Human Resource Management Review 19 (2009) 219–231

Contents lists available at ScienceDirect

Human Resource Management Review

journal homepage: www.elsevier.com/locate/humres

# The use of biodata for employee selection: Past research and future directions[☆]

James A. Breaugh [*]

*College of Business Administration, University of Missouri–St. Louis, St. Louis, MO 63121, United States*

ARTICLE INFO

*Keywords:*
Biodata
Biographical information
Personal history data
Employee selection

ABSTRACT

Although biodata has been shown to be one of the best predictors of employee performance and turnover, a number of important issues remain unresolved (e.g., how broadly or narrowly should biodata be defined?). This paper has three main objectives. The first objective is to provide a selective but representative review of the research that has been conducted on the use of biodata for employee selection. The second objective is to constructively critique this research. This critique is intended to highlight deficiencies of this research that may limit the conclusions that should be drawn. The paper's third objective is to stimulate important future research on biodata that avoids the limitations of past research.



The use of biodata for employee selection has a long history (Stokes, 1999). The results of both individual studies (e.g., Harold, McFarland, & Weekley, 2006) and meta-analyses (e.g., Schmidt & Hunter, 1998) have demonstrated the value of using personal history information as a predictor. In fact, many researchers (e.g., Ployhart, Schneider, & Schmitt, 2006) have concluded that biodata is one of the best selection devices for predicting employee performance and turnover.

Given the substantial evidence documenting its value as a predictor, it is surprising that biodata is not commonly used (at least in any formal way) by most employers. For example, Gatewood, Feild, and Barrick (2008) cited several studies in which human resource (HR) managers in the United States, Europe, and Australia were surveyed concerning their organization's use of biodata. In no study was a usage rate higher than 17% reported. The results of a recent survey of 255 HR professionals (Furnham, 2008) may explain why biodata is not more widely used. Compared to 11 other selection devices including "a personal hunch", biodata was perceived as lacking in terms of its validity (ranked 10th), practicality (ranked 9th), and legality (ranked 10th). Although practitioner concerns about validity, practicality, and legality may explain why biodata is not widely used, in this paper, these concerns are shown to only be relevant for certain types of biodata measures. For example, with regard to practicality, although developing some types of biodata scales can require a large sample size and technical expertise, not all approaches do. In summary, this paper suggests that biodata should be more widely used.

In recent years, it appears that researchers have been giving less attention to biodata in comparison to other selection devices. For example, less than one page was devoted to biodata in Evers, Anderson, and Voskuijl's *Handbook of Personnel Selection* (2005). In contrast, entire chapters addressed personality testing, assessment centers, cognitive ability testing, interviewing, and situational judgment tests. As another indicator of the lack of attention recently given to biodata, a computerized search of the PsycINFO database for English language articles published in 2008 turned up one article, that by Furnham (2008), for the term "biodata". Given its established validity, it is difficult to understand why biodata is not attracting more attention. Conceivably, there is a perception that the important research questions have already been answered. As is shown in this paper, this is not the case.

☆ The author would like to thank Rodger Griffeth for serving as action editor on this paper and three reviewers for their excellent comments on the paper.

* Tel.: +1 314 516 6287; fax: +1 314 516 6420.
*E-mail address:* jbreaugh@umsl.edu.

1053-4822/$ – see front matter 
doi:10.1016/j.hrmr.2009.02.003

In an attempt to stimulate more biodata research and the wider use of biodata by employers, this paper has three goals. The first goal is to provide a review of past research. This review establishes a foundation for addressing the second two goals of this paper. The second goal is to highlight deficiencies of past studies that should limit the conclusions that are drawn from them. The third goal of this paper is to stimulate future research that avoids the limitations of past research and provides a better understanding of why biodata is an effective predictor.

## 1. Biodata research: a selective review of research

Given the number of biodata studies that have been conducted, only a selective review of this research is possible. For the most part, in this review, attention is given to more recent studies. However, in order to provide a sense of early research on biodata and how it has evolved over almost 90 years, it is useful to consider one of the first published studies.

