures one of the four preferences that, taken together, make up the personality of an individual (Myers, 1962). The first preference is between Extraversion and Introversion and is measured by the EI index. The second preference, which is measured by the SN index, is between Sensing and Intuition. The third preference, which is measured by the TF index, is between Thinking and Feeling. The fourth and final preference, which is measured by the JP index, is between Judgment and Perception. Appendix C presents a detailed summary of the characteristics of each of the four MBTI personality preferences.

A comparison of the verbal descriptions of the MBTI personality types (Appendix C) and the HBI behavior dimensions (Appendix A) resulted in the decision to include the MBTI scales in the multitrait-multimethod validity study

of the HBI. The comparison revealed similarities and differences between certain personality types and behavior dimensions. For demonstration purposes, examples of how the similarities between certain MBTI and HBI scales were established is presented in Appendix D. It was the identification of such similarities, as well as differences, that resulted in the selection of the MBTI for this study. The procedure followed to arrive at hypotheses concerning the relationships between MBTI and HBI scales was the same as the one used with the SCII and HBI that will be explained later.

Description of the MBTI. The MBTI is an objective self-report personality test. Form F consists of 166 forced-choice items that, when scored, provide normative measures of the four personality preferences. The subject's responses on the test tend to reflect an habitual choice between opposites. The test is designed for use with high school, college or adult populations.

A number of split-half internal consistency reliability coefficients are reported for the four different preference indices. The split-half coefficients were obtained by calculating product-moment correlation coefficients between X half and Y half continuous scores. These correlations were then corrected by using the Spearman-Brown prophecy formula. The range of coefficients for the four preference

indices for samples from various populations can be reviewed in the latest version of the MBTI manual (Myers & McCaulley, 1985).

Internal consistency reliabilities for each preference index have also been calculated using Cronbach's coefficient alpha. The following ranges have been reported: .74 to .83 for the EI index; .74 to .85 for the SN index; .64 to .82 for the TF index; and .78 to .84 for the JP index (p. 169). Samples used to derive the coefficient alphas can be reviewed in the 1985 MBTI manual (Myers and McCaulley, 1985).

Other types of reliability coefficients were also calculated and have been reported by Myers and McCaulley (1985). Of particular interest are the test-retest product-moment correlations of MBTI continous preference scores because those scores were the type derived and analyzed in this study. The reliabilities of the various continuous preference scores range as follows: .51 to .93 for the EI index; .58 to .93 for the SN index; .45 to .91 for the TF index; and .45 to .89 for the JP index.

The validity of the MBTI has also been studied extensively and the results are reported in great detail in the latest manual (Myers & McCaulley, 1985). Most of the studies have concerned the construct-related validity of the instrument. The authors have included data revealing how

the four preference indices relate to the following: behavior, psychological treatment, psychological treatment modes, careers, aptitudes, interests, achievement, creativity, time, fantasy and imagery, optimism and pessimism, and factors derived from other measures.

The manual reports the ranges of correlations between the various preference indices and other scales with which they would be expected to correlate. The ranges were as follows: Extraversion, -.40 to -.77; Introversion, .40 to .75; Sensing, -.40 to -.67; Intuition, .40 to .62; Thinking, -.40 to -.57; Feeling, .40 to .55; Judging, -.40 to -.59; perception, .40 to .57 (Myers & McCaully, 1985). These data are provided in the manual as further evidence in support of the construct-related validity of the MBTI.

Carlyn (1977), as part of an extensive review of the MBTI, examined studies that assessed the instrument's validity. Carlyn reported that the EI, SN, and TF scales appear to be relatively independent of one another. The JP scale, however, shows a strong relationship to the SN scale.

Procedures

Subjects

The subjects for this study were students attending undergraduate psychology classes at a private liberal arts university during the spring semester of the 1985-86 academic year.

Introduction to the Test Procedures

Five instructors, each teaching sections of various psychology courses, notified students in prearranged sections of their courses that each pupil would be asked in the future to complete certain psychological inventories during some of the regularly scheduled class sessions. The instructors explained that the procedures would be designed to expose the students to the topic of psychological assessment and that the first inventories would be administered at the next regularly scheduled class session.

Administration of the Tests

At the next class session the test administrator (this researcher) was introduced. The administrator read a statement that introduced the nature of the research being undertaken. The statement appears in appendix E. Next, students not wishing to participate were allowed to leave before the testing began. The administrator then indicated that full instructions were written into each inventory explaining how to complete it correctly.

Five different test sessions were needed to gather all of the data required for this study. Each group of subjects was involved in some, but not all, of the test sessions. Each inventory was used in some test sessions but not others. The sequence followed to administer the inventories differed from one test session to another in order to

produce a counterbalancing effect. This counterbalancing of the administrations of the inventories from one test session to the next is recommended because it helps to control for the effects of contigency variables that often tend to introduce systematic, nonrandom error into derived scores (Christensen, 1985). Table 3 indicates which inventories were administered to each of the seven groups. It also indicates the order in which the inventories were given.

During the first test session the subjects in all seven groups (A-G) were given two different forms of the HBI to complete. The situation they focused on while completing the forms was the very first class they attended in the Spring semester. Groups A, B, and F completed the HBI-I first; they then completed the HBI-N. Groups C, D, E and G completed the two forms in the reverse order. The scores obtained from this first session were called <u>test</u> scores because they were used as the test scores in all of the test-retest analyses.

During the second test session, which was designed to gather 7-day retest scores, students in groups A, C, and G completed the two forms of the HBI again. Groups C and G completed the HBI-I first; they then completed the HBI-N. Subjects in group A completed the forms in the reverse order. This session occurred seven days after the first test session.

Table 3

Test Administration Counterbalancing

Group	Order of Administration During Test Session									
	1st Session		2nd Session (7 day retest)		3rd Session (14 day retest)		4th Session		5th Session (28 day retest)	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
A	HBII	HBIN	HBIN	HBII			SCII	MBTI		
B	HBII	HBIN							HBIN	HBII
C	HBIN	HBII	HBII	HBIN			MBTI	SCII		
D	HBIN	HBII							HBII	HBIN
E	HBIN	HBII			HBII	HBIN				
F	HBII	HBIN			HBNI	HBII	SCII	MBTI		
G	HBIN	HBII	HBII	HBIN			SCII	MBTI		

During the third test session, which was designed to collect 14-day retest scores, only groups E and F were tested. Group E took the HBI-I first and the HBI-N second. Group F took the forms in the opposite order. This third test session occurred 14 days after the first test session.

During the fourth test session, Groups A, F, and G took the SCII first, and then the MBTI. Group C took the MBTI first, and then followed up by taking the SCII. These test scores were subsequently used in the multitrait-multimethod construct validity analysis of the HBI. They were also used in the factor analysis procedures.

During the fifth and final test session, which was designed to collect 28-day retest data on the two forms of the HBI, only Groups B and D were tested. Group D took the HBI-I first; the group then took the HBI-N. Group B took the HBI-N first, followed by the HBI-I. This test session occurred 28 days after the first test session.

In all five test sessions the same administrative procedures were followed. The first inventory in the test sequence was distributed to the students along with pencils appropriate for completing the inventory. Once every student had received an inventory, the administrator read aloud the instructions for completing the inventory while the students followed along. Students were given time to fill in all demographic information requested. Students

were encouraged to ask clarifying questions, if necessary, during this part of the administrative procedure. Once all demographic items were completed and all relevant questions answered, the students were instructed to begin filling out the inventory.

After the first inventory was completed the students were asked to place the inventory under their chairs. The second inventory was then distributed. The same procedures, of reading the instructions, providing demographic information, and answering questions, were followed for the second inventory. When everyone was ready, students were advised to begin completing the second inventory. Once both inventories were completed, the students were asked to pass them in to the administrator. The students were then thanked for their participation and dismissed.

Scoring of the Tests

The MBTI forms were scored manually while the SCII forms were machine scored. The HBI ipsative and normative forms were scored by computer at the VPI-SU Learning Resources Center.

Research Questions and Analyses of Data

This study had two major objectives. The first was to investigate and compare the reliability of scores produced by the HBI-I and HBI-N. The second objective was to investigate and compare the validity of scores produced by

these two forms of the HBI. Six research questions pertaining to the two objectives were addressed in this study. Each separate research question is set forth below, along with the type of data analysis procedure that was incorporated in order to generate evidence aimed at answering the question.

1. What is the test-retest reliability of scores produced by the forced-choice ipsative form of the HBI - the HBI-I ?

To answer this question, 7-day, 14-day, and 28-day test-retest reliability coefficients were calculated using all three kinds of ipsative scores (choice, average intensity, and bipolar). Groups A, C, and G were used to gather the 7-day test-retest data; groups E and F were used to gather 14-day test-retest data; and groups B and D were used to gather 28-day test-retest data. The reliability coefficients were computed using the Pearson product-moment correlation procedure.

2. What is the test-retest reliability of scores produced by the single stimulus normative form of the HBI - the HBI-N?

To answer this question, 7-day, 14-day and 28-day test-retest reliability coefficients were calculated using both kinds of normative scores (intensity and bipolar). Groups A, C, and G were again used to gather the 7-day

test-retest data; groups E and F were used to gather 14-day test-retest data; and groups B and D were used to gather 28-day test-retest data. The reliability coefficients were computed using the Pearson product-moment correlation procedure.

3. What is the internal consistency reliability of scores produced by the single stimulus normative form of the HBI - the HBI-N?

To answer this question, three Cronbach coefficient alphas were calculated (Cronbach, 1951). The alphas were computed using the normative T, F, and A intensity scores that are produced by the HBI-N. Subjects in all seven groups (A-G) completed the HBI-N during the first test session. All of the normative test scores gathered during that first session were used to calculate the three alphas.

4. What are the latent dimensions of behavior that underlie the items in the single stimulus normative form of the HBI - the HBI-N?

As mentioned earlier, subjects in all seven groups (A-G) took the HBI-N during the first test session. The test scores of subjects from all of those groups were used in two factor analysis procedures to answer this question. A correlation matrix of the variables (test items) was analyzed using common factor analysis, which is "used primarily to identify underlying factors or dimensions not

easily recognized" (Hair, Anderson, Tatham, & Grablowsky, 1979, p. 221). An oblique factor solution was applied so that the actual correlations between factors could occur (Nunnally, 1978). A principal components factor analysis with varimax (orthogonal) rotation was also carried out for comparison purposes.

5. How do the following relate to each other?
    1) The behavior scores produced by the forced-choice ipsative form of the HBI - the HBI-I.
    2) The behavior scores produced by the single stimulus normative form of the HBI - the HBI-N.
    3) The general occupational theme scores (Holland personality types) produced by the SCII.
    4) The continuous personality preference scores produced by the MBTI?

To answer this question a multitrait-multimethod construct validity matrix was developed to examine the relationships between the different sets of scores listed above. Groups A, C, F, and G were used to gather test scores on the HBI-I and HBI-N during the first test session. Further, all four of these groups took the SCII and MBTI during the fourth test session. Therefore, scores from

these four groups were used to develop the multitrait-multimethod construct validity matrix. An examination was then possible of the convergent validity "demonstrated by high correlations between scores on tests measuring the same trait by different methods" (Allen & Yen, 1979, p. 110). At the same time, discriminate validity could also be observed, and this would be "demonstrated by low correlations between scores on tests measuring different traits" (Allen & Yen, 1979, p. 110).

All of the correlations that were placed into the multitrait-multimethod validity matrix were calculated using the Pearson product-moment correlation procedure. Test-retest reliability coefficients for the SCII and MBTI scores were not calculated; thus, the reliability diagonal of the multitrait-multimethod validity matrix only contained the 7-day test-retest reliability coefficients for the HBI ipsative and normative scores.

A multitrait-multimethod validity matrix was prepared prior to the collection of data that contained the hypothesized direction and magnitude of correlations between scores. A three step procedure was followed in arriving at the hypothesized correlations placed into the matrix. First, verbal descriptions of the scales being correlated were examined for similarities and differences. For the purpose of displaying how this step was carried out,

Appendix D was prepared in order to demonstrate how similarities between verbal descriptions of certain scales were identified. In step two, known relationships were examined between SCII scales correlated with one another, MBTI scales correlated with one another, and SCII scales correlated with MBTI scales. This step frequently altered expectations of how the scales would subsequently correlate. In step three, the affect that the kinds of scales being correlated (bipolar or nonbipolar; ipsative or normative) would have upon one another was considered. The matrix of hypothesized relationships that resulted from following these three steps is presented in Table 4. These hypothesized correlations will be compared to the correlations actually derived in the study.

