

California State University, San Bernardino

CSUSB ScholarWorks

Theses Digitization Project

John M. Pfau Library

2004

Differences by occupation in perceived face validity and procedural justice of a common format application of hard biodata and a written job knowledge test

Anna Maria Forsberg

Follow this and additional works at: <https://scholarworks.lib.csusb.edu/etd-project>



Part of the [Industrial and Organizational Psychology Commons](#)

Recommended Citation

Forsberg, Anna Maria, "Differences by occupation in perceived face validity and procedural justice of a common format application of hard biodata and a written job knowledge test" (2004). *Theses Digitization Project*. 2709.

<https://scholarworks.lib.csusb.edu/etd-project/2709>

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.

59

DIFFERENCES BY OCCUPATION IN PERCEIVED FACE
VALIDITY AND PROCEDURAL JUSTICE OF A
COMMON FORMAT APPLICATION OF HARD
BIODATA AND A WRITTEN JOB
KNOWLEDGE TEST

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Psychology:
Industrial/Organizational

by
Anna Maria Forsberg

June 2004


DIFFERENCES BY OCCUPATION IN PERCEIVED FACE
VALIDITY AND PROCEDURAL FAIRNESS OF A
COMMON FORMAT APPLICATION OF HARD
BIODATA AND A WRITTEN JOB
KNOWLEDGE TEST

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino


by
Anna Maria Forsberg


June 2004

Approved by:


Kenneth S. Shultz, Ph.D., Chair,
Psychology

4/8/04
Date


Donald A. Drost, Ph.D.


Mark D. Agars, Ph.D.

ABSTRACT

This study was designed to evaluate how job applicants from two different professional fields (engineering aides, N = 52; plumbers N = 72) perceive a new type of a common format application of biodata (EVE Background Questionnaire). The job applicants were asked to provide their perceptions of procedural justice and face validity of the EVE Background Questionnaire and a written job knowledge test. The study found significant mean differences in candidates' perceptions of procedural justice and face validity of the two selection tools (EVE Background Questionnaire, written exam). This difference was also found to depend on the type of job. That is, plumbers perceived the written test as more face valid and procedurally just as compared to the EVE Background Questionnaire, whereas the engineering aides' perceptions of procedural justice and face validity were in favor of the EVE Background Questionnaire. The study also looked at the correlation between the applicants' scores on the EVE Background Questionnaire and their scores on the written job knowledge test. The study revealed a significant positive correlation between the biodata score and the written test for plumbers but not for engineering aides, thus indicating that some of the content measured in the

written test for plumbers overlaps with the EVE Background Questionnaire. Implications of the results for human resources selection are also discussed.

ACKNOWLEDGMENTS

Grateful acknowledgement is made to my thesis chair Dr. Kenneth Shultz, who has guided and assisted me through the thesis process by offering his valuable time, scholarly resources, and intellectual expertise. Sincere appreciation is also extended to my committee members Dr. Mark Agars and Dr. Donald Drost for erudite feedback and assistance in the development of my thesis.

Lastly, a warm appreciation is extended to my wonderful son Benjamin. I could not have completed this project without him as my immeasurable source of inspiration and motivation. This thesis is dedicated to him.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGMENTS	v
CHAPTER ONE: INTRODUCTION	1
What is Biodata?	4
Item Taxonomies	6
Keying Items	10
Fakability of Biodata	13
Are Biodata Generalizable Across Jobs and Organizations?	16
Face Validity	20
How do Candidates Perceive Biodata?	22
How do Employers Perceive Biodata?	26
A New Approach to Biodata: Education, Vocational Training, and Experience (EVE) Background Questionnaire	28
Why EVE?	29
Purpose of the Present Study and Hypotheses	31
CHAPTER TWO: METHOD	
Participants	35
Measures	35
EVE Background Questionnaire	35
Procedural Justice/Face Validity Questionnaire	39
Written Exam	39
Procedures	41
Analyses	42

CHAPTER THREE: RESULTS

Data Screening	44
Test of Hypotheses	45
Test of Research Question	47
Post-hoc Analyses	48
CHAPTER FOUR: DISCUSSION	50
Limitations	57
Implications	58
Organizational Resource Preservation	58
Legal Protection	59
Future Research Needs	62
Conclusion	64
APPENDIX A: ITEM TAXONOMIES	66
APPENDIX B: MEASURES	70
APPENDIX C: RESULTS: PERCEPTION OF FACE VALIDITY	78
APPENDIX D: RESULTS: PERCEPTION OF PROCEDURAL JUSTICE	80
APPENDIX E: RESULTS: SCATTERPLOT WRITTEN EXAM AND EVE	82
APPENDIX F: RESULTS: POST-HOC ANALYSIS OF PROCEDURAL JUSTICE ITEMS	84
REFERENCES	86

CHAPTER ONE

INTRODUCTION

Personnel selection is a multifaceted process that involves many choices and decisions prior to actually selecting and hiring new employees. Perhaps the most important consideration for the hiring authority is to choose the most appropriate method of selection for each job classification. Since this decision will affect all candidates and ultimately pave the path to the actual hiring decision, it is crucial that the hiring authority choose an appropriate approach to the selection of job applicants. Commonly, the method of selection is a function of several factors involved in the selection process, such as legal considerations, size of applicant pool, number of openings, minimum requirements, and necessary knowledge, skills, abilities (KSAs) relevant for a particular job. Therefore, choosing the method of selection is not always a clear-cut decision; rather, it is a complex process that involves many important considerations.

Numerous strategies and tools have been developed to assist employers in their pursuit of hiring new employees (e.g., training & experience evaluations [T&E], reference

checks, oral interviews). Typically, the common goal of human resources specialists in using any given selection tool is to strive for a non-discriminatory approach to effectively select the most qualified applicants in the most efficient manner. In this process, employers must evaluate the candidate's background to identify and assess relevant KSAs pertaining to a particular job. This practice commonly includes information regarding academic credentials, vocational training, and relevant on-the-job experience. This type of background assessment is based upon the assumption that previous behavior is a useful indicator of future performance since human behaviors tend to be relatively stable over time (Gatewood & Feild, 2001). This very notion is also consistent with the underlying philosophy of the use of biographical data (a.k.a., biodata) in personnel selection.

A wealth of empirical literature suggests that there is compelling support for the use of biodata in human resource selection. For example, many studies indicate that biodata is capable of predicting criteria such as successful performance on the job, turnover, objective performance measures, and training success, with criterion-related validity coefficients ranging between .20 - .60 (e.g., Asher, 1972; Bliesner, 1996; Mumford &

Whetzel, 1997; Salgado, Viswesvaran, & Ones, 2001; Schmidt & Hunter, 1998; Vinchur, Schippmann, Switzer III, & Roth, 1998).

There are different plausible explanations as to why biodata is an effective predictor in human resource selection. Asher (1972) provided three specific theories explaining why biodata "works": the nonfiction theory, the relevant item theory, and the point-to-point theory. According to the nonfiction theory, biodata has the ability to provide "a systematic, comprehensive collection of factual information about the individual" (p. 259), whereas traditional means (e.g., an unstructured interview) merely produces a "caricature" of the same individual. Furthermore, the relevant item theory suggests that any test battery is composed of "relevant" and "irrelevant" items, whereas biodata commonly adhere to only include items that are shown to be "relevant" to the predicted criterion. Lastly, the point-to-point theory states that "accurate prediction is a function of a point-to-point correspondence between predictor space and the criterion space" (p. 260) and that an increased commonality between the two increases the validity coefficient. The latter theory is closely related to the relevant item theory, as it appears plausible that

relevant items will fall straight into the category of predictor space.

What is Biodata?

The concept of biodata is defined as an assessment and evaluation of "demographic, experiential, or attitudinal variables presumed or demonstrated to be related to personality structure, personal adjustment, or success in social, educational, or occupational pursuits" (Owens, 1976, as cited in Gatewood & Feild, 2001; p. 484). Although this definition appears to convey a detailed and clear picture of what biodata is, reality portrays the field in a light of sharp contrasts, conflicting viewpoints, and a lack of consensus. This poses a concern when a biodata instrument is being constructed as well as a source of confusion for many researchers and practitioners alike. Whereas the researcher may be concerned with what constructs to assess, what items to include, and how to assign weights to items, the practitioner may wonder whether biodata is truly effective, how the candidates will react, and whether biodata will hold up in court. These questions and concerns are warranted since there is no common paradigm

in the field and many conflicting understandings of what constitute biodata still exist.

Additionally, the presence of ambiguity creates problems for researchers and practitioners when they create, discuss, and research biodata, since they are rarely referring to the exact same thing. As a result, the field may experience unnecessary negative connotations and detrimental consequences in research and practice. Hence a consensus of what biodata is and entails would likely aid biodata research, possibly expand its usage in the applied setting, and by doing so also better its reputation.

Mael (1991) provides an extensive discussion of the different attributes and methods that collectively make up biodata and how this complex field should seek to form a current unified paradigm. As of today, there are a myriad of ideas, classifications, inventories, items (Table 1 provides an example), and combinations of items that all fall into one category - biographical items. As one reviews the literature, one will be exposed to biodata items that are closely related to personality assessments, items that merely epitomize behaviors from the past, certain information that can be verified, some items that can be easily faked, items that assess internal values, and some items that are highly subjective etc. Therefore,

I will next discuss some of the most common ways of classifying different types of biodata.

Item Taxonomies

Among the most frequently cited taxonomies in the biodata literature is Asher's (1972) content related classification. Asher's taxonomy explains and categorizes different types of biodata items on eight continuous dimensions. Asher's taxonomy suggests that a biographical item can fluctuate on any of the following dimensions: verifiable-unverifiable, historical-futuristic, actual behavior-hypothetical behavior, memory-conjecture, factual-interpretative, specific-general, response-response tendency, and external event-internal event.

Asher has also identified that a given biodata item may be both objective and verifiable in nature, referred to as "hard biodata." Examples of hard biodata items include "What is your highest level of education?" and "Do you possess a CPA certificate?" By following Asher's logic, there is another item cluster referred to as "soft biodata." Commonly, soft biodata include measures pertaining to a candidate's internal values, attitudes, and interpretations of events. Examples of soft biodata

items include "How much did you enjoy your previous job?" and "Which best describes your ability to cooperate with others?"

As seen in Table 1, a more recent item classification has been proposed by Mael (1991). Mael's categorization of biodata is an extension of Asher's taxonomy with two additional item clusters added. Specifically, Mael has categorized the different items into three separate groups, based on certain commonalities. The first cluster includes the historical items, which he defined as the fundamental domain of biodata (e.g., "Did you work while attending college?"). The second cluster includes items that seek precise information of behaviors (i.e., items that are external, objective, first-hand, discrete, and verifiable) and Mael suggests that the accuracy of the second cluster items increase in the listed order. One example from this cluster is: "How many upper-level science courses have you successfully completed?" The last group of items includes those that reflect internal values, which highly resembles the description of Asher's soft biodata. For example: "Was it important for you to be actively involved in extracurricular activities while attending college?"

