

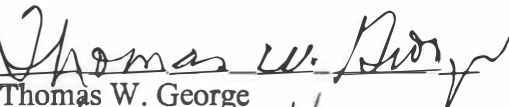
To the Graduate Council:

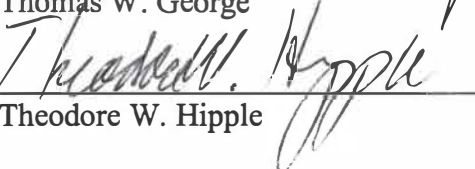
I am submitting herewith a dissertation written by Shannon Kathleen Hays entitled "A computer administered version versus paper-and-pencil administered version of the MMPI-A." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.


R. Steve McCallum, Major Professor

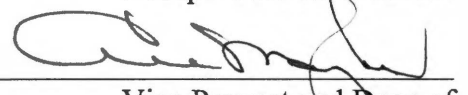
We have read this dissertation
and recommend its acceptance:


William H. Calhoun


Thomas W. George


Theodore W. Hipple

Accepted for the Council:


Vice Provost and Dean of
Graduate Studies

**A COMPUTER ADMINISTERED VERSION VERSUS
PAPER-AND-PENCIL ADMINISTERED VERSION OF THE
MMPI-A.**

**A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**Shannon Kathleen Hays
December 2002**

Copyright © 2002. Shannon K. Hays. All rights reserved.

Thesis
2002b
.H392

DEDICATION

This dissertation is dedicated to my parents, Bill Hays and Martha Hays, for their unwavering support and love through this process and during my life, and to the rest of my family and friends for always believing in me and encouraging me to reach my goals.

ACKNOWLEDGMENTS

I wish to thank all those who have helped me in completing my Doctor of Education. I wish to thank Dr. McCallum for all of his help, support, guidance, and encouragement. I thank Dr. Calhoun, Dr. George, and Dr. Hipple for serving on my committee. I want to thank Nancy Day for all of her selfless help with the statistical portion of my dissertation. I thank Ilene Cornwell for editing my dissertation. I wish to thank Becky Bledsoe and the Educational Psychology and Counseling staff for help with forms and deadlines. I also thank Millie Cheatham for her help with securing computer rooms and keys during the data collecting process. Finally, I would like to thank the Metropolitan Nashville School System and Dr. Julie Williams, Sue Fellman, and Fred Whitley for allowing me to use students in the high school.

ABSTRACT

Within the context of a counterbalanced design, 102 students from a high school and a large university in the southeast were administered two versions of the Minnesota Multiphasic Personality Inventory – Adolescent (MMPI-A): a computer-administered version (CA) and a paper-and-pencil version (PAP). Time between testing sessions was approximately one week. Differences in individual scale means between the CA and PAP were calculated using paired *t*-tests, with the Bonferroni correction procedure; no mean differences were significant ($p. > .05$). To determine if the scale distributions were similar, tests of homogeneity of variance were conducted using Hartley's homogeneity of variance tests; there were no differences in the shapes of the scale distributions ($p. > .05$). Pearson product-moment coefficients were calculated for each scale to determine if the relative rankings were similar; coefficients for every scale were positive and statistically significant ($p. < .01$). Implications of the findings of this study are discussed.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Purpose	1
Rationale	1
Literature Review	2
Establishing Equivalence	2
Anxiety Effects	5
Social Desirability Effects	5
Studies Showing No Difference	7
PAP Mean Scores Higher Than CA Scores	11
CA Mean Scores Higher Than PAP Scores	12
Statement of the Problems	13
Research Questions and Hypotheses	13
II. METHOD	14
Participants	14
Instruments	15
Procedure	18
Data Analyses	21
III. RESULTS	22
Research Question 1	22
Research Question 2	27
Research Question 3	27
IV. DISCUSSION	29
Comparison of the Individual Scale Means	29
Distributions of Scores	31
Rankings of Scores	32
Implications	33
Limitations of the Study and Suggestions for Future Research	35
REFERENCES	38
APPENDICES	45
Appendix A	46
Appendix B	49
Appendix C	52
Appendix D	55
VITA	58

CHAPTER I

INTRODUCTION

Purpose

The purpose of this study was to compare the paper-and-pencil (PAP) version of the Minnesota Multiphasic Personality Inventory - Adolescent (MMPI-A) with a computer-administered (CA) version of the MMPI-A. Comparison is important for test users. Even though CA personality testing is quite popular because of reduced costs, automatic scoring, and lower ambiguity, it cannot be assumed that a test administered by a computer is equivalent to the PAP version. The American Psychological Association has recommended that equivalence between PAP versions and CA versions of tests be addressed empirically and not taken for granted. Studies comparing CA of the adult version of the MMPI with the PAP version of the MMPI are available; however, there are no studies comparing the two administrative formats for the adolescent version. This study will test equivalence of two formats for the MMPI-A.

Rationale

Computer-assisted testing has been around since the 1950's (Ben-Porath and Butcher, 1986). Many authors have noted the advantages of computer-assisted testing. Webster and Compeau (1996) state that computer-assisted testing has several advantages: increased availability of tests, automatic scoring and analysis of data, higher test security, and easier generation of customized questionnaires. Other advantages suggested by Bugbee (1996) include: reduction of test time, ability to obtain more information about test takers, and ease of scheduling. Green (1983) states that computer testing also keeps

testing on target when it adapts to the clients' ability levels in computer-adaptive testing, eliminates messy answer sheets, and allows the client to work at his or her own pace. Computers also permit new types of material. For example, video clips can be added to police academy entrance exams to simulate real life situations (Green, 1983).

Because the MMPI-A is used as a personality test in a clinical setting, it may be more efficient to give the clients this test on a computer. However, it is important to determine the extent to which results from the CA version and the PAP version can be considered equivalent, so that appropriate conclusions can be drawn.

Literature Review

Computer-Administered Versus Paper-and-Pencil-Administered Testing

Establishing Equivalence

In 1973 Meehl noted that a computer program has advantages for clinicians. There are several advantages: 1) a computer is objective; 2) a computer can score and interpret psychological tests far faster; 3) a computer is more cost efficient and; 4) a computer is more reliable. Test-retest reliability is always 1.0 with a computer. The computer will always assign the same scores to the protocol (Ben-Porath and Butcher, 1986).

For the CA test and the PAP test to be considered equivalent, research must be conducted. If differences are found, then Webster and Compeau (1996) suggest that the use of a different administration mode may result in the measure of a different construct. Webster and Compeau (1996) state that the CA test and the PAP test are equivalent if

they “produce equal mean scores, identical distribution and ranking of scores, and correlate to the exact degree with scores on other variables” (p. 567).

Bugbee (1986) discusses APA guidelines for the development, use, and interpretation of computerized testing. CA and PAP are equivalent if “a) rank orders of scores of individuals tested in alternative modes closely approximate each other, and b) the means, dispersions and shapes of the score distributions are approximately the same by rescaling the scores from the computer” (p. 286). If a CA version of a test is intended to serve solely as an alternative form of a test, then it is sufficient that the two test versions yield score means, variances, and correlations with other measures that are approximately equivalent.

There are some general conclusions that can be drawn about CA testing. The first conclusion is that CA and PAP tests can be equivalent, but it is the responsibility of the test developer to show that they are; there is no built-in equivalence between these two forms of administration. Second, equivalence of tests is established by either meeting the criteria for alternative forms or parallel forms. Third, the use of computers can affect test results. For example, according to Sutton (1991), computers may be biased against certain economic classes or race. Researchers should try to learn what these affects are and consider them when drawing conclusions. Fourth, special considerations must be made when computers are used in testing. For example, administrators must be sure that the location accommodates the needs of the test taker. Fifth, users must know psychometric properties of the CA format and have a basic understanding of computer applications to effectively utilize and interpret computer-based testing (Bugbee, 1996). Finally, if

differences between the PAP and CA are found, the researcher should demonstrate that the CA test yields information that is reliable and useful beyond what is found in the PAP (Biskin and Kolotkin, 1977).

There are many explanations for why format differences can affect results. These include test-taker frustration and inability to backtrack during computer testing, increased attention focused on individual items when they are presented singly during computer testing, difficulty in making an initial cursor placement, and purposeful consistency in answers due to backtracking during PAP testing. Also, data collected with computers may be biased for the computer-anxious test taker (Webster and Compeau, 1996).

It is important to conduct multiple analyses when studying CA versus PAP equivalency. For example, as implied earlier, the first analysis may indicate no mean differences; however, reliabilities may differ, as may other concurrent validity measures. It is not appropriate to compare results of tests that have data collected from CA and PAP if the extent of the equivalency has not been established. Researchers should continue to explore individual differences such as computer anxiety and age that might limit the use of a CA test (Webster, and Compeau, 1996).