### *1.1. A study by Goldsmith (1922)*

Goldsmith (1922) examined the ability of nine "personal history" items (e.g., marital status, education, belonging to clubs) to predict the first-year sales of insurance agents. She found that using a person's biodata score would improve the hiring decisions made. Specifically, 58 of the 259 individuals (22%) receiving a score of 4 or above were considered successful (first-year sales was dichotomized in order to allow for expectancy charts to be constructed). In contrast, 11 of the 243 individuals (4%) receiving a score less than 4 were considered successful.

Several aspects of Goldsmith's (1922) study are noteworthy. First, unlike many later studies, only a few biodata items were used. Second, a number of the items used (e.g., age) would likely not be used today. Third, Goldsmith did not report data on the relationship of a given item and sales. Fourth, an explanation was provided for using each item (e.g., it was suggested that individuals whose prior jobs involved interacting with the public were more likely to have the social skills needed to perform well as an agent).

### *1.2. Defining and operationalizing biodata: differences in definitions and the types of items used*

Over the years, the term biodata has been defined in a number of different ways. Some researchers have defined it somewhat narrowly. For example, Nickels (1994) described biodata items as requiring people to describe behaviors and events that occurred earlier in their lives. Others have taken a broader view. For example, many of the biodata scales used have measured "temperament, assessment of working conditions, values, preferences, skills, aptitudes, and abilities" (Mount, Witt, & Barrick, 2000; p. 300). Yet others researchers never defined what they meant by the term.

An examination of the types of biodata items that have been used provides a sense of the divergent views of what biodata entails. For example, some researchers (e.g., O'Connell, Hattrup, Doverspike, & Cober, 2002) have focused on an applicant's work experience by asking questions about time in one's previous position and the number of jobs held in the past 5 years. Other researchers (e.g., Lefkowitz, Gebbia, Balsam, & Dunn, 1999) have used biodata items (e.g., "I like doing things with other people", "grades in math", "My teachers/lecturers regarded me as a sociable boy/girl", "My mother worked outside home when young") to tap a wide range of variables (e.g., educational experiences, preferences, personality, family history).

The failure of researchers to agree on what constitutes biodata is problematic. If biodata is seen as including such things as interests, personality, skills, and values, it becomes difficult to distinguish biodata measures from other measures. In an attempt to clarify matters, Mael (1991) provided a taxonomy of biodata item attributes. He stated that "the core attribute of biodata items is that the items pertain to historical events that may have shaped the person's behavior and identity" (p. 763). Mael argued that items that address such variables as behavioral intentions, self-descriptions of personality traits, personal interests, and ability fell outside of the realm of what biodata encompassed. Given Mael's definition of biodata as being historical in nature, he recommended that biodata items have several other characteristics. For example, he discussed the advantages (e.g., more accurate applicant reporting) of items that addressed discrete events that were verifiable. Mael also addressed how certain types of items (e.g., items that reflected an experience under the applicant's control, items that were seen as job-related) were likely to be viewed more favorably by job applicants.

### *1.3. Methods of gathering biodata*

It is generally agreed that biodata involves self-report data. In most cases, biodata information has been gathered by a paper-and-pencil measure. Mumford (1999) suggested there may be benefits in using a greater variety of data gathering methods. Two studies that involved the use of modern technology have recently been published.

Van Iddekinge, Eidson, Kudisch, and Goldblatt (2003) conducted a study that involved cashier and courtesy clerk positions. Applicants completed a 42-item biodata measure by dialing a toll-free number and responding "yes" or "no" using the telephone keypad. The biodata scores predicted supervisors' ratings of job performance.

Ployhart, Weekley, Holtz, and Kemp (2003) conducted a study that compared the scores on a 20-item biodata measure completed by two groups of applicants who were applying for call center positions. One group completed a paper-and-pencil measure. The second group completed a Web-based version of the measure. Ployhart et al. found the distributional properties of the scores on the Web-based measure were more desirable than the scores on the paper-and-pencil measure. For example, the Web-based group had a lower mean score, which may suggest less faking, and had lower scores in terms of skew and kurtosis.