Since soft biodata items commonly assess internal, subjective areas, such as values and attitudes, one may ask what the difference is between soft biodata and personality. Shultz (1996) analyzed this particular question through a confirmatory factor analysis. His analyses tested four different models that included various combinations of how personality, soft biodata, and hard biodata are related. The results of his study indicate that the most parsimonious model collapsed personality and soft biodata into one factor and the hard biodata items into another. Hence, he suggested that personality is not significantly different from soft biodata and may in fact be included in the same category.

However, this does not mean that hard biodata questionnaires completely exclude information about personality and social adjustment. According to Harvey-Cook and Taffler (2000), soft biodata constructs are often embedded in the hard biodata items. For example, level of education and number of promotions provide employers with implicit information beyond the mere numerical scores, such as a candidate's personality traits and professional attitude. That is, a candidate with an advanced academic degree and several promotions is likely to possess qualities such as being dependable, ambitious,

motivated, and conscientious. However, a candidate that shows evidence of incomplete academic work and a lack of promotions may be more likely to display behavioral traits such as being unreliable, careless, and sloppy in his work.

Since there are apparent differences between types of biodata items, there may also be different indications of overall usefulness. Asher (1972) provided an overview of different studies that evaluated how biographical items fare in comparison with other common assessments (e.g., intelligence, mechanical aptitude, finger dexterity, personality, and spatial relations). Asher concluded that biographical items consistently outperformed the other assessments, in terms of predictive validity evidence, when job proficiency was the criterion. By looking at all items included in his study, he reported that biodata "excelled the intelligence test [items] by 2 to 1" (p. 255) when the cutoff was set at .50. However, he added that this conclusion might be limited to hard biographical items only.

In contrast, Schmidt and Hunter (1998) came to the opposite conclusion. In a meta-analytic assessment of 85 years of personnel research, they reported an average predictive validity of .51 for general mental ability

(GMA) and .35 for biodata when performance on the job was the criterion. The difference was slightly larger when job-related learning was the criterion, in which the average predictive validity of GMA was .56 compared to .30 for biodata. However, Schmidt and Hunter (1998) made no distinction between soft and hard biodata, which possibly explains the discrepancy of the studies.

To sum up, there appear to be two distinct types of biodata - hard and soft. Hard biodata seek to obtain objective, concrete, and verifiable information while soft biodata assess information that tends to be more abstract, subjective, and generally less verifiable. There are also indications that personality may be included in soft biodata items and that hard biodata may be more effective in predicting job performance than soft biodata.

Keying Items

Another way to differentiate between certain biodata devices can be identified in how they are keyed. There are two common approaches to keying biodata items: empirical and rational. Both types of keying have advantages and disadvantages. Empirical keying assigns a weight to items in relation to their ability to predict a certain criterion (e.g., job performance). This means that the

higher predictive value associated with an item, the higher the item weight. The exact weight assigned to each item is derived from a lengthy criterion study that requires large samples and holdout groups to identify, quantify, and cross-validate biodata items predicting success on the job. The major advantage of empirical keying is that each item is directly related to the predicted criterion. Another benefit associated with empirical keying is that it tends to limit socially desired responses because the "ideal" response is not always obvious (Shermis, Falkenberg, Appel, & Cole, 1996).

Perhaps the most obvious disadvantage of using empirically keyed inventories is that they are both time consuming and costly to develop. Another pitfall is the possibility that although an item has a statistical relationship with a particular criterion, it may not appear job related to the candidate, which will result in decreased face validity and the possibility of unfair perceptions of the test. Lastly, empirically keyed inventories have also been criticized for validity shrinkage when applied to a different sample, since included items are highly correlated with the construct within a given sample, but may not be when applied to a new sample (Mael & Hirsch, 1993).

Rational keying, on the other hand, eliminates the need for a time consuming and costly criterion study and assigns weights in accordance to the theoretical relationship between an item and a particular construct (Mael & Hirsch, 1993). By deriving weights based on their conceptual relevance, this also means that the item and item responses must display a rational and clear relationship (i.e., being construct-related), which is likely to generate a greater sense of face validity and job-relatedness to the candidates. Another benefit is that rational keying is associated with less shrinkage. That is, assigned weights may not have a direct empirical correspondence to the construct it seeks to assess. This may of course also be perceived as a disadvantage as it reduces the initial predictive validity of the instrument. However, several studies have shown the validities to be comparable to cross-validated validities from empirical studies.

To conclude, there are obvious benefits and pitfalls accompanying the two primary methods of keying biographical data (rational and empirical). Although the empirical approach can provide evidence of a direct predictive relationship between the item weights and the criterion, it comes with the price of a time and resource

consuming criterion study. Furthermore, empirical relationships are not always conceptually relevant to the job, which may result in negative applicant reactions. In contrast, the rational approach is both time and cost efficient, is also associated with less validity shrinkage, and capitalizes on the conceptual relevance to the criteria, which is likely to result in positive applicant reactions. However, the criterion-related validity coefficients tend to be smaller.

Fakability of Biodata

There are aspects of biodata that have been extensively criticized in the literature. One commonly criticized aspect of biodata is that it may be faked. The main reason why biographical item responses may fall prey to increased fakability is that some biographical items are subjective and unverifiable (e.g., attitudes, internal evaluations, temperaments, i.e., "soft biodata"). Another "invitation" of inaccurate responses stems from the fact that the questions are being asked within a context that may encourage the candidate to answer in a socially desirable way (Shermis et al., 1996).

For example, within the context of the selection process, it is easy for the candidate to understand that

the maximum score on the biodata inventory is based upon the number of "ideal" responses. Thus, job candidates may be tempted to provide answers that better accommodate the demands of the situation rather than a response that is representative of their true job qualifications.

Furthermore, what is considered an "ideal response" can often be obvious in the context of a particular job and the candidate can easily infer that the closer his obtained score is to the "ideal score," the chances of being hired increase substantially. This means that candidates will be reinforced to alter, or even fake, responses (assuming that the candidate wants the job) by responding in a socially desirable way (Shermis et al., 1996).

For example, a candidate who is applying for a position as a preschool teacher may intentionally stress his superb listening skills and ability to work patiently with children. However, the same candidate may convey his aggressive and persuasive communication skills when applying for a sales position, simply because these skills are considered "ideal" within the context of a given job. Although there may be a blurred line between faking and answering in a socially desirable way, biodata has been

criticized for inviting both of these errors into the selection process.

However, all biodata items are not equally fakeable. The literature provides many ideas and suggestions of how to author and structure items to reduce, or limit, the ability of applicants to fake responses. Asher (1972) states that hard biodata items are the most valid items. He suggests that items that are historical and verifiable by nature will provide the employer with the most factual and accurate information. He also states that the desire to maximize accuracy of responses should be accompanied by having the candidate consent to allowing the employer to verify responses. By doing so, the employer will reduce the candidate's willingness to falsify responses and consequently expect higher validity of the result produced. An analogous suggestion is made by Mael (1991), who recommends that researchers and employers should limit their biodata measures to only include items that are external, objective, first-hand, and verifiable.

Becker and Colquitt (1992), who share a similar perspective, state that items that are less prone to being faked are items that are external, current, objective, verifiable, and discrete. They too recommend the inclusion of warning statements against faking on the actual biodata

form. Furthermore, Shermis et al. (1996) suggest that items are least likely to be faked if they are verifiable, empirically keyed (since the ideal answer is not as obvious compared to rationally keyed items), and they also recommend the use of warning statement against faking. Lastly, Harvey-Cook and Taffler (2000) suggest that the biodata inventory should explicitly state that responses are subject to verification through external sources, which is their supporting argument for using verifiable items in biographical surveys.

To sum up the above mentioned recommendations, employers can take precaution and limit fakability of responses by using external, discrete, objective, and verifiable items (i.e., hard biodata). Informing the candidates that their responses will be verified against external sources can also be used to reduce potential faking.

Are Biodata Generalizable Across Jobs and Organizations?

Another source of criticism of biodata is the difficulty of utilizing one specific biodata exam across organizations and job classifications (Hunter & Hunter, 1984; Wilkinson, 1997). The difficulties stem from the fact that biodata exams are commonly specific to a

particular target of interest (i.e., job or criterion), which has lead to the creation of specific biodata inventories that do not generalize into other areas. Furthermore, biodata tools are usually time consuming and expensive to develop, key, and validate, which poses a formidable obstacle when viewed in combination with the limitations of generalizability. The typical concern from the employer's perspective is that its expenses may outweigh the benefits for jobs that are not filled on a continuous basis.

The reported limitation of generalizability in biodata has also lead to difficulties in interpreting the biodata research. Bobko, Roth, and Potosky (1999) correctly pointed out that "different biodata scales measure radically different constructs" (p. 584) and that these differences lead to obvious difficulties in considering biodata as a consistent unidimensional predictor. For this reason, Bobko et al. recommend that future meta-analytic studies report the method and specific criterion associated with each unique biodata scale.

However, there are also indications that biodata can be generalized across job classifications and organizations (e.g., Rothstein et al., 1990; Wilkinson,

1997). The key issue in developing a generalizable biodata exam is to choose a common format, general assessment criterion, or method, that is relevant across different job classifications. For example, Wilkinson (1997) successfully developed and tested a new approach that focused on a specific attribute of a candidate, namely vocational interest. Vocational interest was measured by using Holland's Self Directed Search (SDS), which classifies people into one of Holland's proposed "vocational types": realistic, investigative, artistic, social, enterprising, or conventional. Furthermore, Wilkinson's (1997) biodata inventory used a rational approach in developing the actual instrument and used only items that met certain criteria (i.e., were classified as a unidimensional person-specific attribute, were objective, and not job-specific). Wilkinson concluded that it is possible to develop a biodata exam that is functional across different positions.

Rothstein et al. (1990) provided another successful example of how biodata can generalize across different jobs. Similar to Wilkinson (1997), Rothstein et al. used only items that were general by nature, but also put emphasis on the method of selecting the items. The method of choice in their study was an empirically keyed

inventory that included components from a Supervisory Profile Record (SPR), which is hypothesized to capture general characteristics of individual supervisory potential. All items were first tested on large, diverse samples and the final items were selected only if they were relevant in many different jobs. Rothstein et al. (1990) concluded that the biodata questionnaire was able to capture a general, measurable attribute that can generalize across first-line supervisors from different job categories. The researchers also claimed that the results were stable. Since the consistent method may have played a significant role in their ability to generalize, Rothstein et al. (1990) called for future research on the relationship between item selection and ability to generalize across job classifications.