Hofer (1985) discussed three ways to deal with score differences, should they occur. If the differences are simple mean differences, then a constant can be added to the scores. If the differences are distribution differences, then equipercentile conversion of scores can be done. If the differences are different ranks, then it is likely that the same construct is not being measured.

Anxiety effects

Several types of effects may be present during CA and PAP testing, including client anxiety. Paul (1982) estimated that at least 30% of the community dealing with computers on a daily basis experience some form of anxiety about the computer. Weeter and Halcomb (1986) found that 24% of undergraduate students surveyed reported Computer Anxiety Scale scores that indicated some anxiety toward computer use. George, Lankford, and Wilson (1992) found that high levels of computer anxiety were associated with higher depression scores when measuring this trait by computer.

Two studies have found no anxiety differences. Lushene, O'Neil and Dunn (1974) researched anxiety before and after CA and PAP testing sessions. State anxiety was measured before and after the testing sessions. Higher client anxiety on the CA version was present before testing, but by the end of the testing sessions, no significant differences between the CA and PAP versions were found. Hart and Goldstein (1985) did not find any significant difference in self-reported anxiety, task satisfaction, or electromyograph measures of individuals exposed to short-term CA when compared to individuals exposed to short-term PAP administrations. CA assessment was found to be more speed efficient and more conducive to eliciting openness from clients.

Social desirability effects

Another source of potential difference is the anonymity of the computer. Some authors believe that anonymity may facilitate obtaining more accurate personal information revealed by clients and reduce the tendency to respond in a socially desirable way. Greist and Klein (1980) and Synodinos, Papacostas, and Okimoto (1994) suggested that answers

to sensitive and potentially embarrassing questions can be obtained more readily on CA surveys. This tendency could result in a difference in test results especially on more sensitive tests, such as on alcohol-abuse surveys and personality tests (Webster and Compeau, 1996).

Two studies have found no differences in social desirability effects with CA and PAP tests. Finegan and Allen (1994) examined social desirability issues, examining whether clients are more willing to reveal personal information about themselves if the question is asked by a computer. The authors found no main effects for administration mode on the reaction measures of any social desirability scales. Ford, Vitelli, and Stuckless (1996) studied CA versus PAP administration format and social desirability in a violent male inmate population. Social desirability was measured using the Marlowe-Crowne Social Desirability Scale Form-C. No differences were found between the CA and PAP using this measure.

Several studies have found reduced social desirability effects on CA tests. Lucas, Mullin, Luna, and McEnroy (1977) found that patients at an alcohol treatment center reported more alcohol consumption when given a computerized interview than they did when interviewed by a person. Martin and Nagao (1989) asked clients to play the role of job applicants who were interviewed either by a person or by a computer. They used the Marlowe-Crowne Social Desirability Scale Form-C. They found reduced social desirability effects with the computer. Peterson, Johannsson, and Carlsson (1996) found that the CA version of the Beck Depression Inventory (BDI) caused elevated scores on questions that were determined to be “sensitive” in nature. These items included questions about guilt

feelings, sexual feelings, and thoughts of suicide. The authors concluded that these effects related to reduced social desirability effects. Peters, Clark, and Carroll (1998) compared the CA Composite Internal Diagnostic Interview-Auto (CIDI) with the CIDI administered by a human interviewer. The effects of computer attitudes, computer experience, and social desirability were examined. The results indicate that fewer clients in this study felt embarrassed in revealing information about their symptoms to the computer than to the human interviewer.

Contrary to the aforementioned findings, Lautenschlager and Flaherty (1990) found that the computer enhanced clients' tendency to respond in a socially desirable way on the Paulhus Balanced Inventory of Desirable Responding. They suggested that this might be because a client can easily review his or her answers on a written questionnaire. However, the computer flashes the next question on the screen after a response has been selected.

Studies showing no differences

Thus far, research conducted on CA and PAP equivalence has been inconsistent. Some research shows no differences. In an early study, Hoffman and Lundberg (1976) compared a particular type of computerized system, called a group-response system, with a traditional mode of test administration. Group-response system permits the simultaneous on-line recording of responses of large groups of students. Students were not allowed to see items past the allotted amount of time, and the sequence of the items could not be changed. The two modes did result in equivalent scores and test-taking behavior for true-false and multiple-choice items. Although Greaud and Green (1986) found no overall

effects, they did detect some subtle differences based on the way PAP tests were transposed to a computer setting.

Vansickle, Kimmel, and Kapes (1989) compared CA and PAP for the Strong-Interest Inventory. They found that the PAP and the CA of the Strong-Interest Inventory correlated highly with each other; in addition, clients reported that they were comfortable with the use of the computer.

Two studies have found no differences in CA and PAP questionnaires. Rafaeli and Tractinsky (1989) found no major differences of CAQ in speeded IQ tests, but the presence of a visual cue, in this case an hourglass, promoted fewer unanswered questions. Liefeld (1992) used two versions of PAP scales and three CAQ in a market survey but saw no effects on either scale.

Several studies comparing CA and PAP personality tests have found no differences. Rosenfeld, Dar, Anderson, Kobak, and Greist (1992) examined CA Yale-Brown Obsessive Compulsive Scale with PAP Y-BOCS. They found that the CA version correlated highly with the clinician-administered version. Bader, Hofman, and Kubinger (1993) using a reliability design found no mode of administration differences on the six scales of the Giessen Test, which is a personality test commonly used in German speaking countries.

Rasulis, Schuldberg, and Murtagh (1996) researched the equivalence of CA and PAP of the Rotter Incomplete Sentence Blank College Form. The effects of mode of administration were small and, according to the authors, can be ignored. Nonsignificant format differences were found for standard deviations and internal consistency data based

on maladjustment scores. Merten and Ruch (1996) compared the Eysenck Personality Questionnaire - Revised in CA and PAP form. Questionnaires were divided into two similar halves with each client replying to one-half in a standard way and to the other half at the computer. No systematic differences were found for the PAP and CA.

Some studies conducted on the MMPI and the MMPI-2 have found no differences in modes of administration. Honaker, Harrell, and Buffaloe (1988) evaluated the psychometric and experiential equivalence of the MICROTEST MMPI program developed by National Computer Systems. Clients were administered the PAP MMPI and the CA MMPI in a repeated measures counterbalanced design. Two groups were given both the CA MMPI and the PAP MMPI. A third group was given two CA MMPIs. A final group was given two PAP MMPIs. These results offered tentative evidence that the CA MMPI yields equivalent results to those obtained by the PAP MMPI. For standard validity and clinical scales, as well as several supplementary scales, no significant differences were found across administration modes in mean T-scores, standard deviations, distributions, or rankings. Also, test-retest reliability for the CA was similar to the PAP and was comparable to that reported previously for the PAP MMPI. Clients indicated an overall preference for the CA and tended to view the computer as quicker and more comfortable.

Schuldberg (1988) also used item level-analyses to examine mode of administration effects for the MMPI. This study used all 566 questions and analyzed them individually. Clients were grouped according to their different orders of test administration. Cannot Say responses and blank responses were counted as missing data.

Eight of the clinical scales showed significant differences with regard to time when comparing PAP versus CA. Schuldberg (1988) found that item-level effects of mode of administration occurred, but these effects are relatively small.

Watson, Manifold, Klett, Brown, Thomas, and Anderson (1990) used 200 clients to compare the CA MMPI with the PAP MMPI. The groups were administered the tests in a counterbalanced design. The authors compared differences between the mean scores on the 13 scales by using a Group X Administration, Latin-square multiple analysis of variance, and *t* tests on each scale. They found no across-format differences in mean scores, distributions, or frequency of invalid profiles for this study and support the use of CA.

Pinsonneault (1996) examined the effects of a CA MMPI-2 versus the traditional booklet MMPI-2. He reviewed 12 previous studies comparing the CA MMPI and the PAP MMPI; only two studies reported more than three scales that were different. In fact, four studies reported no scale differences; four studies reported one scale that was different; one study reported two scales that were different; one study reported three scales that differed.

Pinsonneault (1996) selected 32 master/doctoral-level students for his study. The final number of clients was 30. The MMPI-2 softcover booklet form, the CA MMPI-2, and a brief five-point Likert scale attitude survey were the instruments used. IBM computers were used. Clients were randomly assigned to two groups. One group took the CA first, and the other group took the PAP first. Time between testing ranged from three days to two weeks. T-scores were reported. Differences in means were investigated.

