A STUDY OF A KNOWLEDGE-BASED MODEL FOR THE THEORY AND MEASUREMENT OF INTELLIGENCE WITH A WISC-R DATA SAMPLE

By

KARL AUGUSTUS HOFFMAN

Bachelor of Arts
King's College
Wilkes-Barre, Pennsylvania
1979

Master of Arts
Marywood College
Scranton, Pennsylvania
1981

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
Doctor of Philosophy
May, 1986

Thesis
1986D
H699s
cop.2

OKLAHOMA STATE UNIVERSITY LIBRARY

A STUDY OF A KNOWLEDGE-BASED MODEL FOR THE THEORY AND MEASUREMENT OF INTELLIGENCE WITH A WISC-R DATA SAMPLE

Thesis Approved:

Thesis Advisor

Dean of the Graduate College

1258572

PREFACE

This was a study of a proposed approach to the theory and measurement of intelligence entitled the "knowledge propensity" model. The focus of the knowledge propensity model is the conceptualization of human intelligence from the constructual basis of propensity to acquire knowledge. It was concluded that the findings of this study did not fully support the efficacy of the model, but that these findings should be considered heuristically in the overall context of the possibilities of knowledge-based measurement of intelligence.

I wish to thank all those who have provided me with advisement and assistance during my doctoral studies at Oklahoma State University. Most prominent in this regard is Dr. Michael Kerr, my major advisor.

Much acknowledgement and thanks is of course due to the other committee members, Dr. Jo Campbell, Dr. Joseph Pearl, Dr. Thomas Johnston, and Dr. Bill Elsom, the committee chairperson.

I wish also to acknowledge Dr. John Hampton, whose early comments and encouragement eventually led to my pursuit of this topic as a dissertation.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The purpose of this study was to examine an approach to the theory and measurement of intelligence proposed by this writer entitled the "knowledge propensity" model. The focus of this model is the conceptualization of human intelligence from the constructual basis of what is described as "knowledge propensity," or the propensity to acquire knowledge. In the history of modern intelligence theory, various theorists have posited the existence of a general, underlying property of which all aspects of intelligence are functions (Spearman, 1973; Thurstone, 1955; Cattell, 1963; Wechsler, 1958). The property is well known as the general, "*g*" factor of intelligence. What is less well known, as evidenced by the vagueness and variations of its description, is what *g* actually is and what measure or measures of intelligence best function as measures of *g*. Due to these problems of theory and measurement, *g*, along with the accompanying notion that intelligence may be represented by a singular IQ score, has received increasing criticism that has included radical alternatives such as Howard Gardner's (1983) theory of multiple intelligences.

It is proposed in the model that individuals, to varying degrees, have an inherent propensity to acquire knowledge. By "inherent" it is meant that this propensity is not learned or limited to environmental contingencies, such as school achievement. The knowledge propensity model is not a definitive model of intelligence, but is a model based on the assumption of the existence of *g*. The central hypothesis of the knowledge propensity model is that *g* is most directly manifested by propensity to acquire knowledge. This theoretical relationship may potentially serve a basis for further study and understanding of what *g* actually is and what measure or measures of intelligence best function as measures of *g*.

The knowledge propensity model is defined by two criteria for measurement. The first criterion is that a test item measure knowledge as recall of information, such as that measured by test items of general information and vocabulary. The second criterion is that a test item must have a difficulty of .5, or greater among the population for whom the model is to be applied. It is this second criterion that differentiates the knowledge propensity model from existing, generic knowledge-based measurement of intelligence, and it is an attempt to establish theoretical and statistical parameters for such means of measuring intelligence.

In the construction of what may be classified as "range of information," or "breadth of knowledge" measures of intelligence there should be consideration of what knowledge is relevant to a given population, as well as what knowledge is likely to be accessable across all strata of a population. These considerations bear directly on content validity and test bias; two aspects that have proven to be pitfalls with previous knowledge-based measures of intelligence (Jensen, 1980). By only including knowledge held by at least half a sampled population, knowledge represented by items found to fit the knowledge propensity model were assumed to be generally relevant and accessable to all members of that population. In the past, test construction with breadth of knowledge measurement has ultimately involved the subjective judgement of the author of what knowledge is appropiate for inclusion. Though Wechsler (1958, p. 65), in discussion of general information tests, stated that such "... items should call for the sort of knowledge that an average individual with average opportunity may acquire for himself," he reserved judgement for himself of what that "sort of knowledge" was. The criterion of difficulty level was not an arbitrary or dogmatic attempt to establish parameters for the sake of building a model. It was an attempt to employ an empirical basis for defining breadth of acquired knowledge in terms of population characteristics; thus establishing some basis for examining the relation of propensity for knowledge acquisition, as it was conceived in this model, and human intelligence.

A related reason for the criterion of difficulty level of .5 or greater for items found to fit the knowledge propensity model involved what is actually meant by "breadth" of knowledge. Again, in considering Wechsler's (1958, p. 65) statement that such "... items should call for the sort of knowledge that an average individual with average opportunity may acquire for himself," we may detect another problem. In the case of Wechsler's general information test, Jensen (1980) points out that many of the items do not measure knowledge common to all members of a population, and that the test is actually more correlated with level of education than with level of intelligence. The approach to intelligence from the constructual basis of propensity to acquire knowledge includes that this propensity is inherent, and is not learned or limited to environmental contingencies, such as school achievement. Such a perspective, with an emphasis on the inherent, requires the distinction of breadth of acquired knowledge and level of acquired knowledge regarding contents related primarily to level of education.

Certainly, education constitutes a form of life experience, and is a context in which knowledge may be acquired according to an individual's propensity to do so. However, the learned motivational contingencies related to school achievement confound knowledge acquired with knowledge required. Ultimately, except for practical considerations of test construction, items of high difficulty do not contribute to measurement of "breadth" of knowledge, per se.