Harvey-Cook and Taffler (2000) described another successful approach in which they looked at common format applications for selecting entry-level professionals. The researchers used a combination of hard and soft biodata items that assessed educational credentials (e.g., number of completed courses in a specific field), previous job-related experience (e.g., type of job while in school), home life items (e.g., distance between work place and residence), personal items (e.g., demographical

data) and social involvement (e.g., amount of volunteer activity and extra responsibility during college). The authors recommended the common format method to be utilized mainly as a prescreening device for entry-level recruitment.

To sum up, biodata can be made generalizable across job classifications and the key to success in doing so is found in utilizing a common format, or method, that assesses qualifications relevant for many different positions (e.g., educational credentials, previous job-related experience, supervisory potential, vocational interest, social involvement).

Face Validity

A very important aspect of any given selection device is whether its content appears job related and makes intuitive sense to the candidate. Issues surrounding face validity are important from several perspectives, such as applicant's perceptions and the legal perspective. From the applicant's perspective, items should ideally appear valid as it directly relates to their perception of test fairness and job relatedness. Rynes and Connerley (1992) clearly emphasized the importance of face validity when reporting that 390 surveyed applicants indicated that the

tests associated with positive applicant reactions, were those that were face valid.

Smither, Reilly, Millsap, Pearlman, and Stoffey (1993) looked at candidates' perception of biodata face validity and concluded that biodata was generally not perceived as face valid (biodata ranked 12 compared to 14 other common selection tools). This could perhaps be due to the inclusion of soft biodata items that may appear unrelated to the job. However, the fact that biodata is associated with low face validity demands attention since a test that appears unrelated to the job, and thus unfair, can lead to serious and costly legal consequences. Goldman (2001) points out the criticality of the matter and potential risks by reporting a significant negative relationship between level of perceived distributive and procedural justice of a selection tool/decision and a candidate's decision to take legal actions.

Another legal consideration revolves around the problem of effectively communicating statistical jargon in court. Gililand (1993) points out that the allowance of jury trials in discrimination suits emphasizes the significance of face validity in the development and justification to utilize certain selection tools; more specifically, "issues of face validity and perception of

fairness may become a more salient issue with statistically naïve jurors" (p. 695). In other words, face validity may in fact be the best channel of communication to convey job-relatedness to statistically unsophisticated jurors as opposed to impressive magnitudes of empirical validity (i.e., high correlation coefficients).

Furthermore, Smither et al. (1993) state that selection tools with low face validity have been ridiculed in some court cases (e.g., *Vulcan Society v. Civil Service Commission*, 1973) regardless of their empirical validity. One can also hypothesize that although predictive validity is the employers' best defense in a selection related legal case, face validity may very well be the key component to prevent the employer from going to court in the first place. Thus, face validity should be given thorough consideration to limit the negative effects that may be invited when face validity is absent or ignored.

How do Candidates Perceive Biodata?

Many researchers have indicated that a selection tool should have positive qualities beyond those measured through test reliability and validity and that applicant reactions may be an equally important aspect to consider in the development and evaluation of a selection device

(Anderson, Born, & Cunningham-Snell, 2001; Gililand, 1993; Smither et al., 1993). Smither et al. (1993) list three specific reasons why a selection tool cannot claim to be effective without the approval of the candidates: organization attractiveness, legal concerns, and validity measures. Specifically, Smither et al. (1993) explain that negative reactions might damage the reputation of the organization, which may lead potential employees to search for jobs elsewhere and spreading the word to other potential candidates within a professional field. Furthermore, a selection device that is accompanied by negative attitudes is also likely to lead to complaints and possibly costly and time-consuming appeals, which in turn may be difficult to defend. Lastly, unsatisfied candidates may intentionally perform below their capability during the actual assessment, and by doing so, alter the validity and utility of the device.

Gililand (1993) provides similar arguments but adds the ethical consideration of applicant reactions. He states that rejected candidates may experience problems with efficacy, esteem, and overall psychological well-being as a result of taking a test that is perceived in a negative light. Thus, the contribution of the social components (e.g., candidate perception and reactions)

should not be underestimated; rather, these forces need appropriate consideration in the creation and evaluation process of any selection instrument.

There are a number of elements that contribute to candidates' perception of the selection process. Gililand (1993) presents ten distributive and procedural rules that should be considered in the selection process (see Table 2). His rules are clustered into three components (formal characteristics of the selection tool, explanation of process/tool, and interpersonal treatment) that coalesce into candidates' overall judgment of fairness of a given selection tool/process. Perhaps the most important component of the selection tool is perceived job relatedness. That is, the content of a selection tool should be job-related and the overall test should measure a candidate's capability of performing the job. Gililand (1993) cited several independent studies that indicated that perception of fairness was higher when the selection tool was job-related. These studies also indicated that concrete items were perceived as being more job-related compared to abstract items.

Although biodata has been found to be an effective selection device, there is unfortunately little information about how biodata is perceived by candidates

(Elkins & Phillips, 2000). To complicate the picture further, existing research has produced inconclusive results. Kluger and Rothstein (1993) concluded that biodata outperformed General Mental Ability (GMA) assessments in being perceived as a fair selection device by applicants. The supporting arguments favoring biodata over GMA was its ability to incorporate other relevant characteristics, besides intellectual ability, and thus provide a holistic evaluation of the candidate. Kluger and Rothstein also reported a significant correlation between perceived job relatedness and perception of fairness.

In contrast, Smither et al. (1993) found that entry-level managers perceived biodata and other methods involving abstract items (e.g., personality measures) as having low predictive validity compared to other common concrete measures (e.g., math problems, structured interviews, in-basket activities). In fact, biodata was perceived as having the lowest predictive validity out of the fourteen measures with less than 45% of participants indicating that biodata would be a valid, job-related method. Thus, to enhance candidates' perceptions of usefulness of biodata, employers could benefit from acknowledging the importance of the procedural and distributive rules related to applicant reactions and

apply the recommendations from the justice literature (e.g., Gililand, 1993) to the personnel selection process.

How do Employers Perceive Biodata?

Although the effectiveness of biodata as a selection tool has received support in numerous empirical studies, its use in the applied setting is far from common. Two independent studies cited in Hammer and Kleiman (1988) indicate a fairly small number of organizations use biodata as a selection tool. Both studies were done by the Bureau of National Affairs and were conducted in 1976 and 1987. The first study disclosed that only 4% of the organizations included in the sample claim to use biodata for selection purposes and the second survey revealed identical results. More recent trends indicate that its applied usage still remains relatively low.

For example, Salgado, Viswesvaran, and Ones (2001) cited three different surveys (conducted in 1992, 1994, and 1999) where the percentage of organizations claiming to use biodata ranged from 0.4% to 11% within the US and slightly higher numbers in Australia (average 19.1%) and Europe (average 13%). Hence, it seems as if the use of biodata has increased marginally over the last two decades

but it is also apparent that biodata is still underutilized, given its strong empirical support.

From an applied perspective, there are many explanations for the modest use of biodata in organizations. According to Hammer and Kleiman (1988), there are three distinct clusters of causes: (1) lack of knowledge, (2) lack of feasibility, and (3) negative attitudes toward biodata. The three cluster headings had several different subgroups of potential reasons and the study revealed five major explanations. The most recurrent reason why organizations do not use biodata in selection is a lack of resources, such as expertise, funding, and time (62.6%). The remaining explanations were as follows: do not know much about biodata (52.2%), EEOC risks (45.7%), invasion of privacy (39.5%), and lack of statistical/methodological expertise (39.4%).

To sum up the discussed findings, we can conclude that there is clear evidence of the overall usefulness of biodata in personnel selection even though it has been criticized for being potentially faked, context/job specific, and perceived as having low face validity. There is also some evidence of a two-factor theory of biodata that includes two distinct types of items (hard and soft biodata), in which hard biodata is associated with

comparably higher levels of accuracy. We can also conclude that organizations are generally reluctant to utilize biodata in human resource selection due to a lack of knowledge, feasibility, and an overall negative attitude towards biodata. Hence, there seems to be a great need for the biodata field to develop a biodata device that can counteract existing criticism and simultaneously educate organizations about the effectiveness of biodata in the selection process.

A New Approach to Biodata: Education, Vocational
Training, and Experience (EVE) Background
Questionnaire

Based on the problematic aspects associated with biodata and recommendations from the literature outlined above, a new biodata approach (EVE Background Questionnaire) will be introduced. The EVE Background Questionnaire was developed in an attempt to make biodata more "user-friendly" by relying on a systematic approach with clearly labeled parameters related to assessment criteria, item development, and item keying. The EVE Background Questionnaire is a combined evaluation of a candidate's previous Education, Vocational Training, and Experience (hence the acronym EVE) that is hypothesized to provide a measure of general job competence. By following

the recommendations from the literature, the EVE Background Questionnaire only includes objective and verifiable data (i.e., hard biodata) to gain the benefits associated with its use. By only looking at hard biodata, the following gains are hypothesized to emerge: (1) higher accuracy of responses, (2) increased perception of procedural justice and face validity, and (3) less negative reactions from applicants leading to fewer appeals. Lastly, it is also believed that a generalizable, common format application of biodata is likely to enhance the overall efficiency of the selection process.

Why EVE?

From a personnel selection perspective, there are numerous areas that may be beneficial to assess prior to making a hiring decision (e.g., personality, GPA, references, experience, organizational fit). These predictors can also be placed on a continuum ranging from "not relevant information" to "essential information" and it is important that employers only evaluate information deemed critical for successful performance on the job. The EVE approach suggests a common format application of hard biodata that only incorporates areas of information that are considered highly important to a wide variety of

positions, which enables generalization across classifications and levels.

Another key characteristic of the EVE instrument is that it only consists of items that appear job-related to boost overall perception of face validity. That is why educational credentials, vocational training, and relevant experience are key components since these areas are routinely assessed for many different jobs and thus likely to be perceived as having high face validity. Furthermore, information from these categories is often verifiable (e.g., transcripts, certificates, contact previous employers) which enables the employer to rely on the use of self-assessment and save time and resources by delegating the task directly to the candidates (e.g., through mail, email, on-line, or in conjunction with application or employment test). By allowing the candidate to see and score the job-related assessment criteria, the selection process is more likely to be viewed in a favorable light (i.e., increased perceptions of procedural justice), which is likely to augment the reputation of the organization and result in fewer negative outcomes (e.g., appeals).

Purpose of the Present Study and Hypotheses

The primary purpose of this study was to attempt to clarify if the EVE Background Questionnaire taps into similar constructs as those being measured in a traditional written job knowledge test. To clarify this question, the EVE Background Questionnaire was correlated with the written exam to explore whether or not the two selection tools overlap. We were also interested in assessing candidates' beliefs regarding the new selection tool by assessing candidates' perceptions of procedural justice and face validity of the EVE Background Questionnaire and comparing those perceptions to the candidates' perceptions of procedural justice and face validity of the written job knowledge test. Lastly, we explored whether or not differences exist in perceptions of procedural justice and face validity of the two selection tools depending on type of job.