None of the 10 clinical or three validity scales differed by either format or by administration order. Equivalent results were found for the Cannot Say Scale. Possible score distribution differences across formats were investigated using the Bartlett-Box Homogeneity Test. None were found to differ significantly. Test-retest correlations were computed across the variables of administration format and administration order. The two forms were acceptable.

Clients in Pinsoneault's 1996 study reported that they felt more comfortable with the pace of the CA and found it more interesting and less difficult. Results of this study found the two formats to be quite comparable for the MMPI-2. Neither the validity nor the clinical scales differed by format using means, standard deviations, distribution shapes, nor criteria. This study supports the conclusion that a CA MMPI-2 yields scores that do not differ significantly from the traditional format, but it does not address the question of whether small differences might exist that allows the performance of the CA MMPI-2 to be improved if CA norms were developed (Pinsoneault, 1996).

PAP mean scores higher than CA scores

Some studies have reported higher PAP mean scores than CA mean scores on personality tests including the MMPI. Lushene, et al. (1974) found that CA has resulted in significant differences on Hypochondriasis, Psychopathic Deviant, Paranoia, Schizophrenia, Lie, and Defensiveness Scales and that the profiles obtained from the booklet MMPI tended to be more elevated.

Biskin and Kolotkin (1977), in an initial study and a replication, compared the PAP MMPI and the CA MMPI. They selected 126 male undergraduate students for the initial

study and 39 male undergraduates for the replication from introductory psychology courses. At the profile level, differences were assessed by Multivariate Analyses of Variance (MANOVAs) using raw scale scores. Analyses of Variance (ANOVAs) were used for each scale, and chi-square analyses were performed on each item. Significant mean differences were found on the Cannot Say Scale and the Paranoia Scale with higher means found on the PAP group. Differences on the Paranoia Scale were small. Examination of the scale means revealed that the largest differences among the groups occurred on the Cannot Say scale. They concluded that the F ratio associated with the Cannot Say scale was so large that it alone could have accounted for the mean profile differences.

In a meta-analysis which reviewed the results of several studies, Watson, Thomas, and Anderson (1992) found that CA of the MMPI yielded significantly lower estimates on eight of 10 clinical scales when compared to PAP.

CA mean scores higher than PAP scores

One study shows higher mean scores on CA personality tests. Sukigara (1996) found significant differences on the Depression, Paranoia, Psychasthenia, and Schizophrenia scales on the MMPI with the CA scores higher. The nonequivalent scales, Depression, Paranoia, Psychasthenia, and Schizophrenia are not the same as those found to be nonequivalent in studies by previous studies. Three reasons may be responsible for this difference: 1) Japanese people may be less familiar with a keyboard; 2) the kinds of hardware and software could have affected the results; 3) random sampling error across studies might have caused contradictions.

Statement of the Problem

Computers have become a popular way for clinicians and other educators to give tests. For example, administration of the Scholastic Aptitude Test, the Graduate Record Exam, and many licensing exams, such as the licensing exam for nurses, are all available to take on the computer. Since CA has many advantages such as cost effectiveness, speed, and accuracy, it is possible that many psychologists in private practice and hospital settings will be employing this method of administration for clients taking personality tests, and particularly versions of the MMPI such as the MMPI-A. It is important that the equivalence for the PAP and the CA be established before the computerized method is used. There has been no research conducted comparing the MMPI-A CA to the traditional format; consequently, this study is designed to investigate whether differences exist.

Research Questions and Hypotheses

The following research questions and hypotheses were addressed by this study:

- I. Are there differences between the CA means and the PAP means on the specific scales?
- II. Are the distributions of scores for the CA MMPI-A and the PAP MMPI-A equivalent?
- III. Are the relative rankings of scores for the CA MMPI-A and the PAP MMPI-A similar? Are the correlation coefficients statistically significant?

CHAPTER II

METHOD

Participants

Participants numbered 102 and included either 18-year-old undergraduate students enrolled in an introductory psychology class at a large university or high school students from one general-education and three special-education classes, all from the southeast. Participation was voluntary. Participants included: one 14-year-old male, one 14-year-old female, five 15-year-old males, eight 15-year-old females, nine 16-year-old males, five 16-year-old females, seven 17-year-old males, five 17-year-old females, twenty-seven 18-year-old males, and thirty-four 18-year-old females. Of the high school students, there were: 10 learning disabled females, 12 learning disabled males, two mildly mentally retarded females, one hearing impaired male, one health impaired female, one health impaired male, and one language impaired male. Overall, the percentage of the special education students in the sample was 27.45. This number is approximately 20% higher than the national average for school age children receiving special education services (Eggen and Kauchak 2001). The mean age of the participants was 17.14, and the standard deviation was 1.19. The college participants lived at home with their parents and thus qualified to be administered the MMPI-A instead of the MMPI-2. Participants were told that if taking the test causes them to feel uncomfortable or if they needed to talk to someone after the testing sessions, then they could talk to the test administrator, make an appointment with the university counseling center for counseling, or could speak with a guidance counselor at the high school. Importantly, the sample size of 102 is slightly

above the minimum described by Honaker et al (1988) as sufficient to ensure adequate power.

Instruments

The MMPI-2 was designed to assess psychopathology of individuals 18 years old or older. The MMPI-A was developed to assess adolescents' ages 14 to 18, inclusive. An 18-year-old client can be assessed with either the MMPI-2 or the MMPI-A. The MMPI-A manual states that the MMPI-2 should be used for 18-year-olds who are living an independent lifestyle (Butcher, Williams, Graham, Archer, Robert P., Tellegen, Ben-Porath, and Kaemmer, 1992).

The MMPI-A is an instrument designed to assess psychopathology in adolescents. This instrument represents the first revision of the MMPI for use in the evaluation of adolescents. The MMPI-A is designed after the original MMPI. The MMPI-A can be administered to individuals or groups in a hospital or clinical setting (Butcher et al., 1992). The MMPI-A contains 478 items generated from the original MMPI (Archer, 1997).

The MMPI-A has 10 clinical scales. Scale 1 (Hs) is the Hypochondriasis Scale and consists of items selected to characterize respondents with a history of symptomatology characteristic of hypochondriasis. Adolescents with elevations on this scale have preoccupation with health and illness. Scale 2 (D) is the Depression Scale and contains items used to identify respondents with a general dissatisfaction with life, including feelings of discouragement, hopelessness, and low morale. Scale 3 (Hy) is the Hysteria Scale and is used to identify respondents who react to stress situations with hysterical reactions that include sensory or motor problems with no organic base. Scale 4 (Pd) is the

Psychopathic Deviant Scale and is used to identify adolescents with a response pattern of lying, stealing, sexual promiscuity, and alcohol abuse. Scale 5 (Mf) is the Masculinity - Femininity Scale and is used to identify an endorsement of an unusual pattern of stereotypically feminine interests in boys and an endorsement of unusual pattern of stereotypically masculine interests in girls. Scale 6 (Pa) is the Paranoia Scale and is used to identify adolescents manifesting paranoid symptomatology including suspiciousness, feelings of persecution, and rigidity. Scale 7 (Pt) is the Psychasthenia Scale and measures a neurotic syndrome most closely related to obsessive-compulsive disorder. Scale 8 (Sc) is the Schizophrenia Scale and measures adolescents with various forms of schizophrenia such as paranoid and schizoid personality types. Scale 9 (Ma) is the Hypomania Scale and identifies patients with hypomanic symptoms, such as grandiosity, irritability, flight of ideas, egocentricity, elevated mood, and cognitive and behavioral overactivity. Scale 10 (Si) is the Social Introversion Scale and is designed to measure adolescents with extreme scores on social extroversion or social introversion (Butcher et al., 1992). Scale 10 is the only scale developed outside the original Hathaway group and where no psychiatric criterion-group was obtained (Archer, 1997).

The MMPI-A has validity scales used to detect deviant test-taking attitudes or response sets. Validity scales on the MMPI-A include the Cannot Say Scale (?), the Lie Scale (K), the Infrequency Scale (F), and the Defensiveness Scale (K). The Cannot Say Scale includes items omitted or marked both true and false. The Lie Scale is designed to detect attempts of examinees to present themselves in a favorable light, particularly regarding personal ethics or social behavior. The Infrequency Scale is designed to detect

attempts of adolescents trying to present themselves in a bad light. The Defensiveness Scale is designed to detect examinees with severe pathology but who have produced profiles within the normal range (Archer, 1997).

According to the manual, the internal consistency reliability of the clinical scales ranges from .40 - .90, depending on which scale is used. Internal consistency for content scales ranges from .55 - .83 in the norm sample and .63 - .89 in a clinical sample. Test-retest reliability on clinical scales ranges from .65 - .84. Test-retest reliability of the content scales ranges from .62 - .82. The typical standard error of measurement for the clinical scales is four to six T-score points (Butcher et al., 1992).