6. How many, and what kind, of latent constructs underlie the scores produced by the scales of the HBI-N, the SCII, and the MBTI?

Two different factorial validity analyses were performed on the scores from the scales of these instruments in order to explore for latent constructs. First, a common factor analysis that incorporated the Harris-Kaiser method of oblique rotation was carried out while holding for N = 3 factors. Squared multiple correlations (SMCs) were used on the diagonal of the correlation matrix. The second analysis was identical to the first in every respect, excepting that

Table 4

Hypothesized Relationships Between Scales of the HBI, SCII, and MBTI

		HBI-I			HBI-I			HBI-N			HBI-N			SCII						MBTI			
					T	F	A				T	F	A							S	T	E	J
		T	F	A	A	T	F	T	F	A	A	T	F	R	I	A	S	E	C	N	F	I	P
	T	H																					
HBI-I	F	m	H																				
	A	m	m	H																			
	TA	h	L	H	H																		
HBI-I	FT	H	h	L	m	H																	
	AF	L	H	h	m	h	H																
	T	H	l	l	m	H	L	H															
HBI-N	F	l	H	l	L	h	H	L	H														
	A	l	l	H	H	L	h	L	L	H													
	TA	h	l	H	H	m	m	h	L	H	H												
HBI_N	FT	H	h	l	m	H	m	H	h	L	m	H											
	AF	l	H	h	m	m	H	L	H	h	m	m	H										
	R	H	L	L	M	M	m	M	L	M	M	L	m	H									
	I	H	L	L	m	M	l	H	L	L	m	H	l	M	H								
SCII	A	L	L	L	L	m	M	L	H	L	L	m	M	l	M	H							
	S	L	H	L	L	m	M	L	H	M	L	l	L	l	L	L	H						
	E	L	H	H	H	l	h	L	L	H	H	l	m	M	l	l	M	H					
	C	H	L	L	M	l	l	M	L	M	M	M	L	M	L	l	M	H	H				
	SN	m	M	L	m	m	l	m	M	l	m	m	L	L	M	M	L	M	m	H			
MBTI	TF	h	H	L	m	h	H	h	H	L	M	h	M	m	h	M	L	l	m	L	H		
	EI	M	L	h	h	l	H	M	L	h	h	M	M	L	L	L	m	h	l	L	L	H	
	JP	m	M	L	m	l	L	l	L	l	L	l	L	L	L	L	L	L	l	M	L	L	H

Note. H = high positive correlation; M = moderate positive correlation; L = low positive correlation; h = high negative correlation; m = moderate negative correlation; l = low negative correlation.

the number of factors was not restricted. These procedures were used because they could determine the number of underlying factors accounting for the scores and also determine the nature of the intercorrelations among the factors identified in the analyses. The HBI-N first test session scores of all groups were available for these two factor analysis procedures; but only groups A, C, F, and G possessed SCII and MBTI scores that were avaliable for these analyses.

CHAPTER FOUR

Results and Discussion

Demographic Data

A total of 589 subjects participated in this study: 344 females (58.4%) and 245 males (41.6%). All of the subjects were enrolled in psychology courses at a private liberal arts university. Not all subjects were involved in each aspect of the study; therefore, the specific number of subjects involved in each procedure is noted when reporting each set of results. The mean age for subjects was 20.0 with a standard deviation of 3.6 years. Ages ranged from 17 to 57 with approximately 93 percent of the subjects having ages between 17 and 23. Appendix F provides a detailed breakdown of frequencies and percentages concerning subjects' ages and classifications (academic status). Appendix G presents means, standard deviations, and minimum/maximum values for all variables included in the study.

The Test-retest Reliability of HBI-I and HBI-N Scores

The first two research questions concerned the test-retest reliability of the scores of the HBI-I and HBI-N. Reliability coefficients were derived for 7-day, 14-day, and 28-day intervals. Coefficients were derived for the scales under all three scoring methods of the HBI-I (choice, average intensity, and bipolar) and under the two

scoring methods of the HBI-N (intensity and bipolar). All of the reliability coefficients were derived using the Pearson product-moment correlation procedure (Nunnally, 1978). The 7-day, 14-day, and 28-day reliability coefficients appear in Table 5.

7-day Test-retest Reliability

The range of the 7-day test-retest reliability coefficients for the three HBI-I scoring methods - choice, average intensity, and bipolar - was from .68 to .86. For the choice scores the coefficients ranged from .81 to .86; for the average intensity scores the coefficients ranged from .68 to .77; for the bipolar scores the reliability coefficients ranged from .80 to .84. The range of the 7-day test-retest reliability coefficients for the two HBI-N scoring methods - intensity and bipolar - was from .58 to .72. For the intensity scores the coefficients ranged from .58 to .72; for the bipolar scores the coefficients ranged from .69 to .72. Overall, the HBI-I scores had consistently higher 7-day test-retest reliability coefficients than the HBI-N scores.

14-day Test-retest Reliability

The 14-day test-retest reliability coefficients for the HBI-I ranged from .71 to .83. The coefficients for both the choice and bipolar scores were slightly lower than the 7-day test-retest coefficients. However, for the average intensity

Table 5

HBI-I and HBI-N Test-retest Reliability Coefficients

		Interval					
		7-day		14-day		28-day	
Score	Scale	r	n	r	n	r	n
		HBI-I					
Choice			104		68		229
	T	.82		.78		.74	
	F	.86		.78		.74	
	A	.81		.79		.70	
Average Intensity							
	T	.77		.75		.57	
	F	.73		.72		.69	
	A	.68		.74		.65	
Bipolar							
	TA	.80		.71		.75	
	FT	.83		.76		.71	
	AF	.84		.83		.71	
		HBI-N					
Intensity			107		68		236
	T	.58		.36		.48	
	F	.69		.46		.60	
	A	.72		.76		.44	
Bipolar							
	TA	.71		.60		.53	
	FT	.69		.38		.64	
	AF	.72		.71		.55	

scores the coefficients remained stable. The 14-day test-retest reliability coefficients for the HBI-N ranged from .36 to .76. Most of the the HBI-N intensity and bipolar score coefficients were lower than their HBI-I counterparts. A curious fact is that the HBI-N A intensity scores had a higher reliability coefficient for the 14-day interval than for the 7-day interval. Normally, reliability decreases as length of time increases.

28-day Test-retest Reliability

The HBI-I 28-day test-retest reliability coefficients ranged from .57 to .75. The coefficients for the HBI-I choice and bipolar scores were slightly lower than the 14-day test-retest coefficients; they ranged from .70 to .75. The coefficients for the HBI-I average intensity scores, on the other hand, were appreciably lower than the 14-day test-retest coefficients, especially in the case of the T intensity scores ($r = .57$). The HBI-N A intensity scores, as well as the AF and TA bipolar scores, had 28-day reliability coefficients that were lower than their 14-day counterparts, as expected. However, coefficients for the HBI-N T and F intensity scores, and FT bipolar scores, were higher for the 28-day interval than for the 14-day interval. Again, it was curious to occasionally have coefficients increasing in size when normally they should remain stable or decrease in size. Generally, the 28-day test-retest

reliability coefficients for the HBI-N were much lower than the 28-day coefficients for the HBI-I.

The Effectiveness of Test Administration Counterbalancing

Counterbalancing was used in this study in an attempt to ensure that, should multiple-treatment (sequencing) effects occur, they could be controlled by spreading them evenly over all treatments (a treatment here being defined as an administration of one of the inventories used in this study). Practice, treatment-carryover, and demand characteristics are the three general types of multiple-treatment effects that occur in studies where numerous treatments (tests) are administered sequentially (Kiess & Bloomquist, 1985). Independent t-tests were conducted on the HBI-I and HBI-N test and retest scores. The scores of groups who took a particular inventory first in a testing session were compared to the scores of groups who took the inventory last in the testing session. An examination of independent t-test results for the HBI-I test and retest scores revealed no significant differences between groups; nor were any significant differences found between groups for the HBI-N test scores. On the other hand, the t-test results for the HBI-N retest scores revealed significantly higher mean scores on the HBI-N T and A scales for those who took the retest first; whereas those who took the retest second had a higher mean score on the

HBI-N F scale. All of the t-test results are presented in Appendix H.

Uncertainty exists concerning the cause, or causes, of these differences. If the differences were due to practice or treatment-carryover effects, then it would be expected that those taking the HBI-N retest first in the sequence would have higher (or lower) means on all three scales of the HBI-N. This, as mentioned earlier, was not the case. Those taking the HBI-N retest first had significantly higher means on the T (9.86 vs 8.92) and A (7.14 vs 6.34) scales. Those taking the HBI-N retest second had a higher (though not significantly different) mean on the F scale (7.18 vs 7.11).

Apparently one or more factors other than practice or treatment-carryover affected the HBI-N retest scores of the groups. One possibility might be that the subjects who took the HBI-N second were reacting emotionally to the fact that they had just completed the much longer HBI-I retest. This might have caused the F scores to increase, and the T and A scores to decrease. But if so, should not the same results have occurred during the HBI-N initial test session? Possibly not, if we consider that the retest groups had already taken both inventories before, whereas the initial test groups were taking the inventories for the first time.

Another possibility might be that some subtle cue from the experimenter might have influenced either the first or second group of subjects taking the HBI-N retest to alter their behavior in accord with some perceived hypothesis. Still another possibility could be that the first and second groups actually differed at the time of the retest due to mortality that occurred between the test and retest sessions. But any actual differences between the groups should have been reflected in the HBI-I retest scores as well as the HBI-N retest scores, which was not the case.

Some other possible causes that might account for the retest differences in scores include carelessness of subjects in responding to the HBI-N or confusion due to variation in HBI-N administrative instructions. Carelessness of subjects is a very real possibility that cannot be ruled out. With regard to confusion due to administrative instructions, great care was taken to ensure that all administrative instructions were given in a highly standardized fashion. Still, some undetected alteration of instructions could have occurred.

While the cause of the group differences in HBI-N retest scores remains a matter of speculation, it is important to recognize that the differences exist and create a problem concerning the HBI-N test-retest reliability coefficients. Something caused groups that did not have

significantly different HBI-N test scores to have significantly different retest scores; and this fact, to some extent, confounds the interpretation of the HBI-N test-retest reliability coefficients.

The Internal Consistency Reliability of HBI-N Scores

The third research question concerned the internal consistency of HBI-N scores. Coefficient alphas (Cronbach, 1951) were calculated for the T, F, and A intensity scores of the HBI-N. These intensity scores were derived from scales consisting of only five items each. The scores from the first test session for all seven groups (A-G) were used to calculate the coefficient alphas (N = 583).

Coefficient Alphas for the HBI-N

The means, standard deviations, and coefficient alphas for all three HBI-N intensity scores (T, F, and A) are presented in Table 6. The ranges for means and standard deviations of the scores of all three intensity scales were relatively small. The coefficient alphas ranged from .59 (T) to .77 (F); they were relatively large for scores derived from scales that consisted of only five items.

HBI-N Item-scale Correlations

Table 7 presents the HBI-N item-scale correlations. The scale correlations with items within the scale were moderate to high; the range was from .44 to .78. The scale correlations with items outside the scale were low; they

Table 6

Means, Standard Deviations, and Cronbach Coefficient Alphas for the HBI-N Intensity Scores (N = 583)

Scale	Mean	Standard Deviation	Alpha
T	14.63	2.67	.59
F	12.42	3.48	.77
A	11.89	2.99	.67

Table 7

HBI-N Intensity Scales Correlated with Items within the Scales and with Items Outside the Scales (N = 583)

Scale	items within		items outside	
	range	mean	range	mean
T	.44 to .71	.61	.01 to .25	.11
F	.67 to .78	.72	.01 to .17	.09
A	.57 to .75	.66	.01 to .26	.13

Note. Each intensity scale consisted of five items.

ranged from .01 to .26. The internal and external structure patterns found for the scale-item correlations suggest that the scores of the scales are measuring different constructs.

Intercorrelations among HBI-N Summated Scale Scores

The intercorrelations between the HBI-N intensity summated scale scores were all low, suggesting, as did the item-scale correlations, that the HBI-N intensity scores are measuring different constructs. A low correlation occurred between the A (acting) and T (thinking) scores ($r = .25$); an even lower correlation occurred between the A (acting) and F (feeling) scores (.16); and the lowest correlation was between the T (thinking) and F (feeling) scores ($r = .10$).