Job applicants for two types of jobs (Plumber, Engineering Aide) were included in the sample. Candidates from the two positions met specific entrance qualifications in order to compete in the exam. The minimum qualifications for the engineering aide position included a high school diploma, or evidence of equivalent educational proficiency, and a minimum of six months of

experience in drafting that included at least one recognized, major computer-aided design software system. A successful completion of six semester units or their equivalent of college-level courses in architectural or engineering drafting may be substituted for the required experience. The ideal candidate for the engineering aide position has taken college level courses in architectural/engineering drafting, is familiar with algebraic, geometric, and trigonometric procedures, operates computers and Computer Aided Design (CAD) systems, and has a minimum of six months experience on large commercial, government, or school-building structures.

The minimum qualifications for the plumber position included a high school diploma or evidence of equivalent educational proficiency and one year of journey-level experience in the plumbing trade or completion of a plumbing apprenticeship. Additionally, a City Journeyman Plumber License as well as County Registration as a Journeyman Plumber was required. The ideal candidate for the plumber position is knowledgeable in a variety of processes, tools, rules and regulations of the field, and complies with safety rules and regulations pertinent to the plumbing industry.

It was hypothesized that EVE would be positively correlated with the written exam since a better prepared candidate (i.e., higher level of education, more vocational training and experience in the field) is likely to perform better on the job-related content of the written exam, which in turn would be indicative of higher levels of job knowledge and job performance. It was also hypothesized that perception of face validity (whether or not the content of the instrument appears to measure what it intends to measure) would be higher for EVE compared to the written exam. This finding was projected because EVE only asks for highly relevant job preparation and does so in an uncomplicated and less intimidating way compared to the written exam. Additionally, the written exam was a lengthy assessment that included numerous items from several job-related areas, which made the written exam more prone to include some construct-related items that a candidate may perceive as unrelated to their actual ability to do the job, which translates into a lower perception of face validity.

Another advantage of the EVE instrument is that it allows candidates to provide a self-assessment of their background and does so in an uncomplicated and less intimidating way compared to the written exam, which

should translate into increased perceptions of procedural justice. Therefore, it was hypothesized that perceptions of procedural justice of the EVE Background Questionnaire would be higher compared to the written exam. However, due to the differences in the two job classes, we also explored whether differences existed in perceptions of face validity and procedural justice for EVE and the written exam based on job type. These general assumptions were tested through three specific hypotheses and a general research question:

Hypothesis 1: There will be a significant positive correlation between candidates' scores on the EVE background questionnaire and their scores on the written exam for both positions.

Hypothesis 2: The candidates' perception of face validity will be higher for the EVE background questionnaire compared to the written exam.

Hypothesis 3: The candidates' perception of procedural justice will be higher for the EVE background questionnaire compared to the written exam.

Research Question: Will differences occur in perceptions of face validity and procedural justice for EVE and the written exam based on job type (Plumber, Engineering Aide)?

CHAPTER TWO

METHOD

Participants

Data from two employment exams that were administrated by a large public organization were used for this study. The participants consisted of job applicants for two positions (Plumber: N = 72, Engineering Aide: N = 52) that were open to the public and existing employees. Since the participants were competing for regular employment, only job related information was collected to avoid legal problems. Therefore, demographic data were not collected for the study.

Measures

Three measures were used in this study: the EVE Background Questionnaire, the perceptions of procedural justice/face validity questionnaire, and a job-knowledge employment test (written exam) designed to assess core competencies of the plumbing and engineering professions respectively.

EVE Background Questionnaire

The EVE Background Questionnaire is a short biodata questionnaire that generates a score for each applicant based on how the applicant matches up with the assessment

criteria. This score is hypothesized to provide a measure indicative of a candidate's combined quantity and quality of job qualifications. The EVE Background Questionnaire follows a common format that quantifies candidates' educational credentials, vocational training, and previous job-related experience. Although some items are identical as they generalize across a number of different positions (e.g., What type of degree have you earned from an accredited U.S. school or university?), most items are tailored to fit the unique needs and requirements set forth by the position.

For example, items that quantify educational credentials are usually structured to assess 1) what type of degree the candidate holds, 2) what was the degree emphasis, and 3) how many college level units did the candidate complete in areas deemed relevant to the job. This is a useful structure for many positions, but the actual content changes to fit the unique parameters of each particular job. Thus, the two EVE instruments used in this study were very similar in content structure, but dissimilar in content essence (see Appendix B for the two EVE Background Questionnaires used in this study).

The process of constructing and selecting items for the EVE Background Questionnaires were based on several

sources of job relevant information, including archival data (i.e., job analysis material, job bulletin, class description), and statistical information (i.e., based on subject matter experts ratings of criticality and frequency of task performance). Items were only included in the questionnaire if they were deemed objective, verifiable, appeared face valid, and assessed critical components necessary to function effectively in a given position. Thus, the final questionnaires consisted of items that appeared to measure the three EVE components in an objective and verifiable manner, were determined to be critical to the job, and also appeared face valid.

The three EVE elements are conceptualized and operationally defined as follows:

- Education is defined as the *amount* (e.g., 120 units), *level* (e.g., Bachelor of Arts), and *field* (e.g., environmental engineering) of knowledge that a candidate has acquired from an accredited academic institution.
- Vocational training is defined as the *amount* (e.g., number of seminars), *level* (e.g., level of training), and *type* (e.g., certificate) of job-related knowledge that the candidate has acquired from an academic extension or

continuing education program (e.g., workshops and seminars).

- Experience is defined as the *amount* (e.g. number of years), *level* (e.g., supervisor), and *type* (e.g., engineering) of job-related work in which the candidates have applied the necessary knowledge, skills, and abilities required by the position.

The EVE Background Questionnaire uses a rational keying approach, as described by Mael and Hirsch (1993). Each item is in multiple choice format with rational weights assigned to each answer based on criticality and conceptual relevance (e.g., A = 1, B = .75, C = .5, D = .25, E = 0). For items assessing previous experience, the maximum score is consistently set at five years, since job proficiency as a function of job experience reaches its peak at this point and then tends to plateau (Schmidt, Hunter, & Outerbridge, 1986).

The EVE Background Questionnaire can be administered in a number of ways since it can be self-scored (e.g., via email or in conjunction with application and/or employment exam). The EVE Background Questionnaire can be self-scored because it only includes items that are objective and verifiable by nature, which does not completely eliminate

exaggerations, misinterpretations, and/or erroneous responses but its intended purpose is to effectively discourage such attempts. Additionally, each background questionnaire includes a warning statement that indicates that any information provided by the candidate may be verified against external sources and exaggerations and/or false statements may be cause for immediate disqualification.

Procedural Justice/Face Validity Questionnaire

Selected parts of Smither et al.'s (1993) fairness scale were used to assess applicant reactions to the background questionnaire and the written exam. The first part consists of a two-item scale designed to assess procedural justice ($\text{Alpha} = .68$). The second part consists of a five-item scale designed to assess face validity ($\text{Alpha} = .86$). Items on both scales are rated on a 5-point Likert scale on which 5 equals "strongly agree" and 1 equals "strongly disagree" (see Appendix B for a list of specific items).

Written Exam

Candidates competing for both positions were required to take a job-specific employment test designed to assess core competencies of either the plumbing or engineering profession. The content of the written exams reflect job

analysis results and the individual items were authored by subject matter experts in collaboration with human resource specialists.

The written exam for the Plumber position included 100 multiple choice items designed to assess the following areas: interpretation of plumbing blueprints and diagrams (e.g., "What size is the gas line to the A/C unit in room B?" A: 1", B: 2", C: 3", D: 4"), plumbing tools and materials, plumbing practices, and cross-connection knowledge (e.g., "Which of the following is prohibited on combination waste and vent systems? A: Floor drains, B: Water closets, C: Shower drains, D: Floor sinks).

The Engineering Aide test included 75 multiple choice items designed to evaluate the following areas: basic principles of drafting, designing and surveying (e.g., "The vertical distance from the datum plane or surface to the point in question is termed its ____." A: height, B: elevation, C: distance, D: grade), engineering mathematics (e.g., "The tangent multiplied by the cotangent equals ____" A: the sine, B: the cosine, C: unity, D: the same as the cosine divided by the sine), reading comprehension, and interpretation of plans (e.g., "What is the scale of this drawing?" A: 1" = 10'-0", B: 1" = 20'-0", C: 1" = 30'-0", D: 1" = 40'-0").

Both written exams are well-established employment tests that have been used in previous administrations and have been continuously updated and improved. The item analyses for the last test administration revealed a normal distribution of test scores for both tests, appropriate mean difficulty (Plumber exam: 65%, Engineering Aide exam: 55%), and acceptable reliability for both tests (Plumber exam: KR20 = .94, SEM = 4.12; Engineering Aide exam: KR20 = .80, SEM = 3.74).

Procedures

Each candidate completed the regular selection procedure for the position they applied for. The selection strategy was determined by the employer and was based on job analysis material, subject matter experts' opinion, size of applicant pool, number of openings, and future employment needs of the hiring department. A multiple hurdle strategy consisting of a written exam (50%) and an interview (50%) was used for both positions included in this study. Only candidates who were successful on the written exam (i.e., scored above a cut-off score that was set after the test was administered and determined by spread of distribution and number of current and

anticipated future openings) were invited to the interview.

At the day of the written exam, each candidate was asked to fill out the EVE Background Questionnaire, take the written exam, and lastly fill out the perception of procedural justice/face validity questionnaire for the background questionnaire and the written exam respectively. Since the candidates were asked to fill out two separate perception of procedural justice/face validity questionnaires (one for the EVE background questionnaire and one for the written test), the administration was counterbalanced to limit potential carry-over effects. That is, half of the participants were asked to start with the perception of procedural justice/face validity questionnaire assessing the EVE Background Questionnaire, whereas the other half were asked to start with the perception of procedural justice/face validity questionnaire assessing the written exam.

Analyses

The first hypothesis was tested by computing a Pearson correlation coefficient between the candidates' EVE score and the written exam score. Hypotheses 2 and 3

were tested by comparing means using two paired-samples t-test where face validity and procedural justice were the dependent variables and the independent variable was the selection device (EVE, written exam). For the research question, a repeated measures analysis of variance (ANOVA) was performed to test for a potential interaction between the selection device and job type with regard to procedural justice and face validity.

CHAPTER THREE

RESULTS

Data Screening

Prior to analyzing the hypotheses and the research question, SPSS DESCRIPTIVES and FREQUENCIES were used to screen the data for accuracy of data entry, missing values, kurtosis, skewness, and outliers. The following seven variables were included in the analysis: written exam score, EVE total plumber, EVE total engineering aide, perception of face validity EVE, perception of procedural justice EVE, perception of face validity written exam, and perception of procedural justice written exam. Using a criterion of $p < .001$, three distributions of variables were significantly skewed (EVE total plumber $z = 8.06$, perception of procedural justice EVE $z = -4.43$, perception of procedural justice written exam $z = 4.99$) and three kurtotic variables were detected (EVE total plumber $z = 26.52$, perception of procedural justice EVE $z = 5.68$, perception of procedural justice written exam $z = 5.53$). No variables were transformed. By using the same criterion for identifying outliers (z-scores), 2 cases were identified as univariate outliers (two plumbers received a score of zero on the written exam). SPSS SCATTERPLOTS were

analyzed to identify location of the outliers and clarify any other unusual data points (see Appendix D). The two identified outliers were deleted, leaving 124 valid cases to be analyzed (Plumber N = 72, Engineering Aide N = 52).

