Improving The Accuracy Of Biodata Questionnaires

By

Frank Kuschnereit

B.A. (Honours Applied Military Psychology)
Royal Roads Military College, 1990

A Thesis Submitted in Partial Fulfilment of
The Requirements for the Degree of

Master of Arts

In the Graduate Academic Unit of Psychology

Supervisor:	Dr. Barry Spinner, PhD., Psychology Department
Examining Board:	Dr. Frank Szeligo, PhD., Psychology Department
	Dr. Rick Audas, PhD., Faculty of Administration

This thesis is accepted

_____
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK
December 2000

ABSTRACT

Biographical information (biodata) obtained from job applicant forms has been in use since the turn of this century to predict future job success. It has been only relatively recently that there has been a concerted effort to improve biodata's reliability and validity. Two of the major concerns are how to minimize the incidence of faking and how to reduce inaccurate responses due to social desirability. Theoretical and empirical works have identified several factors which might affect the accuracy of biodata responses. Using a 2x2 between-subjects design, this study examined the effects of (1) a warning about the consequences of inaccurate responses (Warning) and (2) the amount of detail requested in responses (Detail). The dependent variables were participants' responses to biodata items that had been previously classified as relevant to the job description or not, and whether or not participants' responses were potentially verifiable. Participants were volunteers registered in Introductory Psychology. It was hypothesized that providing a warning of consequences and requiring greater detail would reduce social desirability of responses, but only for items that were job relevant and not potentially verifiable. The results indicated that none of the factors in this design (the consequences of misreporting, and requesting extended details behind the answers to some items) had a systematic effect on the biodata responses. Results are discussed with reference to the limitations of simulating personnel selection research with undergraduate participants who do not have the same motivations as real job applicants. The characteristics of desirable biodata items and the format in which they are presented as well as implications for future research are also discussed.

TABLE OF CONTENTS

Abstract ii

Table of Contents iii

List of Tables vi

Acknowledgments vii

Preface viii

I. Introduction 1

Definition of Biodata 1

Overview of Mael's Taxonomy 3

Validity of Biodata 6

Utility of Biodata 8

Review of Faking 10

Is Faking a Problem? 11

Methods of Detecting Faking 15

Methods to Reduce Faking 17

Statement of the Problem 23

Hypotheses 24

II. Pilot Study 25

Method 25

Purpose 25

Participants 25

Procedure 26

Measures 26
Fictional Job Advertisement 26
Cover Letter 26
The Biodata Questionnaire Instructions 27
Biodata Questionnaire 28
Debriefing Form 29
Results/Discussion 29
III. Main Study 30
Method 30
Design 30
Participants 30
Procedure 31
Measures 32
Fictional Job Advertisement 32
Cover Letter 33
The Biodata Questionnaire Instructions 33
Biodata Questionnaire 35
Amended Debriefing Form 34
Test Condition Items 35
IV Results 36
Data Preparation 36
Main Analysis 36

Discussion 46

Consistency with past findings 46

Study Limitations 48

Directions for Further Research 50

References 51

Appendices 55

A Pilot Study - Informed Consent Form 56

B Pilot Study - Fictional Job Advertisement 58

C Pilot Study - Covering letter 60

D Pilot Study - Instructions Base Line Group 62

E Instructions with Job Advertisement and Minimum Warning 64

F Instructions with Job Advertisement and Maximum Warning 66

G Biodata Questionnaire for Pilot Study 68

H Debriefing Forms for Pilot Study 77

I Within Subjects Items by Test Condition 79

J Main Study - Informed Consent Form 81

K Amended Fictional Job Advertisement 83

L Amended Coving Letter 85

M Abbreviated Questionnaire with Minimum Warning 87

N Extended Questionnaire with Maximum Warning 100

O Extended Items 114

P Amended Debriefing Form 118

LIST OF TABLES

1. Gender and age frequencies of pilot participants 25
2. Gender and age frequencies of main study participants 31
3. Descriptive statistics of High Verifiability/ Low Relevance items 38
4. HV/LR within+residual correlations with std. Devs. on diagonal 38
5. Results of MANOVA on High Verifiability/Low Relevancy 39
6. Descriptive statistics of High Verifiability/ High Relevance Items 39
7. HV/HR within+residual correlations with std. devs. on diagonal 41
8. Results of MANOVA on High Verifiability/High Relevancy 41
9. Descriptive statistics of Low Verifiability/ Low Relevance items 42
10. LV/LR within+residual correlations with std. devs. on diagonal 42
11. Results of MANOVA on Low Verifiability/Low Relevancy 43
12. Descriptive statistics of Low Verifiability/ High Relevance Items 43
13. LV/HR within+residual correlations with std. Devs. on diagonal 44
14. Results of MANOVA on Low Verifiability/High Relevancy 45
15. Descriptive statistics for check item by test condition 45
16. Tests of significance for faking check item 45

ACKNOWLEDGEMENTS

The thesis was prepared under the supervision of Dr. Barry Spinner, Psychology Department, University of New Brunswick. The author is indebted to Dr. Spinner for his advice and guidance through the course of the study.

I also wish to thank my wife, Glenda for not only typing the manuscript but for giving that most precious of all commodities - time.

Preface

As I was writing this thesis, I was reminded of this story. Years ago a young man was falsely convicted of a crime and sent to a prison on Devil's Island. He was thrown into a cell with an old man. All the young man could do was think of escape. He told the old man of his plan but the old man said nothing. Late one night the young man attempted his escape. Three days later they brought him back to the cell and he was near death. Upon recovering he told the old man about the horrors of the jungle. He told him of the poisonous snakes and spiders. That there was no food, no water, no boats and about the sharks in the ocean. The old man smiled and said "I know all that. I tried to escape many times years ago but it was impossible and I gave up even trying." The young man looked at him and said "I was almost killed! Why did you not tell me?". The old man just shrugged his shoulders and said "Well, no one publishes negative results."

Chapter 1

Introduction

Effective personnel recruitment and selection is one of the most critical steps in the establishment and growth of an organization. Many organizations expend a great deal of time and effort to match the right person to the job as there are many costs associated with hiring personnel who subsequently are not successful. These costs can be measured not only in dollar terms but also by their effect on the lives and careers of the persons involved (Strauss & Sayles, 1972). There are a variety of personnel selection techniques in common use. The interview is one of the most popular (Harris, 1989) but has been criticized as being too subjective (Campion, Pursell & Brown, 1988). The assessment centre approach has the applicant complete a battery of selection tests and work samples but can be very time consuming and expensive (Alder, 1987). One technique that is less subjective, inexpensive and effective involves obtaining biographical information (biodata) from job applicants to predict future job success (Mitchell, 1994).

Definition of Biodata

Lautenschlager (1994) defined biodata as "...factual kinds of questions about life and work experiences, as well as to items involving opinions, values, beliefs, and attitudes that reflect a historical perspective." He notes that since the respondent is limited to replying to the questions posed to him there are elements of both biography and autobiography. The basis of biodata's predictive abilities is the axiom that past behaviour is the best predictor of future behaviour (Mael, 1991). Biographical information is not expected to predict all future behaviours but it is useful in personnel selection in that it can give an indication of probable future behaviours based on an individual's prior learning history (Mumford et al., 1996). Stricker and Rock (1998) state that

biodata instruments have an advantage over personality and interest inventories in that they can "capture directly the past behaviour of a person, probably the best predictor of his of her future actions. These measures deal with facts about the person's life, not introspection's and subjective judgements...".

While resembling the standard job application form, a typical biodata questionnaire seeks to obtain more extensive and detailed information on significant events in applicants' lives and especially on their past behaviours. Nichols (1994) stated that "Biodata measures may predict performance across so many aspects of behaviour as well as they do because responses to biodata items may serve to capture previous manifestations of the constructs and mechanisms that ultimately determine predictive relationships with criteria.".

While biodata items were in wide use, there was some debate about what items should be classified as biodata. Asher (1972) stated that typical biodata items could vary widely on a number of dimensions such as: verifiable - unverifiable; historical - futuristic; actual behaviour - hypothetical behaviour; memory - conjecture; factual - interpretive; specific - general; response - response tendency; and external - internal event. He also divided the different types of items into two categories: soft and hard items. He described hard biodata items as historical and verifiable items. Soft biodata items are more abstract in nature and tap value judgments, aspirations, motivations, attitudes and expectations.

Winter (1986) in his survey of the Australian, United Kingdom, Canadian and US militaries' selection procedures discovered that there continued to be a wide variation in what the different militaries considered to be biodata. He ended up having to define biodata as autobiographical accounts of past events derived from the sorts of items found on a biographical data blank or

application form. Winter admits that this is scarcely helpful but it shows that the term lacks precise definition. Winter hypothesizes that those items that Asher described as hard items tend to be more reliable (and so potentially more valid) than the soft items perhaps because they are less prone to faked response or ambiguous interpretation by the applicant. He also concludes that so long as their psychometric deficiencies are recognized, soft items can add greatly to the range of predictors used in selection as they permit investigation of otherwise inaccessible aspects of the applicant.

Overview of Mael's Taxonomy

Mael (1991) noted that while previous authors have described various attributes of biodata, none had systematically defined biodata or their attributes. Mael developed a taxonomy to identify the key characteristics of biodata. Mael preferred not to consider the 'soft' biodata items as true biodata. He defined true biodata items as those that are: historical, external, verifiable, objective, first-hand, discrete, controllable, equal access, job relevant, and noninvasive.

Historical items are those that refer to events that have already occurred. It is hypothesized that historical items may be more reliable and accurate then items asking about intentions, hypothetical situations, or intentions. These, and other types of non-historical items, are not considered to be biodata items.