The Number of Latent Dimensions Present in the HBI-N

The fourth research question concerned the number of latent dimensions of behavior present in the HBI-N. Hinkle, Wiersma, and Jurs (1979) suggested that the items in a single measure or test can be factor analyzed in order to determine if the test measures the number of constructs hypothesized. Hair, Anderson, Tatum, and Grablowsky (1979) noted that the appropriate factor analysis procedure to use when seeking to determine the number of latent constructs underlying a set of variables is common factor analysis using an oblique rotation of the initial factors. Nunnally (1978) suggested, in his thorough discussion of common

factor analysis, that using squared multiple correlations (SMCs) for communality estimates on the diagonal of the correlation matrix is a "sensible approach" (p. 411).

It is important to note that Nunnally (1978), Zeller and Carmines (1980), and others cautioned that there are certain disadvantages that go along with the use of common factor analysis, especially when an oblique rotation is used. One is that the factor loadings cannot be interpreted as item-factor correlation coefficients. A second is that the eigenvalues "lose their integrity" (Zeller & Carmines, 1980, p. 44). These disadvantages are not present when one conducts a principal components analysis using orthogonal rotation. However, with principal components analysis, followed by orthogonal rotation, one loses the ablility to determine if the rotated factors are correlated. Nunnally (1978) stated that since both approaches are "mathematically legitimate, which is used boils down to a matter of taste" (p. 376).

Because of differences in opinion surrounding which approach to use, both a common factor analysis using oblique rotation and a principal components analysis using an orthogonal rotation were conducted in this study for comparison purposes.

A common factor analysis of the HBI-N items was conducted using squared multiple correlations (SMCs) on the

diagonal of the correlation matrix and using the Harris-Kaiser method of oblique rotation. The solution was restricted to three factors. The decision to use N = 3 factors was made because Hutchins has hypothesized that there are three components of behavior in the TFA system - thinking, feeling, and acting - and he designed the HBI to measure these.

The principal components analysis of the HBI-N items was conducted using the Varimax method of orthogonal rotation. The Varimax orthogonal rotation has been continually improved by Kaiser and others and is widely accepted as the best method of orthogonal rotation (Nunnally, 1978).

Eigenvalues

The eigenvalues of all three factors produced by the common factor analysis using the Harris-Kaiser method of oblique rotation were greater than 1.0: factor 1 = 2.165; factor 2 = 1.708; factor 3 = 1.483. The eigenvalues of all three factors produced by the principal components analysis using the Varimax method of orthogonal rotation were greater than 2.0: factor 1 = 2.722; factor 2 = 2.267; factor 3 = 2.008.

Scale item-factor loadings

Nunnally (1978) suggested that factor analysis is useful for investigating the "internal structures of

variables hypothesized to be related to a construct" (p. 107). An examination of the scale-item loadings of the 15 items (variables) of the HBI-N was carried out, therefore, in order to determine how the items related to their hypothesized constructs (T, F, and A). The criterion suggested by Hair, Anderson, Tatham, and Grablowsky (1979) was adopted to judge how well the HBI-N items related to their respective constructs. According to their criterion, a factor loading should account for "approximately 10 percent of the variance of a particular variable" (p. 234). Thus, a scale item-factor loading of .30 is considered an important loading according to this rule of thumb. Further, a loading of .40 is of greater importance, while a loading of .50 or more is regarded as very important.

Fourteen of the fifteen HBI-N items loaded as hypothesized in both the common factor analysis using the Kaiser-Harris oblique rotation and the principal components analysis using the Varimax orthogonal rotation. Table 8 reveals the only item that failed to load significantly in both the common and principal components analyses was the thinking word "curious". This word loaded with the feeling words on the F factor (.24) stronger than with the thinking words on the T factor (.17).

The results of the common factor analysis strongly supported the hypothesis that the HBI measures the three

Table 8

HBI-N Scale Item-factor Loadings Derived while Restricting the Solution to Three Factors (N = 583)

Scale	Word-Item	Factor 1:T	Factor 2:F	Factor 3:A
Common factor analysis: Harris-Kaiser oblique rotation				
T				
	logical	.55	.04	.29
	contemplative	.42	.11	-.01
	curious	.17	.24	.13
	rational	.60	.00	.26
	analytical	.57	.02	.13
F				
	sensitive	.06	.64	.05
	compassionate	-.01	.64	.19
	emotional	-.11	.55	.01
	caring	.10	.70	.19
	concerned	.10	.59	.11
A				
	initiating	.23	.03	.55
	decisive	.32	.05	.43
	spontaneous	.04	.24	.51
	assertive	.21	.07	.66
	doing	.25	.23	.50
Principal components analysis: Varimax orthogonal rotation				
T				
	logical	.62	-.01	.26
	contemplative	.63	.15	-.16
	curious	.21	.30	.09
	rational	.70	-.05	.21
	analytical	.73	.00	.03
F				
	sensitive	.06	.74	-.04
	compassionate	-.08	.71	.16
	emotional	-.13	.67	-.05
	caring	.05	.75	.12
	concerned	.09	.67	.03
A				
	initiating	.09	-.06	.68
	decisive	.27	-.02	.52
	spontaneous	-.14	.21	.63
	assertive	.03	-.02	.78
	doing	.14	.21	.59

dimensions of behavior - thinking, feeling, and acting. Factor 1 had five items that loaded above .30 (range = .32 to .60): four of these were thinking words that loaded above .40 (range = .42 to .60); and one was the acting word "decisive" (loading = .32). Factor 2 had five items that loaded above .50 (range = .55 to .70). All five items were feeling words. Factor 3 had five items that loaded above .40 (range = .43 to .66). All five of these items were acting words.

The results produced by the principal components analysis also strongly supported the hypothesis that the HBI assesses the thinking, feeling, and acting dimensions of behavior as hypothesized by Hutchins. Factor 1 had four items that loaded above .60 (range = .62 to .73). All four words were thinking words. Factor 2 had five items that loaded above .60 (range = .67 to .75). All five items were feeling words. Factor 3 had five items that loaded above .50 (range = .52 to .78). All five items were acting words. The only word that did not load on to any factor was, again, the thinking word "curious".

Inter-factor correlations

The significant advantage of the common factor analysis with oblique rotation over the principal components analysis with orthogonal rotation was that the oblique rotation allowed the correlations among the final factors to become

apparent. When such correlations are low, they serve as evidence that the factors represent different constructs (Nunnally, 1978). In this study the oblique rotation revealed virtually no relationship between the T factor and the F factor ($r = .07$); a low relationship between the T factor and the A factor ($r = .35$); and a low relationship between the F factor and the A factor ($r = .19$).

The Relationships Between Scores Derived from Scales of the HBI-I, HBI-N, SCII, and MBTI

The fifth research question concerned the relationships between the scores of the various scales of the HBI-I, HBI-N, SCII, and MBTI.

The Multitrait-Multimethod Validity Matrix

All possible pairs of scores from the scales of the various instruments were correlated and the resulting coefficients were placed into a multitrait-multimethod validity matrix. The layout of the multitrait-multimethod validity matrix is specifically designed so that reliability, convergent validity, and discriminant validity can be easily observed between different sets of scores. Reliability coefficients appear on the diagonal of the matrix. High correlations between scores from scales of different tests measuring the same trait are evidence of convergent validity. Low correlations between scores from

scales of different tests that measure different traits are evidence of discriminant validity (Allen & Yen, 1979).

As noted previously, using ipsative scores for variables in correlational analysis is not recommended (Guilford, 1952, 1954). However, it was decided that in developing the multitrait-multimethod validity matrix for this study correlations of the HBI-I ipsative scores with other scores would be included for three reasons. First, it was thought that including the HBI-I ipsative score correlations could provide additional evidence concerning the validity of the HBI. Secondly, the chance would exist to note how the correlations of the HBI-N scores with one another compared to the correlations of the HBI-I scores with one another. Finally, the inclusion in the matrix of correlations of the HBI-I ipsative scores with other scores would demonstrate some of the statistical idiosyncrasies encountered when ipsative scores, with their inherent psychometric problems, are correlated with other scores. However, these correlations involving ipsative scores must be interpreted with caution.

Table 9 presents the multitrait-multimethod validity matrix developed using the Pearson product-moment correlations between all possible pairs of the respective scores. For interpretive purposes it was decided that high correlations would be equal to or greater than +.70 or -.70;

Table 9

A Multitrait-Multimethod Validity Matrix Using Scales from the HBI-I, HBI-N, SCII, and MBTI

		HBI-I			HBI-I		
		T	F	A	TA	FT	AF
HBI-I	T	.82					
	F	-.60	.86				
	A	-.16	-.65	.81			
HBI-I	TA	-.76	.01	.68	.80		
	FT	.82	-.90	.34	-.30	.83	
	AF	-.32	.92	-.84	-.25	-.69	.84
HBI-N	T	.36	-.25	-.03	-.26	.34	-.16
	F	-.43	.61	-.36	.08	-.59	.55
	A	-.28	-.20	.51	.50	.00	-.37
HBI-N	TA	-.52	.03	.46	.63	-.26	-.19
	FT	.60	-.68	.28	-.23	.71	-.56
	AF	-.16	.65	-.66	-.29	-.49	.72
SCII	R	.07	-.21	.20	.09	.19	-.21
	I	.14	-.11	.01	-.03	.16	-.03
	A	.02	.03	-.09	-.04	.00	.09
	S	-.06	.04	.00	.11	-.05	.04
	E	.17	.06	.10	.19	-.11	.01
	C	-.08	.02	.07	.10	-.07	-.03
MBTI	SN	.11	-.06	-.04	-.04	.14	.00
	TF	-.27	.38	-.21	.05	-.35	.32
	EI	.22	-.02	-.21	-.30	.11	.08
	JP	-.01	.03	-.04	-.03	.02	.03

Note. Continued on the following page. N = 583 for the HBI-I and HBI-N; N = 134 for the MBTI; N = 128 for the SCII.

Table 9 continued

A Multitrait-Multimethod Validity Matrix Using Scales from the HBI-I, HBI-N, SCII, and MBTI

		HBI-N			HBI-N		
		T	F	A	TA	FT	AF
HBI-N	T	.58					
	F	.11	.69				
	A	.25	.15	.72			
HBI-N	TA	-.56	.05	.67	.71		
	FT	.55	-.77	.03	-.40	.69	
	AF	-.08	.72	-.58	-.43	-.65	.72
SCII	R	-.02	-.25	.04	.05	.20	-.24
	I	.20	-.13	-.01	-.14	.24	-.11
	A	.24	.08	.11	-.08	.09	-.01
	S	.06	-.01	.02	-.03	.05	-.02
	E	.13	.04	.17	.04	.05	-.08
	C	.01	.01	.16	.12	.00	-.11
MBTI	SN	.22	.01	-.04	-.19	.14	.04
	TF	-.25	.25	-.05	.14	-.38	.24
	EI	-.05	-.07	-.31	-.21	.03	.16
	JP	.02	-.01	.00	-.01	.02	-.01

Note. N = 583 for the HBI-I and HBI-N; N = 134 for the MBTI; N = 128 for the SCII.

Table 9 continued

A Multitrait-Multimethod Validity Matrix Using Scales from the HBI-I, HBI-N, SCII, and MBTI

		SCII						MBTI			
		R	I	A	S	E	C	SN	TF	EI	JP
SCII	R	-									
	I	.58	-								
	A	.15	.44	-							
	S	.27	.39	.41	-						
	E	.24	.33	.33	.41	-					
	C	.24	.31	.23	.37	.64	-				
MBTI	SN	.11	.26	.43	.16	.06	-.07	-			
	TF	-.35	-.27	.13	.07	-.05	-.13	-.05	-		
	EI	.08	.03	-.04	-.10	-.27	-.04	-.08	-.12	-	
	JP	.08	.03	.11	-.10	.01	-.16	.31	.18	-.07	-

Note. N = 134 for the MBTI; N = 128 for the SCII.

moderate correlations would fall in the ranges from +.20 to +.69 and -.20 to -.69; and low correlations would fall in the range from +.19 to -.19.

A total of 243 correlation coefficients were computed and placed into the matrix. Of these, 189 were in the direction hypothesized. Ninety-one correlations were of the size and direction hypothesized. An extensive summary of the correlations between the scores of the various scales of the inventories can be reviewed in Appendix I.