### 1.4. Strategies used for developing biodata scales

Over the years, a number of strategies have been used for developing biodata scales (Hough & Paullin, 1994). Although several different names have been attached to these strategies, four labels (i.e., empirical, behavioral consistency, rational, and factorial) capture their essence. For presentation purposes, these strategies are treated as if they represent independent approaches. However, in reality, researchers sometimes use a combination of them.

An empirical strategy for scale development is based upon establishing a statistical relationship between a biodata item and the criterion of interest. Frequently, a large pool of items is used, and those items that were predictive are chosen for use (ideally, a cross-validation study would be conducted). This subset of items is used to arrive at an overall biodata score. With an empirical approach, understanding why biodata items predict the criterion is not a central concern. Given this fact, the term "dust bowl empiricism" is sometimes used to describe this strategy (e.g., Schoenfeldt, 1999).

A second biodata strategy that has been used, behavioral consistency, is based on the adage that past behavior is the best predictor of future behavior. With this strategy, a researcher selects or develops items that are consistent with the criterion of interest. For example, given they were interested in predicting turnover, Barrick and Zimmerman (2005) utilized a biodata item that asked how many months an applicant had been in his/her most recent job. Although there is a certain inherent logic as to why past behavior should predict future behavior (e.g., work ethic may be stable), researchers using a behavioral consistency strategy typically have not investigated these underlying causal variables (Dean & Russell, 2005).

A third strategy for developing biodata measures has been called the rational or deductive approach (Schmitt, Jennings, & Toney, 1999). This approach generally involves conducting a job analysis to determine the knowledge, skills, abilities, and other characteristics (KSAOs) that are important for the criterion of interest (e.g., for a new employee to successfully perform a job). Having determined these KSAOs, a researcher uses items that are thought to reflect them. For example, in attempting to reduce voluntary turnover, Barrick and Zimmerman (2005) thought that having a realistic view of the job being applied for was important. They hypothesized that applicants who knew people who worked for the organization would be more likely to have realistic job expectations. Therefore, they asked applicants whether they had such acquaintances. This biodata item did, in fact, predict voluntary turnover.

A commonly used strategy (e.g., Schmitt et al., 1999) for developing biodata scales is the factorial approach. Typically, this strategy involves the use of principal axis factor analysis or principal components analysis to extract the dimensions underlying the biodata items used. The hope is that the factors extracted will shed light on the constructs that explain why biodata scales predict the criterion of interest (Hough & Paullin, 1994).

In an attempt to assess the relative effectiveness of these biodata scale development strategies, a few studies (e.g., Stokes & Searcy, 1999) have compared them. In this regard, a study by Schoenfeldt (1999) is instructive. He compared four strategies (i.e., empirical, factorial, *a priori* rational, and *post hoc* rational) in terms of their validities for a validation sample ($N=475$) and a cross-validation sample ($N=252$). This study involved a concurrent validity design and focused on a customer service position. Given the results for the two rational scales were similar, only the results for the *a priori* scale are discussed. Schoenfeldt's study involved five criterion measures: supervisory ratings of customer service orientation and overall performance and days absent, times absent, and lateness.

With regard to the empirical strategy, scores on the 240 biodata items administered were entered into separate multiple regression analyses to predict each of the five criteria. Based upon a variety of decision rules Schoenfeldt (1999) used (e.g., a correlation of .10 or more with one of the criteria), 68 items were selected to comprise the final empirical biodata scale. This scale predicted all five criteria for the validation sample. For the cross-validation sample, only two criteria were predicted at a $p<.05$ level.