External refers to events or actions that have occurred in the past. This would exclude internal attitudes or thoughts but is not necessarily a publicly seen event. Mael uses the example of someone who has never actually served in the military but may have had their attitudes and behaviours shaped through watching war movies, reading history books, or collecting weapon models. One advantage of using external items is that there is less chance of someone faking their answer. Individuals may unconsciously distort their responses when trying to assess a personal

characteristic but an external item would require someone to misremember or consciously lie in order to 'fake good'.

A *verifiable* item can refer to archival information such as school marks or information that could be verified through a third party. It is not necessary for the response to be actually checked as long as the individual believes that there is a potential for it to be verified.

Mael describes *objective* as that which only requires the faculty of recall. For example, asking a baseball player about their batting average would be objective, while asking them how well they play baseball would call for a subjective assessment of their own performance.

First-hand information is that which the respondent knows for a fact from their own experience and does not require an assessment of what others might have thought. For example, asking someone about their university GPA would be considered first-hand, while asking what a professor thought of their academic performance would not.

Mael states that *discrete* items that measure a single behaviour are preferred to summary items in that they tend to require some form of estimation or computation which could lead to errors. Mael admits that with regularly performed behaviours such as levels of physical activity or time spent studying, a summary average could be more indicative of capabilities than just a snapshot of the events on a particular day.

Controllable: When constructing biodata instruments, there are certain ethical and legal considerations. Mael argues that it is not ethical to measure and use information about events over which the respondent has no control, such as birth order. Also a biodata item should not inquire about activities to which not everyone has *equal access*. Stricker and Rock (1998) also noted that biodata items that concern opinions, internal states, and subjective judgements that are not under the

applicant's control or involve opportunities and resources that are not available to most people are problematic, both psychometrically and ethically. Mael notes that while controllability and equal access are preferable for moral and ethical reasons they are not intrinsically unfair or unethical. Just because one respondent was a captain on a sports team and another respondent went to a small school where there were no sports teams does not diminish the first respondent's accomplishments. A balance has to be struck when using such items.

Although almost any past experience could be considered *job relevant* in some way, the courts have taken a dim view of the gathering of information that is not visibly job relevant. Mael also makes a distinction between 'signs' or extrapolated inferences from an attitude, disposition or ability and 'samples' which are point-to-point examples of previous behaviour. He notes that while most biodata items span the continuum between signs and samples, samples bear a stronger relationship to performance criteria than do signs and therefore may be more valid predictors.

Mael states that biodata items should be *non-invasive*. He postulates that intrusive questions such as those dealing with sexual behaviour, criminal behaviour or religion are "likely to incur resistance and resentment, and thereby encourage willful faking, random responding, or other behaviour aimed at foiling the testers." Not only do intrusive questions make the respondent uncomfortable, they may result in a lawsuit. For example, while studies have shown that married Canadian military recruits have a much higher attrition rate than do single recruits (Winter, 1986), it is illegal in Canada to discriminate against anyone on the basis of race, religion, or marital status (Cronshaw, 1986). The US Privacy Protection Study Commission in 1997 has also questioned the fairness of collecting certain types of personnel history information for personnel selection purposes

such as family background, credit history, criminal record and military service. (Stone & Jones, 1997)

Mael summarised his taxonomy by stating that the only necessary attribute to define an item as a biodata item is that it be historical. He stated that the other attributes are desirable as they could reduce faking or would be seen as more fair and legally defendable. Considering the number of studies and papers that have quoted Mael's taxonomy of biodata items, he has had quite an impact on subsequent biodata research.

Validity of Biodata

Over the years, personnel selection has relied on the standardized psychological tests (Riley & Chao, 1982). The five major categories for these tests are intellectual abilities, spatial and mechanical abilities, perceptual accuracy, motor abilities and personality tests. The mean coefficient for a standardized test is .35. A review of 58 studies on biodata found coefficients that ranged from .32 to .46 with an mean validity of .35. The mean validity of interviews was found to be .19. Harris (1989) reported a validity coefficient of .29 for unstructured interviews and .31 for structured interviews but admits that interview results can be affected by interviewer biases and have been challenged in a number of different court cases.

While all job-application forms collect some form of biographical information, and have been in use since the turn of this century, it has been only relatively recently that there has been a concerted effort to improve biodata's reliability and validity. Shaffer, Saunders & Owens (1986) examined the accuracy of biodata through a long term test-retest design and also by collecting observer ratings. They administered a biodata questionnaire to 1,264 first year university students and then readministered the test to 237 subjects randomly selected from the first group. They were

also able to collect usable observations from at least one parent from 108 participants from the second group. The results indicated a high level of accuracy for the objective or hard biodata items. The subjective or soft items were not as stable over the five-year interval and there was less correlation between the observer ratings and the student ratings. Many factors can affect the validity of a biodata instrument. For example, the Life Insurance Marketing and Research Association (LIMRA) first developed and validated the life insurance agent aptitude index in 1922. This instrument has had to undergo a number of revisions throughout the years due to some unexpected factors. Hughes, Dunn & Baxter (1956) described the experience of a large insurance company where a biodata selection tool was scored at a central location and found to be predictive and valid. Once the field managers were allowed access to the scoring keys it was found that the instrument became less predictive. Examination of the data indicated that managers were coaching the applicants on how to answer certain questions. Thayer (1977) also found that some life insurance branch managers that are under pressure to produce have been found to coach applicants on how to complete the test or in extreme cases have offered to take the test for the applicant. He notes that while the original biodata instrument may be valid when first created, the insurance industry has found that the test must be continually updated and re-validated due to changes in society, economy, and organizational dynamics.

Pannone (1984) conducted a field study to examine the validity of a biodata questionnaire. The subjects were 221 electrician applicants. Pannone based his study on the behavioural consistency model which is based axiom that the "best indicator of future performance is past performance". He notes that it's also similar to the "point to point theory" in which the more points in common between the predictor and the criterion, the higher the validity coefficient. The

applicants were required to rate their previous work experience with each task listed on a 4-point scale. The results of the biodata questionnaire were compared to a written test of proficiency in electrical knowledge to years of electrical training and years of work experience. The results indicated that the biodata questionnaire, when compared to the written test of proficiency, was a better indicator than years of electrical training or years of work experience.

Utility of Biodata

The first recorded instance of a company using a biodata form was a life insurance company in 1889 (McManus & Kelly, 1999). It was found that collecting information on an applicant's life history in a standardized fashion could be useful in predicting their potential for success in the future.

Melamed (1992) examined the use of biodata for predicting academic success and found that both high school leaving averages and general biodata showed a moderate predictive validity. She did note that the validity of the biodata items reduced over time. Asher (1972) notes that while high school marks do very well at predicting performance in College or University the University GPA does a very poor job of predicting success after University.

As noted above, biodata measures have been used extensively by the insurance industry to select sales representatives, by the military, and to predict academic performance. Biodata measures have also been effective in selecting vocational employees (Brown, 1981). In addition to selecting the right person for the job, biographical information is useful for finding the right job for the person. Brush and Owens (1979) surveyed 2,140 hourly waged employees from a large oil company using a 263-item biodata form. The aim of this study was to determine if personnel could be grouped according to their personal characteristics and if this related to their success in job

categories. The results of their study indicated that "a model incorporating the use of employee subgroups and job clusters has real utility for accomplishing this goal." This method allows the person to be offered a job for which they are most suited and with which they will be likely to be most satisfied.

Pannone (1994) looked at the applicability of using biodata instruments to select "blue collar workers". He describes "blue collar workers" as "hourly workers involved in the manufacturing, maintenance, and repair of machinery, equipment, and building structure". He noted that the nature of the functions performed by the blue collar worker are generally more narrow in range that professional and technical occupations. He recommends that biodata items should be based on a job analysis for blue collar workers and include very specific questions. He states that based on the ease of collecting background data as well as the benefits of standardization, background data forms are a viable alternative to traditional testing programs.

Ellis & Spinner(1997) attempted to develop and validate a biodata instrument that would be predictive of success on the Canadian Forces Basic Officer Training course. The participants consisted of 314 officer candidates. Their inventory consisted of 220 items. The instructions included a warning that answers might be verified for accuracy. The criteria were the course grades and an instructor performance rating scale. The results indicated that the biodata instrument accounted for a further 16% of the variance over and above the selection measures currently in use with a validity coefficient of .55 with the derivation sample and .38 with the cross-validation sample.

Wilkinson (1997) conducted a study that showed the generalizability of a biodata instrument to the selection of managers. Carlson and Scullen (1999) demonstrated that biodata instruments do

not have to be specially developed for a particular organization. While it is preferable for a biodata instrument to be developed based on a job analysis (Fine & Cronshaw, 1994) they showed that validity could be generalized to other organizations and even to other industries.

Melamed (1992) states that one of the problems with biodata is the lack of a basis in theory. An item may be very predictive but if there is no underlying theory to explain why it predicts so well, it could just be a statistical fluke. He relates an example of chance correlation where in one study, the best predictor of being a successful door-to-door salesperson was attendance at a circus show. Another example was related by Asher (1972). The item "did you ever build a model air plane that flew?" was found to be almost as good as predictor of success in flight training during World War II as the remainder of the entire United States Air Force Battery. It is hard to disregard items that are very predictive even though the underlying theory is not readily apparent. This is sometimes referred to as "blind empiricism" since there is no understanding of why the items predict (Mael & Hirsch, 1993). Winter (1986) also states that biodata research does tend to suffer from "blind empiricism" at its most extreme and that the statistical methods used tend to exploit chance fluctuation in the data.