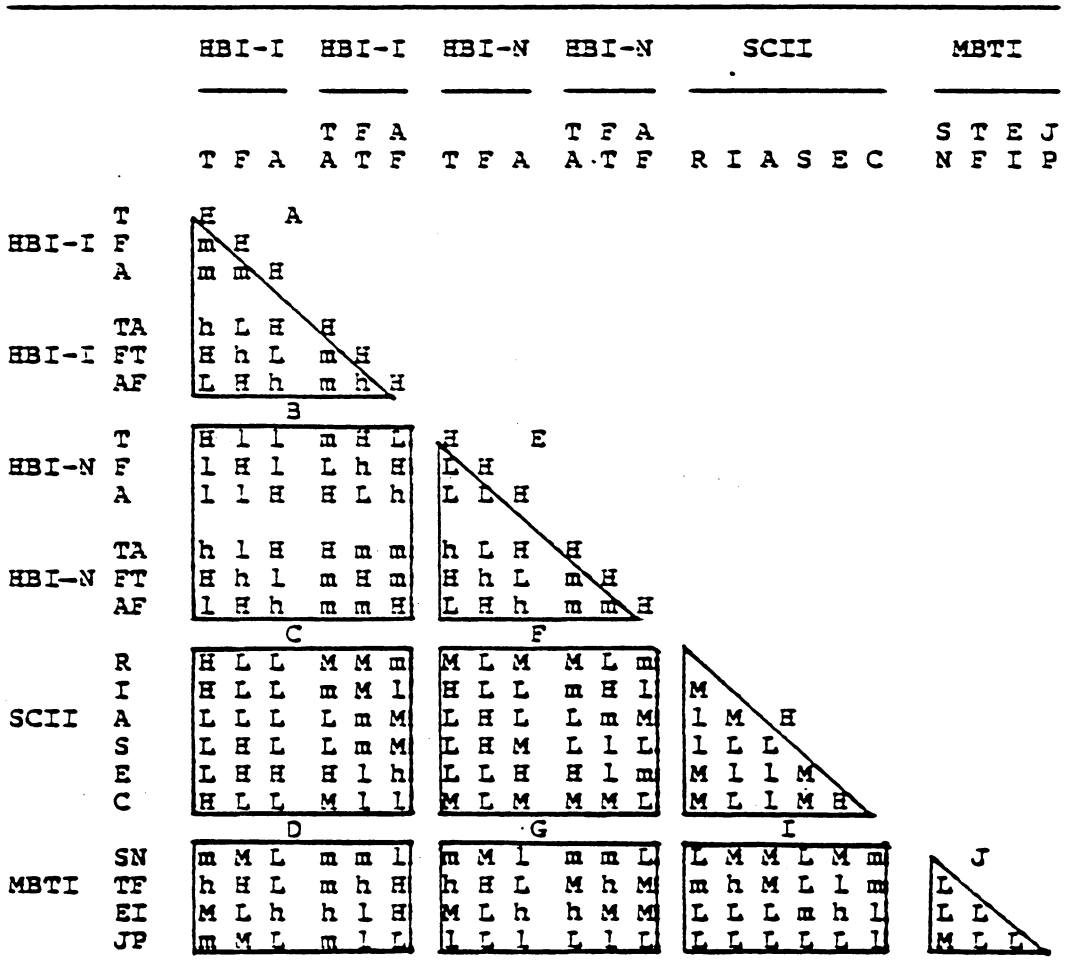
Reliability Coefficients on the Diagonal. Examining the diagonal of Table 9, we note that nine of the 12 expected high correlations for 7-day test-retest reliability were obtained for the scores from the various scales of the HBI-I and HBI-N. These 7-day test-retest correlations were discussed previously. All of the HBI-I choice and bipolar score coefficients were high and three of the HBI-N intensity and bipolar score coefficients were high. The other three coefficients, one HBI-N bipolar score coefficient and two HBI-N intensity score coefficients, were moderate. The 7-day test-retest reliability coefficients for the scores from the SCII and MBTI were not calculated in this study and do not appear on the diagonal of the matrix.

Convergent and Discriminant Validity Evidenced Among Scores from Scales of the HBI-I, HBI-N, SCII, and MBTI. In order to understand the discussion that follows concerning the many convergent and discriminant validity coefficients presented in the multitrait-multimethod validity matrix, the reader will need to refer to Table 9 and Table 10 repeatedly. Table 10 is a reproduction of Table 4, which presented the size and direction predicted for each of the correlations among the various pairs of scores derived from the scales of the HBI-I, HBI-N, SCII, and MBTI. In Table 10, boxes of different shapes have been drawn around certain sets of the hypothesized correlations. These boxes have been labeled (A-J) so that they can be referred to in the text, and so that the reader can find quickly the hypothesized sizes of correlations being discussed. The sets of predicted correlations in Table 10 will be referred to, box by box, in the text; the reader will need to compare each set of predicted correlations to the corresponding actual correlations contained in table 9.

Correlations among the scores derived from the various scales of the HBI-I. Each predicted size for correlations among scores derived from the various scales of the HBI-I appear in box A of Table 10. Comparing these to the actual correlations in Table 9 reveals that five of the seven high

Table 10

Hypothesized Relationships for Establishing Convergent and Discriminant Validity Between Scales of the HBI-I, HBI-N, SCII, and MBTI (Boxes Labeled for Interpretive Purposes)



Note. Correlations: H = high positive (.70 +); M = moderate positive (.20 to .69); L = low positive (.00 to .19); h = high negative (-.70 +); m = moderate negative (-.20 to -.69); l = low negative (.00 to -.19).

correlations expected for convergent validity were obtained. The other two correlations were very high moderate; .68 and -.69. One of three low correlations expected for discriminant validity occurred.

Correlations among scores of the HBI-I and HBI-N. Each predicted size for correlations between scores of the HBI-I and HBI-N appear in box B of Table 10. Again comparing predicted and actual correlations reveals that two of the 16 hypothesized high correlations expected for convergent validity among scores of the HBI-I and HBI-N were obtained. Ten more of the correlations were in the high to very high moderate range (.50 to .69) and in the direction hypothesized. Only four of the correlations fell below the .50 level. Six of the 12 low correlations anticipated for discriminant validity among scores of the HBI-I and HBI-N were obtained. Four more were in the .20 to .28 very low moderate range. These results suggest that the scores of the HBI-I and HBI-N that were hypothesized to be measuring the same constructs are indeed doing so; also, that scores of the two instruments hypothesized to be measuring different constructs are doing so.

Correlations among scores of the HBI-I and SCII. Box C of Table 10 contains each predicted size for correlations between scores of the HBI-I and SCII. It was noted earlier that certain scales of the HBI and SCII have similar verbal

descriptions, while other scales have dissimilar verbal descriptions. For instance, the HBI T and SCII I scales have similar verbal descriptions, as do the HBI A and SCII E scales. The reader can observe the apparent similarities and differences between scales by examining the verbal descriptions of the HBI and SCII scales in Appendices A and B, respectively. Furthermore, Appendix D presents some examples of how the similarities between scales were established.

Comparing predicted and actual correlations between scores of the HBI-I and SCII revealed a surprising finding; none of the eight high correlations predicted for convergent validity among scores derived from scales with similar verbal descriptions occurred. For example, the HBI-I T and SCII I correlation was low; the HBI-I AF and SCII E correlation was low; and the HBI-I F and SCII S correlation was low. In fact, all eight correlations were low. Thus, in spite of the similarity of verbal descriptions between certain HBI and SCII scales, the scales appear to be producing scores that measure different constructs. Every one of the 18 correlations manifesting discriminant validity among scores of the HBI-I and SCII were low as hypothesized.

Correlations among scores of the HBI-I and MBTI. Box D of Table 10 contains each predicted size for correlations between scores of the HBI-I and MBTI. It was noted earlier

that certain scales of the HBI and MBTI also have similar verbal descriptions, while other scales have dissimilar verbal descriptions. The reader can observe these apparent similarities and differences by examining the verbal descriptions of the HBI and MBTI scales in Appendices A and B, respectively. Furthermore, Appendix D presents some examples of how the similarities between certain scales were established.

Comparing predicted and actual correlations reveals that similarity of verbal descriptions among scales of the HBI and MBTI did not prove to be a good predictor of strong relationships between scores of those scales. None of the seven high correlations expected for convergent validity among scores of the HBI-I and MBTI were obtained. However, some moderate correlations did materialize in the hypothesized directions (eg., HBI-I F and MBTI TF, r = .38; HBI-I FT and MBTI TF, r = -.35; HBI-I TA and MBTI EI, -.30). Thus, the personality preferences for behaving in certain ways as measured by the MBTI appear to differ from behavior as measured by the HBI. Seven of the eight low correlations predicted for discriminant validity among scores of the HBI-I and MBTI occurred. The one correlation that was not low was very low moderate (-.21). Thus, the evidence provided by these correlations suggests that while certain

scales of the HBI and the MBTI have similar verbal descriptions they do not measure the same constructs.

Correlations among scores derived from the various scales of the HBI-N. Box E of Table 10 contains each predicted size for correlations between scores of the HBI-I. Comparing predicted and actual correlations reveals that two of the six high correlations expected for convergent validity among scores of the HBI-N were obtained. The other four correlations were high moderate (between .55 and .67, ignoring signs). Five of the six low correlations expected for discriminant validity among scores of the HBI-N occurred. These correlations provide support for the results of the internal consistency analysis of the HBI-N scales, which indicated that the HBI-N scales measure different constructs.

Correlations among scores of the HBI-N and SCII. Box F of Table 10 contains each predicted size for correlations between scores of the HBI-N and SCII. Comparing predicted and actual correlations reveals that none of the high correlations expected for convergent validity among scores of the HBI-N and SCII occurred. Indeed, all but two of the correlations were low. Fourteen of the 17 low correlations predicted for discriminant validity among scores of the HBI-N and SCII occurred. These results, derived using HBI-N (normative) scores, are quite similar to the results found

when the HBI-I (ipsative) scores were correlated with the SCII scores. This provides further evidence that certain HBI and SCII scales may possess similar verbal descriptions yet measure different constructs.

Correlations among scores of the HBI-N and MBTI. Box G of Table 10 contains each predicted size for correlations between scores of the HBI-N and MBTI. Comparing predicted and actual correlations reveals that none of the five high correlations anticipated for convergent validity among scores of the HBI-N and MBTI were obtained. However, all of the correlations were moderate and in the predicted direction. All ten of the low correlations expected for discriminant validity occurred. These results, obtained using HBI-N normative scores, agree with the results produced when the HBI-I ipsative scores were correlated with the scores of the MBTI. The HBI-N/MBTI correlation coefficients provide further evidence to support the notion that scores of certain scales of the HBI and MBTI measure different constructs even though the verbal descriptions of the scales are similar.

Correlations among scores of the SCII, MBTI, and between scores of the SCII and MBTI. Appendix J contains a discussion of the results of the correlations among scores of the SCII, as well as a discussion of the results of correlations among scores of the MBTI. Also included is a

discussion of the results of correlations of SCII scores with MBTI scores. These correlations are presented in Appendix J in order to demonstrate that the various correlations among SCII and MBTI scores obtained in this study compare very favorably to prior correlations identified between these scores.

The Latent Dimensions Underlying Scores from Scales of the HBI-N, SCII, and MBTI

The sixth research question concerned the kind of latent dimensions that would be revealed by a factor analysis of the HBI-N, SCII, and MBTI scale scores. Two different common factor analyses, both using the Harris-Kaiser method of oblique rotation, were conducted. The first analysis was carried out while controlling for N = 3 factors. In the second analysis, for comparison purposes, the number of factors was not restricted.

Eigenvalues

All three factors produced by the first factor analysis had eigenvalues greater than 1.0; but only three of the four factors produced by the second analysis met this criterion. The eigenvalues produced by both procedures appear in Table 11.

Scale Item-factor Loadings

The criterion recommended by Hair, Anderson, Tatham, and Grablowsky (1979) for interpreting the item-factor

Table 11

Eigenvalues for Factors Produced Using Scores from the HBI-N, SCII, and MBTI

Factor	Eigenvalue
N = 3 factors	
1	2.052
2	1.756
3	1.618
Not restricting the number of factors	
1	2.062
2	1.886
3	1.388
4	0.879

Note. N = 583 for the HBI-N scores; N = 134 for the MBTI scores; N = 128 for the SCII scores.

loadings was applied to the results of these analyses also. Thus, a loading needed to be .30 or greater to be considered significant. For both analyses, a review of Table 12 reveals that many of the various scores (variables) derived from the scales of the different instruments loaded significantly on two or even three factors.

Both analyses produced three factors with eigenvalues greater than 1.0; however, the second analysis also produced an additional fourth factor with an eigenvalue of .88. No clear interpretation of the factors was apparent. Since the correlations between the HBI scores and the scores of the SCII and MBTI were generally quite low, it is not surprising that no clearly discernable factors emerged to account for the variance in the variables being analyzed. The evidence seems to suggest that the various instruments are measuring basically different constructs.