In order to derive a factorial scale, Schoenfeldt (1999) entered scores on the 240 biodata items into a principal components analysis. Both varimax and promox rotations were used, with the promax solution chosen given it resulted in more interpretable results. A 10-factor solution resulted (only 19% of the item variance was accounted for). The 10 highest loading items on each factor were used to create factor scores. The factor scores on the 10 factors were summed and used to predict the criteria. All five criteria were successfully predicted ($p<.05$) for both the validation sample (median $r=.24$) and the cross-validation sample (median $r=.23$).

Schoenfeldt's (1999) *a priori* rational biodata scale consisted of 68 items that were written to reflect 15 dimensions that a job analysis had shown to be important. A total biodata score was used to predict each of the five criteria. The rational scale score predicted ($p<.05$) each of the criteria for the validation sample (median $r=.31$) and the cross-validation sample (median $r=.25$). These results are roughly equivalent to those for the factorial scale.

In concluding this abbreviated treatment of biodata scale development strategies, two additional points should be noted. First, one should not generalize too readily from the results of Schoenfeldt's (1999) study. In other studies, carefully developed empirical biodata scales have often outperformed scales developed via other strategies (Schmitt et al., 1999). Second, some researchers have discussed subgrouping as an additional biodata scale development strategy. The rationale underlying subgrouping is that different groups may have different patterns of constructs that underlie their responses to biodata items. In this regard, Steinhaus and Waters (1991) discussed research that found the types of biodata items that best predicted military suitability for high school graduates differed from those that best predicted suitability for non-graduates. Mumford and Owens (1987) have provided a description of how different developmental experiences can lead to such subgroup differences.

### 1.5. The reliability of biodata scales

In those situations in which a biodata scale is designed to focus on a given construct, measuring internal consistency is appropriate. A study by Allworth and Hesketh (2000) represents such a case. These researchers developed an eleven-item biodata scale to tap past experience interacting with people. They reported a coefficient alpha of .89 for this scale. Internal consistency estimates can vary depending upon such things as the nature of the items used and the sample involved (Mumford, Whetzel, Murphy, & Eubanks, 2007). In past studies, internal consistency estimates ranging from .50 to .80 have been reported.

In many cases, biodata scales have been comprised of items that tap a variety of different variables. For example, the items that comprised the biodata scale used by O'Connell et al. (2002) included marital status, age, schooling completed, number of jobs held in the past 5 years, and living arrangement (i.e., rents, owns home, lives with parents). Given such varied items, computing coefficient alpha is not appropriate. Instead, as exemplified in a study by Shaffer, Saunders, and Owens (1986), reporting retest reliability may be appropriate. They used 118 biodata items that reflected demographic, experiential, and attitudinal variables. Schaffer et al. reported an average retest reliability of .77, but found differences by item type with scores on verifiable items being more stable. In those few studies that have reported retest reliability, the estimates range from .60 to .90.

### 1.6. The validity of biodata scales

Although attention should be given to reliability when appropriate, for many studies, the primary emphasis has been on criterion-related validity. For the most part, the results of recent studies have shown biodata to be a good predictor of job performance (e.g., O'Connell et al., 2002) and voluntary turnover (Barrick & Zimmerman, 2005). The results of somewhat dated meta-analyses by Schmidt and Hunter (1998), Vinchur, Shippmann, Switzer, and Roth (1998), and Bliesener (1996) also suggest that biodata is a good predictor. For example, Schmidt and Hunter estimated the predictive validity of biodata to be .35 for job performance and .30 for performance in a training program.

However, a caveat should be offered before one draws too strong a conclusion from the results of meta-analyses. These results represent averages. They do not shed light on such important issues as whether certain types of items (e.g., items focused on prior work experience) predict better than other types of items (e.g., items focused on school experiences). Nor, with the exception of the study by Bliesener (1996) which is discussed later in this paper, do the meta-analytic findings consider the fact that most biodata studies have involved concurrent validity designs (i.e., results for current employees may differ from those of job applicants).