Review of Faking

Since a biodata instrument is essentially a self reported set of responses gathered for a specific purpose, there has been some concern about verifying the accuracy of these responses. Lautenschlager describes three ways of operationally defining accuracy. He states that correlational accuracy is where evidence from an external source is compared with the self-reported responses. An alternative method would be to ask the same type of question at two different times to see how stable the response is. He notes that some use level or mean differences to measure accuracy of

biodata studies. In this situation the average response for a group of subjects is compared to the average response of another group or compared under two different administration conditions. He also refers to absolute accuracy. The author notes while under this condition some individual responses may be seen as inaccurate the result may not have any effect on keyed scores or one combined with the entire group scores. Lautenschlager states that there is evidence that warnings of possible verification can reduce faking. He notes that there may be some distortion of responses that this may not be a bad thing if the pattern of distortion is uniform across the entire pool of subjects. Therefore the type of faking that would be of most concern would be when individuals distorts their responses and this affects the rank order of applicants to improve their chances at selection. Lautenschlager notes that there is evidence from previous studies that there is a relationship between faking and an applicant's age or level of experience. He states that accuracy can be improved by using historical, transparent and verifiable items, through the use of warnings, faking scales or false items and by examining reaction time or other statistical methods. He recommends that future research should determine if absolute accuracy is required or worth the time and effort to achieve it.

Is Faking a Problem?

In an early study on faking, Dunnette et. al., (1962) administered a biodata test to 62 salesmen with a large company and directed half the group to answer honestly while directing the other half to answer as if a successful salesman would answer it and try and beat the test. The results indicated that those under the fake condition could score higher than those who answered honestly.

Abrahams et al., (1971) conducted a study to determine if the incidence of faking in a real life selection situation compared to the results that had been found in laboratory studies. He noted that previous studies had merely attempted to determine whether or not the test could be faked. The authors looked for a group of participants who might have an incentive to distort their responses. At that time the US Navy was using a commercially developed biodata instrument to select applicants for a full college scholarship. It was hypothesised that the potential benefits of receiving a full college scholarship could be a strong inducement to fake. The authors compared their responses of two groups of applicants. The first group had previously written the biodata measure in high school as part of a routine testing, they then retook the measure as part of the scholarship selection procedure. The second group completed the biodata measure as part of the selection procedure and then later rewrote the test as part of normal college testing. The rank-order correlation between the mean profiles for the first group was .95 and supported the conclusion that the applicants did not fake. The analysis of the second group showed a correlation of .98. The authors concluded from these results that even though the measure had been shown to be fakeable in the laboratory, that in practice there is no evidence to support the hypothesis that applicants were faking good to be selected in actual selection situations. They also conclude that simulated faking designs do not provide "a particularly appropriate estimate of what occurs in selection, instead they provide only an indication of how much a test can be faked."

Zalinski & Abrahams (1979) also attempted to place their participants in a situation where they would be motivated to do better on a test, rather than just instruct them to fake good. Their participants were 129 undergraduate psychology students. Half of the group were instructed to complete a biodata measure accurately and honestly. The other half were instructed that the

measure would provide an indication of their potential to become psychologists. They completed an abbreviated version of the measure which contained items directly related to predicting success as a psychologist. As predicted, the second group scored significantly higher. The authors concluded that "experimentally induced motivations can be created leading examinees to increase typical performance measure scores without using instructions to fake in a specific fashion." They also found that extracting items from a test that was related to a particular scale and administering these items separately could affect the way an applicant scored. Removing these items from the context of the complete test may make the items more transparent as to the purpose of the test. If the applicants are highly motivated to do well on the test, they maybe tempted to try to improve their answers on these items.

LoBello & Sims (1993) found that prison inmates could better their scores on a pre-employment integrity test when instructed to fake good. The authors note that the tendency to fake good is not activated unless the individual is instructed to do so. They state that response distortion is not a problem if reasonable precautions are taken such as test administrators only giving the verbatim test instructions and not coaching test takers.

Thumin & Barclay (1993) administered a biodata test to a 161 job applicants and 73 graduate and undergraduate business students. Both groups were instructed to describe themselves honestly and accurately. The authors expected that the job applicants' scores would be higher under the selection condition and that they might be expected to fake good to some degree. The authors noted that while the employment group had smaller standard deviations than the student group on seven of the 11 scales, the overall profiles were virtually identical. The authors conclude that even if the individuals were faking good it had no negative impact on the reliability of the test. The results of this

study were consistent with those of Hough et al. (1990) who also examined the effects of faking on biodata validity and found the effects to be minimal. These studies cast doubt on the common assumption that faking is a serious threat to the validity of the biodata instruments, however more research is needed.

Mitchell (1994), in his review of the literature, states that biodata can be highly resistant to exaggeration or faking by applicants. He notes that biodata questions do not have right or wrong answers and that scoring keys are most often developed using criterion-based techniques to weight each possible answer to each question in direct reference to a job-relevant criterion such as job performance or turnover. He also notes that applicants are much less likely to exaggerate potentially verifiable and factual data. He notes that biodata measures offer a cost effective means to add to the validity and accuracy of personnel selection programs by capturing and predicting unique criterion variance and increases the overall accuracy of these selection programs when combined with other procedures. He also notes that biodata measures offer a fair and valid mechanism for decreasing the adverse impact of selection programs.

Becker & Colquitt (1992) stated that one of the limitations on previous biodata-faking research was that the participants were usually college students and they are "unlikely to adequately represent the population of individuals in the work force or labour market." They conducted two studies using actual employees or job applicants. The first study looked at general faking ability and actual faking of a biodata form as well as the effect of faking on selection decisions. The second study examined the relationship between the type of questions and the extent of faking in practice. The subjects were 58 employees of a large company who were a part of the first study and the second set of subjects consisted of 231 job applicants of which 50 were later hired by the company.

The current employees were divided into two groups. The first group was instructed to answer the 25 item biodata questionnaire as accurately as possible. The second group was told to fake the test and make themselves to look as good as possible. The group of real applicants were instructed to answer the questions accurately and that they were warned that the information they provided may be verified with other data collected during the employment process. The results indicated that the subjects instructed to fake scored an average of 5.7 points higher than those who were instructed to answer honestly. The faking group was also 2.6 points higher than the applicants who were later hired. However when compared to the entire group of applicants there was not a significant difference between the group that answered honestly and the real applicants. The authors concluded from this that while the test could be faked in artificial conditions in actual practice there appears to be much less faking occurring in an actual selection situation. The authors also noted that in the faking situation, the items that were most likely to be faked were the most job related and the more objective and verifiable items are less prone to faking in practice. They concluded that even if there was faking occurring that it would not have a significant impact on the selection decisions.

Methods of Detecting Faking

One way of attempting to detect faking is to use a check item. For example, Pannone (1984) included an item that referred to the use of a piece of electrical equipment that, in reality, did not exist. One third of the subjects indicated that they had experience with this piece of equipment. Pannone took this as an indication that this type of questionnaire was highly susceptible to faking and that the faking had a tendency to distort the validity coefficient. When the fakers were examined as a group, it was found that their mean score was significantly higher than the mean score for the non fakers.

Shermis et al. (1996) describes the construction and validation of a faking detector scale for an US Air Force officer selection biodata instrument. They modelled their scale on the Minnesota Mutiphasic Personality Inventory (MMPI) Lie and K scales. The Lie scale has been found to be effective "... in detecting gross response distortion, the K scale was effective in identifying more subtle response distortion and clinical defensiveness". The biodata faking-scale consisted of 12 multiple choice questions. The factor analysis revealed that the scale was unidimensional and accounted for 59% of the total variance and the test-retest reliability coefficient was only fair (.65). The correlations of the results to the two criterion indicators were low and insignificant. The authors attribute this to a possible restriction of range and the homogeneity of the participants as all of the participants had been already admitted into the Air Force and did not represent the target population. They also note that the scale may also misidentify a truly outstanding candidate.

Another method of detecting faking involves the use of a computer to measure the amount of time a participant takes to respond to an item. Holden(1993) theorized that someone who is making a conscious effort to fake would take a longer time to answer than someone answering honestly due to the requirement of accessing a number of schema. He administered a 158-item true/false biodata questionnaire to 64 first year undergraduate students. The subjects were randomly assigned to either a standard or fake group. The standard group was instructed to answer honestly while the faking group was instructed to maximize their chances for being hired. The subjects entered their responses on a personal computer which measured the amount of time it took for them to answer (response latencies). Holden found no difference in the response latencies between the honest and faking groups. He then replicated the study using 100 unemployed adults. The results indicated that those in the faking group did take longer to answer questions. The mixed results of

this study and the results of a previous study by Kluger, Reilly & Russell (1991) which found no significant difference in the response latencies between the test conditions indicates that more research is required.

In sum, there is a variety of methods to detect faking. A check item on a non-existent event or responses latencies measured by a computer have been shown to be able to give an indication of faking. Attempts to develop a faking detector scale have had less success.

Methods to Reduce Faking

Mael (1991) states that intrusive questions are "likely to incur resistance and resentment, and thereby encourage willful faking, random responding, or other behaviour aimed at foiling the testers." Some items that appear to be intrusive may be acceptable to applicants if seen to be relevant or fair. This should be taken into consideration when constructing biodata instruments. Mael (1996) had a panel of professionals and a group of potential applicants rate a set of biodata items for their perceived level of invasiveness. The results indicated that there was a very high level (.95) level of consensus between the two groups. By identifying potentially invasive items test authors may reduce the likelihood of faking.