The computerized version is given by the MICROTEST Q computer program. Once the program has been initiated, the computer asks for the client's demographic data including name, identification number, birth date, test date, sex, and setting. The F1 key is the help key and explains how to select answers. The F2 key is used to return to previous questions. This key will back up as far as the client wishes. The F3 key is used to skip a question or to erase an answer. The F4 key is used to continue with the next screen. After the test administrator enters the demographic data, the client sees a screen with instructions. The instructions tell the client that the computer will present questions. To answer a question, the client can double click with the mouse on an answer or can hit the space bar to highlight an answer, then press the enter key to select an answer. An example item is presented. After the client answers the sample item, he or she uses the mouse to double click on "OK" to continue. Then the instructions for the test are presented on the computer's monitor. The client is to answer true if the item applies mostly to him or her.

The client is to answer false if the item does not apply mostly to him or her. The client is encouraged to answer as many questions as he or she can. Then the client uses the mouse to double-click on "OK" to continue. Then the first item is presented. When the client reaches the end of the test, the computer instructs the client to contact the test administrator (K.M. Bartels, personal communication, August 11, 2000).

Procedure

College participants were recruited by requesting volunteers from the campus of a large southeastern university; participants were given extra credit in a large introductory section of psychology. The researcher's e-mail address and telephone number were provided if any potential participants had questions.

A computer lab on campus contained approximately 25 computers in rows of five. Each computer was an IBM loaded with Windows 98 or above. The researcher loaded the MICROTTEST Q software on each computer before the participants were scheduled to arrive. Since the MICROTTEST Q program has many test-administration options which required much computer memory and slowed item presentation, the researcher loaded only the specific test (MMPI-A) into the computer.

All participants were assigned to the same room, whether they received PAP or CA. The researcher and a proctor randomly assigned each participant to a group. This was accomplished by assigning the first person to the CA group first and the second person to the PAP group first and alternating participant assignment until everyone had been assigned to a group. Each participant was given two identical copies of a consent form, one to return to the researcher and one to keep. The researcher then explained the nature

of the study. The participants were told that if any of the questions made them feel uncomfortable, then they could talk to the test administrator or make an appointment at the university's counseling center. The participants were told that they would not receive their test results. The participants were allowed to ask questions.

Then the researcher distributed the PAP version of the MMPI-A to the PAP group and explained how to complete the front page of the answer form. Each participant was given a code to write on their protocol or enter into the computer to show if they were answering the CA first or the PAP first. The researcher explained to the CA group how to enter demographic data and answer questions on the computer to the CA group. The CA group was told to let the researcher know when they finished answering, so that the researcher could save the data. Participants were allowed to ask questions.

When each participant finished the test, the researcher either looked over the PAP protocol or saved the CA data. The participant then wrote down his or her name, social security number, psychology class instructor, and signed up for a time for the second testing session. All participants were urged to call or e-mail the researcher if they had questions.

To maintain test security and participant anonymity, the data from each CA participant was saved to a disk using the transfer data procedure of the MICROTTEST Q program. After all data was saved and transferred to a disk, the researcher deleted the MICROTTEST Q program from each of the computers.

For the second testing session, the researcher again loaded the MICROTTEST Q program before the participants were scheduled to arrive. When a participant entered the

testing room, he or she was asked which administration version had already been taken.

Then the participant was assigned to answer the opposite administration version.

Instructions for completing the PAP form and entering data of the CA version were repeated. Participants were allowed to ask questions. After the participants completed the test, they were told that their names and social security numbers would be sent to their class instructors, so they would receive their extra credit. Again, the MICROTEST Q program was deleted from the computers after the data had been saved.

High school participants were recruited by the researcher speaking to three special-education classes and one-general education class. The researcher introduced herself to the class and explained her role at the school. The researcher told the class why she needed participants for this study. The teacher of the special-education classes informed the students that the researcher would supply pizza and soft drinks on a Friday morning near the end of the school year for participants. The teacher of the general-education class informed the students that participation in this study would allow them a perfect quiz grade score (100) added to the end of their semester averages. Parental permission forms were explained to them. All interested students were given a copy of the parental consent form and an assent form, and were asked if they had questions. The students were also informed that they would need to arrange to stay after school to participate. After parental consent forms and assent forms were turned in, the researcher allowed students to sign up for times to participate.

A computer lab in the high school was secured by speaking to the teacher who uses the lab. The computer lab contained approximately 15 computers in rows of five.

Each computer was an IBM loaded with Windows 98 or above. The researcher loaded the MICROTTEST Q software on each computer before the participants were scheduled to arrive. The procedure and the directions previously used for the college participants were repeated for the high school participants. There was not a second proctor present during the high school participants testing sessions.

All participants took both the CA and the PAP versions of the MMPI-A in a counterbalanced order. Time between testing was about one week. Participants were assured that their responses would be kept confidential.

Data Analyses

The following statistical procedures were performed:

- I. The individual scale mean differences from the PAP MMPI-A and the CA MMPI-A were investigated by utilizing paired t tests with the Bonferroni correction procedure.
- II. The distributions of individual scale scores for the CA MMPI-A and the PAP MMPI-A were investigated by using Hartley's homogeneity of variance test.
- III. The relative rankings of scores for the CA MMPI-A and the PAP MMPI-A were investigated by using Pearson product-moment correlation coefficients.

CHAPTER III

RESULTS

Evidence necessary to address the general research questions, specific statistical questions, and hypotheses is found in the results of this study. The first question focuses on whether or not specific scale differences exist between the CA means and the PAP means. Next, the similarity of the distributions of scores for the CA MMPI-A and the PAP MMPI-A were examined. Finally, the relationships of the distributions of the two scores were investigated by evaluating the rankings of scores for the CA MMPI-A and the PAP MMPI-A via Pearson product-moment correlation coefficients. Data analyses for each question are preceded by presentation of initial descriptive statistics (See Tables 1-4).

Research Question 1

The CA means of the specific scales were compared to the PAP means of the specific scales. Paired t tests were used to calculate the t values. For question 1, none of the scale mean differences were significant, using the Bonferroni correction procedure. Means and standard deviations for each scale are presented in Table 1. Means for the CA 10 clinical and three validity scales ranged from 48.52 to 54.82. Standard deviations for the CA 10 clinical and three validity scales ranged from 8.64 to 14.30. Means for the PAP 10 clinical and three validity scales ranged from 48.25 to 54.29. Standard deviations for the PAP 10 clinical and three validity scales ranged from 8.52 to 14.53. The T-scores for the scale means were close to the population mean of 50 for both the CA and the PAP. t and p values for each scale are presented in Table 2. There is no evidence to support the hypotheses that CA scale score means differ statistically from PAP scale score means.

Table 1

Means, Standard Deviations, and *t* Values for CA and PAP MMPI-A Scales

Scale	CA Mean	CA SD	PAP Mean	PAP SD	<i>t</i> Value
Hs	53.62	10.56	53.51	10.55	.07
D	54.52	9.46	54.03	9.53	.37
Hy	53.44	10.50	53.10	10.52	.23
Pd	51.50	8.64	51.98	8.52	.40
Mf	54.82	11.42	54.29	11.37	.33
Pa	51.99	12.07	52.09	11.81	.06
Pt	49.53	10.64	49.70	10.73	-.11
Sc	51.49	12.62	52.08	13.41	-.32
Ma	50.90	11.29	51.77	11.32	-.55
Si	48.52	10.34	48.25	10.03	.19
K	53.33	10.15	53.28	10.31	.03
L	54.16	12.53	53.14	11.95	.59
F	54.06	14.30	53.93	14.53	.06

Note 1. Hs is Hypochondriasis; D is Depression; Hy is Hysteria; Pd is Psychopathic

Deviant; Mf is Masculinity - Femininity; Pa is Paranoia; Pt is Psychasthenia; Sc is

Schizophrenia; Ma is Hypomania; Si is Social Introversion; K is Defensiveness; L is Lie; F

is Infrequency; CA is computer-administrated; PAP is paper-and-pencil

Note 2. None of the *t* values were statistically significant.