It is implied here that due to lack of theoretical basis and concern for measurement parameters, that breadth of acquired knowledge has not met its potential as a means for estimating or studying intelligence.

"Breadth" of acquired knowledge has specialized meaning in this model, and this meaning reflects the focus of the model on intelligence from the perspective of propensity to acquire knowledge. On one hand, the meaning of "breadth" of knowledge is substantially broader than traditional usage; with not only knowledge of facts included by the way of general information types of measurement, but also knowledge of words by the way of vocabulary and picture vocabulary types of measurement. On the other hand, the meaning is substantially narrower than traditional usage; with only knowledge held by at least half a sampled population to be considered content valid and non-biased in measurement. The actual denotation in the model of breadth of acquired knowledge, as well as the theoretical conceptualization of propensity to acquire knowledge as a perspective by which the nature of intelligence may be better known, is not based directly on previous empirical findings and is a hypothetical assumption to be examined. The model itself is offered more as a heuristic tool, than as a description or explanation in its presented form. It is also important to note that the criterion of a difficulty level of .5 or greater for all items found to fit the knowledge propensity model is problematic. This criterion is established in a theoretical attempt to identify what knowledge contributes to the valid measurement of intelligence. However, such a restriction on difficulty of items necessarily limits the variance of measures found to fit the model according to this criterion of difficulty. There is, then, a trade-off in which the criterion that, in theory, identifies what knowledge contributes to the valid measurement of intelligence also limits the variance of that measurement.

The relation of intelligence and propensity to acquire knowledge can also potentially be studied under controlled, experimental conditions. Under these conditions, equal exposure to measured controlled knowledge can occur. Such means of studying propensity to acquire knowledge would not include concerns about validity of the knowledge measured, or what knowledge is "generally relevant and accessable." The present study of the knowledge propensity model is directed toward judging the potential utility of the model in the practical measurement of intelligence. The criterion of item difficulty may or may not ultimately prove advantageous in this utility and will receive further investigation, revision, or rejection, accordingly; as will the knowledge propensity model, itself.

Though the focus of this model on human intelligence from the perspective of propensity to acquire knowledge is a new conceptualization, the intent of formulating a knowledge-based orientation to human intelligence does have antecedant basis. Knowledge acquisition has been acknowledged from diverse sources as being the most essential and basic aspect of human mental functioning (R.I. Evans, 1975; C. Evans, 1979; Bloom, 1956). As will be reviewed in the following section, consideration of such orientations in this context is accompanied by related evidence that knowledge-based measures of intelligence, such as information, vocabulary, and picture vocabulary measures, are generally better estimates of general intelligence than are other measures of intelligence.

In summary, it is proposed in the knowledge propensity model that individuals, to varying degrees, have an inherent propensity to acquire knowledge; and that this propensity is not learned or limited to environmental contingencies, such as school achievement. The knowledge propensity model is based on the assumption of the existence of *g*, the general factor of intelligence. The central hypothesis of the knowledge propensity model is that *g* is most directly manifested by propensity to acquire knowledge; and that this theoretical relationship may potentially serve as a basis for further study and understanding of *g*. The knowledge propensity model is defined by two criteria for measurement. The first criterion is that a test item measure knowledge as recall of information, such as that measured by test items of general information and vocabulary. The second criterion is that a test item must have a difficulty level of .5 or greater among the population for whom it is to be applied.

In this study, two hypotheses that test the knowledge propensity model in the context of a conventional, standardized measure of intelligence were examined in two separate analyses. The measure of intelligence used in this study was the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974). The rationale for testing the model in this manner stems from the central hypothesis of the model, that *g* is most directly manifested by propensity to acquire knowledge; as well as from an interest in judging the potential utility in the practical measurement of intelligence. The WISC-R is considered for the purpose of this study to be a valid estimate of general intelligence. The hypotheses test the extent to which WISC-R items found to fit the knowledge propensity model contribute, individually and collectively, to this measure. These items, then, are not analyzed as separate measures from the WISC-R in whole, but as a constituent part of that general measure.

In the first analysis, the hypothesis tested was that items found to fit the knowledge propensity model from the WISC-R will yield a significantly higher median discrimination index on sampled full-scale WISC-R IQ scores than WISC-R items found not to fit the model.

In the second analysis, two research hypotheses and one null hypothesis were tested to examine the general hypothesis that both high and low IQ classification groups will differ significantly in raw score distributions of correctly answered

WISC-R items found to fit the knowledge propensity model from the overall sample from which these groups were derived. One research hypothesis was that a high IQ classification group will differ significantly in raw score distribution of correctly answered WISC-R items found to fit the knowledge propensity model from the overall sample from which this group was derived. Similarly, the other research hypothesis was that a low IQ classification group will differ significantly in raw score distribution of correctly answered WISC-R items found to fit the knowledge propensity model from the overall sample from which this group was derived. Also, a null hypothesis was that an average IQ classification group will not differ significantly in raw score distribution of correctly answered WISC-R items found to fit the knowledge propensity model from the overall sample from which this group was derived.

CHAPTER II

REVIEW OF LITERATURE

The knowledge propensity model is a knowledge-based model of human intelligence. Its development stemmed not from any one particular antecedent model or theoretical perspective, but from a synthesis of several areas; including theories of *g* and prior knowledge-based measurement of intelligence. In the model it was proposed by this writer that the property that has been traditionally identified as the general, *g* factor of intelligence is actually propensity for knowledge acquisition. It was also proposed in the model that measures of breadth of acquired knowledge that fit the model may be examined as particularly valid singular estimates of propensity for knowledge. With respect to the scope of the proposed model, this review begins with a summary of the historical development and present state of the art of intelligence theory. This review then progresses to theories of *g* and prior treatment of knowledge acquisition and knowledge-based measurement in the theory and measurement of intelligence.