Stone & Jones (1997) conducted a study to determine if "individuals will perceive the collection of biodata items as more fair when the data are collected for career tracking purposes than when the data are collected for personnel selection purposes." The second hypothesis was that "females are more likely to perceive that requests for biodata are unfair than are males." The authors employed a 2 x 2 experimental design, which assessed the effects of the purpose of the request for biographical data and the gender of the applicant on perceptions regarding the fairness of biodata items. The subjects were 86 graduate and senior undergraduate students (38 males/48

females). The average age was 22.58 and approximately 70% indicated that they were actively seeking work. They were randomly assigned into two groups. All of the participants were asked to complete the same biodata items. Only the purpose of the inquiry and the gender of the respondent varied.

One group was asked to play the role of a job applicant. The other group was told that the information was for research purposes only. The biodata questionnaire consisted of 39 items relating to family background, parental relationships, educational achievement, and leisure activities. After completing the biodata questionnaire they were asked to rate each item's fairness.

The results indicated that the participants viewed the 'fairness' of the items differently depending on the purpose for which the data were collected and the gender of the applicant. The authors noted that biodata items that were more directly related to an individual's job knowledge or skills, such as grades, academic standing, job preferences and behaviour when angry, were not rated differently by the two groups and received relatively high mean fairness ratings across the four conditions. Items such as leisure activities and family background were perceived as very unfair when asked for the purposes of personnel selection. The authors recommend that organizations use personnel selection methods that minimize intrusiveness such as only using biodata items that are perceived to be directly related to job performance and have been based on a careful analysis of the job in question. The results also indicated that females did not view a request for biodata as more unfair than males. Only the item concerning marital status was viewed as less fair by females than males.

Kluger, Reilly & Russell (1991) note that studies reporting problems with faking used item-keying(IK) strategies and that there were other studies that reported fewer problems with faking that

used an option-keying strategy. Item-keying is where you assign a numeric value to an option and assume that the other responses have a linear relationship to the criterion. For example, if there were four possible responses to an item the first might be scored as one, the second as two, the third as three and the fourth is worth four points. An option keying approach recognizes that not all responses may correlate with the criterion in a linear fashion and may be a negative value where an alternate response may be a positive value. For example, a Likert-type biodata item response of one may be scored minus one and response of four may be scored plus two with the other two responses scored a zero depending on how they have been chosen by high and low criterion groups. Kluger states that "option-keying methods offer the potential advantage of capturing both linear and nonlinear relationships between the item scale and the criterion."

The authors administered a 25 item Likert-type biodata questionnaire using computer administered testing to 85 graduate students. They used a 2 X 2 X 2 cross over design with one within-subjects factor and two between-subjects factors. The subjects were tested twice. They were first asked to complete the questionnaire honestly for research purposes and the second administration they were asked to simulate being a job applicant and were told that they "really wanted the job".

The results indicated that those applicants that were told to simulate applying for a job responded in a more socially-desirable way than did the subjects responding honestly. It was noted that the item-keyed scores were inflated by the faking but not the option-keyed scores. Not only did the scores not rise but they actually decreased under the faking condition (Kluger, Reilly & Russell, 1991).

Stokes, Hogan & Snell (1993) described two types of social desirability responding: self-deception and impression management. While self-deception involves an unconscious tendency to see oneself in a positive way, impression management is a conscious attempt to present oneself in a favourable light. It was noted that impression management in some occupations may not be a bad thing. It has been found to be the best single predictor for success as a salesman for example. The aim of the Stokes et. al. study was to measure the extent of impression management exhibited in applicant and incumbent samples when responding to a biodata form. Their participants were 2,262 incumbent sales representatives and 2,726 applicants for sales positions. A 171-item biodata questionnaire was administered that covered work experiences, preferences, education, background, personnel history and leisure time activities in a multiple choice format. The criterion used was the turnover rate after 12 months. Both item-keyed and option-keyed instruments were developed and validated on the same samples. It was found that the items that had the most significant validities within their respective validation samples had very little overlap between the applicant and the incumbent samples. This led the authors to caution about using incumbent samples for the development of biodata instruments as they do not appear to generalize to an applicant population. Their results also indicated there was a higher level of socially desirable responding in the applicant sample than there was in the incumbent sample.

Contrary to the results found by Kluger et al. (1991), there was no difference between the item-keying and option-keying scoring methods. Hogan (1994) notes that it appears option-keyed instruments are less vulnerable to score inflation than item-keyed instruments primarily when the item-criterion relationships are significantly nonlinear. He also notes that when the item-criterion

relationships are linear there is an advantage to using item-keyed methods as they would provide the same results and have the advantage of requiring smaller sample sizes and be easier to develop.

Mael & Hirsch (1993) noted that one method of reducing susceptibility to faking is to design items that are objective and easily verifiable but this can be problematic in that "objective actions tend to be heterogeneous and therefore difficult to attribute to unitary constructs". Their study used a 73-item biodata questionnaire. The items were limited to historical, external, objective, first-person, and verifiable attributes to minimize faking. The participants were 2,565 cadets at a US military academy. The results indicated that both approaches could successfully predict success at the military academy and were found to add incremental validity over and above the selection method that was currently being used. The results appeared to corroborate previous findings that historical, objective, discrete and verifiable items are less amenable to faking. The authors noted, however, that by using these types of items you give up the richness of information that can be found through less objective items.

Schrader & Osburn (1977) examined the effects of warning on reducing faking. Their subjects were 148 college students who were instructed to complete a biodata questionnaire and make themselves to look as good as possible. Half the group was warned that a faking scale might be included. Between one and two weeks later the same form was readministered to all subjects under honest instructions. The results indicated that instructing the subjects to make themselves look good increased their score by at least one standard deviation. The group that was warned that there might be a faking scale increased their score by only one half of a standard deviation.

Kluger and Colella (1993) studied the effects of warning on biodata test faking behaviour in a field study. They randomly warned 214 of 429 applicants for a nurse's assistant position against

faking. While the warning mitigated the propensity to fake, the specific warning effects depended on item transparency. For transparent items, warning reduced the extremeness of item means and decreased item variances. For non-transparent items, warning did not have an effect on item means and increased item variances. These faking effects were strongest when transparency was operationalized in terms of item-specific job desirability in addition to the item's general social desirability.

Roy (1994) examined the effect of a warning on biodata-faking behaviour. She hypothesized that respondents would be more likely to fake when they were not warned about potential consequences and that the effect of warnings would be strongest when the information was job relevant and not easily verified. Her subjects were 128 graduate and undergraduate business students who were instructed to role-play as job applicants. The results indicated that although there was a significant (at $\alpha = .10$) difference between warning levels within the High Job Relevance/Low Verifiability condition as predicted, the pattern of mean differences between the No Warning, Minimum Warning and Maximum Warning conditions was not interpretable. The author concludes that the differences were likely to be occurring by chance and that more research is required under more realistic conditions.

In general, the research suggests that to reduce the likelihood of faking, a well constructed biodata questionnaire should avoid intrusive questions and give the impression of being fair. The information being requested should be on discrete events and have the potential of being verified. Option-keyed biodata instruments may be less vulnerable to score inflation than item-keyed instruments but the results vary depending on how linear the item-criterion relationships are. There

is also some evidence that warning applicants on the consequences of faking can mitigate the propensity to fake. The warning effects varied, however, depending on the degree of job relevancy.

Statement of the Problem

Effective personnel recruitment and selection is one of the most critical steps in the establishment and growth of an organization. Many organizations expend a great deal of time and effort to match the right person to the job as there are many costs associated with hiring personnel who subsequently are not successful. These costs can be measured not only in dollar terms but also by their effect on the lives and careers of the persons involved (Strauss & Sayles, 1972).

One process that is frequently used in personnel selection is the collection of selected biographical information (biodata) from job applicants, which is then scored and used to make hiring decisions. Biodata forms are widely used, relatively inexpensive and have proven to be effective in predicting future job success (Mitchell, 1994). However, responses to biodata items are self-reports that are provided by applicants in order to accomplish their own goals: obtaining a job offer. Unless the information is verified, there is no assurance that the self-reports are accurate, and inaccurate information may reduce the validity of the selection system. As a result, there is concern about the extent to which applicants inaccurately report biographical information in a way that they believe will increase their chances of being hired. Nevertheless there has been relatively little research on the factors that might affect the accuracy of respones to a biodata form. The purpose of the present study is to examine some of the characteristics of biodata forms, and of biodata items that might affect the accuracy of applicants' self-reports.

Hypotheses

1. Based on previous findings by Schrader & Asburn (1977), Kluger & Colella (1993) and Roy (1994), it was expected that administering a forceful caution prior to the completion of a biodata questionnaire (versus the absence of a warning) would reduce the incidence of faking good. Based on the predictions of Mael (1991) and Lautenschlager (1994), it was expected that these effects would be observed particularly on items that were relevant to the job and which were not verifiable. Participants were expected to make the least accurate (most favourable) self-reports when there was no warning and items were job-relevant and not verifiable. Responses to these items were expected to be more accurate when a warning was present. The effects of the warning were expected to be smaller or absent on items that were not relevant to the job, or that were easily verified, or both.
2. Based on the predictions of Mael (1991), Lautenschlager (1994), it was expected that requesting additional details about specific items (versus not requesting details) would reduce faking. It was expected that the effects of requesting additional details would be strongest on items that were relevant to the job and which were otherwise not verifiable. Requesting additional information was expected to have less or no effect on items that were not job relevant, or that were easily verified.

Chapter II

Pilot Study

Method

Purpose

In the midst of the data collection phase, a change in the author's circumstances prompted a number of changes. The study had started as a research apprenticeship and it was decided to redo the questionnaire in order to complete the requirements for a Master's thesis. This first study was employed as a pilot study to help identify design weaknesses and modify the biodata questionnaire.

Participants

For the pilot study, 54 undergraduate and graduate business students volunteered to participate. One individual started but did not complete the questionnaire leaving a final N = 53. Participants' ages ranged from 18 - 56 years with a mean of 21.5 years (SD = 6.9).