Table 2Variances and F_{max} Values from CA and PAP MMPI-A Scales from Hartley's Test

Scale	Variance CA	Variance PAP	F_{max}
Hs	111.51	111.36	1.00
D	89.58	90.88	.99
Hy	110.35	110.74	.97
Pd	74.63	72.57	.97
Mf	129.24	130.31	.99
Pa	145.71	139.45	.96
Pt	113.20	115.03	.98
Sc	159.34	179.82	.87
Ma	127.50	128.06	1.00
Si	106.98	100.65	.94
K	103.00	106.36	1.03
L	156.95	142.76	1.10
F	204.51	211.06	1.03

Note 1. Hs is Hypochondriasis; D is Depression; Hy is Hysteria; Pd is Psychopathic

Deviant; Mf is Masculinity - Femininity; Pa is Paranoia; Pt is Psychasthenia; Sc is

Schizophrenia; Ma is Hypomania; Si is Social Introversion; K is Defensiveness, L is Lie, F

is Infrequency; CA is computer-administrated; PAP is paper-and-pencil

Note 2. None of the F_{max} values were statistically significant

Table 3Agreement on Presence or Absence of Clinical Elevations

	Agree Elevated	Agree Not Elevated	Elevated PAP	Elevated CA
Count	186	2426	7	13
Percentage	7.01	91.47	.26	.49
Percentage Agreement	98.48			
Percentage Disagreement			.75	

Table 4

Pearson Product Moment Correlation Coefficients (*r*) and *p* Values Showing the

Relationship Between CA and PAP MMPI-A Administration

Scale	Pearson <i>r</i>	<i>p</i> <
Hs	.44	.01
D	.43	.01
Hy	.47	.01
Pd	.50	.01
Mf	.51	.01
Pa	.52	.01
Pt	.42	.01
Sc	.44	.01
Ma	.50	.01
Si	.54	.01
K	.60	.01
L	.38	.01
F	.65	.01

Note 1. Hs is Hypochondriasis; D is Depression; Hy is Hysteria; Pd is Psychopathic Deviant; Mf is Masculinity - Femininity; Pa is Paranoia; Pt is Psychasthenia; Sc is Schizophrenia; Ma is Hypomania; Si is Social Introversion; K is Defensiveness; L is Lie; F is Infrequency; CA is computer-administered; PAP is paper-and-pencil

Research Question 2

To determine whether the distributions of scores for the CA AND PAP MMPI-A were equivalent, tests of homogeneity of variance were conducted using Hartley's homogeneity of variance tests; 10 scales were compared. Variances for CA and PAP and F_{max} values for each scale are presented in Table 2. CA variances ranged from 89.78 to 204.51. PAP variances ranged from 90.88 to 211.06. F_{max} values ranged from .87 to 1.10. None of the variance estimates were statistically significantly different as a function of administration format. There is no evidence to support the hypotheses that CA scale score distributions differ statistically from PAP scale score distributions.

Some of these examinees assigned to the so called "clinically significant" range ($T > 70$) via the PAP version are not the same as those assigned via the CA version. That is, there may not be complete agreement of scores within the elevated positions of the scale. Correlational data showing moderate but not extremely strong relationships provide additional evidence that there may be less than perfect agreement in this critically important range. Scores in this range are used to help confirm diagnoses, and the need for subsequent treatment - consequently, it is important to evaluate the extent to which the two administration forms produce agreement. Elevations on the individual scales and agreement for each participant are presented in Table 3.

Research Question 3

To determine the extent to which the relative rankings of scores for the CA MMPI-A and the PAP MMPI-A are similar a Pearson product-moment coefficient was calculated for each scale. Pearson product-moment coefficients are presented in Table 4.

Pearson product-moment coefficients for the 10 clinical scales ranged from .38 to .65 and are all statistically significant, indicating that the scales are correlated. Pincus (1996) stated that an important consideration that relates to the distribution of scores is the clinical interpretations that are yielded by the profiles. Apparently, the relative rankings of scores for the CA MMPI-A and the PAP MMPI-A are somewhat similar.

Even though the Pearson product-moment coefficients were statistically significant, they were not as large as expected and as found in previous studies (Lushene et al., 1974; Pincus, 1996; Sukigara, 1996). Inspection of scatterplots revealed homogeneity within the distributions, i.e., few outliers. All participants in this study were primarily healthy individuals with little pathology, which might have caused the correlation coefficients to be lower than anticipated.

CHAPTER IV

DISCUSSION

This study was designed to compare the CA version of the MMPI-A with the PAP of the MMPI-A. This was accomplished by administering the PAP and CA versions of the MMPI-A in a counterbalanced order. Discussion focuses first on the findings from each research question and how these findings relate to previous theory and research. Second, implications for practical use of the versions are discussed. Finally, limitations of the study and suggestions for future research are discussed.

Comparisons of the Individual Scale Means

Research question 1 addressed whether the individual scale means between the CA and PAP differed. Based on results from multiple *t* tests, using the Bonferroni correction procedure, no differences were found. As previously discussed, there are no data in the literature to address the extent to which computer and noncomputer administration differences exist on the MMPI-A. However, several studies conducted with other instruments have found no differences. For example, no differences in composite scores were found by: Vansickle et al. (1989) using the Strong-Interest Inventory; Rosenfeld et al. (1992) using the Yale-Brown Obsessive-Compulsive Scale; and Rasulis et al. (1996) using the Rotter Incomplete Sentence Blank College Form. No individual scale differences were found by: Bader et al. (1993) using the Giessen Test (personality test); and Merten and Ruch (1996) using the Eysenck Personality Questionnaire - Revised.

Similarly, several studies conducted with the MMPI and MMPI-2 found no individual scale differences, e.g., Schuldberg (1988); Honaker et al (1988); Watson et al.

(1990); and Pinsonneault (1996). Because the two forms yielded scores that are not different, the authors suggest that either of the two formats can be used with confidence.

On the other hand, some investigators have found individual scale mean differences on the MMPI as a function of format. In a review of studies, Pinsonneault (1996) reported 12 studies, eight of which reported at least one scale difference. In an early study, Lushene et al. (1974) found that CA administration resulted in significantly lower differences on Hypochondriasis, Psychopathic Deviant, Paranoia, Schizophrenia, Lie, and Defensiveness Scales. Because these differences were not large (less than 2 T-score points) and are probably not clinically significant, Lushene et al. (1974) concluded that the CA MMPI was as valid as the PAP MMPI.

Similarly, Biskin and Kolotkin (1977) found significant differences on the Cannot Say Scale and the Paranoia Scale with lower means found for the CA group. They stated that preliminary MANOVA results indicated significant differences, but when the Cannot Say scale was excluded, no differences were found. They concluded that the large *F* ratio associated with the Cannot Say scale could have accounted for the mean profile differences. Watson et al. (1992) found that computer administration of the MMPI yielded significantly lower means on eight of 10 clinical scales when compared to PAP; they suggested that the differences were meaningful and that a small correction in norms might be helpful.

On the other hand, Sukigara (1996) found significant differences on the Depression, Paranoia, Psychasthenia, and Schizophrenia scales on the MMPI with the CA scores higher. Sukigara (1996) suggested that the contradictory findings may be due to

random sampling error, the unfamiliarity of Japanese people with a keyboard, and the kinds of hardware and software used.

Of the individual scales that differed, it appears that the Cannot Say scale was found to be significant in four studies reviewed by Pinsoneault (1996). Pinsoneault (1996) suggested that by forcing the respondent to choose an answer might resolve the differences in Cannot Say scale scores.

In summary, research studies have not been consistent, and not all investigators used standard scores (as opposed to raw scores), similar equipment, and like designs. These differences may have accounted for the contrary findings. For example, Lushene et al. (1974); Biskin and Kolotkin (1977); and Sukigara (1996), all of whom reported differences as a function of administration format, conducted data analyses on the raw scores instead of T-scores and used antiquated equipment. Biskin and Kolotkin (1977) and Sukigara (1996) did not use a counterbalanced design. In addition, Sukigara's examinees used a translated version of the MMPI (Japanese Kanji characters), which may have also contributed differences across formats.

Distributions of the Scores

Research question 2 addressed whether the distribution of scores differed across the administration formats. The distribution of scores was examined via tests of homogeneity of variance; distributions did not differ between the CA and PAP MMPI-A. Similar distributions mean that proportions of examinees are similar across the two distributions, i.e., the PAP and CA administrations assign similar numbers of examinees to low, medium, and high portions of the distribution. Table 3 compared the percentage of

agreement of examinees in the clinical range. The percentage of agreement was high (98.48) and means that diagnoses assigned to examinees will be relatively consistent across formats. Honaker et al. (1988); Watson et al. (1990); and Pincus et al. (1996) all found no differences in the distribution of scores for the MMPI and MMPI-2. Pincus et al. (1996) stated that an important consideration that relates to the distribution of scores is the clinical interpretations that are yielded by the profiles. The same diagnosis or diagnoses should be assigned to examinees regardless of administration format. Pincus et al. (1996) and Honaker et al. (1988) stated that although the distribution of scores did not differ, definitive conclusion should not be made because the power of their analyses was lower than the accepted range.