Intelligence Theory

Intelligence has been a perennial subject matter in attempts to understand human nature and behavior. The concept of intelligence has been traced classically to Cicero, and in modern psychological thought to Spenser (in Torrance and White, 1969). Since the advent of modern psychology the question of the nature of human intelligence and its measurement has proven to be a perplexing problem. In 1921 the editors of the _Journal of Educational Psychology_ published a fourteen paper symposium concerning the question. The results of this symposium were fourteen separate and conflicting accounts of intelligence and its measurement by such figures as Thorndike (1921), Terman (1921), Colvin (1921), and Pinter (1921). Despite the passing of more than 60 years and much theorizing, the same atmosphere of diverse viewpoint exists today. In Chaplin's _Dictionary of Psychology_ (1975, p. 263) three definitions are given as common:

1. The ability to meet and adapt to novel situations quickly and effectively.
2. The ability to utilize abstract concepts effectively.
3. The ability to grasp relationships and to learn quickly.

The dilemma of defining intelligence in more than one manner is not merely a theoretical problem, but is ultimately a problem concerning the practice of assessing human intelligence. The issue of test bias in assessment of intelligence thrives on the fact that psychology can not take a united, defensible stance on what intelligence is and how it may best be measured (Herrnstien, 1982; Burrill and Wilson, 1980; Jensen, 1980). Chaplin (1975) cites the conflict of different definitions as actually definitions focusing on particular aspects of intelligence. The notion of "focusing" on particular aspects of intelligence was also used by Eysenck (1953) in similar rationalization of multiple and conflicting definitions of intelligence. Wechsler (1975) has defended lack of understanding of the nature of intelligence by taking the position that intelligence is relative to sociocultural contexts, and ultimately unknowable in an absolute sense.

Theories of *g*

Movement towards reducing the complexity and deliniating the structure of intelligence has been made to some extent by factor analytic studies. Besides identifying distinct classes of mental abilities, the factor analysis of intelligence has, except in certain instances (Eysenck, 1953; Guilford, 1967), indicated the presence of an underlying general property, or factor, of which all aspects of intelligence are functions. Thurstone, for example, eventually postulated a second-order factor common to his primary factors which is conceptually related to *g* (in Sattler, 1974). Though the notion of a general factor of intelligence is basically a statistical concept, it has most often been considered by theorists to be a dynamic property by which demonstrable elements of intelligence are made manifest. Cattell (1959, 1963) described this dynamic property as "fluid intelligence;" which, in interaction with cultural environment, is the basis for "crystallized intelligence." Similarly, Spearman (1973) earlier described *g* as "psychic energy" that generates the "mechanisms" of specific, "*s*" factors in the form of mental abilities. Vernon (1961) has also described *g* in a corresponding manner. The theoretical implication of *g* is that intelligence, despite its ultimate complexity, has a definable nature and is an attribute caused by a specific phenomenon or property basic to human mental functioning.

Knowledge Acquisition and Knowledge-Based Measurement of Intelligence

Computer models of intelligence have also offered the compelling theoretical analogy of artificial intelligence to human intelligence in terms of information acquisition and processing (C. Evans, 1979). Bloom (1956) identified knowledge acquisition as the most basic function in the cognitive domain. A more direct indication of the role of knowledge acquisition in human intelligence is the evidence supporting the efficacy of knowledge-based measures of intelligence. Three such measures that have received attention, and are well documented in the literature, are what have been known as general information (or range of information), vocabulary and picture vocabulary tests.

Prior to the advent of modern standardized tests, tests of "range of information" were used by psychiatrists to estimate the intellectual functioning of patients (Wechsler, 1958). This usage, however, was not continued by psychologists with the development of the first intelligence scales. Even when information items emerged with the development of the first group tests, reasons for use were based in practicality, rather than theory (Wechsler, 1958). The Army Alpha Battery, considered to be antecedant to group tests of intelligence (Shertzer and Linden, 1979) provides the first available data on an information measure of mental functioning. Wechsler (1958) reported that the information component of the Alpha gave a better curve distribution than did other components of the test, with less extreme scores. Wechsler (1958) believed the findings to be indicative of the inherent potential of general information as a measure of intelligence.

At present, the Wechsler series of intelligence scales (1967, 1974, 1981) is the primary source on general information as a measure of intelligence. In general, the information subtests of the Wechsler series have shown to be the best, or near, best estimate of general intelligence (full-scale Wechsler scores) at all age levels of the subtests, as is reported in Table I.

TABLE I

AVERAGE r OF INFORMATION SUBTESTS
WITH WECHSLER FULL-SCALE IQ SCORES
ACROSS ALL AGE LEVELS OF EACH SCALE

SCALE	VERBAL	PERFORMANCE	FULL-SCALE	SUBTESTS YIELDING HIGHER r WITH FULL-SCALE
WPPSI	.75	.56	.70	---
WISC-R	.74	.56	.70	vocabulary (.74) similarities (.71)
WAIS-R	.79	.62	.76	vocabulary (.81)

It is worth noting that on the WAIS (Wechsler, 1955), the information subtest yielded the highest correlation with full-scale IQ in all but one of the age groups of that scale. In addition, the Wechsler information subtests tend to correlate higher with performance subtests than they do with the digit-span subtest, which is a measure of short-term memory (Wechsler, 1958). The finding that tests of general information correlate highly with non-verbal performance tests has been noted by both Wechsler (1958) and Jensen (1980), but the possible implications of this noted relationship between these seemingly separate measures of intelligence has not been pursued.

Despite the importance of the information subtests of the Wechsler series, their usage is generally discounted by reviewers. One criticism of the information subtests is that it is inappropriate to equate the range of information held by an individual with intelligence because acquisition of isolated facts does not itself indicate capacity for appropriate or effective use of these facts (Taylor, 1961). Cohen (1959) states that information, despite being a good measure of *g*, is not independently interpretable in terms of intellectual functioning. Jensen (1980, p. 147) notes that breadth of knowledge measures such as the information subtest in the Wechsler series are the "most problematic to define" and are "less theoretically defensible" than are other measures.