Table 1

Gender and Age Frequencies of Pilot Participants

	n	% of non missing responses
Gender		
Male	24	45.3
Females	25	47.2
Unknown	4	7.5

Procedure

Once the nature of the study was described to the participants and they agreed to take part, they completed an informed consent form which is in appendix A. They were then each given a package of materials. They were instructed to read the instructions and complete the questionnaire.

Measures

The questionnaire package consisted of a fictional job advertisement, an acceptance letter, and the Biodata questionnaire. After completing the questionnaire, the participants were given a debriefing form.

Fictional Job Advertisement. The participants in the two role-playing groups were requested to role-play the part of job seekers. Participants in a third group were not given any role-play instructions. A logo of a fictional company called Master Corporation was placed prominently on all the forms to add realism. The advertisement indicated that management trainees were required for a wide range of entry level positions. The nature of the job and the remuneration was made to be very attractive. The following desired job applicant qualifications were specified:

> "We are looking for highly energetic, assertive, self-starting individuals. Working along with members of your team, you should be able to demonstrate excellent interpersonal and communication skills. You should be a highly organized individual with a record of success in achieving quality results. Bilingualism, computer literacy and flexibility to travel would be definite assets. Valid drivers' license required."

The entire form is in appendix B.

Cover Letter. After reading the job advertisement, participants read a letter informing them that the company considered them potentially suitable and now required further information as part of the

selection process. The applicant was requested to complete the enclosed biodata questionnaire. The cover letter is in appendix C.

The Biodata Questionnaire Instructions. There were three versions of the instructions. The first was for the base line group that was not part of the role-playing exercise. Therefore there was no simulated job application, no cover letter, and no forceful warning (Appendix D). The instructions read:

> "Thank you for agreeing to completing this questionnaire. The following questionnaire is being developed as part of a personnel selection research project. We are interested in obtaining an accurate sample so please respond to the questions as honestly as you can."

The second group received the simulated job advertisement, the cover letter and the following instructions (Appendix E):

> "As part of the job application procedure, here at Master Corp. Inc., we request that all candidates complete the following biographical information questionnaire. If you wish, you may refer back to the job advertisement.
>
> By responding to this questionnaire, you will provide the selection committee with a view of your background. All selection decisions will be based upon the information you provide.
>
> Please tick the appropriate response and/or fill in the blank with the requested information."

The third group received the simulated job advertisement, the cover letter, the instructions as above and the following forceful warning printed in bold face (Appendix F):

> "By using biographical questionnaires, our company can more accurately verify your information. Master Corp. may choose to check any information provided for accuracy and completeness.
>
> Statements of incorrect, misleading or incomplete information are grounds for dismissal."

To ensure that they had indeed read the instructions, all participants were requested to mark a box indicating they had read and understood the instructions.

Biodata Questionnaire. Due to the proprietary nature of commercial biodata instruments, a new biodata questionnaire had to be constructed. The items were constructed based on the recommendations found in Russell (1994) and Mumford (1996). They were also based on Mael's (1991) taxonomy of biodata item attributes which described prototypical items as historical, external, verifiable, objective, first-hand, discrete, controllable, equal access, job relevant, and noninvasive. A number of items that were deliberately hard to verify and/or not job related were included for contrast purposes. The initial draft of the questionnaire was reviewed by a panel of professors and graduate psychology students. They were asked to comment on face validity and rate each item as to its degree of job relatedness and verifiability. Based on the results of the panel, a total of 37 items tapping into several aspects of the respondents' past were selected for the pilot questionnaire. Some of the items (High Relevancy) were selected as they related back to desirable characteristics identified in the job advertisement. Others were selected to be less relevant to the job (Low Relevancy). Similarly, some of the items were selected to be easily verifiable (High Verifiability) while others were not easily verified (Low Verifiability). Item job relevancy and

verifiability were confirmed by having 48 participants (not in the current study) rate them on each of these dimensions (Appendix G).

Debriefing Form. Upon completing the questionnaire, the participants were provided with a debriefing form which explained the purpose of the study. They were requested not to talk about the study to anyone else as there would be further testing in the future. If they wished a copy of the research findings, they were invited to leave an address (Appendix H).

Results and Discussion

The pilot study was beneficial from the point of view that it demonstrated some of the weaknesses of the initial design. It was apparent that there was no significant difference in the results of the baseline group and the role-playing group with the minimum warning so it was decided to drop the baseline group from the main study. It was noted that the participants were not reading the job advertisement and covering letter closely enough which prompted the inclusion of more direct instructions at the top of each page. The forceful warning was placed on the first page of the questionnaire and highlighted in a box to increase its visibility and impact. Feedback from the raters indicated that the verifiability of the items depended on how much detail was requested. This prompted the creation of an extended version of the questionnaire where certain items asked for additional specific details such as names and dates. It was also decided that the only way to increase the number of participants was to employ Introductory Psychology students who could volunteer in order to receive a credit for research participation. This resulted in having to change the role-playing scenario from the participants applying for full-time jobs to applying for a summer internship program.

Chapter III

Main Study

Method

Design

The study consisted of a 2 (warning) x 2 (versions) between-subjects multivariate design, with four separate sets of dependent measures. The between-subjects factors were (1) the consequences of inaccurate responses (warning present or absent) and (2) the amount of detail requested in responses (extended detail or no detail). The biodata items were divided into four sets of dependent measures, reflecting the different combinations of (a) relevance to the advertised job (Relevant or Not Relevant) and (b) whether or not the information they requested could be independently verified (Verifiable or Not Verifiable). A separate 2x2 MANOVA was conducted on each of the fours sets of items (Relevant and Verifiable; Relevant and Not Verifiable; Not Relevant and Verifiable; Not Relevant and Not Verifiable). The selection of items to be analyzed in each MANOVA was based on the verifiability and relevance ratings of 48 people who did not participate in the current study (Appendix G). The items used within each MANOVA are listed in Appendix I.

Participants

The participants were 212 Introductory Psychology students. These students volunteered to participate in return for receiving credit toward their course research participation marks. 65 were male and 147 were female. Participants' ages ranged from 17 - 45 years (M = 19.3, SD = 2.8)(See Table 2). Using a formulation proposed by Green (1991) regarding power

of analysis, it was determined that the sample size exceeded the minimum required to conduct multivariate regression analyses.

Table 2

Gender and Age Frequencies of Main Study Participants

	n	% of Nonmissing Responses
Gender		
Male	66	31
Females	147	69
Age		
17	3	1.4
18	106	49.8
19	51	23.9
20	21	9.9
21	12	5.6
22	5	2.3
23	6	2.8
24	2	.9
25	1	.5
26	1	.5
27	1	.5
34	1	.5
35	1	.5
45	1	.5

Note. N = 213 M = 19.3, SD = 2.8

Procedure

Due to the large size of the groups, the testing sessions were conducted in a large lecture hall. Once the nature of the study was described to the participants and they agreed to take part, they received a receipt for their research participation credit and then completed an informed

consent form which is in appendix J. Each was then given a package of materials. They were instructed to read the instructions and complete the questionnaire.

Measures

The questionnaire package consisted of a fictional job advertisement, an acceptance letter, and the Biodata questionnaire. After completing the questionnaire, the participants were given a debriefing form.

Fictional Job Advertisement. The Job advertisement was on a one page job poster complete with a logo of the fictional company called Master Corporation. The text indicated that management trainees were required for a wide range of entry level positions. The nature of the job and the remuneration was made to be very attractive. The advertisement was modified based on observations from the pilot study. It was decided to put instructions on the job advertisement to ensure the participants understood what was required. (In bold)

> "You are taking part in a simulated job application process. Read the following job advertisement closely so that you have a clear idea of the job you will be applying for."

Since the participants were now first year students, the position was changed from management trainee to a management internship program that would give them summer employment leading to a job on graduation. As a further incentive, they were offered a salary plus full scholarship for the remainder of their university program. The advertisement is in appendix K.

Cover Letter. To reinforce the role-playing instructions the following was placed at the top of the cover letter. (In bold)

> "You are very interested in this job. You sent a letter indicating your interest but did not include a resume. You receive the following letter in response:"

The letter informed the research participant that the company considered him or her to be potentially suitable and now required further information as part of the selection process. The applicant was requested to complete the enclosed questionnaire. The cover letter is in appendix L.

The Biodata Questionnaire Instructions. The questionnaire instructions were placed at the top of the first page of the questionnaire. There were two versions of the instructions. The "No Warning" group received the simulated job advertisement, the cover letter and the following instructions:

> "As part of the job application procedure, here at Master Corp. Inc., we request that all candidates complete the following biographical information questionnaire. If you wish, you may refer back to the job advertisement.
>
> By responding to this questionnaire, you will provide the selection committee with a view of your background. All selection decisions will be based upon the information you provide. Please tick the appropriate response and/or fill in the blank with the requested information."

The "Warning" group received the simulated job advertisement, the cover letter, the instructions as above and the following forceful warning highlighted by being bolded and in a box:

"Please read the following carefully: By using biographical questionnaires, our company can more accurately verify your information. Master Corp. may choose to check any information provided for accuracy and completeness. Statements of incorrect, misleading or incomplete information are grounds for dismissal."

Prior to starting the questionnaire, all participants were requested to place a mark in a box to ensure that they had indeed read the instructions.

Biodata Questionnaire. A number of new items were added to the pilot version of the biodata questionnaire which raised the total number of items to 65 with some having multiple response options. There were two versions of the biodata questionnaire. The abbreviated version asked yes/no and multiple-choice type questions (appendix M). The items in extended version (appendix N) required the respondent to provide further details such as names and dates to increase the potential verifiability of some of the items. The extended items are listed at appendix O.