Prior to analysis, assumptions of normality of sampling distributions, homogeneity of variance, and independence of errors were checked. Normality of sampling distributions was met (degrees of freedom (error) > 20). Homogeneity of variance was also satisfactory as the ratio between the largest and smallest within cell variance was small (< 10:1 ratio) and the sample sizes were relatively equal (< 4:1 ratio). The assumption of independence of error was partly met as each individual was analyzed independently of one another. However, the sample was not completely randomly selected as it consisted of a specific group of self-selected job-candidates from a specific professional field within a restricted geographical area.

Test of Hypotheses

SPSS CORRELATION was used to calculate a Pearson correlation coefficient and Spearman's rho for "EVE score total plumber" and "written exam" and "EVE score total engineering aide" and "written exam." Since the data was

skewed, Spearman's rho was computed to see if the Pearson correlation and rho were comparable.

Hypothesis 1 was partly supported as a significant positive correlation was found between the written exam score and the EVE score for plumbers ($r = .34$, $r^2 = .11$, $\rho = .37$, $p < .05$). Hence, 11% of the variance in the EVE score was associated with the written test. For engineering aides, however, the written exam scores and the EVE scores were not significantly correlated ($r = .01$, $\rho = .06$, $p > .05$).

SPSS PAIRED SAMPLES T-TEST was used to test the second hypothesis. The results of the analysis found significant mean differences in perceptions of face validity as a function of type of selection device (EVE, written exam), $t(123) = -2.86$, $\eta^2 = .06$, $p < .01$. Six percent of the variance in perceptions of face validity was accounted for by selection device (EVE, written exam). The candidates' perceptions of face validity of the EVE were lower ($M = 18.99$) than their perceptions of face validity of the written exam ($M = 19.87$) [see Table 3 for descriptive statistics]. Since candidates perceived the written exam as more face valid compared to the EVE Background Questionnaire, the second hypothesis was not supported.

SPSS PAIRED SAMPLES T-TEST was also used to test the third hypothesis to see whether mean differences were found in perceptions of procedural justice. The results of the analysis discovered significant mean differences in perceptions of procedural justice as a function of type of selection device (EVE, written exam), $t(123) = -2.29$, $\eta^2 = .04$, $p = .05$. Four percent of the variance in perception of procedural justice was accounted for by selection device (EVE, written test). The candidates reported lower perceptions of procedural justice for the EVE Background Questionnaire ($M = 7.69$) compared to their perception of procedural justice for the written exam ($M = 8.00$) [see Table 3]. Hence, the third hypothesis was not supported.

Test of Research Question

SPSS REPEATED MEASURES was used to analyze the research question. The research question explored the possibility of differences occurring in perceptions of face validity and procedural justice for EVE and the written exam depending on job type (Plumber, Engineering Aide). The analysis revealed that the significant mean differences in perception of procedural justice, as a function of selection tool (EVE Background Questionnaire,

written exam) does depend on job type (Engineering Aide, Plumber), $F(1,122) = 33.82, p < .05$. Twenty-two percent of the variance in the differences in perceptions of face validity and procedural justice was accounted for by job type.

A simple effects analysis revealed that the interaction was due to a significant difference in how the plumbers perceived the two selection tools. More specifically, we found a significant mean difference in how the plumbers perceived face validity of EVE and the written exam ($t(71) = -6.156, p < .05$) as well as how plumbers perceived the perception of procedural justice of EVE and the written exam ($t(71) = -5.379, p < .05$) [see Appendix C and D for a graphical representation of the interaction effect and Table 3 for means].

Post-hoc Analyses

Based on the impression that the two-item scale used in this study appeared to measure both "general fairness" and "procedural justice," a post-hoc analysis was conducted. SPSS REPEATED MEASURES was used to test whether significant mean differences exist in perception of procedural justice of the written exam as well as for the EVE Background Questionnaire as a function of the two

different items. The results of the post-hoc analysis revealed that there were significant mean differences in perception of justice of the written exam as a function of the two items in the procedural justice scale ("general fairness," "procedural justice"), $F(1, 122) = 18.41$, $p < .05$). Thirteen percent of the variance in perceptions of procedural justice of the written exam was accounted for by item type ("general fairness", "procedural justice"). As seen in Appendix F, engineering aides' perception of "general fairness" was lower ($M = 3.58$) than their perception of "procedural justice" ($M = 3.77$) of the written exam. For plumbers, on the other hand, this relationship was reversed. Plumbers' perception of "general fairness" of the written exam was slightly higher ($M = 4.25$) than their perception of "procedural justice" ($M = 4.22$) [see Table 3 for descriptive statistics].

For perception of procedural justice of the EVE Background Questionnaire, the results of the post-hoc analysis found no significant mean differences in perception of justice of the EVE Background Questionnaire as a function of the two items in the procedural justice scale ("general fairness", "procedural justice"), $F(1, 122) = .124$, $p > .05$ (see Appendix F).

CHAPTER FOUR

DISCUSSION

The results of the study indicate that the score on the EVE Background Questionnaire is positively correlated with the score on the written exam for plumbers ($r = .34$, $p < .05$) but not for engineering aides. The fact that the written exam scores and the EVE scores are positively correlated is not a big surprise per se. In fact, this relationship was anticipated since the written exam score essentially is a manifestation of the content of the EVE Background Questionnaire. More specifically, an EVE item that asks for the candidates' highest degree assigns a higher score to candidates with higher degrees, meaning that the more knowledge a candidate has acquired from an academic institution, the higher the EVE score will be. In similar fashion, a logical link between a higher degree and amount of job knowledge also exist, which in turn would transpire into a higher score on the written employment test. Thus, EVE and the written exam should conceptually be correlated and this relationship was partly supported in this study.

The significant correlation is also consistent with previous research. As previously discussed, biodata has

been found to be an effective and consistent predictor of job-related criteria such as objective performance measures and training success (e.g., Asher, 1972; Bliesner, 1996; Mumford & Whetzel, 1997; Salgado, Viswesvaran, & Ones, 2001; Schmidth & Hunter, 1998; Vinchur, Schippman, Switzer III, & Roth, 1998). However, the lack of a significant positive correlation between the EVE Background Questionnaire and the written test for engineering aides was both surprising and contradictory to previous biodata studies. This is difficult to interpret since EVE should conceptually be more suitable for engineering aides since they acquire job knowledge through all three elements of the EVE Background Questionnaire. Thus, there should be a strong relationship between the EVE score and the written exam for engineering aides.

This conceptual link does not exist between the EVE-components and the competencies required by plumbers. In fact, it is both possible and common that an exceptional plumber may have minimal education and/or vocational training since the trade commonly is learned on the job. The key component in job performance and knowledge of the plumbing trade is practice, which makes experience the one variable that should account for most variance in job knowledge/performance. Vocational training

(e.g., apprenticeship program, trade school) should also account for some of the variance in job knowledge/performance, but education may or may not contribute at all. By running the EVE components separately, this relationship was indeed supported as vocational training and experience were both significantly correlated with the written score (vocational training $r = .295$, $p < .05$; experience $r = .249$, $p < .05$) but education was not correlated with the written exam score ($r = -.031$, $p > .05$). By running the EVE components separately for engineering aides, none of the individual components were significantly correlated (education $r = .103$, vocational training $r = -.021$, experience $r = .129$, $p > .05$), which may explain why the EVE component combined did not correlate with the written exam.

The results of the study also revealed significant mean differences in perceptions of procedural justice and face validity as well as a significant interaction. By taking a closer look at these results, it is apparent that the differences lie in how the plumbers perceive the two selection tools. More specifically, plumbers had higher perceptions of the traditional written exam over the EVE Background Questionnaire for both procedural justice and

face validity. This finding is quite interesting as the plumbers' perception of EVE is contrary to how well the score actually correlated with their score on the written job knowledge test. Engineering aides, on the other hand, favored EVE over the written exam for both perception of procedural justice and face validity, yet their EVE score did not correlate significantly with their written exam score. Although this result seems to indicate that engineering aides like EVE and dislike the written test, and that the opposite holds true for plumbers, this notion is likely to be spurious and should thus be avoided. The significant mean differences that were found in perceptions of face validity and procedural justice reveal nothing in regards to whether the candidates liked or disliked the two selection tools.

We also found a significant interaction, which sheds light on the appropriateness of using EVE across job classifications and the conditions in which EVE may or may not be an effective selection tool. The significant interaction suggests that EVE may be more accepted and lead to more positive applicant reactions (i.e., perceived as face valid and procedurally just) with jobs that require specific academic credentials, vocational training, and experience (e.g., engineering aide) as

opposed to non-academic, entry-level trades such as plumbing. The query behind the interaction analysis was intentionally stated as an exploratory research question due to the lack of a clear expectation of the outcome. As previously discussed, some studies have shown that it is possible to develop a biodata inventory that generalizes across job classifications (Rothstein et al., 1990; Wilkinson, 1997) and that the key to successful generalization lies in the utilization of a common format or method, such as the method used in the development and scoring of EVE, but that there are also several studies that have found the opposite (Bobko, Roth, and Potosky, 1999; Hunter & Hunter, 1984; Wilkinson, 1997). The results of this study seem to suggest that EVE should not be used across classifications and levels because there are apparent differences in perceptions of face validity and procedural justice depending on job type.

The study suggests that engineering aides favor EVE over a written job knowledge test and that plumbers favor a written job knowledge test over EVE. This piece of information may be better understood by closely examining the nature of the two jobs. For example, engineering aides are both trained academically and vocationally and can easier understand that as quantity and quality of

education, vocational training, and experience accumulates, so does the level of job performance. Furthermore, it is also easier to see how the components measured in EVE are directly related to performance on the job since the actual work of an engineering aide involves most of the areas that are measured in EVE. Therefore, engineering aides should conceptually accept the EVE Background Questionnaire as a procedurally just and face valid measure of their ability to perform the duties and responsibilities associated with the engineering aide position.

This is not true for the plumbers. The plumbing trade is very much a hands-on profession that is commonly learned on the job. Little education and/or vocational training is available for plumbers and the available training is typically offered through an apprenticeship, in other words, a hands-on experience under the supervision of a journey-level plumber on the job. Hence, it is understandable that plumbers perceive a background questionnaire as not only being procedurally unjust but also unrelated to the actual job. The comparably lower level of perception may also stem from the fact that EVE asks specific questions about education, which not only turned out to be negatively correlated with performance on

the written exam, but from a plumber's point of view, may have little or no relevance to actual job performance.