While the use of information measures of intelligence have not received wide acceptance, vocabulary measures have. In the case of the Wechsler test series (1967, 1974, 1981), the vocabulary subtest has most often yielded the highest correlation with full-scale IQ of all the Wechsler subtests across the respective age levels and scales of the series. Wechsler (1958) designated the vocabulary subtest as an alternative test in the original Wechsler-Bellevue Scale to be omitted with illiterative and native speakers of foreign languages; but found its "merits" warranted full inclusion in subsequent scales. Cohen (1958) found the Wechsler vocabulary subtest to be essentially a measure of *g*; and that, by itself, it can be used to estimate intelligence in research and clinical screening situations.

Vocabulary measures of intelligence have received much attention and acceptance (Wechsler, 1958; Cohen, 1959; Dunn and Dunn, 1981). Indeed, it can be stated that, except for certain "culture-free" tests (Cattell, 1959; Raven, 1960), vocabulary has been relied on more heavily than any other means of measuring intelligence in traditional usage. The existence of at least several picture vocabulary tests (Dunn, 1959, 1965; Dunn and Dunn, 1981; Ammons and Ammons, 1948, 1962) is evidence of the general belief that verbal fluency is central to intellectual functioning. Several other indivicual intelligence tests, in past and present usage, include picture vocabulary components (Terman, 1916; Terman and Merrill, 1937, 1960; French, 1964; Kaufman and Kaufman, 1983).

The use of picture vocabulary as a measure of intelligence began with Binet in his early work to produce an adequate intelligence scale for use in the educational system of France. Dunn and Dunn (1981) trace the findings of Binet and Terman regarding the picture vocabulary test of the Binet series to support the inference that such tests are particularly valid measures of intelligence. In the original Stanford version of the Binet scale, Terman (1916) found the picture vocabulary test of the Stanford revision of the Binet-Simon scale produced intelligence quotients only "negligibly" different from those derived from the entire scale. In the second edition of the Stanford-Binet, Terman and Merrill (1937) later reaffirmed the position that vocabulary is the best singular measure of intelligence.

Terman's position on the subject of vocabulary as a measure of intelligence was preceded by that of Binet and Simon (in Dunn and Dunn, 1981).

Summary

Again, as described in the previous chapter, the knowledge propensity model was defined by two criteria for measurement. The first criterion was that a test item measure knowledge as recall of information, such as that measured by test items of general information and vocabulary. The second criterion was that a test item must have a difficulty of .5, or greater among the population for whom the model is to be applied. In general, the types of presently existing measures of intelligence that meet the first criterion of the model under examination have been presented as being highly valid. Of course, no presently existing measure of intelligence also meets the second criterion of the knowledge propensity model; so, the model is essentially unprecedented in the literature. This review, in whole, is better understood in its purpose as one that presents the role of knowledge-based measures in view of the more theoretical problems of defining intelligence and identifying its basic constitution.

CHAPTER III

METHODOLOGY

In certain respects, the model under study is difficult to describe empirically, due to its unprecedented status and theoretical concern with the nature of intelligence; which, as discussed in the previous section, is not an issue that has been amenable to definition or agreement. In this study, two hypotheses that test the validity of the model were tested in two separate analyses. This study is based on sample test data derived from the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1967), with particular attention to items that were found to fit the knowledge propensity model by meeting the two model item criteria. Although such usage was not the original intent of the measure used, or of its items found to "fit" the model, the purpose was to test the model in the context of a conventional intelligence test.

Instrumentation

The measure of general intelligence used in this study was the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1967). The WISC-R is the most recent edition of the Wechsler series of individual intelligence tests for children six through sixteen years of age. Among individual intelligence tests, the Wechsler series has received attention and acceptance paralleled only by the Stanford-Binet series (Sattler, 1974). Wechsler (1974) reported a full-scale split-half reliability of .96 for the WISC-R across all age levels, and a test-retest reliability of .95. Wechsler (1974) also reported an average correlation of the WISC-R with the Stanford-Binet of .75 across all age levels. However, Wechsler did not explicitly present this correlation as concurrent validity, and no other validity information was presented. Specifically, knowledge propensity model items were derived in this study from the information and vocabulary subtests of the WISC-R.

Subjects

The WISC-R protocol forms for 180 subjects were drawn from the files of a psycho-educational clinic in Pennsylvania. These protocols were selected from the most recently administered 200 forms at the clinic; ranging in administration dates from 1981 through 1983, the time at which they were drawn. Of the 20 rejected forms, 17 forms were not used due to extreme low scores, under a full-scale IQ of 76. The IQ cutoff of 76 was utilized because such scores are indicative of retardation, or possibly various emotional, learning, or physical disabilities that are potentially interfering factors in the test situation; and which were, in fact, identified reasons why a number of the children in the sample were tested.

All subjects were tested by interns under supervision in the clinic's school psychology internship program. Due to concerns about confidentiality by program administrators, the reasons for testing the individual subjects were not disclosed. However, approximately half of the subjects were referred for testing by local schools because of identified learning or behavioral difficulties. Some subjects were also referred to be tested for giftedness. Other subjects were enlisted by student interns and program administrators as volunteers in order for the interns to meet the required number of test administrations in their practicums.

The sample mean IQ was 101 with a standard deviation of 9.9. The relatively low standard deviation was due primarily to the IQ cut-off of 76. In the sample there were 117 boys and 63 girls. The mean age of the sample was 10 years, 6 months; with an age range of 6 years, 1 month to 16 years, 10 months. The sample, then, was of a heterogenous population of school referral cases and volunteers for testing at the psycho-educational clinic from which the WISC-R protocol forms were drawn.