Inter-factor Correlations

The first factor analysis, which was restricted to three factors, produced inter-factor correlations that ranged from .21 to .38. The second analysis, which did not restrict the factors, produced inter-factor correlations that ranged from .40 to -.26. Table 13 presents the inter-factor correlations produced by both of these analyses. It should be remembered, while examining the correlations for the second analysis, that the eigenvalue

Table 12

HBI-N, SCII, and MBTI Scale Item-factor Loadings

Inventory	Score	Factor 1	Factor 2	Factor 3	Factor 4
		N = 3 factors			
HBI-N					
	T	0.14	0.36	0.07	
	F	0.06	0.11	-0.35	
	A	0.26	0.17	-0.14	
MBTI					
	EI	-0.23	-0.16	0.23	
	SN	0.03	0.60	0.14	
	TF	-0.07	0.05	-0.53	
	JP	-0.13	0.31	-0.05	
SCII					
	R	0.32	0.25	0.69	
	I	0.45	0.52	0.65	
	A	0.42	0.71	0.15	
	S	0.54	0.38	0.27	
	E	0.77	0.33	0.20	
	C	0.72	0.13	0.28	
	Not restricting the number of factors				
HBI-N					
	T	0.09	0.29	-0.18	0.46
	F	0.02	0.03	0.34	0.29
	A	0.20	0.08	0.02	0.51
MBTI					
	EI	-0.18	-0.08	-0.18	-0.38
	SN	0.01	0.58	-0.07	0.14
	TF	-0.08	-0.01	0.65	-0.01
	JP	-0.14	0.29	0.10	0.05
SCII					
	R	0.37	0.37	-0.62	-0.22
	I	0.49	0.62	-0.55	-0.07
	A	0.41	0.70	-0.02	0.21
	S	0.56	0.44	-0.12	-0.04
	E	0.75	0.35	-0.14	0.24
	C	0.73	0.18	-0.22	0.05

Note. N = 583 for the HBI-N scores; N = 134 for the MBTI scores; N = 128 for the SCII scores.

Table 13

Inter-factor Correlations Between Factors Produced Using Scores from the HBI-N, SCII, and MBTI

Holding for N = 3 factors

	Factor 1	Factor 2	Factor 3
Factor 1	1.00		
Factor 2	0.38	1.00	
Factor 3	0.29	0.21	1.00

Not holding the factors

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1.00			
Factor 2	0.40	1.00		
Factor 3	-0.26	-0.23	1.00	
Factor 4	0.14	0.15	0.17	1.00

Note. These inter-factor correlations were produced using a common factor analysis procedure with the Harris-Kaiser oblique rotation method. Squared multiple correlations were placed on the diagonal of the correlation matrix. N = 583 for the HBI-N scores; N = 134 for the MBTI scores; N = 128 for the SCII scores.

for the third factor produced by that analysis was less than 1.0 (0.88 to be exact).

Summary

This chapter presented basic demographic data concerning subjects who participated in this study. Also presented were the results pertinent to the six research questions investigated.

The results of research questions one and two concerned the 7-day, 14-day, and 28-day test-retest reliability of the scores of the HBI-I and HBI-N. The reliability coefficients for the HBI-I choice and bipolar scores were in the .70 to .86 range, remaining relatively stable over the three time intervals. The reliability coefficients for the HBI-I average intensity scores were in the .68 to .72 range initially but were much lower for the 28-day interval. The reliability coefficients for the HBI-N scores did not attain as high an initial level of reliability as did the HBI-I scores; nor did they maintain their stability. They ranged from .58 to .72 for the 7-day test-retest, with some coefficients as low as .36 and .38 for the 14-day interval. It was noted that, for reasons unknown, the HBI-N retest scores differed between those subjects taking the retest first as opposed to last. This fact confounded the interpretation of the HBI-N test-retest reliability coefficients to some extent.

The third research question concerned the internal consistency of the HBI-N test scores that were gathered during the first test session. The coefficient alphas computed for the HBI-N intensity scores were quite high considering that each set of scores was derived from scales consisting of only five items. The alphas ranged from .59 to .77. The analysis of the internal consistency of the HBI-N scores also provided some evidence concerning the construct-related validity of the HBI scales. The patterns of item-scale correlations provided evidence of good internal structure among the items measuring their respective constructs. The intercorrelations among the scales of the HBI-N provided external structure evidence suggesting that the scales measure different constructs.

The fourth research question concerned the construct-related validity of the HBI. Two different factor analyses were conducted on the items of the HBI-N. The analyses produced almost identical item-factor loading patterns that provided strong evidence that the HBI, as hypothesized, measures three different dimensions of behavior. Of the fifteen words that make up the HBI-N, only the thinking word "curious" failed to load on to a factor. The inter-factor correlations provided additional evidence that the factors measured different constructs.

The fifth research question again concerned the construct-related validity of the HBI. The development and analysis of a multitrait-multimethod construct validity matrix was undertaken in order to help determine the relationships among the scores of the HBI-I, HBI-N, SCII, and MBTI. The analysis revealed good convergent and discriminant validity among scores of the HBI-I, and also among scores of the HBI-N. Good convergent and discriminant validity were again obtained when the scores of the HBI-I and HBI-N were correlated with one another. The convergent validity for HBI-I and HBI-N scores correlated with SCII and MBTI scores never occurred, which provided evidence that the HBI scores are measuring constructs other than those measured by the SCII and MBTI scores. The correlations of the HBI-I ipsative and HBI-N normative scores with scores of the SCII and MBTI were generally very similar in nature.

The last research question also concerned the construct-related validity of the HBI. Two separate factor analyses of the HBI were conducted using scores from the HBI-N, SCII, and MBTI inventories. Each analysis produced three factors that had eigenvalues greater than 1.0. Interpreting the factors was difficult because many items loaded significantly on to two or three factors. The totality of the factorial validity evidence supported the

notion that the HBI scores measure constructs that differ from the constructs measured by the scores of the SCII and MBTI.

CHAPTER FIVE

Summary, Conclusions, Implications, and Recommendations

Summary of the Results

The purpose of this study was to investigate and compare the measurement properties of two forms of the Hutchins Behavior Inventory (HBI). There were two major objectives related to this purpose: the first was to investigate and compare the reliability of scores produced by the HBI-I (ipsative) and HBI-N (normative) forms of the HBI; the second was to investigate and compare the validity of scores produced by these two instruments.

Six research questions were posed that related to these objectives. Research questions one and two dealt with the test-retest reliability of scores derived from scales of the HBI-I and HBI-N. Research question three related to the internal consistency of scores derived from the HBI-N. This question was unique in that it pertained to both the reliability and construct-related validity of the HBI-N. Research questions four, five, and six concerned the construct-related validity of the HBI-I and HBI-N. Questions four and six dealt with the factorial validity of the HBI-N. Question five pertained to the relationships among scores of the HBI-I, HBI-N, SCII, and MBTI.

Reliability and Validity of the HBI-I

Reliability. Only test-retest reliability was investigated for the HBI-I. The stability coefficients for the HBI-I choice and bipolar scores were high enough to warrant the use of these scores for research purposes (range = a high of .86 for the 7-day retest to a low of .70 for the 28-day retest). The instability of HBI-I average intensity scores requires that caution be exercised when interpreting these scores since their cofficients were much lower (range = a high of .77 for the 7-day retest to a low of .57 for the 28-day retest). It is not surprising to find that the average intensity scores were the least stable of the HBI-I scores. The HBI was designed to measure situationally specific behavior. The subjects' ability to recall the exact intensity of their behavior during the situation might be expected to change with the passage of time. Further, the subjects were focusing on one occasion of a situation that was ongoing for them throughout the study. The situation they focused on was the very first class they attended during the spring semester. But they attended the class many times after that first occassion. The intensity of their behavior in the class might have actually changed over time, and then the instability of the average intensity scores would have been a reflection of the actual changes.

Validity. The content-related validity of the HBI-I was developed by Walker (1984) in her study of the ipsative form of the HBI. Walker carried out procedures necessary to ensure that good content-related validity was obtained.

In this study a multitrait-multimethod validity matrix was developed to investigate the construct-related validity of the HBI-I scores. The high convergent and low discriminant validity coefficients expected between the various scores of the HBI-I, and between the scores of the HBI-I and HBI-N, generally occurred as hypothesized. These coefficients provided evidence that the HBI-I scores are measuring the thinking, feeling, and acting dimensions of behavior as hypothesized by Hutchins.

The fact that high convergent validity coefficients did not occur between the scores of the HBI-I and SCII, nor between the scores of the HBI-I and MBTI, provides evidence that the HBI-I scores measure constructs other than those measured by the scores of the SCII and MBTI. The lack of convergent validity was surprising since the verbal descriptions of certain scales of the HBI were similar to the verbal descriptions of certain SCII and MBTI scales.

The large number of discriminant coefficients that occurred as hypothesized between scores of the HBI-I and scores of these other instruments also provided evidence supporting the inference that the scales of the HBI measure

constructs other than those measured by the scales of the SCII and MBTI.

As mentioned earlier, these correlation coefficients involving HBI-I scores should be interpreted cautiously. They should only be interpreted within the context of all other evidence regarding the validity of HBI-I ipsative scores.

Reliability and Validity of the HBI-N

Reliability. Two different kinds of reliability coefficients were calculated for the scores of the HBI-N. First, test-retest reliability coefficients were calculated for 7, 14, and 28 day time intervals. The coefficients ranged from a high of .72 (7-day retest) to a low of .36 (14-day retest). Two factors should be mentioned that may have affected the stability of these test-retest coefficients. One factor was that while initial test scores did not differ for subjects who took the HBI-N first as opposed to second, retest scores did differ for subjects who took the HBI-N first versus second. This difference in retest scores confounded the interpretation of the HBI-N test-retest reliability coefficients to some extent. The second factor was that the HBI-N is a very short 15 item instrument. Nunnally (1978) noted that an instrument's reliability usually increases as the number of items in the instrument increases; thus, the overall low test-retest

reliability of the HBI-N scores may also be partially due to the brevity of the test.

The second kind of reliability coefficients calculated had to do with the internal consistency of the HBI-N scores. Coefficient alphas produced for the HBI-N T, F, and A scores ranged from .59 to .77. These coefficients were very respectable considering that each scale of the HBI-N consisted of just five items.

Obtaining relatively strong internal consistency reliability coefficients while failing to obtain good test-retest reliability coefficients is not a paradox since the different kinds of coefficients are affected by different sources of measurement error. The internal consistency coefficients show the extent to which the items of an instrument correlate with one another and measure the same construct(s) (Nunnally, 1978). Thus, internal consistency coefficients are primarily affected by the nature of the sample of items placed into the test. The internal consistency coefficients are not affected by changes in the individual, nor by changes in the administration of the test, nor by changes in the test environment from one test administration to another, as can often be the case with test-retest coefficients. So both types of reliability coefficients provide important, but different information about the scores of a test.

Internal consistency is the more important estimate of reliability when using the domain sampling model because it sets an upper limit for the reliability of a test's scores. If the scores of a test have relatively high estimates of internal consistency, then further efforts to develop the test may prove to be worthwhile (Nunnally, 1978).

Validity. The content-related validity of the HBI-N was not investigated in this study because the HBI-N was derived directly from the HBI-I. Only the final step in the process of developing a test, in which the items are placed into a testable format (Carmines & Zeller, 1979), was altered in order to create the HBI-N items from the already existing HBI-I. Therefore, the HBI-N was assumed to possess the same degree of content-related validity possessed by the HBI-I.

The construct-related validity of the HBI-N was investigated by carrying out four different analyses. The first analysis was the study of the internal consistency of the HBI-N already alluded to under the discussion of the reliability of HBI-N scores. Two aspects of the internal consistency analysis were of critical importance to the construct-related validity of the HBI-N. First, the HBI-N item-scale correlations provided strong internal structure evidence for the inference that the scales of the HBI-N each measure a different dimension of behavior. Secondly, the

HBI-N scale intercorrelations provided equally strong external structure evidence for inferring that the scales of the HBI-N measure different constructs of behavior.