### 1.7. Adverse impact

A concern that has been raised with using biodata is adverse impact against members of protected groups (e.g., Sharf, 1994). Given some of the items that have been used (e.g., age, educational level), this concern seems appropriate. In particular, biodata items that reflect cognitive ability (e.g., college grade point average) are likely to result in adverse impact. On the other hand, Hough, Oswald, and Ployhart (2001) suggested that, compared to other selection devices, the use of biodata has modest adverse impact. The findings of a few recent studies (e.g., Barrick & Zimmerman, 2005) support Hough et al.'s assertion. Given there is not a lot of research regarding adverse impact, it seems prudent for an organization to examine each biodata item it is considering using. With careful item screening, adverse impact should be minimized (Stokes, 1999).

### 1.8. Applicant reactions to biodata

Although several researchers have discussed job applicant reactions to the use of biodata, studies typically have not involved applicants. Rather, students or current employees have been asked their reactions to different types of items. Although data are lacking on actual applicant reactions, Ryan and Huth (2008) have provided an excellent framework for estimating applicant reactions. They suggested that applicants were likely to react negatively to items that were perceived as lacking job relatedness, that were perceived as fakable, and that were perceived as overly personal in nature.

### 1.9. The incremental validity of biodata

In a few studies, biodata has been shown to provide incremental validity. For example, Mount et al. (2000) reported that biodata added unique variance in predicting supervisory ratings of performance beyond that accounted for by tenure, general mental ability, and the Big Five personality traits. Allworth and Hesketh (2000) found their biodata scale accounted for unique variance in performance ratings when added after a cognitive ability test. In a study that looked at performance in a training program by Dean (2004), a biodata measure provided incremental validity beyond that predicted by a cognitive ability test. These and other studies (Schmidt, 1988) suggest biodata is likely to provide incremental validity.

### 1.10. The accuracy of biodata

Gatewood et al. (2008) have provided a thorough review of the limited research that has looked at the accuracy of reported biographical data. For the most part, the results of studies involving students in an educational setting have found fairly accurate reporting (e.g., external verification from parents supports the self-reported student data). In contrast, the results of studies conducted in a selection context are mixed. Given the types of items commonly used, biodata is susceptible to faking. For example, McFarland and Ryan (2000) demonstrated that individuals who were instructed to fake "good" could do so on a biodata scale.

Although relatively little data on their effectiveness exist, three strategies (i.e., the use of certain types of items, notifying individuals that data will be verified, and requiring that elaboration be provided for biodata answers) have been recommended for increasing the accuracy of the biodata gathered. For example, the results of a few studies (e.g., Graham, McDaniel, Douglas, & Snell, 2002; Harold et al., 2006) suggest that biodata data items that are historical and verifiable are less prone to faking than items that lack these attributes. Lautenschlager (1994) demonstrated that warning individuals that their responses may be verified reduced intentional distortion. A third strategy proposed for reducing faking involves requiring individuals to elaborate on their answers to biodata items (i.e., provide written support). Schmitt and Kunce (2002) demonstrated that requiring such item elaboration lowered scores on items in comparison to when elaboration was not required. Schmitt et al. (2005) conducted a follow-up study on item elaboration. They too found that requiring elaboration reduced biodata scores. Taken as a whole, it appears that all three strategies hold promise in terms of reducing biodata item distortion. In the future, it is important that these strategies receive more attention with regard to whether lower scores are related to higher biodata predictive validity.

### *1.11. Computing a biodata scale score: unit weighting versus differential weighting*

Regardless of the type of biodata items that are used (e.g., verifiable or non-verifiable) or the strategy that is employed for selecting/developing items (e.g., behavioral consistency or factorial), a key issue is how to weight them. A detailed treatment of this topic is beyond the scope of this paper (see England, 1971; Mumford & Owens, 1987). However, two of the most commonly used methods (i.e., correlational and differential regression) are briefly described.