Rankings of Scores

Research question 3 addressed whether the rankings of scores differed across administration formats, using Pearson product-moment correlations. The rankings of the scores were significantly related and moderately strong, which suggests that the scores are similarly arrayed within the two distributions produced by the CA and PAP MMPI-A formats. Others have found similar results with other MMPI and MMPI-2 distributions. For example, Honaker et al. (1988) and Pincus et al. (1996) found that the ranking of scores between the CA and PAP MMPI and MMPI-2 were significantly related, although the magnitude of the coefficients were generally higher, a difference in magnitude of .16 or greater was reported for their studies. A possible explanation of this difference could be that the age of participants in this study were considerably lower than the ages reported by Honaker et al. (1988) and Pincus et al. (1996). As stated with regard to the distributions,

these authors concluded that the results might not be conclusive because of the low power of their analyses. In fact, Honaker et al.(1988) suggested that at least 100 participants per group should be used in future studies to increase the power to a sufficient level.

Implications

There are advantages of the CA format of the MMPI and MMPI-2 reported in the literature. Participants in this study did not report anxiety effects to the researcher. Several researchers have investigated anxiety effects. Lushene et al. (1974); and Hart and Goldstein (1985) found that no anxiety effects were present. However, Paul (1982); Weeter and Halcomb (1986); George et al. (1992); found anxiety effects with regard to administration format. No social desirability effects were found as a function of format. Finegan and Allen (1994) and Fórd et al. (1996) found no differences in social desirability. However, Lucas et al.(1977); Martin and Nagao (1989); Peterson et al. (1996); and Peters et al. (1998) found reduced social desirability effects with a computer. Lautenschlager and Flaherty (1990) found that the computer enhanced clients' tendency to respond in a socially desirable way. Honaker et al. (1988) reported that their participants showed an overall preference for the computer and viewed the computer as faster. Pinsoneault (1996) also found that his participants reported that they felt more comfortable with the computer, found it more interesting, and found it less difficult. Other advantages of the CA that are found in the literature include: less time for test taking, accuracy of scoring, cost efficiency, and the objectivity of the computer (Ben-Porath and Butcher, 1986). In addition, after the data has been saved, it can be scored and reported quickly without

manual entry of the test items. Data also can be transferred rather easily to different computers.

On the other hand, several disadvantages of the CA format for the MMPI and MMPI-2 are reported in the literature. For example, Honaker et al. (1988) reported that their participants had an overall education level that was higher than the education level of the population and suggested that clients with less education may respond negatively to the computer. Biskin and Kolotkin (1977) reported that some individuals with psychiatric diagnoses, such as paranoid schizophrenia, might react to the computer differently than a typical respondent would. This researcher found that the CA administration of the MMPI-A was reported by many participants to be difficult to use, even by the college students, many of whom had previously used computers. Participants had difficulty understanding how to enter the background information prior to answering the test questions. Another practical concern is finding computers at any given site that can successfully run the MICROTTEST Q program, because it requires a large amount of computer memory to operate. Finally, computers are sometimes unreliable. For four participants, the computer locked during the process of saving the data, and the participant had to reenter all 478 responses. This problem caused fatigue in these participants.

The PAP format has advantages and disadvantages as well. The booklet format is almost foolproof. None of the participants reported any concerns or confusion with the PAP administration. Second, the booklets can be reused and transported easily between sites. Third, no special equipment is required. Fourth, the PAP can be given anywhere. Finally, there are no concerns with data being lost during the saving process. In this study,

this researcher found one disadvantage of the PAP format. If the examiner intends to use the computer scoring program, all items must be entered into MICROTEST Q. This process is time-consuming and requires that the psychologist either carefully check the entry or reenter each response to verify that the items were entered correctly. However, once the data has been entered into the computer program, it can be easily transferred from computer to computer. For the psychologist traveling between sites, the PAP administration makes more sense because its ease of transportation and administration.

Limitations of the Study and Suggestions for Future Research

There are some limitations associated with the design and implementation of this study. Characteristics of the sample may limit the generalizability of the findings. Watson et al.(1992) discussed the need for future research to include more minority populations as participants. In addition, participants in this study did not match the sample used during the original norming of the MMPI-A. For example, although the majority of the participants were primarily Caucasian (66%), the sample of Caucasians was not as large as in the normative group. Of the 1625 adolescents used in the normative sample, 76% were Caucasian adolescents. 12% of the normative group were African-American adolescents versus 34% in this sample. 12% of the normative group were adolescents from other ethnic groups (Archer 1997). Additionally, the sample for this study was selected from two cities in the southeastern part of the United States and does not represent the entire United States. Data collection for the original normative sample included participants from

Minnesota, Ohio, California, Virginia, Pennsylvania, New York, North Carolina, and Washington (Archer, 1997).

There are other limitations of this study as well. One limitation was the unreliability of the two computers that locked during the process of saving the data. Four participants had to reenter their data. Another limitation is also the computers used in the study. While all of the computers were IBM and had Windows 98 or above as the primary operating system, the computers were of different makes and models from one setting to another, which may have produces subtle differences in administration characteristics (e.g. latency times). Also, there was no proctor during the testing at the high school setting, which could have produced differences due to administrator personality. Also, validity scales for the CA and PAP were not used in data analyses. This was because of a limitation with the MICROTTEST Q software. The program would not score any profile with missing data; therefore, every question must be answered.

Another characteristic is that some of the adolescents who participated in this study had special education handicaps. The proportion of special education adolescents in this study may limit the generalizability of the findings. Approximately 27.45 %of the adolescents from the high school participants had a special education handicap. The MMPI-A manual does not list whether or not the normative sample included any adolescents with special education handicaps (Butcher et al. 1992). Approximately 10% of school age children receive special education services, according to the 1997 United States Department of Education (Eggen and Kauchak 2001).

Biskin and Kolotkin (1977) and Watson et al. (1990) recommended that future research focus on clients with psychiatric diagnoses, since the majority of studies conducted have included mostly undergraduate college students or volunteers from the community.

REFERENCES

REFERENCES

- Archer, Robert P. (1997). *MMPI-A: Assessing Adolescent Psychopathology Second Edition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Bader, P., Hofman, K., and Kubinger, K.D. (1993). Zur brauchbarkeit der normen von papier-bleistift-tests für die computer-vorgabe: Ein experiment am beispiel des Giessen-tests. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 14, 129-135.
- Bartels, K.M., personal communication, August 11, 2000
- Ben-Porath, Yossef S., and Butcher, James N. (1986). Computers in personality assessment: A brief past, an ebullient present, and an expanding future. *Computers in Human Behavior*, 2, 167-182.
- Biskin, Bruce H., and Kolotkin, Ronette L. (1977). Effects of computerized administration on scores of the Minnesota Multiphasic Personality Inventory. *Applied Psychological Measurement*, 1(4), 543-549.
- Bugbee, Alan C. Jr. (Spring 1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education*, 28(1), 282-299.
- Butcher, James N., Williams, Carolyn L., Graham, John R., Archer, Robert P., Tellegen, Auke, Ben-Porath, Yossef S., & Kaemmer, Beverly (1992). *MMPI-A: Minnesota Multiphasic Personality Inventory for Adolescents. Manual for Administration, Scoring, and Interpretation*. Minneapolis, MN: The University of Minnesota Press.
- Eggen, Paul, and Kauchak, Don (2001). *Educational Psychology Windows on Classrooms, 5th Edition*. Upper Saddle River, NJ: Prentice Hall.