The second analysis conducted to gather evidence pertaining to the construct-related validity of the HBI-N consisted of two factor analyses of the items of the HBI-N. One was a common factor analysis using the Harris-Kaiser method of oblique rotation with squared multiple correlations (SMCs) on the diagonal of the correlation matrix. The other was a principal components analysis using the Varimax method of orthogonal rotation. The results of the two factor analyses were virtually identical and strongly supported the hypothesis that the HBI-N scales measure three different constructs. In both analyses, all but one of the fifteen HBI-N items (the one being the thinking word "curious") loaded significantly on their hypothesized factors, creating three distinct factors with eigenvalues greater than 1.0. Furthermore, the common factor analysis with the oblique rotation of the initial factors produced inter-factor correlations that were low, adding to the evidence that supports the inference that the scales of the HBI-N measure different dimensions of behavior.

The third analysis conducted to investigate the construct-related validity of the HBI-N scores was an

analysis of a multitrait-multimethod validity matrix containing coefficients for correlations between scores from the HBI-I, HBI-N, SCII, and MBTI. Convergent and discriminant validity coefficients for correlations between the various scores of the HBI-N were supportive of the notion that the three HBI-N scales are measuring different constructs. The lack of any high convergent validity coefficients between the scores of the HBI-N and the scores of the SCII and MBTI suggested that the HBI-N scores are measuring constructs other than those measured by the SCII and MBTI. This finding was surprising since the verbal descriptions of certain scales of these instruments appeared to be quite similar.

The HBI-N behavior scores correlated higher with the MBTI scores than with the SCII scores. This might be due to the possibility that the MBTI personality type scores are more strongly related to preferences for certain kinds of behavior than are the SCII personality type scores. If so, then the HBI behavior scores would be expected to correlate more highly with the MBTI personality type scores that indicate a preference for certain kinds of behavior.

The results of the correlations of HBI-N scores with HBI-I scores have already been discussed. It only remains to be mentioned that the correlations obtained were strong evidence in support of the inferences that the scores of the

HBI-N and HBI-I that are supposed to measure the same dimensions of behavior do so; and that scores of the two instruments that are supposed to measure different dimensions of behavior do so as well.

The final analysis conducted to investigate the construct-related validity of the HBI-N was a factorial validity analysis of the scores of the HBI-N. The HBI-N scores, along with the scores of the SCII and MBTI, were subjected to two different common factor analyses. The first analysis was restricted to N = 3 factors; the second analysis did not restrict the factors. The analyses produced very similar results. Both produced three factors with eigenvalues greater than 1.0 (a fourth factor in the second analysis had an eigenvalue of .88). Both had numerous variables which loaded significantly (.30 or greater) on to two or three different factors. Both analyses also produced relatively low inter-factor correlations (range = .38 to .21 for the first analysis; range = .40 to -.26 for the second analysis). The three factors with eigenvalues greater than 1.0 produced by the second analysis corresponded very closely to the three factors produced by the first analysis.

Interpretions of the factors produced by the two analyses were difficult. The overall evidence seemed to best support the notion that the scores of the HBI-N are

measuring constructs that are different from those measured by the SCII and MBTI.

Conclusions

Reliability

1. The test-retest reliability of HBI-I choice and bipolar scores indicates that they can be used for research, as well as clinical, purposes.

2. The test-retest reliability of HBI-I average intensity scores is respectable for the 7-day interval, but becomes increasingly unstable for longer intervals of time. Therefore, these scores are not recommended for use of any kind at this time.

3. The internal consistency reliability of the HBI-N scores is encouragingly high and suggests that the instrument may hold potential for future development. However, the test-retest reliability of the HBI-N scores is so unstable that HBI-N scores should not be used for research, or any other purpose, at this time.

Validity

4. The content-related validity of the HBI-I was established by Walker (1985) and supports the inference that both the HBI-I and HBI-N measure the thinking, feeling, and acting dimensions of behavior hypothesized by Hutchins.

5. The evidence gathered in this study of the construct-related validity of the HBI-N and, through logical

deduction and generalization, the HBI-I, suggests that these instruments measure the thinking, feeling, and acting dimensions of behavior as hypothesized by Hutchins.

6. The multitrait-multimethod validity matrix provided evidence that supports the inference that the HBI measures constructs that are different from the constructs measured by the SCII and MBTI.

Implications

1. On the basis of all the evidence gathered in this study it appears that the ipsative form of the HBI, the HBI-I, is superior to the normative form, the HBI-N. This means that users should be free to administer the HBI-I and interpret its scores in spite of the inherent psychometric limitations that those ipsative scores possess. The condition, established by experts such as Clemans (1966) and Hicks (1970), that the validity of ipsative scores should be shown to equal or exceed the validity of scores produced by a companion normative form, has been satisfied. If we grant, in the case of the HBI, that the ipsative and normative instruments have equal validity, then it is the superior reliability of the HBI-I that makes it the preferred instrument.

2. The correlations of various HBI-I scores with one another were quite different from the correlations of various HBI-N scores with one another. However, the

correlations of the HBI-I scores with the scores of the SCII and MBTI were very similar in nature to the correlations derived between HBI-N scores and scores of the SCII and MBTI. Therefore, if future users of the HBI-I decide to subject the HBI-I ipsative scores to analyses involving correlation techniques, caution should be exercised when interpreting the results.

3. Some experts recommend that ipsative scores should only be used to compare traits within an individual. This is because ipsative scores, by definition, measure intra-individual differences, not inter-individual differences (Clemans, 1966; Guilford, 1952). Other experts claim that the actual results obtained when comparing inter-individual traits using ipsative scores are often very similar to the results obtained using normative scores. Because no HBI related evidence is available that actually bears directly upon this issue, users of the HBI-I scores need to be cautious when using the scores to make inter-individual comparisons of behavior patterns.

4. Users of the HBI should inform test takers that ipsative scores are, as Clemans (1966) explains, "relative scores" (p. 52). This means that a person obtaining a low ipsative score on a particular behavior scale may possibly possess a greater degree of the behavior than someone who obtains a much higher score on the behavior.

5. Because HBI-I scores are situationally specific, users should be careful to interpret the scores within the context of the specific situation to which they relate.

Recommendations for Further Research

1. Further efforts should be made to determine the relationship of HBI-I scores to constructs that are similar, and dissimilar, to the thinking, feeling, and acting behavior constructs measured by the HBI-I.

2. Future efforts to investigate the validity of the HBI-I scores should focus on approaches such as: the measurement of group differences; changes in performance; and investigations of the test taking process itself. Use of correlational and internal consistency investigations should be carried out with caution because of the ipsative nature of the HBI-I scores.

3. Efforts should be continued toward the development of a reliable single stimulus normative form of the HBI. Scores from such a form would permit inter-individual comparisons of thinking, feeling, and acting behaviors that would have a sound psychometric basis for interpretation.

4. Efforts should be made to determine if the current form of the HBI-I can be used with populations other than the college and university populations studied thus far.

5. Efforts should be made to identify various HBI-I TA, AF, and FT behavior patterns. Research into HBI-I behavior profile analysis should also be undertaken.

6. Research into the way that HBI behavior patterns may affect the counselor/client and other interpersonal relationships should be undertaken.

REFERENCES

Allen, M. J., & Yen, W. M. (1979). Introduction to measurement theory. Monterey, CA: Brooks/Cole.

Anastasi, A. (1976). Psychological testing (4th ed.). New York: Macmillan.

Auerbach, A., & Johnson, M. (1977). Research on the therapists' level of experience. In A. Gurman & A. Razin (Eds.), Effective psychotherapy (pp. 84-102). Elmsford, NY: Pergamon Press.

Barclay, J. R. (1983). Moving toward a technology of prevention: A model and some tentative findings. School Psychology Review, 12 (3), 228-239.

Barclay, J. R. (1984). Primary prevention and assessment. The Personnel and Guidance Journal, 62, 475-478.

Bartlett, C. J., Quay, L. C., & Wrightsman, L. S., Jr. (1960). A comparison of two methods of attitude measurement: Likert-type and forced-choice. Educational and Psychological Measurement, 20, 699-704.

Blalock, H. M. (1970). Estimating measurement error using multiple indicators and several points in time. American Sociological Review, 35, 101-111.

Brammer, L. M. (1969). Eclecticism revisited. Personnel and Guidance Journal, 48, 192-197.

Brammer, L. M., & Shostrom, E. L. (1982). Therapeutic psychology: Fundamentals of counseling and psychotherapy (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Broverman, D. M. (1961). Effects of score transformations in Q and R factor analysis techniques. Psychological Review, 68 (1), 68-80.

Broverman, D. M. (1962). Normative and ipsative measurement in psychology. Psychological Review, 69 (4), 295-305.

Bruce, P. (1984). Continuum of counseling goals: A framework for differentiating counseling strategies. The Personnel and Guidance Journal, 62, 259-263.

Campbell, D. P. (1974). Manual for the Strong Vocational Interest Blank. Stanford, CA: Stanford University Press.

Campbell, D. P., & Hansen, J. C. (1981). Manual for the SVIB-SCII (3rd ed.). Stanford, CA: Stanford University Press.

Carlyn, M. (1977). An assessment of the Myers-Briggs Type Indicator. Journal of Personality Assessment, 41, 461-473.

Carmines, E. G. & Zeller, R. A. (1979). Reliability and validity assessment. Beverly Hills, CA: Sage Publications.

Cattell, R. B. (1944). Psychological measurement: Normative, ipsative, interactive. Psychological Review, 51, 292-303.

Christensen, L. B. (1985). Experimental methodology (3rd ed.). Boston: Allyn and Bacon.

Clemans, W. V. (1966). An analytical and empirical examination of some properties of ipsative measures. Psychometric Monograph, no. 14.

Corey, G. (1986). Theory and practice of counseling and psychotherapy (3rd ed.). Monterey, CA: Brooks/Cole.

Cormier, W. H., & Cormier, L. S. (1979). Interviewing strategies for helpers: A guide to assessment, treatment, and evaluation. Monterey, CA: Brooks/Cole.

Cronbach, L. J. (1946). Response sets and test validity. Educational and Psychological Measurement, 6, 475-494.

Cronbach, L. J. (1950). Further evidence on response sets and test design. Educational and Psychological Measurement, 10, 3-31.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure tests. Psychometrika, 16, 297-334.

Cronbach, L. J., & Meehl, P. (1955). Construct validity in psychological tests. Psychological Bulletin, 52 (4), 281-302.

Cronbach, L. J., & Snow, R. E. (1977). Aptitude and instructional methods. New York: Irvington Publishers.

Davis, T. M. & Chissom, B. (1981). Factor analysis (R-technique) of ipsative data may be misleading. Psychological Reports, 49 (2), 643-647.

Diagnostic and statistical manual of mental disorders (3rd ed.). (1980). Washington, D.C.: American Psychiatric Association.

Downing, L. N. (1975). Counseling theories and techniques: Summarized and critiqued. Chicago: Nelson-Hall.

Egan, G. (1975). The skilled helper: A model for systematic helping and interpersonal relating. Monterey, CA: Brooks/Cole.

Egan, G. (1982). The skilled helper (2nd ed.). Monterey, CA: Brooks/Cole.

Egan, G. (1986). The skilled helper (3rd ed.). Monterey, CA: Brooks/Cole.

Ellis, A. (1982). Major systems. The Personnel and Guidance Journal, 61, 6-7.

Eysenck, H. J., Wakefield, J. A., & Friedman, A. F. (1983). Diagnosis and clinical assessment: The DSM-III. Annual Review of Psychology, 34, 167-193.

Fiedler, F. E. (1950a). Factor analyses of psychoanalytic, nondirective, and Adlerian therapeutic relationships. American Psychologist, 5, 324-325.

Fiedler, F. E. (1950b). The concept of an ideal therapeutic relationship. Journal of Consulting Psychology, 14, 239-245.

Frey, D. H., & Raming, H. E. (1979). A taxonomy of counseling goals and methods. Personnel and Guidance Journal, 58, 26-33.

Garfield, S. L. (1980). Psychotherapy: An eclectic approach. New York: Wiley.

Garfield, S. L., & Kurtz, R. (1976). Clinical psychologists in the 1970s. American Psychologist, 31, 1-9.

Gleser, L. J. (1972). On bounds for the average correlation between subtest scores in ipsatively scored tests. Educational and Psychological Measurement, 32, 759-765.

Gottfredson, G. D., Holland, J. L., & Ogawa, D. K. (1982). Dictionary of Holland occupational codes. Palo Alto, CA: Consulting Psychologists Press.

Guilford, J. P. (1952). When not to factor analyze. Psychological Bulletin, 49, 26-37.

Guilford, J. P. (1954). Psychometric methods (2nd ed.). New York: McGraw-Hill.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Grablowsky, B. J. (1979). Multivariate data analysis. Tulsa, OK: Petroleum Publishing Company.