Procedures

In the first phase of the study, the WISC-R items that met the two knowledge propensity item criteria were identified. Items which met the first criterion, that of measuring recall of literal or conceptual information, were identified as all items of the information and vocabulary subtests. The difficulty levels of the items of the information and vocabulary subtests were then determined to identify which items of these two subtests met the second criterion, that of difficulty level of at least .5. The difficulty levels for these items were determined by compiling the number of sampled subjects correctly answering the respective items.

In the second phase of the study, all items from the WISC-R subtests were analyzed for discrimination between extreme high and low scores of subjects on the sample distribution of WISC-R full-scale IQ scores. The identification of the groups of subjects with the extreme highest and extreme lowest IQ scores in the sample was based on the identification of the 27% highest and 27% lowest IQ scores of the sample distribution; the method suggested by Kelly (1939) for determining discrimination indexes of test items. From the sample of 180 subjects, then, there were 49 subjects in both the extreme highest and extreme lowest IQ score groups. The extreme highest IQ score group ranged in scores from 143 to 109, with a mean score of 120. The extreme lowest IQ score group ranged in scores from 89 to 76, with a mean score of 85.

For the purpose of this analysis, an item was considered as being answered correctly if a subject received full or partial credit for his or her answer. An item was considered being answered incorrectly if no credit was given for an answer, or if the item was not administered due to the subtest being discontinued with successive incorrect responses to previous items of that subtest. Partial credit was considered to be a correct answer because this is the case in the standard administration of the WISC-R. Not included in this analysis were the digit span and maze subtests, which are optional tests of the WISC-R. Also not included in this analysis was the coding subtest, for which all subjects received partial credit due to its structure as a speed test of a singular performance task.

In a final phase of the study, the raw score sample distribution of the total WISC-R items found to fit the knowledge propensity model was also compiled. The raw scores for the subjects were of course the total number of correctly answered WISC-R items found to fit the knowledge propensity model.

Design

Data in this study was to be analyzed in two designs. In one, the discrimination indexes on the sample distribution of full-scale WISC-R IQ scores for all WISC-R items included in this study were compiled and arranged in two groups. The one group consisted of the discrimination indexes of WISC-R items found to fit the knowledge propensity model. The other group consisted of the discrimination indexes of all other WISC-R items found not to fit the model.

In the other design, the raw score distribution of correctly answered WISC-R items found to fit the model was compiled. This distribution was then partitioned according to three IQ category groups derived from the sample (high, average, and low). Subjects in the high and low IQ category groups were the same as in the procedure used to determine item discrimination indexes. Resulting from the partitioning, for the purposes of analysis, were raw score distributions of each of the three IQ category groups of subjects, as well as the total raw score distribution of the entire sample.

CHAPTER IV

RESULTS

In the sample, a total of 27 WISC-R items were found to meet both criteria of the knowledge propensity model. These 27 items consisted of 12 items from the information subtest and 15 items from the vocabulary subtest.

In the first analysis, a Mann-Whitney U test was used to determine whether there was a significant difference between the median discrimination indexes of WISC-R items that met the model criteria and WISC-R items that did not. This non-parametric test was used rather than a T-test for independent samples due to the non-normalcy of the distribution of the data. The results of this test are presented in Table II.

A significant difference in the median discrimination indexes on full-scale WISC-R scores between the model and non-model item groups at the .05 confidence level was not indicated. Therefore, the hypothesis that such a difference would be demonstrated with WISC-R items was not supported. The median discrimination index of WISC-R items found to fit the knowledge propensity model was .16. The median discrimination index of all other, non-model, WISC-R items included in this analysis was .22.

TABLE II

MANN-WHITNEY U ANALYSIS OF DISCRIMINATION INDEXES BETWEEN MODEL AND NON-MODEL WISC-R ITEMS

STATISTIC	MODEL ITEMS	NON-MODEL ITEMS
No. of cases	27	140
R	2574.5	11370
U	2196.5	
Z	1.33*	

*Not significant

Three chi-square goodness-of-fit tests were calculated for the three sample IQ category groups (high, average, and low) to determine whether significant differences existed between obtained and expected raw score frequencies of correct responses to WISC-R items found to meet knowledge propensity model item criteria. The expected raw score frequencies for each IQ group in the goodness-of-fit tests were equal to the expected frequencies for these groups in a contingency table of the entire sample. The obtained and expected frequencies of WISC-R knowledge propensity item raw score totals for the sample IQ groups are presented in Table III.

TABLE III

OBTAINED AND EXPECTED FREQUENCIES OF MODEL WISC-R ITEM RAW SCORE TOTALS FOR SAMPLE IQ CATEGORY GROUPS

	RAW SCORE TOTALS							
GROUP	27	26	25-24	23-22	21-20	19-17	16-8	TOTAL
HIGH IQ (143-109)								
OBTAINED	22	8	8	3	3	4	1	49
EXPECTED	10.62	7.08	8.44	5.17	5.72	4.90	7.08	
AVERAGE IQ (109-90)								
OBTAINED	15	13	13	11	12	7	11	82
EXPECTED	17.77	11.84	14.12	8.65	9.57	8.20	11.84	
LOW IQ (89-76)								
OBTAINED	2	5	10	5	6	7	14	49
EXPECTED	10.62	7.08	8.44	5.17	5.72	4.90	7.08	
SAMPLE N	39	26	31	19	21	18	26	180

In the three chi-square goodness-of-fit tests, obtained χ^2 values (See Table IV) were significant at the .05 level for the high and low IQ category groups; but not significant for the average IQ category group. The two research hypotheses and the null hypothesis tested in this analysis were supported; as was the general hypothesis that propensity for knowledge acquisition will differ significantly across sampled IQ category groups.