Another part of the study that may have impacted the overall results and conclusions is the scale for procedural justice. The study measured perceptions of procedural justice using an already established two-item scale developed by Smither et al. (1993). If the two-item scale measures the same construct (procedural justice), then we will expect a high alpha-value. However, as seen in Appendix B, the two-item scale reports an alpha of .68, which is a marginal value for a scale, yet common for measures having only two items. By taking a closer look at the individual items in the two-item scale (see Appendix B), it is noticeable that the first question ("Overall, I believe that the Background Questionnaire/Written Exam was fair") is general and may conceptually appear to tap into a construct labeled "general fairness" instead of "procedural justice." However, the second question ("I felt good about the way the Background Questionnaire/Written Exam was conducted and administered") is more specific and appears to tap directly into the construct "procedural justice." Therefore, a post-hoc analysis was conducted to see how, if at all, the two items differed. The post-hoc revealed

significant mean differences in perceptions of procedural justice of the written test but not for the EVE Background Questionnaire, which may imply that the two items may not measure the exact same construct as intended.

Limitations

A major limitation of this study was that EVE is still in a stage of infancy and has not been under thorough evaluation before, which makes the entire study exploratory by nature. For this reason, it may be difficult to interpret the results, since there are no previous studies to compare to. The findings in this study may have been explained differently and/or in greater detail if the instrument had a history of previous results.

Another limitation is that the EVE is a common format instrument, which means that although the overall structure of the EVE Background Questionnaire is similar for engineering aides and plumbers, the exact make-up of the actual items were different. This may pose a threat to the ability of accurately comparing the results. However, although the items are dissimilar, they are close to equivalent when viewed in the context of each profession. For example, all items assessing work experience are

created from the duty statements that have been determined to be highly critical to job performance and/or are frequently performed. That is, an item is only included in the background questionnaire if it is deemed critical to the job through job analysis results and/or the collective opinion of several subject matter experts.

Another source of concern is that EVE is being compared to a written job knowledge test under the assumption that the written test is an effective measure of job knowledge/job performance. Although the well-established written tests for the plumbers and the engineering aides are likely to capture some of the variability in job knowledge/ job performance, the two written tests have not gone through a criterion related validity study. This poses a concern when the written test is used as a point of reference since the true validity of the written test is unknown.

Implications

Organizational Resource Preservation

Although the EVE Background Questionnaire accounted for only 11% of the variance in the written exam, a significant positive correlation may translate into a number of potential benefits. The most obvious benefits

are related to organizational resource preservation, which is a historical and transcultural objective for any given organization as it directly or indirectly leads to organizational success (e.g., profit, effectiveness, efficiency). In the limited scope of this study, a significant positive correlation between a short background questionnaire and a comparably longer written exam, could have a direct influence on the time, cost, and personnel resources involved in the selection process. More specifically, the time it takes to develop and administer the EVE Background Questionnaire is substantially less compared to the time it takes to develop and administer a written exam. This piece of information is obviously directly linked to the actual cost of the selection device, as cost commonly is positively correlated with the time it takes to go through the selection process. In similar fashion, the amount of personnel resources required in the process is linked to time and cost as well. , a significant correlation between the two test parts inevitably favors the test device that is less time consuming and costly to develop (i.e., EVE).

Legal Protection

Another highly sought quality and priceless aspect of a selection device is having the ability to proactively

guard against legal allegations and effectively fend off accusations in court. Under the protection of Title VII, candidates have the right to file suit if the selection process is perceived as unfair (e.g., disparate treatment, adverse impact). Consequently, candidates can easily transform a selection process into a costly journey through the legal system that possibly results in holding the organization liable for pricey compensatory and punitive damages. Thus, to limit expensive and time-consuming appeals, organizations can and should take proactive steps to prevent legal predicaments in every possible way.

The best route to safe guard against legal problems in personnel selection is to make the process and test parts as job related as possible and statistically validate the inferences that are made from each selection tool. Although the statistical validation is the ultimate proof of job relatedness, applicant reactions may be seen as a manifestation of the degree of perceived job relatedness and should thus not be ignored.

In this study, perception of procedural justice and perception of face validity were used to address applicant reactions of job relatedness and it was apparent that there was a difference in how the two tests were

perceived. Whereas the engineering aide candidates perceived EVE as having higher face validity and being more procedurally just compared to the written exam, the plumber candidates, who favored the traditional written exam, did not share this view.

The logic behind the emphasis on face validity and procedural justice is two-fold. First, if a selection device is perceived as face valid and procedurally just, the entire selection process is likely to generate positive applicant reactions, which consequently will lead to fewer appeals. Secondly, attaining positive applicant reactions is a proactive measure against legal problems. Although statistical validity is the best tool to tackle legal problems once an organization is faced with a law suit, selection tools that are perceived as job related (face valid) and procedurally just may be the preventive defense that will keep the employer from going to court in the first place. Additionally, overall positive applicant reactions are likely to lead candidates (regardless of success in the process) to perceive the entire organization in a favorable light, making the candidate more prone to apply again and/or recommend the employer to others, which ties neatly into the key objective of selection - to attract and retain the best candidates.

Future Research Needs

Since the EVE Background Questionnaire has never been explored before, there are many general and specific questions left unanswered that would benefit from future exploration. Some of the more general questions revolve around the use of a common format application that only includes hard biodata items. The literature has recognized that biodata items can be categorized as either hard or soft and that the hard biodata items tend to be more effective. However, the idea of only including hard biodata items in biodata instruments has not been looked at in great detail. As it is today, most biodata instruments consist of a mixture of hard and soft items and it is very unclear to what extent the different types of items, or combinations of different types of items, influence the ability of a biodata device to capture the construct it seeks to assess. Therefore, future research should explore this area further.

In the same manner, future studies should also be directed towards the use of common format applications and common methods to develop and score biodata instruments. It makes intuitive sense that specific parameters are necessary to develop a biodata tool that is both effective and efficient. By relying on specific guidelines, it will

be easier to systematically evaluate how well the instrument works while simultaneously identify areas in need of improvement. Additionally, a standardized process with specific guidelines will also allow researchers to assess how a particular instrument works across different job classifications and better understand how job applicants from different professional fields perceive it.

Some of the more specific questions that would benefit from additional exploration concern the findings of this study. For example, it would be interesting to isolate the exact reasons why the engineering aides favored EVE over the written exam when the plumbers held an apposing view. It would also be interesting to see if perceptions of the two selection tools would change after the hiring decision was made (i.e., distributive justice) and if so, how. In a similar way, it would also be interesting to see if perceptions of face validity and procedural justice would change once the candidates have become used to the use of background questionnaires in personnel selection. As discussed earlier, biodata is underutilized as compared to some of the more traditional selection tools and it is possible that extended exposure may change candidates' perceptions of the background questionnaire. Many candidates may expect to go through a

written test, a performance test, and/or an interview, and may view the background questionnaire as an odd hiring strategy and consequently have a negative opinion as a function of lack of exposure.

Another area in need of further exploration is alternative methods to measure the construct procedural justice. This study assessed perception of procedural justice by using a two-item scale with a marginal alpha-value. Therefore, it may be questionable whether or not both items actually measure procedural justice or whether it measures "general fairness" and "procedural justice" combined. If a similar study would be conducted again, it would perhaps be wise to develop a different scale, or use another existing scale, to measure perception of procedural justice, preferably with a few more items and higher reliability.

Conclusion

This study aimed to clarify how well a common format application of biodata (EVE Background Questionnaire) correlates with a written job knowledge test and how the job applicants perceive the new selection tool from the standpoint of perception of face validity and perception of procedural justice. It can be concluded that EVE did

measure some of the job-related criteria that is assessed through the written test for the plumbers but that it was completely orthogonal to the assessment criteria of the written test. It can also be concluded that job applicants from two different job classifications (Plumbers, Engineering Aides) differed in their perceptions of procedural justice and face validity of the two selection tools (written exam, EVE Background Questionnaire). The results of the study indicate that plumbers report significantly higher levels of perceptions of procedural justice and face validity of the written exam, as compared to the EVE Background Questionnaire, and that engineering aides report slightly higher, but not significantly higher level of perceptions of procedural justice and face validity of the EVE Background Questionnaire, as compared to the written exam.

APPENDIX A
ITEM TAXONOMIES

TABLE 1
Mael's Taxonomy of Biodata

Historical

What was your college major?

External Event

Did your parents help you with your homework?

Objective

Did you pass the California Bar Exam?

First-Hand

How well do you perform at work?

Discrete

Do you know how to replace a flat tire?

Verifiable

Do you have a college degree?

Controllable

How many statistics courses have you completed?

Equal Access

Were you involved in sports?

Job Relevant

How many years of experience do you have in auditing?

Noninvasive

Were you actively involved in a youth organization?

Futuristic

Where do you think you will be working in the future?

Internal Event

How do you feel about working in a team?

Subjective

Would you describe yourself as an assertive person?

Second-Hand

What do you think your supervisor thinks about your level of performance?

Summative

How many tires do you change per week?

Nonverifiable

Did you like your previous job?

Noncontrollable

Did you attend a public elementary school?

Nonequal access

Did you compete in beauty pageants?

Not Job Relevant

Can you play the piano?

Invasive

How many times have you been married?

TABLE 2
Gililand's Model of Applicants' Reaction to Selection Systems

Formal Characteristics

Job Relatedness

Opportunity to Perform

Reconsideration Opportunity

Explanation

Feedback

Selection Information

Honesty

Interpersonal Treatment

Interpersonal Effectiveness

Two-way Communication

Propriety of Questions

TABLE 3
Descriptive Statistics

Measure	Mean	SD
Procedural Justice		
EVE	7.69	1.51
Written Test	8.00	1.54
Engineering Aide (EVE)	7.75	1.41
Engineering Aide (Written Test)	7.35	1.71
Plumber (EVE)	7.65	1.58
Plumber (Written Test)	8.47	1.21
Face Validity		
EVE	18.99	3.42
Written Test	19.87	3.93
Engineering Aide (EVE)	19.15	3.39
Engineering Aide (Written Test)	18.37	3.87
Plumber (EVE)	18.87	3.47
Plumber (Written Test)	20.96	3.62

APPENDIX B

MEASURES

Perception of Fairness Questionnaire
(from Smither et al., 1993)

Procedural Justice (2 items, Alpha = .68)

1. Overall, I believe that the examination* was fair.
2. I felt good about the way the examination* was conducted and administered.

Face Validity (5 items, Alpha = .86)

1. I did not understand what the examination* had to do with the job (R).
2. I could not see any relationship between the examination* and what is required on the job (R).
3. It would be obvious to anyone that the examination* is related to the job.
4. The actual content of the examination* was clearly related to the job.
5. There was no real connection between the examination* that I went through and the job (R).