Heilbrun, A. B., Jr. (1963). Evidence regarding the equivalence of ipsative and normative personality scales. Journal of Consulting Psychology, 27, 152-156.

Helmstadter, G. C. (1964). Principles of psychological measurement. New York: Appleton-Century-Crofts.

Hicks, L. E. (1970). Some properties of ipsative, normative, and forced-choice normative measures. Psychological Bulletin, 74 (3), 167-184.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1979). Applied statistics for the behavioral sciences. Chicago: Rand McNally.

Holland, J. L. (1973). Making vocational choices: A theory of careers. Englewood Cliffs, NJ: Prentice-Hall.

Holland, J. L. (1978). Vocational Preference Inventory: Manual. Palo Alto, CA: Consulting Psychologists Press.

Holland, J. L. (1985). Making vocational choices: A theory of vocational personalities and work environments (2nd ed.). Engelwood Cliffs, NJ: Prentice-Hall.

Hutchins, D. E. (1979). Systematic counseling: The T-F-A model for counselor intervention. The Personnel and Guidance Journal, 57, 529-531.

Hutchins, D. E. (1982). Ranking major counseling strategies with the T-F-A/matrix system. The Personnel and Guidance Journal, 60, 427-431.

Hutchins, D. E. (1984). Improving the counseling relationship. The Personnel and Guidance Journal, 62 (10), 572-575.

Hutchins, D. E., & Cole, C. G. (1986). Helping relationships and strategies. Monterey, CA: Brooks/Cole.

Impara, J. C., & Stoker, H. W. (1985). Determining reliability and validity of licensure examinations. In J. C. Fortune (ed.), Understanding testing in occupational licensing: Establishing links between principles of measurement and practices of licensing (pp. 65-86). San Francisco: Jossey-Bass.

Ivey, A. E. (1980). Counseling 2000: Time to take charge! Journal of Counseling Psychology, 8 (4), 12-15.

Ivey, A. E., & Authier, J. (1978). Microcounseling (2nd ed.). Springfield, MO: Thomas.

Ivey, A. E., & Matthews, W. J. (1984). A meta-model for structuring the clinical interview. Journal of Counseling and Development, 63, 237-243.

Ivey, A. E., & Simek-Downing, L. (1980). Counseling and psychotherapy: Skills, theories, and practice. Englewood Cliffs, NJ: Prentice-Hall.

Karr, C. (1958). A comparison of EPPS scores obtained from the standard forced-choice procedure and a rating-scale procedure. (Doctoral dissertation, University of Washington, 1958). Dissertation Abstracts International, 18, 3382-3383.

Kiess, H. O, & Bloomquist, D. W. (1985). Psychological research methods. Boston: Allyn and Bacon.

Knapp, R. R. (1964). An empirical investigation of the concurrent and observational validity of an ipsative versus a normative measure of six interpersonal values. Educational and Psychological Measurement, 24, 65-73.

Kottler, J. A., & Brown, R. W. (1985). Introduction to therapeutic counseling. Monterey, CA: Brooks/Cole.

L'Abate, L. (1981). Classification of counseling and therapy theorists, methods, processes, and goals: The E-R-A model. The Personnel and Guidance Journal, 59, 263-265.

Landman, J. T., & Dawes, R. M. (1982). Psychotherapy outcome. American Psychologist, 32, (5), 504-516.

Lazarus, A. A. (1984). Multimodal therapy. In R. J. Corsini (ed.), Current Psychotherapies (pp. 491-530). Itasca, IL: Peacock.

Maranell, G. M. (1974). Scaling: A sourcebook for behavioral scientists. Chicago: Aldine.

Merritt, S. L., & Marshall, J. C. (1984). Reliability and construct validity of ipsative and normative forms of the learning style inventory. Educational and Psychological Measurement, 44, 463-472.

Myers, I. B. (1962). The Myers-Briggs Type Indicator: Manual. Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B., & McCaulley, M. H. (1985). Manual: A guide to the development and use of the Myers-Briggs Type Indicator. Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B., & Myers, P. B. (1980). Gifts differing. Palo Alto, CA: Consulting Psychologists Press.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Palmer, J. O. (1980). A primer of eclectic psychotherapy. Monterey, CA: Brooks/Cole.

Patterson, C. H. (1973). Theories of counseling and psychotherapy (2nd ed.). New York: Harper and Row.

Patterson, C. H. (1986). Theories of counseling and psychotherapy (4th ed.). New York: Harper and Row.

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

Prochaska, J. O., & DiClemente, C. C. (1984). The transtheoretical approach. Homewood, IL: Dow Jones-Irwin.

Radcliffe, J. A. (1963). Some properties of ipsative score matrices and their relevance for some current interest tests. Australian Journal of Psychology, 15 (1), 1-11.

Rorer, L. G. (1965). The great response-style myth. Psychological Bulletin, 63 (3), 129-156.

Scott, W. A. (1968). Comparative validities of the forced-choice and single-stimulus tests. Psychological Bulletin, 70, 231-244.

Smith, D. (1982). Trends in counseling and psychotherapy. American Psychologist, 37 (7), 802-809.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.

Standards for educational and psychological testing (1985). Washington, D.C.: Committee to Develop Standards for Educational and Psychological Testing of the American Educational Research Association, The American Psychological Association, and the National Council on Measurement in Education.

Thorne, F. C. (1967). Integrative psychology. Brandon, VT: Clinical Psychology Publishing Company.

Travers, R. M. W. (1951). A critical review of the validity and reliability of the forced-choice technique. Psychological Bulletin, 48, 62-70.

Walker, J. L. (1984). Assessing thinking, feeling and acting components of behavior. Unpublished master's thesis, Virginia Polytechnic and State University, Blacksburg, VA.

Ward, D. E. (1983). The trend toward eclecticism and the development of comprehensive models to guide counseling and psychotherapy. The Personnel and Guidance Journal, 62, 154-157.

Wright, C. E. (1961). A factor dimension comparison of normative and ipsative measurements. Educational and Psychological Measurement, 21, 433-444.

Zeller, R. A., & Carmines, E. G. (1980). Measurement in the social sciences. Cambridge: Cambridge University Press.

Appendix A

Hutchins' Descriptions of the HBI TFA Behavior Dimensions

Dimension	Description
Thinking.	Generally, thinking persons are characterized by intellectual, cognitively oriented behavior. They tend to behave in logical, rational, deliberate, and systematic ways. They are fascinated by the world of concepts, ideas, theories, words, and analytical relationships. The range of behavior in this category runs from minimal thought to considerable depth in quality and quantity of thinking. Organization of thoughts ranges from scattered to highly logical and rational. Hutchins, 1984, p. 573)
Feeling.	Feeling persons generally tend to behave in emotionally expressive ways. They are likely to go with their feelings in making decisions: "if it feels good, do it!" The expression and display of emotions, feelings, and affect provide clues to people with a primary feeling orientation. A person's mood can range from angry, anxious, bitter, hostile, or depressed to one of elation, joy, or enthusiasm. One's emotional energy level can vary from low to high. (Hutchins, 1984, p.573)

Note. These descriptions are quoted from Hutchins' 1984 article on the TFA system.

Appendix A continued

Hutchins' Descriptions of the HBI TFA Behavior Dimensions

Dimension	Description
Acting.	Acting persons are generally characterized by their involvement in doing things and their strong goal orientation. They are frequently involved with others, and tend to plunge into the thick of things. Action types get the job done, one way or another. To them, doing something is better than doing nothing: thus, they are frequently involved in a variety of activities. Their behavior may range from loud, aggressive, and public-oriented to quiet, subtle, and private. (Hutchins, 1984, p. 573)

Note. These descriptions are quoted from Hutchins' 1984 article on the TFA system.

Appendix B

Summarized Descriptions of the Holland Personality Types Measured by the SCII General Occupational Theme Scales

Type	Description
Realistic	Physically strong, robust, rugged, practical; like to build things with tools; uncomfortable in social settings; good motor coordination, weak verbal and interpersonal skills; often athletic; stable, persistent; preferring concrete to abstract problems; aggressive; mechanically inclined; conventional with regard to political and economic goals; natural, direct; rarely creative in the arts or sciences.
Investigative	Strong scientific orientation; task oriented; introspective and asocial; need to understand the physical world; prefer to think through rather than act out problems; enjoy ambiguous tasks; analytical, scholarly, intellectual, curious; independent and reserved; dislike repetitive tasks; unconventional; lack leadership and persuasive skills.
Artistic	prefer unstructured situations; emphasize self-expression; introspective and asocial, nonconforming; value aesthetic qualities; impulsive, original, intuitive; avoid highly structured problems; independent, creative, disorderly.

Note. Continued on the next page.

Appendix B continued

Summarized Descriptions of the Holland Personality Types Measured by the SCII General Occupational Theme Scores

---

Social: Sociable, responsible, humanistic; often religious; like to work in groups; have good verbal and interpersonal skills; avoid intellectual problem-solving; avoid physical exertion; enjoy being central in a group; avoid highly ordered activities; see themselves as understanding, idealistic, and helpful; solve problems through feelings and interpersonal manipulation; enjoy training, informing, curing, developing, enlightening.

Enterprising: Verbal skils suited to selling, dominating and leading; avoid long periods of intellectual effort; aggressive, popular, self-confident, cheerful, and sociable; high energy level; dislike scientific activities; strong drive to attain organizational goals; like ambiguous social tasks; seek power and status.

Conventional: Like well-ordered environments; conscientious, efficient, obedient, calm, orderly and practical; are usually conforming and prefer subordinate roles; identify with power; value material possessions and status; good at well structured tasks; dislike ambiguity; dislike interpersonal problems; dislike physical tasks.

Note. The descriptions in this table have been taken from the latest SCII manual (Campbell & Hansen, 1981).

Appendix C

Summarized Descriptions of the Personality Preferences Measured by the Four MBTI Indices

Preference	Description
Extraversion:	A person with this preference is primarily oriented toward the outer world or external environment; they focus perception and judgement upon things and people, usually do their best work externally, in action, and like to get down to practical applications of theory.
Introversion:	A person with this preference is primarily oriented to the inner world of concepts and ideas; they do their best work reflectively in their head, are more interested in the insight a theory provides for understanding others or themselves, and concentrate perception and judgement upon ideas.
Sensing:	A person with this preference becomes aware of things through the five senses; they are interested in the actuality around them; they give scant attention to ideas coming faintly out of nowhere; they focus on what is specifically written on a piece of paper.

Note. This table is continued on the next page.

Appendix C continued

Summarized Descriptions of the Personality Preferences Measured by the Four MBTI Indices

---

Intuition: A person with this preference uses indirect perception by way of the unconscious; they rely on unconscious contributions that range from mere hunches to brilliant examples of creativity or scientific discovery; they get absorbed in studying possibilities and seldom look at actualities; they read between the lines hoping to identify possibilities.

Thinking: A person with this preference relies on logical processes; they arrive at impersonal findings; they seek to know if something is true or false; they desire to judge whether an idea is consistent and logical, or not; they are good at organizing facts and ideas.

Feeling: A person with this preference is concerned with whether an idea is pleasing or displeasing; they are sensitive to whether an idea is supporting or threatening; they bestow a personal, subjective value upon things; they are more mature when handling human relations.

Note. This table is continued on the next page.

Appendix C continued

Summarized Descriptions of the Personality Preferences Measured by the Four MBTI Indices

Preference	Description
Judgment	A person with this preference wants to agree or disagree quickly; they want to arrive at a conclusion; they shut off perception at a certain point; they reach a point where additional evidence is irrelevant and immaterial; they seek to arrive at a verdict; they like to order their life.
Perception	A person with this preference remains open-minded; they delay judgment; they wait to see if all the evidence is in; they expect new developments to occur; they hesitate to make the irrevocable commitment; they enjoy living their life rather than ordering it.

Note. The descriptions in this table have been taken from the definitions of the preferences provided by Myers (1962) and by Myers and Myers (1980).

Appendix D

Examples of Similarities in Verbal Descriptions Between HBI Scales and Scales of the SCII and MBTI

HBI Thinking	SCII Investigative
fascinated by concepts, fascinated by theories, systematic	scientific orientation
cognitively oriented	introspective
fascinated by theories, rational	need to understand the physical world
intellectually oriented, cognitively oriented	think through rather than act out problems, intellectual
fascinated with analytical relationships	analytical
fascinated by ideas	scholarly, curious
HBI Thinking	**SCII Conventional**
logical, deliberate	prefer order
systematic	systematic verbal activities systematic numerical activities
logical	good at well structured tasks
deliberate	consciencious
systematic	orderly
HBI Feeling	**SCII Artistic**
Go with feelings when making decisions	prefer unstructured situations
emotionally expressive	emphasize self-expression
go with feelings	impulsive, intuitive, avoid structured problems

Note. This table is continued on the next page.