Certain job relevant items were directly related to the requirements outlined in the fictional job advertisement. In addition, a single item was included which asked participants if they had experience with a (nonexistent) computer software package whose name was fabricated for this study (see item 49e at appendix M). This was a check on faking behaviour that was based on a study by Pannone (1994).

Amended Debriefing Form. As the participants were first year Introductory Psychology students, some references were added in case they would like to learn more about biodata measures. The amended debriefing form is at appendix P.

Test Condition Items. Embedded within the biodata questionnaire were items which would be analysed separately for the four test conditions: high verifiability /low relevance items (HV/LR), high verifiability /high relevance items (HV/HR), low verifiability /low relevance items (LV/LR), and low verifiability /high relevance items. There were six items for HV/LR (items 4a, 5c, 5d, 6, 7, 8) which asked about changing a major area of study, language proficiency in Spanish and German, and activity in student politics or organizations. There were four items analysed for HV/HR (items 5b, 36, 37, 41), which asked about language proficiency in French, days absent in the last 6 months, hours worked, and customer service experience. For LV/LR there were six items (42, 43, 59, 60, 61, 62) which asked about experience dealing with rude customers, keeping a personal budget, discussing personal issues, allowing time for leisure activities, number of books read in three months, and recreational sports activities. The final condition, LV/HR, asked about preferring individual over group participation and the degree of difficulty they encountered in planning, carrying out plans, monitoring progress and completing a task on a project. All of the items by condition are listed in full at appendix I.

Chapter IV

Results

Data Preparation

Prior to analysis, the data were examined for univariate and multivariate outliers. Using the criteria suggested by Tabachnick & Fidell (1989) defining univariate outliers as being more than three standard deviations from the mean and severely discontinuous from the rest of the distribution, no univariate outliers were found. The data were then examined for multivariate outliers defined as cases with a large and discontinuous Mahalanobis distance. Based on this criterion, no multivariate outliers were found in the data set.

Although, there was no indication from the literature to expect a difference due to gender, initial analyses were conducted using a series of ANOVAs with gender, warning condition and version as the independent variables. Results confirmed there were no significant gender effects and this variable was discarded for the main analysis.

The data set was examined for normality and linearity. No transformations were done to correct for skewness as the sample sizes were large enough to be robust to violations (Tabachnick & Fidell, 1989). Bivariate plots within and across sets of measures did not show any violation of linearity.

Main Analysis

Four separate 2 x 2 SPSS MANOVAs were conducted. In each, the independent variables were Level of Warning (high or low) and Version (Extended or Abbreviated). Alpha for significance was set at .05 per effect.

The first MANOVA examined the effects of Warning and Version on the high verifiable low relevance items. An examination of Box's M indicated the data had not violated the assumption of homogeneity of dispersion matrices ($p < .10$). The descriptive statistics for the items are in table 3.

The correlations among the measures are presented in Table 4 and indicate that the responses to the 6 items are largely unrelated to each other except for the three items on participation in student politics (items 6, 7 and 8). With the use of Pillai's criterion it was determined that there were no significant main or interaction effects at an alpha of .05 (see Table 5).

Table 3

Means (SD) of High Verifiability/ Low Relevance Items by Condition

		Version			
		Extended		Abbreviated	
Item	Label	Warning	No Warning	Warning	No Warning
0.17	Change Major	.18 (0.4)	.23 (0.4)	.23 (0.4)	.12 (0.3)
5c	Spanish	.78 (1.5)	1.00 (1.7)	.56 (1.4)	.81 (1.4)
5d	German	.64 (1.4)	.81 (1.7)	.46 (1.1)	.57 (1.5)
6	Student Politics	.40 (0.5)	.38 (0.5)	.31 (0.5)	.49 (0.5)
7	Run for Organizations	.40 (0.5)	.38 (0.5)	.31 (0.5)	.27 (0.5)
8	Elected to Organizations	.36 (0.5)	.35 (0.5)	.19 (0.4)	.31 (0.5)
	Sample Size (N)	50	48	52	49

Note. Item 4a Changed major in university (1=yes; 2 = no)
5b Spanish proficiency (0 = none; 10 = fluent)
5c German proficiency (0 = none; 10 = fluent)
6 Student politics (1=yes; 0 = no)
7 Run for org (1=yes; 0 = no)
8 Elected org (1=yes; 0 = no)

Table 4

WITHIN+RESIDUAL Correlations with Std. Devs. on Diagonal

Items	4A	5C	5D	6	7	8
4A	.395					
5C	-.099	1.495				
5D	-.015	.146	1.428			
6	-.063	.177	.217	.489		
7	.000	.154	.081	.464	.474	
8	-.065	.198	.113	.526	.740	.458

Note. Item 4a Changed major in university (1=yes; 0 = no)
5c Spanish proficiency (0 = none; 10 = fluent)
5d German proficiency (0 = none; 10 = fluent)
6 Student politics (1=yes; 0 = no)
7 Run for org (1=yes; 0 = no)
8 elected org (1=yes; 0 = no)

Table 5

Results of MANOVA on High Verifiability/Low Relevancy

Version type (extended or short versions)	F(6, 190) = 1.05, p < .39
Warning condition (maximum or minimum warning)	F(6, 190) = .88, p < .51
Interaction of version type and warning condition	F(6, 190) = .96, p < .45

Note. Biodata item numbers are 4a, 5c, 5d, 6, 7, 8 and are listed in appendix I.

The second MANOVA examined the effects of Warning and Version on the high verifiable, high relevance items. An examination of Box's M indicated the data had violated the assumption of homogeneity of dispersion matrices p < .001, indicating the results of significance tests may be somewhat less precise and should be interpreted with caution. The descriptive statistics for the items are in Table 6.

Table 6

Means (SD) of High Verifiability/ High Relevance Items by Condition

		Version			
		Extended		Abbreviated	
Item	Label	Warning	No Warning	Warning	No Warning
5b	French	6.13 (3.0)	6.27 (2.5)	6.39 (2.6)	6.54 (2.8)
36	Days Absent	2.08 (3.2)	2.00 (3.0)	1.61 (2.8)	0.98 (1.3)
37	Hrs/ Week Worked	26.59 (14.0)	33.16 (14.0)	24.89 (16.1)	27.52 (13.6)
41	Deals With Public	.89 (0.3)	.86 (.4)	.94 (.3)	.95 (.2)
	Sample Size (N)	46	44	46	48

Note. Item 5b French proficiency (0 = none; 10 = fluent)
36 Days absent (____ days)
37 Hours per week worked (___ hours)
41 Deal with public (1=yes; 0 = no)

The correlations among the measures are presented in Table 7 and indicate that the responses to the 4 items are largely unrelated to each other. With the use of Pillai's criterion it was determined that there were no significant main or interaction effects at an alpha of .05 (see Table 8). The effect for additional information(extended or short version) did come very close to being significant ($p < .053$) and may merit further research. Examination of the group means indicates that participants in the extended version condition reported a slightly lower level of French proficiency, slightly more days absent and that they were slightly less likely to have dealt with the public in the past. These findings are consistent with expectations as they suggest that those who were asked to provide more detailed information gave less job-favourable responses. However, participants in the extended version condition also reported more hours worked per week. This trend is contrary to expectations as the request for extended information appears to have resulted in somewhat more favourable self-reports on this measure. These differences do not meet the criterion set for significance and therefore should be viewed as suggestive.

The third MANOVA examined the effects of Warning and Version on the low verifiable, low relevance items. An examination of Box's M indicated the data had violated the assumption of homogeneity of dispersion matrices $p < .014$, indicating the results of significance tests may be somewhat less precise and should be interpreted with caution. The descriptive statistics for the items are in Table 9.

Table 7

WITHIN+RESIDUAL Correlations with Std. Devs. on Diagonal

Items	5b	36	37	41
5b	2.717			
36	.103	2.640		
37	- .053	- .027	14.436	
41	.026	.046	-.145	.282

Note. Item 5b French proficiency (0 = none; 10 = fluent)
36 Days absent (____ days)
37 Hours per week worked (___ hours)
41 Deal with public (1=yes; 0 = no)

Table 8

Results of MANOVA on High Verifiability/High Relevancy

Version type (extended or short versions)	$F(4,177) = 2.38, p < .053$
Warning condition (maximum or minimum warning)	$F(4,177) = 1.43, p < .230$
Interaction of version type and warning condition	$F(4,177) = 0.40, p < .810$

Note. Biodata item numbers are 5b, 36, 37, 41 and are listed in appendix I.

The correlations among the measures are presented in Table 10 and indicate that the responses to the 6 items are largely unrelated to each other. With the use of Pillai's criterion it was determined that there were no significant main or interaction effects at an alpha of .05 (see Table 11).