TABLE IV

χ^2 VALUES FOR SAMPLE IQ GROUPS

IQ GROUPS	N	χ^2	*df*	*p*
HIGH IQ	49	19.92	6	$p < .05$
AVERAGE IQ	82	2.12	6	$p > .05$
LOW IQ	49	15.58	6	$p < .05$

To test for significant differences between obtained and expected raw score frequencies for the entire distribution, a 3x7 chi-square analysis was performed. The obtained x^2=37.62, *df*=12, was significant at the .05 level. The seven raw score categories used in the chi-square analyses were representative of the range of raw scores obtained by the subjects of 27 to 8, with five of the categories consisting of two or more raw score values collapsed together in order to allow for expected frequencies of at least five in at least eighty percent of the cells.

CHAPTER V

CONCLUSIONS

The development of a general model of intelligence is an ambitious endeavor; and the complete examination and validation of such a model cannot possibly occur in one, initial study. The fact that there has not to this point been a fully adequate and acceptable general model of intelligence suggests that attempts to this end should be conducted with modest expectations. In the case of the knowledge propensity model, which has been presented with this study, two basic hypotheses that tested the validity of the model were examined in two separate analyses. The implications of this study in respect to the model tested and to knowledge-based measurement of intelligence, in general, will be considered.

The sample in this study was derived from existing WISC-R test data on 180 children drawn from the files of a psycho-educational clinic in Pennsylvania. Data on an additional 17 children were not used due to extreme low IQ scores under 76. Approximately half of the subjects were referred for testing by local schools because of identified learning or behavioral difficulties. Some subjects were also referred to be tested for giftedness. Other subjects were enlisted by student interns and program administrators as volunteers in order for the interns to meet the required number of test administrations for their practicums. The sample, then, was of a heterogenous population of school referral cases and volunteers at the psycho-educational clinic from which the WISC-R protocol forms were drawn. With the atypical nature of the sample population it must be noted that the findings of this study can not be considered generalizable to other populations, and conclusions from the findings must be considered with that limitation.

Data on the knowledge propensity model in this study came in the form of items from the Wechsler Intelligence Scale for Children-Revised (WISC-R) that met the knowledge propensity model item criteria. The purpose of generating data for the model in this manner was to examine the model in the context of conventional intelligence measurement. Specifically, this context was the testing of these items in their constituent and conglomorate contribution to the standardized measure of intelligence from which they were derived. Although this means of study was appropriate as an initial examination of the model, it was found not to be without its limitations.

In the preparation of this study, it was discovered that the Wechsler series of intelligence scales was unique among other major standardized measures of intelligence in that it contained items that fit the model. Yet, in this study, only 27 WISC-R items were found to meet the knowledge propensity item criteria. A main point of the limitation of these items was that 11 of these items had difficulty levels of .9, or above; and were relatively non-contributive in the discrimination and sample distribution analyses. The problem of overall item discrimination is inherent with the model item criterion of .5 or greater difficulty level for all items. However, this problem could be reduced with a large, unconstrained item sample in a situation in which a measure based on this model was generated through normal procedures of item analysis and selection.

In the first analysis of this study, the results did not support the hypothesis that items that met the knowledge propensity item criteria would discriminate to a higher degree on a sample distribution of IQ scores than items that did not. In the second analysis, the results did support the general hypothesis that both high and low IQ classification groups will differ significantly in raw score distributions of correctly answered WISC-R items found to fit the knowledge propensity model from the overall sample from which these groups were derived. It can be stated that such a difference should occur with any items, and that the results do not demonstrate that these items are superior in any way to other types of items. However, these results may be seen as encouraging due to limitations in the item sample because they suggest that, despite the presence of these limitations, adequate score distributions may possibly occur with a workable number of knowledge propensity model items.

The criterion of .5 difficulty or greater for all items included in the knowledge propensity model plays a paradoxical role. On one hand, it serves as a theoretical parameter by which breadth of acquired knowledge may be defined and studied in its relation to general intelligence. On the other hand, the criterion presents limitations on distribution of scores derived from measures based on this model. With only items of lesser difficulty included in this model, items answered correctly by at least half of a sample population, lack of variance should be expected at the upper limits of a score distribution derived from model measures. In the 27 model items of this study there was a definite skewdness towards the upper limits of the raw score distribution; with the mode, in fact, being a raw score of 27. As mentioned earlier, the relation of intelligence and propensity to acquire knowledge can also potentially be studied under controlled, experimental conditions. Under these conditions, equal exposure to measured controlled knowledge can occur, thus eliminating the rationale in the knowledge propensity model for an item difficulty level of .5 or greater. Results obtained under such experimental conditions may not be completely generalizable to propensity to acquire knowledge across the full range of life experiences in the course of one's life. However, such means of research is definitely a viable alternative to the methods of this study, and it is recommend in future investigation of this topic.

The ultimate test in assessing the value of this model would be the development of a standardized measure of intelligence based on the knowledge propensity item criteria. By the nature of the item criteria itself, especially the criteria of only measuring knowledge held by at least half a sampled population, technical aspects such as an adequate sample distribution of scores make developing such a measure a prohibitive task. The practical problems involved in developing a measure of intelligence based on the knowledge propensity model were another reason why already existing items were used in this study. By the results of this study, it would not appear called for to develop such a measure without further indication of the performance of the model.

Finally, whatever results found in this study or in future studies of this model must be considered in terms of the full milieu of intelligence theory; and the problems and limitations thereof. The construct of knowledge propensity was a proposed model for the theory and measurement of intelligence based on the premise that knowledge is the most basic function of intelligence, and that knowledge-based measurement of intelligence is inherently the best singular means of estimating general intellectual functioning. As explained earlier, the model is an attempt to establish an empirical basis for the parameters of knowledge in the measurement of intelligence. The findings of this study, as well as any findings on this model in the future, should be considered heuristically in the overall context of the possibilities of knowledge-based measurement of intelligence. These possibilities include radical revisions in this model and the generation of new, refined models.