- Finegan, Joan E., and Allen, Natalie J. (1994). Computerized and written questionnaires: Are they equivalent? *Computers in Human Behavior*, 10(4), 483-496.
- Ford, B. Douglas, Vitelli, Romeo, Stuckless, Noreen (1996). The effects of computer versus paper-and-pencil administration on measures of anger and revenge with an inmate population. *Computers in Human Behavior*, 12(1), 159-166.
- George, C.E., Lankford, J.S., and Wilson, S.E. (1992). The effects of computerized versus paper-and-pencil administration on measures of negative effect. *Computers in Human Behavior*, 8, 203-209.
- Greud, V.A., and Green, B.F. (1986). Equivalence of conventional and computer presentation of speed tests. *Applied Psychological Measurement*, 10, 23-24.
- Green, Bert F. (1983). Adaptive testing by computer. In R.B. Ekstrom (Ed.), *Measurement, Technology, and Individuality in Education* (pp. 5-11). San Francisco: Josey-Bass.
- Greist, J.H., and Klein, M.H. (1980). Computer programs for patients, clinicians, and researchers in psychiatry. In J.B. Sidowski, J.H. Johnson, and T.A. Williams (Eds.), *Technology in Mental Health Care Delivery Systems*, 161-182.
- Hart, Ronald R. and Goldstein, Michael A. (1985). Computer-assisted psychological assessment. *Computers in Human Behavior*, 1(3), 69-75.
- Hofer, P.J. and Green, B.F. (1985). The challenge of competence and creativity in computerized psychological testing. *Journal of Consulting and Clinical Psychology*, 53, 826-838.
- Hoffman, Kaaren I., and Lundberg, George D. (1976). A comparison of computer-

- monitored group tests with paper-and-pencil tests. *Educational and Psychological Measurement*, 36, 791-809.
- Honaker, L. Michael, Harrell, Thomas, and Buffaloe, Janis D. (1988). Equivalency of Microtest Computer MMPI administration for standard and special scales. *Computers in Human Behavior*, 4, 323-337.
- Lautenschlager, G.J. and Flaherty, V.L. (1990). Computer administrations of questions: more desirable or more social desirability?. *Journal of Applied Psychology*, 75, 310-314.
- Liefeld, J.P. (1988). Response effects in computer-administered questioning. *Journal of Marketing and Research*, 25, 405-409.
- Lucas, R.W., Mullin, P.T., Luna, C.B.X., and McEnroy, D.C. (1977). Psychiatrists and a computer as interrogators of patients with alcohol-related illnesses: A comparison. *British Journal of Psychiatry*, 131, 160-167.
- Lushene, Robert E., O'Neil, Harold F. Jr., and Dunn, Thomas (1974). Equivalent validity of a completely computerized MMPI. *Journal of Personality Assessment*, 38, 353-361.
- Martin, C.L. and Nagao, D.H. (1989). Some effects of computerized interviews on job applicant responses. *Journal of Applied Psychology*, 74, 72-80.
- Merten, Thomas and Ruch, Willibald (1996). A comparison of computerized and conventional administration of the German versions of the Eysenck Personality Questionnaire, and the Carroll Rating Scale for Depression. *Personality and Individual Differences*, 20(3), 281-291.

- Paul, L. (1982). Research on cyberphiliacs, cyberphobics reveals 30% of workers fear computers. *Computer World*, 16, 14.
- Peters, L. Clark, D, and Carroll F. (1998). Are computerized interviews equivalent to human interviewers? CIDI-Auto versus CIDI in anxiety and depressed disorders. *Psychological Medicine*, 28, 893-901.
- Peterson, Leif, Johansson, Valur, and Carlsson, Sven G. (1996). Computerized testing in a hospital setting: Psychometric and psychological effects. *Computers in Human Behavior*, 12(3), 339-350.
- Pinsonneault, Terry B. (1996). Equivalency of computer-assisted and paper-and-pencil administered versions of the Minnesota Multiphasic Personality Inventory-2. *Computers in Human Behavior*, 8, 291-300.
- Rafaeli, S. and Tractinsky, N. (1989). Computerized tests and time: Measuring, limiting, and providing visual cues, for response time in on-line questioning. *Behaviour and Information Technology*, 8, 335-351.
- Rasulis, Richard Jr., Schuldberg, David, and Murtagh, Michael (1996). Computer-administered testing with the Rotter Incomplete Sentences Blank. *Computers in Human Behavior*, 12(4), 497-513.
- Rosenfeld, Rochelle, Dar, Reuven, Anderson, Ron, and Greist, John H. (1992). A computer-administrated version of the Yale-Brown Obsessive-Compulsive Scale. *Psychological Assessment*, 4(3), 329-332.
- Schuldberg, David (1988). The MMPI is less sensitive to the automated testing format

than it is to repeated testing: Item and scale effects. *Computers in Human Behavior*, 4, 285-298.

Sukigara, Masune (1996). Equivalence between computer and booklet administrations of the new Japanese version of the MMPI. *Educational and Psychological Measurement*, 56(4), 570-584.

Sutton, R.E. (1991). Equity and computers in the schools: A decade of research. *Review of Educational Research*, 61(4), 475-503.

Synodinos, Nicolaos, E., Papacostas, C.S., and Okimoto, Glenn M. (1994). Computer-administered versus paper-and-pencil surveys and the effect of sample selection. *Behavior Research Methods, Instruments, and Computers*, 26(4), 395-401.

Vansickle, T.R., Kimmel, C., and Kapes, J.T. (1989). Test-retest equivalence of computer-based and paper-and-pencil versions of the Strong Interest Inventory. *Measurement and Evaluation in Counseling Development*, 22, 88-93.

Watson, Charles G., Manifold, Victor, Klett, William G., Brown, Jack, Thomas, Delores, and Anderson, Douglas (1990). Comparability of computer- and booklet-administered MMPIs among primarily chemically dependent patients. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2(3), 276-280.

Watson, Charles G., Thomas, Delores, and Anderson, Patricia E.D. (1992). Do computer-administered MMPIs underestimate booklet-based scores? *Journal of Clinical Psychology*, 48(6), 744-748.

Webster, Jane and Compeau, Deborah (1996). Computer-assisted versus paper-and-pencil

administration of questionnaires. *Behavior Research Methods, Instruments, and Computers*, 28(4), 567-576.

Weeter, R. and Halcomb, C. (1986). Factors influencing the negative perceptions of computers. Paper presented at the 32nd annual meeting of the Southwest Psychological Association, Ft. Worth, Texas.

APPENDIX A
PROPOSAL TO SCHOOL ADMINISTRATOR

RESEARCH PROPOSAL

A computer-administered version versus paper-and-pencil-administered version of the MMPI-A.

SUBMITTED TO: Metropolitan Nashville Public Schools
Nashville, Tennessee 37204

Attention: Dr. Bob Crouch, Director
Research and Evaluation

SUBMITTED BY: Shannon K. Hays, Graduate Student, University of Tennessee,
Knoxville
225 John Rice Blvd. A-4
Murfreesboro, TN 37129
Telephone: (615) 217-6706
(615) 594-2800

INTRODUCTION: Computers have become a popular way for clinicians and other educators to give tests. For example, the Scholastic Aptitude Test, the Graduate Record Exam, and many licensing exams, such as the licensing exam for nurses, are available to take via the computer. Since computer-administration (CA) has many advantages such as cost effectiveness, speed, and accuracy, it is possible that many psychologists in private practice and hospital settings will be employing this method of administration for clients taking personality tests, and particularly versions of the MMPI such as the MMPI-A. It is important that the equivalence for the paper-and-pencil (PAP) and the CA be established before the computerized method is used. There has been no research conducted comparing the MMPI-A computer-administration to the traditional format; consequently, this study is designed to investigate whether differences exist.

PURPOSE: The purpose of this study was to compare the PAP version of the Minnesota Multiphasic Personality Inventory - Adolescent (MMPI-A) with a CA version of the MMPI-A. Comparison is important for test users. Even though CA personality testing is quite popular because of reduced costs, automatic scoring, and lower ambiguity, it cannot be assumed that a test administered by a computer is equivalent to the PAP. The American Psychological Association has recommended that equivalence between PAP and CA versions of tests be addressed empirically and not taken for granted. Studies comparing the CA version of the adult version of the MMPI with the PAP version of the MMPI are available; however, there are no studies comparing the two administrative formats for the adolescent version. This study will test equivalence of two formats for the MMPI-A.

SAMPLE: Approximately 100 boys and girls from Hunters Lane High School, ages 14-18, will be needed for this study. All students will be volunteers.

PROCEDURE: (1) All parents of children who volunteer participants in this study will be given a letter explaining the experiment and allowing the parent to give permission for his or her child to participate. In the permission letter, an explanation of the MMPI-A will be given and an invitation to read the MMPI-A will be given. (2) The experiment will be conducted on Hunters Lane Campus when convenient for the parents and students. It is anticipated that all experiments will take place after school hours and on the weekends. (3) All participants will complete both the CA and the PAP versions of the MMPI-A in a counterbalanced order, with one randomly assigned group completing the PAP version first then the CA version second, and the second randomly assigned group completing the CA version first and the PAP version second. Time between testing will be approximately two weeks. Participants will be assured on their consent forms that their responses will be kept in a locked filing cabinet on the University of Tennessee campus in Claxton Room 535 and that their results would be kept confidential.

SCHEDULE: The plan will be completed by the end of the 2001-2002 school year. If the proposed plan is approved by Metropolitan Nashville Public Schools, I would like to begin collecting data as soon as possible.