Table 9

Means (SD) of Low Verifiability/ Low Relevance Items by Condition

		Version			
		Extended		Abbreviated	
Item	Label	Warning	No Warning	Warning	No Warning
42	Arrogant/rude	.71 (.5)	.84 (.4)	.93 (.3)	.88 (.3)
43	Discuss personal issues	2.77 (1.1)	2.93 (1.1)	3.25 (1.0)	2.84 (.9)
59	Budget	.61 (.5)	.53 (.5)	.54 (.5)	.54 (.5)
60	Leisure activities	3.86 (.7)	3.98 (.9)	4.02 (.9)	4.10 (.8)
61	How many books read	2.78 (3.3)	1.67 (2.0)	1.84 (2.8)	1.96 (2.6)
62	Physical activities	.78 (.4)	.76 (.4)	.82 (.4)	.82 (.4)
	Sample Size (N)	51	45	44	50

Note. Item 42 Experience with rude customers (1=yes; 0 = no)
43 Discuss personal issues (1 = never; 5 = always)
59 Personal budget (1=yes; 0 = no)
60 Make time for leisure (1 = never; 5 = always)
61 How many books read (___ books)
62 Physical activities/12 months (1=yes; 0 = no)

Table 10

WITHIN+RESIDUAL Correlations with Std. Devs. on Diagonal

Items	42	43	59	60	61	62
42	.378					
43	.107	1.034				
59	- .065	-.008	.501			
60	- .062	.159	-.099	.812		
61	.030	.053	.195	-.011	2.732	
62	.015	-.035	-.032	.135	.064	.407

Note. Item 42 experience with rude customers (1=yes; 0 = no)
43 Discuss personal issues (1 = never; 5 = always)
59 Personal budget (1=yes; 0 = no)
60 Make time for leisure (1 = never; 5 = always)
61 How many books read (___ books)
62 Physical activities/12 months (1=yes; 0 = no)

Table 11

Results of MANOVA on Low Verifiability/Low Relevancy

Version type (extended or short versions)	F(6, 181) = 1.55, p < .16
Warning condition (maximum or minimum warning)	F(6, 181) = .67, p < .68
Interaction of version type and warning condition	F(6, 181) = 1.45, p < .20

Note. Biodata item numbers are 42, 43, 59, 60, 61, 62 and are listed in appendix I.

The fourth MANOVA examined the effects of Warning and Version on the low verifiable high relevance items. An examination of Box's M indicated the data had not violated the assumption of homogeneity of dispersion matrices (p < .76). The descriptive statistics for the items are in Table 12.

Table 12

Means (SD) of Low Verifiability/ High Relevance Items by Condition

		Version			
		Extended		Abbreviated	
Item	Label	Warning	No Warning	Warning	No Warning
10	Projects	3.19 (1.0)	3.27 (.8)	3.17 (1.0)	3.06 (1.0)
20	Planning/organizing	2.54 (.6)	2.71 (.7)	2.76 (.8)	2.68 (.7)
21	Carrying out Plans	2.50 (.6)	2.71 (.6)	2.65 (.6)	2.56 (.6)
22	Monitoring Progress	2.25 (.8)	2.30 (.6)	2.15 (.7)	2.24 (.7)
23	Completing the Task	2.46 (.7)	2.63 (.7)	2.41 (.6)	2.44 (.6)
39	Turned down request	2.96 (1.7)	2.80 (2.0)	2.85 (1.9)	3.08 (2.0)
	Sample Size (N)	48	44	46	50

Note. Item 10 Projects (1 = prefers individual projects; 5 = prefers group projects)
20 Planning (1 = not responsible; 4 = difficult)
21 Carrying out Plans (1 = not responsible; 4 = difficult)
22 Monitoring Progress (1 = not responsible; 4 = difficult)
23 Completing the Task (1 = not responsible; 4 = difficult)
39 Turned down request (1 = quit; 6 = Persisted)

The correlations among the measures are presented in Table 13 and indicate that the responses to the 6 items are largely unrelated to each other. With the use of Pillai's criterion it was determined that there were no significant main or interaction effects at an alpha of .05 (see Table 14).

A separate 2 x 2 ANOVA was conducted for the single item designed to check faking. While 19.2% of the participants indicated they had experience with this software, there were no differences as a function of either between-subject factor (warning and extended versions). Descriptive statistics on this item are displayed in Table 15 and the ANOVA results are displayed in Table 16.

Table 13

WITHIN+RESIDUAL Correlations with Std. Devs. on Diagonal

Items	10	20	21	22	23	39
10	.948					
20	.012	.708				
21	-.008	.141	.616			
22	.075	.129	.145	.704		
23	.052	.098	.234	.248	.624	
39	-.013	.069	.110	.107	.077	1.889

Note. Item 10 Projects (1 = prefers individual projects; 5 = prefers group projects)
20 Planning (1 = not responsible; 4 = difficult)
21 Carrying out Plans (1 = not responsible; 4 = difficult)
22 Monitoring Progress (1 = not responsible; 4 = difficult)
23 Completing the Task (1 = not responsible; 4 = difficult)
39 Turned down request (1 = quit; 6 = Persisted)

Table 14

Results of MANOVA on Low Verifiability/High Relevancy

Version type (extended or short versions)	F(6, 179) = .64, p < .70
Warning condition (maximum or minimum warning)	F(6, 179) = .26, p < .96
Interaction of version type and warning condition	F(6, 179) = .91, p < .49

Note. Biodata item numbers are 10, 20, 21, 22, 23, 39 and are listed in appendix I.

Table 15

Descriptive Statistics for Check Item by Test Condition

Condition	Mean	Std Dev	Minimum*	Maximum	N	n**	%***
Abbreviated - Min Warning	.82	1.7	1	6	49	10	20.3
Abbreviated - Max Warning	.51	1.3	1	5	47	8	17.1
Extended - Min Warning	1.21	2.5	2	10	47	12	25.6
Extended - Max Warning	.79	1.7	1	6	52	11	21.1

Note. Scale range was 0 = no experience with Intrude 4.x, 5 = a working knowledge and 10 = a great deal of expertise. *Min score > 0, n** scores > 0, %*** percentage of participants indicating some experience.

Table 16

Tests of Significance for Faking Check Item

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	669.64	191	3.51		
Version type	5.53	1	5.53	1.58	.211
Warning condition	6.48	1	6.48	1.85	.175
Interaction	.17	1	.17	.05	.825

Discussion

The purpose of this study was to determine if a forceful caution could reduce the incidence of faking good on a biodata questionnaire by participants who were role-playing job applicants. In addition, this study explored the possibility that asking for additional details in response to multiple-choice type items would increase the accuracy of responses.

Consistency With Past Findings

None of the factors in this design (item verifiability, item-job relevance, the consequences of misreporting, and requesting extended details behind the answers to some items) had a systematic effect on the biodata responses. These findings are consistent with some other studies that employed university students as participants (Abrahams et al., 1971; Hough et al., 1990; Holden, 1993; Thumin & Barclay, 1993; Roy, 1994). For example, Roy (1994) attempted to simulate a personnel selection situation by having university students role-play as job applicants. She attempted to reduce faking by varying the levels of warning given to the participants. Although one of her conditions produced significant results (at $\alpha = .10$), she concluded that the absence of a pattern in the direction of the results indicated that it was likely due to chance rather than a real effect.

However, not all studies using student samples have failed to find significant results. Schrader and Osburn (1977) found that warning participants that their answers would be verified had an effect on faking with college students who were instructed to try to fake good. The only effect the warning had was to reduce the degree of faking in the group that had been warned that their answers may be checked. Becker & Colquitt (1992) concluded that college students are "unlikely to adequately represent the population of individuals in the work force or labour

market." They noted that most studies had tried to determine if a test could be faked and not if the applicants were actually faking in real life. Their study found no significant difference between a group instructed to answer honestly and a group of real applicants who could have been expected to be motivated to fake good. They concluded that there appears to be very little faking occurring outside of artificial situations. If they are correct, it could account for the failure to find significant effects in the present study, as participants were not encouraged to fake good, and in fact were encouraged by the instructions to answer honestly.

Holden (1993) found no difference in response latencies of undergraduate students who were instructed to answer honestly and others who were instructed to fake good. He did, however, find the differences in response latencies when he administered the same protocol to unemployed adults in a real world situation. Thumin & Barclay (1993) expected real job applicants to score higher than college students who were instructed to answer honestly. Their results showed no difference between the two groups. This would seem to contradict Becker & Colquitt's assertion that college students do not represent the work force, and suggests that both populations tend to respond honestly unless they are instructed to do otherwise.

There are some studies that indicate that there is faking occurring and that levels of warning do have the effect of reducing faking and thereby increasing the accuracy of the test. Kluger & Colella (1993) found that warning applicants for a nursing assistant position mitigated the propensity to fake but the strength of the effect varied. They found that faking effects were strongest when transparency was operationalized in terms of item-specific job desirability and that warning did not have an effect for non-transparent items. Lautenschlager (1994) states that there is evidence that warnings of possible verification can reduce faking but also notes that there

appears to be a relationship between faking and age or level of experience. The current study found some suggestion that asking for detailed information on items that are both more relevant to the job, and more easily verified, may increase accuracy in responses. However, this effect did not quite meet the criterion for significance, and the pattern of means was not entirely clear in their meaning.

Study Limitations

The most likely reason for the absence of significant effects in the current study is the nature of the sample and the setting used for the study. It is difficult to determine whether the results would have been different if the participants had been older and more experienced as job seekers. Although there is some indication in the literature that university students can effectively role play as job seekers it makes intuitive sense that first year students are not ideal as they are less likely to have extensive job-seeking experience.

In addition, this population may have little motivation to fake good in role-play situations. This is consistent with the pattern of correlations between biodata items. Responses to the various biodata items were generally not correlated with each other. If some participants were consistently faking good across a number of items, one would expect responses to these items to be correlated with each other in spite of the independence of their content (i.e., the same people would score relatively high on all of them). The absence of correlations between responses to the different items suggests that the same participants were not consistently faking their responses across a number of items.

Participants in this study completed the materials in a large-group setting. As a result the role-playing may not have been realistic and may have made it more difficult for participants to

get into their roles. In retrospect, a smaller room could have been set up to better represent a business location rather than have large groups of participants complete the instrument in an amphitheatre.

The level of warning failed to have an effect on the results. This could have been an artifact of the warning not being meaningful to the participants. The same could be said for the failure of Relevancy and Verifiability to be significant. The effectiveness of these manipulations relied upon participants "getting into their role", and taking the warnings seriously.