LITERATURE CITED

Ammons, R.B., and H.S. Ammons. The Full-Scale Picture Vocabulary Test. New Orleans: R.B. Ammons, 1948.

Ammons, R.B., and C.H. Ammons. "The Quick Test: Provisional Manual." Psychological Reports, 11 (1962), 111-161.

Bloom, B.S. (Ed.) Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain. New York: Longmans, Green, and Co., 1956.

Burrill, L., and R. Wilson. Test Service Notebook 36: Fairness and the Matter of Bias. New York: The Psychological Corporation, 1980.

Cattell, R.B. Handbook for the Culture Fair Intelligence Test: A Measure of "g". Champaigne, Ill.: Institute for Personality and Ability Testing, 1959.

__________. "Theory of Fluid and Crystallized Intelligence: A Critical Experiment." Journal of Educational Psychology, 54 (1963), 1-22.

Chaplin, J.P. Dictionary of Psychology. New York: Dell, 1975.

Cohen, J. "The Factoral Structure of the WISC-R at ages 7-6, 10-6, 16-6." Journal of Consulting Psychology, 23 (1959), 285-299.

Colvin, S.S. "Intelligence and its Measurement." Journal of Educational Psychology, 12 (1921), 136-139.

Dunn, L.M. Peabody Picture Vocabulary Test Manual. Minneapolis: American Guidance Service, 1959.

__________. Expanded Manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service, 1965.

__________. and L.M. Dunn. Manual for the Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service, 1981.

Evans, C. The Micro Millenium. New York: Pocket Books, 1979.

Evans, R.I. Jean Piaget: The Man and His Ideas. New York: E.P. Dutton & Co., Inc., 1975.

Eysenck, H.J. Uses and Abuses of Psychology. Great Britain: Penguin, 1953.

French, J.L. Manual: Pictorial Test of Intelligence. Boston: Houghton Mifflin, 1964.

Gardner, H. Frames of Mind. Basic Books, Inc., 1983.

Guilford, J.P. "Intelligence: 1965 Model." American Psychologist, 21 (1966), 20-26.

__________. The Nature of Intelligence. New York: McGraw-Hill, 1967.

Herrnstien, R.J. "IQ Testing and the Media." The Atlantic, August (1982), 68-74.

Jensen, A. Bias in Mental Testing. New York: The Free Press, 1980.

Kaufman, A.S. and N.L. Kaufman. Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service, 1983.

Kelly, T.L. "The Selection of Upper and Lower Groups for the Validation of Test Items." Journal of Educational Psychology. 30 (1939), 17-24.

Pinter, R. "Intelligence and its Measurement." Journal of Educational Psychology, 12 (1921), 139-143.

Raven, J.C. Guide to the Standard Progressive Matrices. London: Lewis, 1960.

Sattler, J.M. Assessment of Children's Intelligence; Philadelphia: W.B. Saunders Co., 1974.

Shertzer, B. and J.D. Linden. Fundamentals of Individual Appraisal. Boston: Houghton Mifflin Co., 1979.

Spearman, C. The Nature of Intelligence and the Principles of Cognition. New York: Arno Press, 1973.

Taylor, E.M. Psychological Appraisal of Children with Cerebral Defects. Cambridge, MA: Harvard University Press, 1961.

Terman, L.M. The Measurement of Intelligence. Boston: Houghton Mifflin Co., 1916.

_________. "Intelligence and its Measurement." Journal of Educational Psychology, 12 (1921), 127-133.

Terman, L.M. and M.A. Merrill. Measuring Intelligence. Boston: Houghton Mifflin Co., 1937.

_________. Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin Co., 1960.

Thorndike, E.L. "Intelligence and its Measurement." Journal of Educational Psychology, 12 (1921), 124-127.

Thurstone, L.L. The Differential Growth of Mental Abilities. Chapel Hill, NC: University of North Carolina, Psychometric Laboratory, 1955.

Torrance, E.P. and W.F. White (Eds.) Issues and Advances in Educational Psychology. Itasca, IL: F.E. Peacock Publishers, Inc., 1969.

Vernon, P.E. The Structure of Human Abilities. London: Methuen, 1961.

Wechsler, D. Wechsler Adult Intelligence Scale Manual. New York: Psychological Corporation, 1955.

_________. The Measurement and Appraisal of Human Intelligence. Baltimore: The Williams and Wilkins Company, 1958.

_________. Manual for the Wechsler Preschool and Primary Scale of Intelligence. New York: Psychological Corporation, 1967.

_________. Manual for the Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation, 1974.

_________. "Intelligence Defined and Undefined: A Relativistic Appraisal." American Psychologist, 30 (1975), 135-139.

_________. Manual for the Wechsler Adult Intelligence Scale-Revised. New York: Psychological Corporation, 1981.

APPENDIX

THE RAW DATA

Subjects	Age	Sex	IQ	Model Item Score
001	7	F	82	13
002	8	F	71	16
003	10	F	81	14
004	16	F	105	27
005	8	F	101	21
006	8	F	119	26
007	14	F	101	27
008	8	F	130	27
009	10	F	40	10
010	8	F	100	22
011	15	F	73	26
012	7	F	98	14
013	10	M	106	27
014	7	M	90	17
015	11	M	105	27
016	13	M	111	27
017	12	F	86	24
018	11	M	84	25
019	6	M	103	19
020	13	F	131	27
021	12	F	82	27
022	11	M	121	27
023	7	M	97	17
024	9	M	123	27
025	12	M	104	27
026	12	M	107	27
027	7	M	119	19
028	7	F	107	23
029	8	M	101	24
030	12	M	87	26
031	15	F	107	26
032	12	M	100	27
033	11	F	119	26
034	8	M	115	25
035	8	M	103	21
036	8	M	115	27
037	6	F	85	12
038	9	M	74	16