REPORTING THE RESULTS: The results of this study will be reported in the form of a dissertation and probably in professional research journals. Copies of this dissertation will be provided to the Research and Evaluation Director of Metropolitan Nashville Public Schools, the principal of the participating school, and any other interested parties.

PARENTAL CONSENT FORM

A computer-administered version versus paper-and-pencil administered version of the MMPI-A.

Dear Parents:

My name is Shannon Hays, and I am a school psychologist with Metropolitan Nashville Public Schools. One of the schools I serve is Hunters Lane High School. I am interested in conducting a research study for the completion of my doctoral dissertation (University of Tennessee, Knoxville). The purpose of this study is to compare the paper-and-pencil form (PAP) version of the Multiphasic Personality Inventory for Adolescents (MMPI-A) with a computer-administered (CA) version of the MMPI-A. The MMPI-A is multiple-choice test of personality.

Your child has expressed an interest in participating in the study. Participants will complete the MMPI-A on school grounds after school hours or on the weekends, whichever is more convenient for you. During the first testing session, participants will be randomly assigned to two groups. The first group will complete the CA version first and then the PAP version during the next session. The second group will complete the PAP version first then the CA version during the next testing session. Time between testing sessions will be approximately two weeks. The amount of time per person will be approximately two-and-one-half hours per session. One hundred participants will be participating in this research.

The information in the study records will be kept confidential. Data will be stored securely and will be made available only to persons conducting the study unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports which could link your child to the study. Data will be maintained without identifiers indefinitely for further research projects by the principal investigator. Consent forms will be stored for three years past the completion of the study in Claxton Room 535, University of Tennessee, Knoxville campus.

Your child's participation is voluntary and has no reflection on your child's performance at school. You may decline for him/her to participate with no penalty. Your child may choose to withdraw from the study at any time. If your child withdraws from the study before data collection is completed, your child's data will be returned to you or destroyed. The risks to participants are minimal to nonexistent.

The nature of questions on the MMPI-A ranges from questions such as "I like car magazines" to "I often am bothered by thoughts about sex." (These examples were reworded from the original items.) I will make the MMPI-A available to you

to read if you have any concerns or questions. I can be reached at 615-885-8816 on Fridays or at Hunters Lane High School on Tuesdays and Thursdays in the Guidance office (615-860-1407).

Please sign below if you understand the conditions of the study and agree to allow your child to participate if he/she is interested.

Name of Child (please print) _____

Parent's Signature _____

Date: _____

APPENDIX C
CONSENT FORM

The University of Tennessee

Office of Research Research Compliance Services

INFORMED CONSENT FORM

A computer-administered version versus paper-and-pencil administered version of the MMPI-A.

You are invited to participate in a research study. The purpose of this study is to compare the paper-and-pencil form (PAP) of the Minnesota Multiphasic Personality Inventory - Adolescent (MMPI-A) with a computer-administered (CA) version of the MMPI-A. The MMPI-A is a multiple-choice test of personality.

INFORMATION

Participants will complete the MMPI-A inside a computer lab in the Claxton Education Building. Participants will receive extra credit for their participation. During the first testing session, participants will be randomly assigned to two groups. The first group will complete the CA version first and then the PAP version during the next session. The second group will complete the PAP version first and then the CA version during the next testing session. Time between testing sessions will be approximately two weeks. The amount of time per person will range about two-and-one-half hours per session. One hundred participants will be participating in this research.

RISKS

No physical and social risks are expected. Minimal psychological risks are expected. The MMPI-A is used for assessment purposes only and cannot be used alone to diagnose any psychological condition. Some of the questions on the MMPI-A are of a sensitive nature and may cause some participants to feel discomfort, but to protect the participants, they can make an appointment with the University of Tennessee Counseling Center for counseling. Permission for students to contact the Counseling Center was obtained from Dr. Gary Klukken, Director of the Counseling Center. Students will be informed as to the nature of the study. Students will be informed that they may withdraw at any time without penalty. If a participant withdraws, then his or her data will be destroyed. The principal investigator with access to the data will sign a confidentiality statement. Students will be provided with the principal investigator's telephone number and e-mail address so that they may express any questions or concerns.

Participant's Initials _____

BENEFITS

The incentives will be extra credit as determined by the University of Tennessee professor. The rationale for using these incentives is that extra credit is desired by most students. Since computer administration has many advantages such as cost effectiveness, speed, and accuracy, it is possible that many psychologists in private practice and hospital settings will be employing this method of administration for clients taking personality tests, and particularly versions of the MMPI such as the MMPI-A. It is important that the equivalence for the PAP and the CA be established before the computerized method is used. There has been no research conducted comparing the MMPI-A computer administration to the traditional format; consequently, this study is designed to investigate whether differences exist. Ultimately, it is hoped that this study will lead to further research comparing the two versions of the MMPI-A.

CONFIDENTIALITY

The information in the study records will be kept confidential. Data will be stored securely and will be made available only to persons conducting the study, unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports which could link you to the study. Data will be maintained without identifiers indefinitely for further research projects by the principal investigator. Consent forms will be stored for three years past the completion of the study in Claxton Room 535, University of Tennessee campus.

COMPENSATION

For participating in this study you will receive extra credit.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the researcher, Shannon K. Hays, by phone at 615-594-2800 or by email at drh2b@aol.com.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at anytime without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed, your data will be returned to you or destroyed.

CONSENT

I have read the above information and agree to participate in this study. I have received a copy of this form.

Participant's name (print) _____

Participant's signature _____

Date _____

The University of Tennessee

Office of Research Research Compliance Services

INFORMED ASSENT FORM

A computer-administered version versus paper-and-pencil administered version of the MMPI-A.

You are invited to participate in a research study. The purpose of this study is to compare the paper-and-pencil form (PAP) of the Minnesota Multiphasic Personality Inventory - Adolescent (MMPI-A) with a computer-administered (CA) version of the MMPI-A. The MMPI-A is a multiple-choice test of personality.

INFORMATION

Participants will complete the MMPI-A on Hunters Lane High School Campus after school hours. During the first testing session, participants will be randomly assigned to two groups. The first group will complete the CA version first and then the PAP version during the next session. The second group will complete the PAP version first, and then the CA version during the next testing session. Time between testing sessions will be approximately two weeks. The amount of time per person will range about two-and-one-half hours per session. One hundred participants will be participating in this research.

RISKS

No physical and social risks are expected. Minimal psychological risks are expected. The MMPI-A is used for information only and cannot be used alone to diagnose any psychological condition. Some of the questions on the MMPI-A are of a sensitive nature and may cause some participants to feel discomfort, but to protect the participants, they may talk to the researcher about their concerns. Students will be informed as to the nature of the study. Students will be informed that they may withdraw at any time without penalty. If a participant withdraws, then his or her data will be destroyed. The principal investigator with access to the data will sign a confidentiality statement. Students will be provided with the principal investigator's telephone number so that they may express any questions or concerns.

Participant's Initials _____

BENEFITS

The incentives will be either pizza and soft drinks to be delivered after each testing session is completed or extra credit as determined by the high school teacher. The rationale for using these incentives is that pizza and soft drinks are liked by most people and easy to obtain and extra credit is desired by most students.

Since computer administration has many advantages such as cost effectiveness, speed, and accuracy, it is possible that many psychologists in private practice and hospital settings will be employing this method of administration for clients taking personality tests, and particularly versions of the MMPI such as the MMPI-A. It is important that the equivalence for the PAP and the CA be established before the computerized method is used. There has been no research conducted comparing the MMPI-A computer administration to the traditional format; consequently, this study is designed to investigate whether differences exist. Ultimately, it is hoped that this study will lead to further research comparing the two versions of the MMPI-A.

CONFIDENTIALITY

The information in the study records will be kept confidential. Data will be stored securely and will be made available only to persons conducting the study unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports which could link you to the study. Data will be maintained without identifiers indefinitely for further research projects by the principal investigator. Consent forms will be stored for three years past the completion of the study in Claxton Room 535, University of Tennessee campus.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the researcher, Shannon K. Hays, by phone at 615-860-1407.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at anytime without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed, your data will be returned to you or destroyed.

ASSENT

I have read the above information and agree to participate in this study. My parents have turned in their signed permission form. I have received a copy of this form.

Participant's name (print) _____

Participant's signature _____

Date _____

VITA

Shannon Kathleen Hays was born in Knoxville, TN on November 15, 1972. She was raised in Knoxville, TN and graduated from Farragut High School in 1991. From there, she went to the University of Tennessee, Knoxville and received a B.A. in English Literature with a minor in Psychology and German in 1995.

Shannon is currently a school psychologist with Metropolitan Nashville Schools, Nashville, Tennessee where she is in her second year of practice.