The items were measured using a five-point Likert scale on which 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree.

* The words "Written Exam" and "Background Questionnaire" replaced "examination" in the perception of fairness questionnaire (see Appendix B).

EVE Background Questionnaire: Engineering Aide

(Item weights in parenthesis and item cluster correlation with written exam in bold.)

EDUCATION

$r = .103, p > .05$

1. What type of degree have you earned from an accredited U.S. school or university? (If you have obtained your degree from a foreign university, it must be translated, evaluated, and deemed equivalent to the standards and curriculum of an accredited U.S. university.)
 - A) Doctorate, Masters of Arts/Science (1.00)
 - B) Bachelor of Science (.75)
 - C) Bachelor of Arts (.50)
 - D) Associate of Arts/Science (.25)
 - E) I have no college degree (0)
2. Have you successfully completed academic courses in any of the following areas: **design, AutoCAD, blueprint reading**?
 - A) I have completed at least one course in each area. (1.00)
 - B) I have completed at least one course in two areas. (.66)
 - C) I have completed at least one course in one area. (.33)
 - D) I have not taken academic courses in any of the above areas. (0)
3. How many college level courses in **engineering or architecture** have you completed?
 - A) More than 6 courses (1.00)
 - B) 5 – 6 courses (.75)
 - C) 3 – 4 courses (.50)
 - D) 1 – 2 courses (.25)
 - E) I have not taken any courses. (0)

Vocational Training

$r = -.021, p > .05$

4. Do you possess a valid **Engineer in Training (EIT) Certificate**?
 - A) Yes (1.00)
 - B) No (0)
5. Do you possess a current license to be a **professional engineer or architect**?
 - A) Yes (1.00)
 - B) No (0)

6. In addition to your formal education, how many documented training hours (e.g., workshops, seminars) related to engineering or architecture have you completed?
- A) At least 30 hours (1.00)
 - B) 20 to 29 hours (.75)
 - C) 10 to 19 hours (.50)
 - D) Less than 10 hours (.25)
 - E) I have not received any training related to engineering or architecture (0)

Experience

$r = .129, p > .05$

7. How many years of on-the-job experience do you have performing **routine drafting or tracing, sketching, lettering, and/or delineating**?
- A) Five or more years (1.00)
 - B) Greater than three years but less than five years (.75)
 - C) Greater than one year but less than three years (.50)
 - D) One year or less (.25)
 - E) No experience (0)
8. How many years of on-the-job experience do you have **taking measures in the field and preparing sketches and field notes showing dimensions and locations of buildings and ground areas**?
- A) Five or more years (1.00)
 - B) Greater than three years but less than five years (.75)
 - C) Greater than one year but less than three years (.50)
 - D) One year or less (.25)
 - E) No experience (0)
9. How many years of on-the-job experience do you have **performing measurements (i.e., calculation involving the use of algebra, geometry, and trigonometry)**?
- A) Five or more years (1.00)
 - B) Greater than three years but less than five years (.75)
 - C) Greater than one year but less than three years (.50)
 - D) One year or less (.25)
 - E) No experience (0)

EVE Background Questionnaire: Plumber

(Item weights in parenthesis and item cluster/individual item correlation with written exam in bold. The asterisk flags a significant result.)

EDUCATION

$r = -.031, p > .05$

1. What type of degree have you earned from an accredited U.S. school or university? (If you have obtained your degree from a foreign university, it must be translated, evaluated, and deemed equivalent to the standards and curriculum of an accredited U.S. university.)
 - A) AA/AS or higher (1.00)
 - B) Some college (.66)
 - C) High School Diploma (.33)
 - D) I have no degree/diploma (0)

$r = -.121$

2. How many job-related courses have you successfully completed at a trade technical college or occupational center in the following areas: **trade theory, welding, blueprint reading, estimating, and/or trade practice?**
 - A) I have completed at least one course in four or more areas. (1.00)
 - B) I have completed at least one course in three areas. (.75)
 - C) I have completed at least one course in two areas. (.50)
 - D) I have completed at least one course in one area. (.25)
 - E) I have not completed courses in any of the above areas. (0)

$r = .032$

Vocational Training

$r = .295, p < .05$

3. Do you possess a valid **LA City Journeyman Plumber License?**
 - A) Yes (1.00)
 - B) No (0)
4. Do you possess a current **LA County Certificate of Registration as a Journeyman Plumber?**
 - A) Yes (1.00)
 - B) No (0)

$r = .421^*$

5. Have you received job-related training through an **apprenticeship program**?
- A) Yes, I have successfully completed an apprenticeship program. (1.00)
 - B) Yes, I have/am receiving job-related training through an apprenticeship but I have not completed the program. (.50)
 - C) No, I have not received job-related training through an apprenticeship program. (0)

$r = .017$

Experience

$r = .249, p < .05$

6. How many years of journey-level experience do you have **installing, replacing, maintaining, and repairing plumbing systems, equipment, appliances, controls, and fixtures**?
- A) Five or more years (1.00)
 - B) Greater than three years but less than five years (.75)
 - C) Greater than one year but less than three years (.50)
 - D) One year or less (.25)
 - E) No experience (0)

$r = .106$

7. How many years of journey-level experience do you have **installing, maintaining, and repairing gas piping systems**?
- A) Five or more years (1.00)
 - B) Greater than three years but less than five years (.75)
 - C) Greater than one year but less than three years (.50)
 - D) One year or less (.25)
 - E) No experience (0)

$r = .154$

9. How many years of journey-level experience do you have **surveying gas safety devices and other plumbing equipment that may need repair or replacement**?
- A) Five or more years (1.00)
 - B) Greater than three years but less than five years (.75)
 - C) Greater than one year but less than three years (.50)
 - D) One year or less (.25)
 - E) No experience (0)

$r = .370^*$

10. How many years of journey-level experience do you have **performing heavy physical labor related to the installation of plumbing equipment and piping** (e.g., digging ditches, climbing ladders, and lifting heavy material)?

- A) Five or more years (1.00)
- B) Greater than three years but less than five years (.75)
- C) Greater than one year but less than three years (.50)
- D) One year or less (.25)
- E) No experience (0)

r = .175



The following questions assess your opinion of the Background Questionnaire. The questions are optional and confidential and your answers will not influence your chances of getting a job with us. This information is collected for research only.

INSTRUCTIONS:

Think specifically about the Background Questionnaire when you answer the following questions. Respond to each question by circling the number that best represents your opinion about the Background Questionnaire.

1 = Strongly Disagree
2 = Disagree
3 = Neutral
4 = Agree
5 = Strongly Agree

1. I did not understand what the Background Questionnaire had to do with the job.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

2. I could not see any relationship between the Background Questionnaire and what is required on the job.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

3. It would be obvious to anyone that the Background Questionnaire is related to the job.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

4. The actual content of the Background Questionnaire was clearly related to the job.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

5. There was no real connection between the Background Questionnaire that I went through and the job.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

6. Overall, I believe that the Background Questionnaire was fair.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

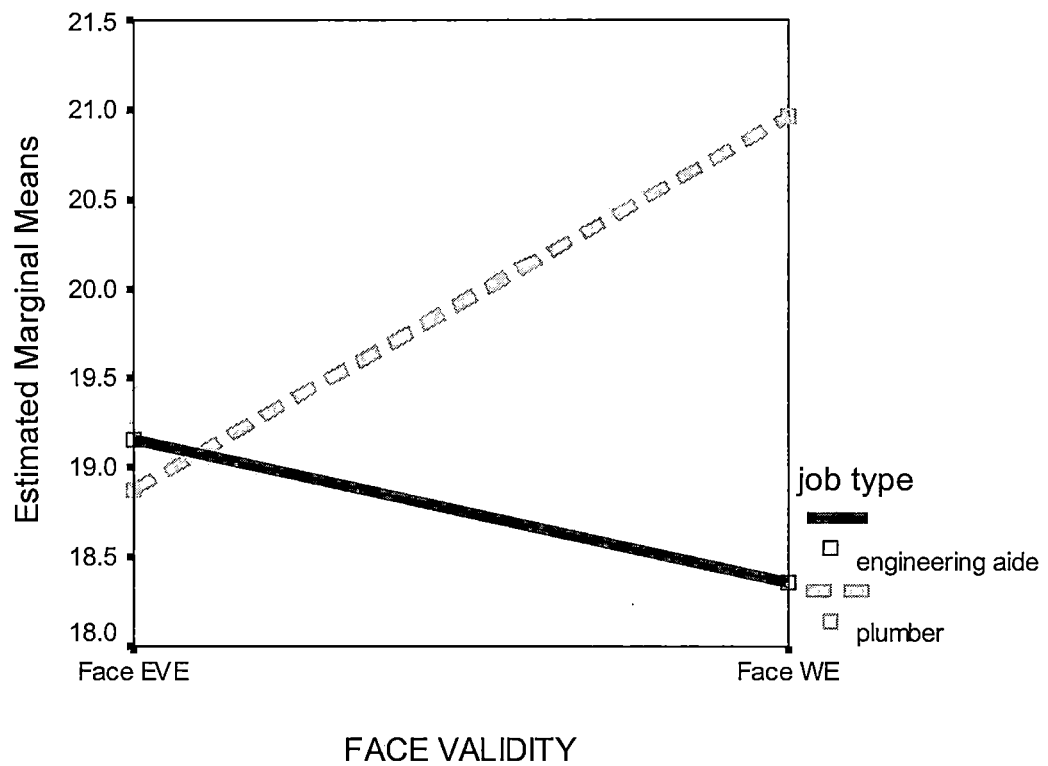
7. I felt good about the way the Background Questionnaire was conducted and administered.