Subjects	Age	Sex	IQ	Model Item Score
039	8	M	93	21
040	11	F	83	23
041	6	M	92	13
042	7	M	121	24
043	12	M	74	22
044	9	F	98	23
045	7	M	83	15
046	9	F	134	26
047	6	M	115	18
048	10	M	110	23
049	10	M	118	27
050	8	F	118	25
051	9	M	88	25
052	11	F	80	23
053	8	M	94	23
054	12	M	87	26
055	8	M	105	21
056	11	M	80	25
057	9	F	99	25
058	10	F	106	27
059	7	F	98	15
060	10	F	130	27
061	11	M	72	24
062	13	M	120	27
063	11	M	85	25
064	10	M	100	23
065	9	M	133	27
066	8	F	105	26
067	9	F	78	16
068	6	M	96	14
069	11	M	133	27
070	11	M	101	26
071	9	M	91	24
072	8	M	106	24
073	15	M	101	27
074	7	M	123	27
075	7	M	100	19
076	8	M	110	26
077	8	M	87	20
078	7	F	98	15
079	9	M	93	24
080	14	M	127	27

Subjects	Age	Sex	IQ	Model Item Score
086	6	F	96	16
087	8	F	114	26
088	9	F	70	22
089	8	F	79	21
090	7	F	108	22
091	13	F	82	13
092	7	F	76	13
093	10	F	100	25
094	7	F	76	12
095	9	M	89	26
096	13	M	107	27
097	9	M	112	25
098	10	F	86	26
099	11	M	96	26
100	12	M	102	26
101	10	M	88	24
102	7	M	101	20
103	11	M	101	25
104	15	F	109	27
105	8	F	110	27
106	7	M	67	13
107	11	M	98	26
108	6	F	100	22
109	7	M	92	19
110	8	F	76	15
111	8	M	96	20
112	9	M	85	21
113	9	F	88	23
114	9	F	98	26
115	9	M	76	20
116	8	M	97	18
117	8	F	123	26
118	12	M	87	22
119	8	M	88	0
120	6	M	123	21
121	6	M	131	21
122	6	M	107	17
123	12	M	92	14
124	12	F	116	27
125	11	M	121	27
126	10	M	100	24
127	10	F	81	22
128	9	M	101	21
129	8	F	115	24
130	8	F	103	20

Subjects	Age	Sex	IQ	Model Item Score
131	8	F	81	19
132	10	M	108	26
133	7	M	123	22
134	15	F	54	24
135	7	M	89	17
136	7	M	84	15
137	6	M	96	14
138	9	M	108	24
139	8	M	109	24
140	14	M	90	27
141	6	M	112	14
142	13	M	112	26
143	9	F	92	23
144	13	M	107	26
145	11	F	133	27
146	10	M	69	25
147	7	M	64	9
148	15	M	81	27
149	12	F	113	27
150	13	F	59	19
151	8	M	99	24
152	7	F	98	21
153	10	M	114	26
154	6	M	87	12
155	11	M	91	24
156	9	F	93	20
157	11	F	89	25
158	6	M	101	12
159	10	M	143	27
160	7	M	84	20
161	11	M	131	27
162	9	F	91	26
163	8	M	90	8
164	10	M	91	27
165	7	M	84	19
166	15	M	112	27
167	7	M	71	13
168	7	M	86	14
169	10	M	98	21
170	9	M	111	22
171	7	M	93	16
172	6	F	89	18
173	7	M	134	25
174	7	F	75	17
175	13	M	98	25

Subjects	Age	Sex	IQ	Model Item Score
176	9	M	115	24
177	7	M	111	17
178	9	M	77	21
179	7	F	107	21
180	9	M	102	22
181	13	M	76	25
182	11	F	42	16
183	9	M	93	26
184	8	M	83	18
185	8	M	95	22
186	7	M	67	11
187	6	M	119	18
188	11	F	94	26
189	6	M	88	19
190	12	M	89	25
191	8	M	100	20
192	6	M	80	15
193	7	M	112	23
194	14	M	103	27
195	12	F	87	25
196	11	F	40	11
197	13	F	86	26
198	16	M	101	27
199	6	M	103	19
200	8	M	105	21

2

VITA

Karl Hoffman

Candidate for the Degree of

Doctor of Philosophy

Thesis: A STUDY OF A KNOWLEDGE-BASED MODEL FOR THE THOERY AND MEASUREMENT OF INTELLIGENCE WITH A WISC-R DATA SAMPLE

Major Field: Applied Behavioral Studies

Biographical:

Personal Data: Born in Newark, New Jersey, September 19, 1957. Resident of Pennsylvania since 1971.

Education: Graduated from Crestwood Area High School, Mountaintop, Pennsylvania, in 1975; received Associate of Science Degree in Social Studies from Luzerne County Community College, Nanticoke, Pennsylvania, in 1977; received Bachelor of Arts Degree in English and Sociology from King's College, Wilkes-Barre, Pennsylvania, in 1979; received Master of Arts Degree in Psychology from Marywood College, Scranton, Pennsylvania, in 1981; completed requirements for the Doctor of Philosophy Degree at Oklahoma State University in May, 1986.

Professional Experience: Residential Program Supervisor, United Cerebral Palsy of Luzerne County, Wilkes-Barre, Pennsylvania, April, 1980, to August, 1981; Research Associate, Department of Applied Behavioral Studies, Oklahoma State University, August, 1982, to September, 1983; Residential Program Supervisor, Pan Am Group Homes, Inc., Bethlehem, Pennsylvania, November, 1983, to September, 1984; Psychological Services Associate II, Bradford and Sullivan Counties Mental Health/Mental Retardation Program, Towanda, Pennsylvania, September, 1984, to present.

Professional Affiliation: Member of the American Educational Research Association.