Appendix D Continued

Examples of Similarities in Verbal Descriptions Between HBI Scales and Scales of the SCII and MBTI

HBI Acting	SCII Enterprising
involved with others, public oriented	good at selling, dominating, leading, popular, sociable
goal oriented	drive to attain goals, drive to attain economic gain
get the job done	power
aggressive	aggressive
plunge into the thick of things	self confident
HBI Thinking	**MBTI Introversion**
intellectually oriented, fascinated by concepts	oriented to inner concepts
fascinated by ideas	oriented to inner world of ideas
cognitively oriented	reflective, use their head
rational	like insight in order to understand self, others
HBI Thinking	**MBTI Thinking**
logical	relies on logical processes, judge whether ideas are logical
rational	impersonal findings, seek to know true/false
systematic	good at organizing facts, ideas, judge whether ideas are consistent

Note. This table is continued on the next page.

Appendix D Continued

Examples of Similarities in Verbal Descriptions Between HBI Scales and Scales of the SCII and MBTI

HBI Thinking	MBTI Sensing
rational	interested in actuality, ignore ideas out of nowhere
logical	focus on specific facts, information
HBI Feeling	**MBTI Intuition**
go with feelings	use of unconscious, indirect perception, read between the lines, possibilities
if it feels good, do it	seldom look at actualities
HBI Feeling	**MBTI Feeling**
go with feelings	is idea pleasing
go with feelings when making decisions	sensitive to whether an idea is threatening or supportive bestow personal, subjective value on things
HBI Acting	**MBTI Extraversion**
public oriented	oriented toward the outer world and the environment
involvement with others	focus on people
do things	focus on things
get the job done, plunge into the thick of things	best in action, like practical applications of theory

Appendix E

Statement Read to Subjects to Introduce the Research Project

As part of your exposure to psychological assessment in this course, you are going to be completing psychological inventories during this semester. Some of the inventories are being studied for their usefulness. Thus, you are also being introduced to the topic of psychological research, since completing the inventories will constitute participation in a research project. You will be able to receive feedback concerning the results of most of the completed inventies later in the semester. The feedback will consist of your scores on certain of the inventories, as well as a general explanation of the results provided by those inventories. Should you desire an individual interpretation of any of these inventories, arragements can be made through either the advising or counseling centers of the university.

Appendix F

Frequencies for Age and Classification

	Frequency	Cum Freq	Percent	Cum Percent
Age				
U	81	-	-	-
I	1	1	0.175	0.175
17	13	14	2.269	2.443
18	170	184	29.668	32.112
19	150	334	26.178	58.290
20	89	423	15.532	73.822
21	57	480	9.948	83.770
22	32	512	5.585	89.354
23	23	535	4.014	93.368
24	9	544	1.571	94.939
25	9	553	1.571	96.510
26	3	556	0.524	97.033
27	3	559	0.524	97.557
28	2	561	0.349	97.906
29	1	562	0.175	98.080
30	2	564	0.349	98.429
33	1	565	0.175	98.604
35	1	566	0.175	98.778
38	1	567	0.175	98.953
39	1	568	0.175	99.127
40	2	570	0.349	99.476
43	1	571	0.175	99.651
47	1	572	0.175	99.825
57	1	573	0.175	100.000
Classification[a]				
U	521	-	-	-
Fresh	33	33	24.812	24.812
Soph	21	54	15.789	40.602
Junior	36	89	26.316	66.917
Senior	39	128	29.323	96.241
Special	5	133	3.759	100.00

Note. U = unknown. I = incorrectly marked.

[a]Classification data were only gathered by way of the MBTI.

Appendix G

Means, Standard Deviations, and Minimum/Maximum Values for Variables Included in the Study

Variable	N	Mean	Standard Deviation	Values Minimum	Values Maximum
First Administration of the HBI-N and HBI-I					
N-intensity T	583	9.6	2.7	1.0	15.0
N-intensity F	583	7.4	3.5	0.0	15.0
N-intensity A	583	6.9	3.0	0.0	14.0
N-bipolar FT	583	2.2	4.2	-12.0	14.0
N-biploar TA	583	-2.7	3.5	-15.0	8.0
N-bipolar AF	583	0.5	4.2	-13.0	14.0
I-choice A	583	20.3	9.6	0.0	49.0
I-choice F	583	21.7	12.2	0.0	49.0
I-choice T	583	32.3	9.5	3.0	50.0
I-intensity A	576	1.6	0.5	1.0	3.0
I-intensity F	576	1.7	0.5	1.0	3.0
I-intensity T	583	1.9	0.4	1.0	3.0
I-bipolar AF	583	12.9	7.3	0.0	25.0
I-bipolar FT	583	16.1	6.4	0.0	25.0
I-bipolar TA	583	7.8	5.4	0.0	25.0
Second Administration of the HBI-N and HBI-I					
N-intensity T	454	9.3	2.8	0.0	15.0
N-intensity F	454	7.1	3.5	0.0	15.0
N-intensity A	454	6.7	3.0	0.0	15.0
N-bipolar FT	454	2.3	4.4	-15.0	13.0
N-bipolar TA	454	-2.6	3.4	-14.0	10.0
N-bipolar AF	454	0.4	4.1	-13.0	14.0
I-choice A	456	19.7	9.5	0.0	47.0
I-choice F	456	21.6	12.6	0.0	50.0
I-choice T	456	32.5	10.2	0.0	50.0

Note. This table is continued on the next page.

Appendix G continued

Means, Standard Deviatons, and Minimum/Maximum Values for Variables Included in the Study

Variable	N	Mean	Standard Deviation	Value Minimum	Value Maximum
Second administration of the HBI-I and HBI-N					
I-intensity A	447	1.6	0.4	1.0	3.0
I-intensity F	450	1.7	0.5	1.0	3.0
I-intensity T	455	1.9	0.4	1.0	3.0
I-bipolar AF	456	13.3	7.4	0.0	25.0
I-bipolar FT	456	16.3	6.9	0.0	25.0
I-bipolar TA	456	7.5	5.5	0.0	25.0
First and only administrations of the SCII and MBTI					
MBTI E-I	134	96.2	24.1	49.0	149.0
MBTI S-N	134	87.3	22.0	45.0	143.0
MBTI T-F	134	106.3	21.8	35.0	141.0
MBTI J-P	134	93.3	23.9	45.0	155.0
SCII R	129	42.4	9.8	29.0	66.0
SCII I	129	41.3	9.6	23.0	68.0
SCII A	129	44.9	9.5	24.0	65.0
SCII S	129	54.0	9.2	29.0	71.0
SCII E	129	46.1	9.2	28.0	72.0
SCII C	129	46.9	9.1	23.0	73.0
Demographics					
Age	573	20.0	3.6	1.0[a]	57.0
Classification	133	2.7	1.2	1.0	5.0

[a]One subject incorrectly marked his/her age (age = 1). The actual minimum age was 17.

Appendix H

Independent T-tests on HBI Scores of Counterbalanced Groups

Scale	Admin	Group Order	N	Mean	Std Dev	T	DF	Prob
				HBI-N				
T	1st	1	332	9.68	2.8	0.88	566.6	0.378
		2	251	9.49	2.5			
F	1st	1	332	7.59	3.6	1.34	581.0	0.178
		2	251	7.20	3.4			
A	1st	1	332	7.09	2.9	1.95	581.0	0.052
		2	251	6.61	3.1			
T	2nd	1	214	9.86	2.5	3.70	450.0	0.000**
		2	239	8.92	2.9			
F	2nd	1	214	7.01	3.4	-0.51	451.0	0.608
		2	239	7.18	3.7			
A	2nd	1	214	7.14	2.9			
		2	239	6.34	3.1	2.86	451.0	0.005**
				HBI-I				
AF	1st	1	249	12.61	7.4	-0.96	581.0	0.337
		2	334	13.20	7.3			
FT	1st	1	249	16.31	6.2	0.51	581.0	0.611
		2	334	16.03	6.5			
TA	1st	1	249	8.07	5.3	0.95	581.0	0.345
		2	334	7.64	5.5			
AF	2nd	1	240	13.44	7.0	0.55	454.0	0.585
		2	216	13.07	7.7			
FT	2nd	1	240	15.89	6.8	-1.24	454.0	0.215
		2	216	16.69	6.9			
TA	2nd	1	240	7.65	5.5	0.65	454.0	0.514
		2	216	7.32	5.4			

**$p<.01$, two-tailed.

Appendix I

Summary of Correlations Between Scores of the HBI-I, HBI-N, SCII, and MBTI that Appear in the Multitrait-Multimethod Validity Matrix

Scores	Number of r's	Number of r's with the predicted direction	Number of r's with the predicted size and direction	Number of r's with the predicted size - ignoring direction
HBI-I Choice HBI-I Choice	6	6	5	5
HBI-I Choice HBI-I Bipolar	9	6	6	6
HBI-I Choice HBI-N Intensity	9	9	1	1
HBI-I Choice HBI-N Bipolar	9	8	1	2
HBI-I Choice SCII	18	13	7	10
HBI-I Choice MBTI	12	7	1	4
HBI-I Bipolar HBI-I Bipolar	6	6	5	5

Note. This table is continued on the next page.

Appendix I continued

Summary of Correlations Between Scores of the HBI-I, HBI-N, SCII, and MBTI that Appear in the Multitrait-Multimethod Validity Matrix

Scores	Number of r's	Number of r's with the predicted direction	Number of r's with the predicted size and direction	Number of r's with the predicted size - ignoring direction
HBI-I Bipolar HBI-N Intensity	9	7	3	4
HBI-I Bipolar HBI-N Bipolar	9	9	7	7
HBI-I Bipolar SCII	18	16	6	7
HBI-I Bipolar MBTI	12	8	1	4
HBI-N Intensity HBI-N Intensity	6	6	3	3
HBI-N Intensity HBI-N Bipolar	9	8	4	5
HBI-N Intensity SCII	18	13	5	7

Note. This table is continued on the next page.

Appendix I continued

Summary of Correlations Between Scores of the HBI-I, HBI-N, SCII, and MBTI that appear in the Multitrait-Multimethod Validity Matrix

Scores	Number of r's	Number of r's with the predicted direction	Number of r's with the predicted size and direction	Number of r's with the predicted size - ignoring direction
HBI-I Intensity MBTI	12	6	2	7
HBI-N Bipolar HBI-N Bipolar	6	6	5	5
HBI-N Bipolar SCII	18	10	3	9
HBI-N Bipolar MBTI	12	8	2	5
SCII SCII	15	11	7	7
SCII MBTI	24	22	15	17
MBTI	6	2	2	6
Totals	243	189	91	126

Appendix J

Correlations Among SCII Scores, MBTI Scores, and Between the Scores of the SCII and MBTI Obtained in this Study

		SCII						MBTI			
		R	I	A	S	E	C	SN	TF	EI	JP
SCII	R	-									
	I	.58	-								
	A	.15	.44	-							
	S	.27	.39	.41	-						
	E	.24	.33	.33	.41	-					
	C	.24	.31	.23	.37	.64	-				
MBTI	SN	.11	.26	.43	.16	.06	-.07	-			
	TF	-.35	-.27	.13	.07	-.05	-.13	-.05	-		
	EI	.08	.03	.11	-.10	-.27	-.04	-.08	-.12	-	
	JP	.08	.03	.11	-.10	.01	-.16	.31	.18	-.07	-

SCII scores: Fourteen of the correlations among scores of the SCII were moderate in size and one was low. All of the correlations were positive. Eleven of the 15 correlations among scores of the SCII obtained in this study compare quite favorably to those produced in the latest version of the SCII manual (Campbell & Hansen, 1981).

MBTI scores: Five of the correlations among scores of the MBTI were low and one was moderate. All six correlations among the MBTI scores compared very favorably with correlations among scores of the MBTI obtained for college and university student populations reported in the latest MBTI manual (Myers & McCaully, 1985).

SCII and MBTI scores: Nineteen of the correlations among scores of the SCII and MBTI were low and the other five were moderate. Eighteen of the 24 correlations among scores of the SCII and MBTI obtained in this study compare very favorably to those reported between the scores of the SCII and MBTI in the latest MBTI manual (Myers & McCaulley, 1985)

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2