1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
---------------------------	---------------	--------------	------------	------------------------

APPENDIX C

RESULTS: PERCEPTION OF FACE VALIDITY

Estimated Marginal Means of Face Validity



APPENDIX D

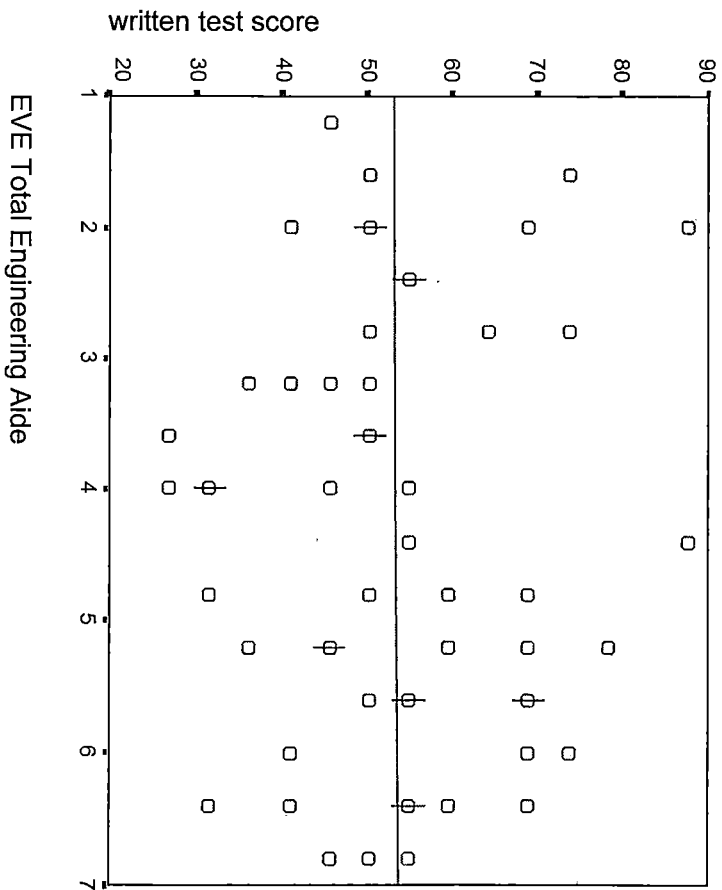
RESULTS: PERCEPTION OF PROCEDURAL JUSTICE

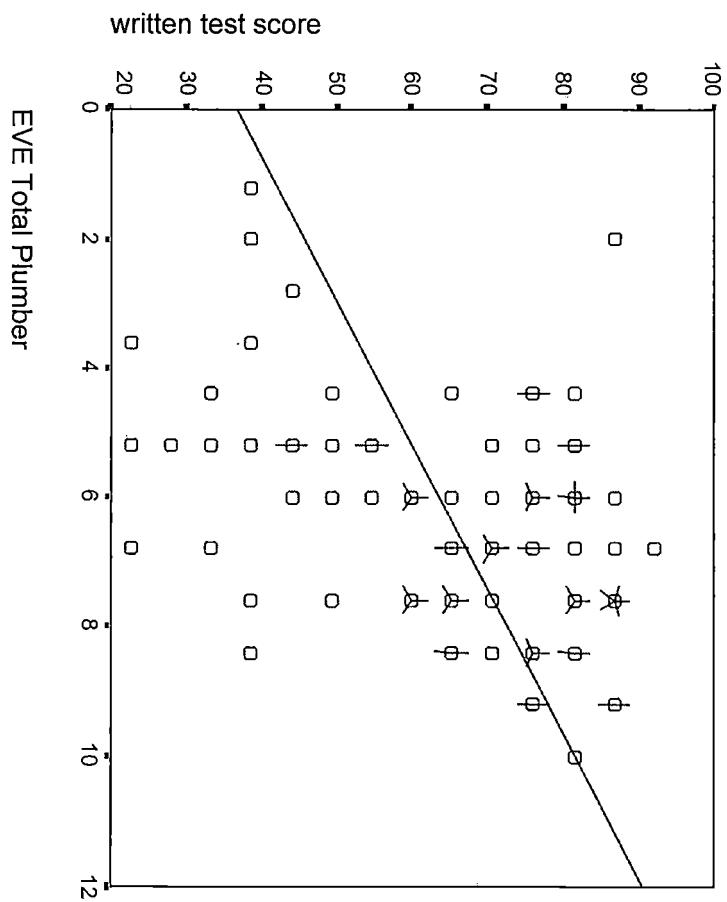
Estimated Marginal Means of Justice



APPENDIX E

RESULTS: SCATTERPLOT WRITTEN EXAM AND EVE



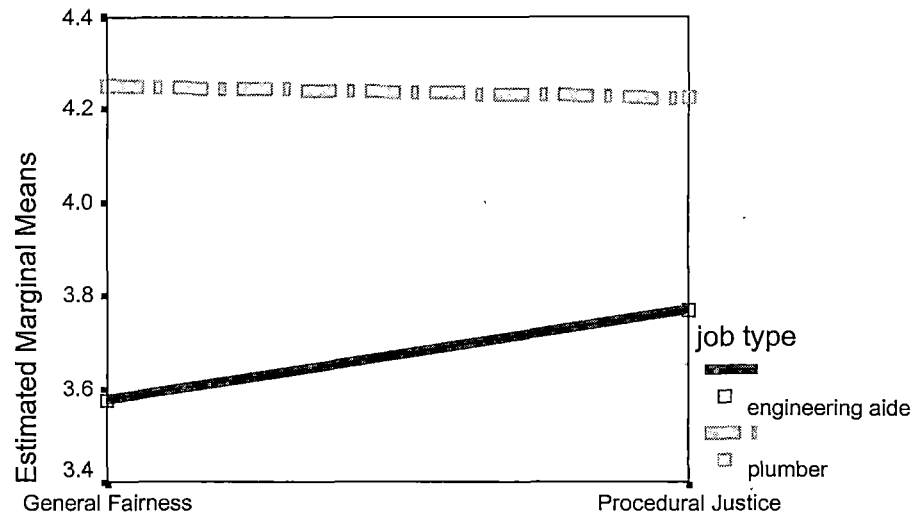


APPENDIX F

RESULTS: POST-HOC ANALYSIS OF PROCEDURAL
JUSTICE ITEMS

Estimated Marginal Means:

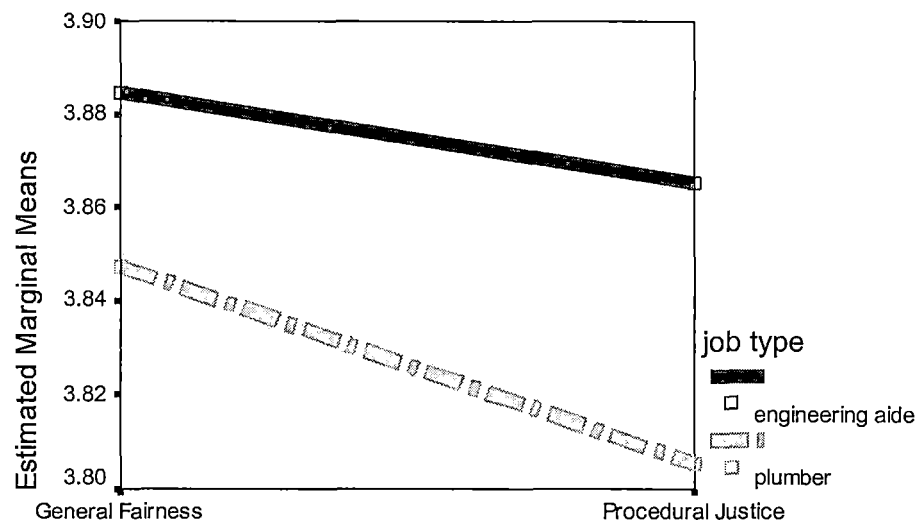
Procedural Justice Items Written Exam



PROCEDURAL JUSTICE ITEMS WE

Estimated Marginal Means

Procedural Justice Items EVE



PROCEDURAL JUSTICE ITEMS EVE

REFERENCES

- Anderson, N., Born, M., & Cunningham, N. (2001).
Recruitment and selection: Applicant perspectives and
outcomes. *Handbook of Industrial, Work and
Organizational Psychology* (Vol 1). Thousand Oaks, CA:
Sage Publications Ltd 200-218.
- Asher, J. J. (1972). The biographical item: Can it be
improved? *Personnel Psychology*, 25, 251-269.
- Becker, T., & Colquitt, A. (1992). Potential versus actual
faking of a biodata form: An analysis along several
dimensional item types. *Personnel Psychology*, 45,
389-406.
- Bliesner, T. (1996). Methodological moderators in
validating biographical data in personnel selection.
*Journal of Occupational and Organizational
Psychology*, 69, 107-120.
- Bobko, P., Roth, P., & Potosky, D. (1999). Derivation and
implications of meta-analytic matrix incorporating
cognitive ability, alternative predictors, and job
performance. *Personnel Psychology*, 52, 561-589.

- Elkins, T. J., & Phillips, J. S. (2000). Job context, selection decision outcome, and the perceived fairness of selection tests: Biodata as an illustrative case. *Journal of Applied Psychology* 85, 479-484.
- Gatewood, R. D., & Feild, H. S. (2001). *Human resource selection* (5th ed.). Mason, OH: South-Western.
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18, 694-734.
- Goldman, B. M. (2001). Toward an understanding of employment discrimination claiming: An integration of organizational justice and social information processing theories. *Personnel Psychology*, 54, 361-386.
- Hammer, E. G., & Kleiman, L. S. (1988, May). Getting to know you: Biographical information that typically does not appear on the resume can help determine a job candidate's qualifications. *Personnel Administrator*, 86-92.

- Harvey-Cook, J. E., & Taffler, R. J. (2000). Biodata in professional entry-level selection: Statistical scoring of common format applications. *Journal of Occupational and Organizational Psychology*, 73, 103-118.
- Hunter, L. E., & Hunter, R. F. (1984). The validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Kluger, A. N., & Rothestein, H. R. (1993). The influence of selection test type on applicant reactions. *Journal of Business and Psychology*, 8, 3-25.
- Lefkowitz, J., & Gebbia, M. I. (1999). Dimensions of biodata items and their relationships to item validity. *Journal of Occupational Psychology*, 72, 331-351.
- Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology*, 44, 763-792.
- Mael., F. A., & Hirsch, A. C. (1993). Rainforest empiricism and quasi-rationality: Two approaches to objective biodata. *Personnel Psychology*, 46, 719-739.

- Mumford, M., & Whetzel, D. (1997). Background data. In D. Whetzel, & G. Wheaton (Eds.), *Applied measurement methods in industrial psychology* (pp. 207-240). Palo Alto, CA: Davies-Black Publishing.
- Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. A., & Sparks, P. C. (1990). Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology*, 75, 175-184.
- Rynes, S. L., & Connerley, M. L. (1992). Applicant reactions to alternative selection procedures. *Journal of Business Psychology* 7(3), 261-278.
- Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2001). Predictors used for personnel selection: An overview of constructs, methods, and techniques. *Handbook of Industrial, Work and Organizational Psychology* (Vol 1). Thousand Oaks, CA: Sage Publications, 165-199.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact on job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71, 432-439.
- Shermis, M. D., Falkenberg, B., Appel, V. A., & Cole, R. W. (1996). Construction of a faking detection scale for a biodata survey instrument. *Military Psychology*, 8, 83-94.
- Shultz, K. (1996). Distinguishing personality and biodata items using confirmatory factor analysis of multitrait-multimethod matrices. *Journal of Business and Psychology*, 3, 263-288.
- Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, 46, 49-76.
- Vinchur, A. J., Schippman, J. S., Switzer III, F. S., & Roth, P. L. (1998). A meta-analytical review of predictors of job performance for salespeople. *Journal of Applied Psychology*, 83, 586-597.

Wilkinson, L. J. (1997). Generalizable biodata? An application to the covational interests of managers. *Journal of Occupational & Organizational Psychology*, 70, 49-61.