One possible explanation for the lack of differences between conditions is that participants generally responded honestly to all items and therefore conditions designed to reduce faking would have had no effect. While this may have happened to some extent (i.e., participants were not motivated enough to fake good), it is unlikely that all answers were honest and accurate since more than 19% of the participants reported having some experience with a computer program called Intrude 4.0. This was a check item and there is no such program in existence. Clearly the responses must be some form of error or may be evidence of "faking good" behaviour. The manipulations in the study had no effect on this item as there were no significant differences between the groups. This may suggest that neither the threat of verification nor the warning had any effect and participants responded inaccurately in spite of these conditions. Alternately it may be that the participants who falsely claimed to have experience with the non-existent program were not aware they were being inaccurate.

Directions for Future Research

Unfortunately the absence of significant results prevents any conclusions about the factors investigated in this study. The effect for additional information (extended or short version) did come close to being significant and merits further research. Future research should be oriented toward investigating these factors in a more realistic setting and ideally with participants who are actually applying for a position which they want to obtain.

References

Abrahams, N. M., Neumann, I., & Githens, W. H. (1971). Faking vocational interests: Simulated versus real life motivation. Personnel Psychology, 24, 5-12.

Adler, S. (1987). Toward the more efficient use of assessment centre technology in personnel selection. Journal of Business and Psychology, 2 (1), 75-93.

Asher, J. J. (1972). The biographical item: Can it be improved? Personnel Psychology, 25, 251-269.

Baehr, M. E., Jones, J. W., Baydoun, R. B., & Behrens, G. M. (1994). Proactively balancing the validity and legal compliance of Personal background measures in personnel Management. Journal of Business and Psychology, Volume 8 (No. 3).

Becker, T. E., & Colquitt, A. L. (1992). Potential versus actual faking of a biodata form: an analysis along several dimensions of item type. Personnel Psychology, 45, 389 - 406.

Brush, D. H., Owens, W. A. (1979). Implementation and evaluation of an assessment classification model for manpower utilization. Personnel Psychology, 32, 369-383.

Campion, M., Pursell, E. D., & Brown, B. K. (1988). Structured Interviewing: Raising the psychometric properties of the employment interview. Personnel Psychology, 41, 25 - 42.

Carlson, K. D., & Scullen, S. E. (1999). Generalizable biographical data validity can be achieved without multi-organizational development. Personnel Psychology, 52(3), 731 - 756.

Christiansen, N. D., Richard D, G., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. Personnel Psychology, 47, 847-861.

Cronshaw, S. F. (1986). The status of employment testing in Canada: A review and evaluation of theory and professional practice. Canadian Psychology, 27(2), 183-195.

Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behaviour on a forced-choice self description checklist. Personnel Psychology, 15, 13-24.

Ellis, R. T., & Spinner, B. (August 1997). Experimental predictors of BOTC performance. Ottawa, Ontario: Canadian Forces Personnel Research Team.

Fine, S. A., & Cronshaw, S. (1994). The Role of Job Analysis in Establishing the Validity of Biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata Handbook (pp. 39-64). Palo Alto, California: CPP Books.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? Multivariate Behavioral Research, 26 (3), 499-510.

Harris, M. M., & Schaubroeck, J. (1988). A Meta-Analysis Of Self-Supervisor, Self-Peer, and Peer-Supervisor Ratings. Personnel Psychology, 41, 43-62.

Holden, R. R. (May 1993). Response latency detection of fakers on personnel tests. Paper presented at the Canadian Psychological Association Annual Convention, Montreal, Quebec, Canada.

Hughes, J. F., Dunn, J. F., & Baxter, B. (1956). The validity of selection instruments under operating conditions. Personnel Psychology, 9, 321-324.

Kilcullen, R. N., White, L. A., Mumford, M. D., & Mack, H. (1993). Assessing the construct validity of rational biodata scales. Military Psychology, 7(1), 17-28.

Kluger, A. N., Reilly, R.R., & Russell, C.J. (1991). Faking biodata tests: are option-keyed instruments more resistant? Journal of Applied Psychology, 76(6), 889-896.

Kluger, A. N., & Colella, A. (1993). Beyond the mean bias: The effect of warning against faking on biodata item variances. Personnel Psychology, 46, 763-861.

Lautenschlager, G. J., & Shaffer, G. S. (1987). Reexamining the component stability of Owen's biographical questionnaire. Journal of Applied Psychology, 72(1), 149-152.

LoBello, S. G., & Sims, B. N. (1993). Fakability of a commercially produced pre-employment integrity test. Journal of Business and Psychology, 8(No 2).

Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items. Personnel Psychology, 44, 763-927.

Mael, F. A., & Hirsch, A. C. (1993). Rainforest empiricism and quasi-rationality: Two approaches to objective Biodata. Personnel Psychology, 46, 719-738.

Mael, F. A. (1994). If past behaviour really predicts future, so should biodata's. In M. G. Rumsey, C. Walker, & J. Harris (Eds.), Personnel Selection and Classification (pp. 273-292). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Mael, F. A., Connerley, M., & Morath, R.A. (1996). None of your business: Parameters of biodata invasiveness. Personnel Psychology, 49, 613-650.

McManus, M. A., & Kelly, M. L. (1999). Personality measures and biodata: Evidence regarding their incremental predictive value in the life insurance industry. Personnel Psychology, 52, 137-148.

Melamed, T. (1992). Use of biodata for predicting academic success over thirty years. Psychological Reports, 71, 31-38.

Mitchell, T. W. (1994). The utility of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata Handbook (pp. 485-516). Palo Alto, California: CPP Books.

Mumford, M. D., Costanza, D.P., Connelly, M.S., & Johnson, J.F. (1996). Item generation procedures and background data scales: Implications for construct and criterion-related validity. Personnel Psychology, 49, 360-398.

Nickels, B. J. (1994). The nature of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata Handbook (pp. 1-16). Palo Alto, California: CPP Books.

Pannone, R. D. (1984). Predicting test performance: A content valid approach to screening applicants. Personnel Psychology, 37, 507-514.

Pannone, R. D. (1994). Blue collar selection. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata Handbook (pp. 261-274). Palo Alto, California: CPP Books.

Reilly, R. R., & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. Personnel Psychology, 35, 1-62.

Roy, N. R. (1994). The Effects of Differing Instructions on Biographical Information. Unpublished honours theses. University of New Brunswick.

Schrader, A. D., & Osburn, H. G. (1977). Biodata faking: Effects of induced subtlety and position specificity. Personnel Psychology, 30, 395-404.

Shaffer, G. S., Saunders, V., & Owens, W. A. (1986). Additional evidence for the accuracy of biographical data: Long-term retest and observer ratings. Personnel Psychology, 39, 791-809.

Sharf, J. C. (1994). The impact of legal and equal employment opportunity issues on personal history inquiries. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata Handbook (pp. 351-390). Palo Alto, California: CPP Books.

Shermis, M. D., Falkenberg, B., Appel, V. A., & Cole, R. W. (1996). Construction of a Faking Detector Scale for a biodata survey instrument. Military Psychology, 8(2), 83-94.

Stokes, G. S., Hogan, J. B., & Snell, A. R. (1993). Comparability of incumbent and applicant samples for the development of biodata keys: The influence of social desirability. Personnel Psychology, 46, 739-761.

Stone, D. L., & Jones, G. E. (1997). Perceived fairness of biodata as a function of the purpose of the request for information and gender of the applicant. Journal of Business and Psychology, 11(3), 313-323.

Strauss, G., & Sayles, L. (1972). Personnel: The Human Problems of Management. Englewood Cliffs, NJ., Prentice-Hall, Inc.

Stricker, L. J., & Rock, D. A. (1998). Assessing leadership potential with a biographical measure of personality traits. International Journal of Selection and Assessment, 6(3), 164 - 184.

Tabachnick, B. G., & Fidell, L. S. (1989). Using multivariate statistics (2nd ed.). New York: Harper Collins.

Thayer, P. W. (1977). Something old and something new. Personnel Psychology, 30, 513-524.

Thumin, F. J., & Barclay, A. G. (1993). Faking Behaviour and Gender Differences on a New Personality Research Instrument. Consulting Psychology, 45(4), 11-22.

Wilkinson, L. J. (1997). Generalizable biodata? An application to the vocational interests of managers. Journal of Occupational and Organizational Psychology, 70, 49-60.

Winter, B., & Dennison. D. (1986). Comparative selection measures: Utilization of biographical information. London: The Technical Co-operation Programme Subgroup U Technical Panel 3 (Military Manpower Trends).

Young, M. C., & White, L. A. (1995). Assessment of Background and Life Experiences (ABLE): U.S. Army Research Institute for the Behavioural and Social Sciences.

Zalinski, J. S., & Abrahams, N. M. (1979). The effects of item context in faking personnel selection inventories. Personnel Psychology, 32, 161-166.

APPENDICES

Appendix A
Pilot Study - Informed Consent Form

Appendix B
Pilot Study - Fictional Job Advertisement

Appendix C
Pilot Study - Covering letter

Appendix D
Pilot Study - Instructions Base Line Group

Appendix E
Instructions with Job Advertisement and Minimum Warning

Appendix F
Instructions with Job Advertisement and Maximum Warning

Appendix G
Biodata Questionnaire for Pilot Study

Appendix H
Debriefing Forms for Pilot Study

Appendix I
Within Subjects Items by Test Condition

Appendix J
Main Study - Informed Consent Form

Appendix K
Amended Fictional Job Advertisement

Appendix L
Amended Coving Letter

Appendix M
Abbreviated Questionnaire with Minimum Warning

Appendix N
Extended Questionnaire with Maximum Warning

Appendix O
Extended Items

Appendix P
Amended Debriefing Form