The correlational method for weighting biodata items can take one of two approaches (Hogan, 1994). The first approach involves computing the simple correlation between an item and the criterion and using this value to weight the item (i.e., more highly correlated items receive higher weights). The second approach involves selecting all biodata items that are significantly related to the criterion and unit weighting them. With the differential regression method, weights are derived for biodata items that maximize the criterion variance explained (Hogan, 1994). Given regression analysis can produce unstable weights when used with several predictors and/or a modest sample size (e.g., Schoenfeldt, 1999; used 242 items with a sample size of 475), this method only should be used when a large sample is available and the number of items is not large.

In concluding this overview of item weighting, two additional points should be noted. The first point is that the correlational and the differential regression methods tend to provide comparable results (Gatewood et al., 2008). The second point is that most experts (e.g., Hogan, 1994) agree that the differential weighting of items is more beneficial when the correlations among the items are low, there are relatively few items, and there is a large sample.

### *1.12. The generalizability of biodata scales*

Given the time and the cost of that can be involved in developing a biodata scale, two important questions are: "Will a biodata scale developed in one organization be valid if applied in another organization?" and "Does biodata validity hold up over time?"

With regard to transportability, research suggests that biodata scales can be developed so as to be useful in different organizations. For example, Brown (1981) found a biodata scale developed for selecting insurance agents predicted sales volume across 12 companies. Rothstein, Schmidt, Erwin, Owens, and Sparks (1990) investigated the cross-organizational validity of a biodata scale comprised of two types of items (i.e., self-concept/evaluation and work values/orientation). Data were gathered from employees who worked for 79 organizations. Rothstein et al. found their biodata scale predicted the performance of supervisors across organizations. A study by Carlson, Scullen, Schmidt, Rothstein, and Erwin (1999), that used similar items to those used by Rothstein et al. (i.e., generic in nature for supervisors), also examined the transportability of a biodata scale. They found that a biodata score was a valid predictor of the rate of promotions across 24 organizations. In developing scales that are transportable, a key factor is that the biodata items are relevant to a given job (e.g., insurance agent, supervisor) regardless of the organization.

Some data also exist which suggest that a biodata scale that has been found to be valid in one country will have value if used in other countries. For example, Laurent (1970) reported that a biodata measure that had been validated on a group of managers working in the United States was also valid in predicting management success in company affiliates located in Denmark, Norway, and the Netherlands. Similarly, Dalessio, Crosby, and McManus (1996) demonstrated that the same scoring key for a biodata instrument used in the United States to select insurance agents could be used with equal effectiveness in the United Kingdom and Ireland.

A second generalizability question concerns whether the validity of a biodata scoring key is likely to remain valid over time. To address this question, Brown (1978) examined whether the scoring key for a 10-item biodata measure developed in 1933 for selecting insurance agents would hold up when used with a sample of agents hired between 1969 and 1971. He found that the 1933 key applied to the 1969–1971 data predicted survival and performance. Results supporting the potential stability of biodata validity also were reported by Rothstein et al. (1990) and Carlson et al. (1999). Rothstein et al. found that for the studies included in their meta-analysis, the earliest of which was conducted in 1974 and the latest of which was conducted in 1985, validity coefficients were similar. Carlson et al. reported that the scoring key of the Manager Profile Record yielded valid scores up to 11 years after the key was developed. Given the results of these studies, it appears that a well-developed biodata instrument will retain its predictive power over a sizable period of time. This stability is likely due to researchers using items that were developed so as to be generic (e.g., apply to most managers) and the fact that attributes of the jobs tapped by the biodata items have not changed greatly.

## 2. Past biodata research: three potential concerns

Taken as a whole, the results of biodata studies are impressive. For example, biodata has been established as a good predictor of employee performance. Furthermore, it appears a biodata scale may retain its validity over several years. The positive findings that have been reported have resulted in some strong conclusions being drawn by researchers. For example, Salgado, Viswesvaran, and Ones (2001) stated that the results of meta-analyses "suggest that biodata are one of the most valid predictors of personnel selection, and that their validity can generalize across organizations, occupations, and samples" (p. 182). Despite the substantial evidence supporting the value of using biodata, a careful reading of past studies raises three potential concerns (i.e., the heavy reliance on concurrent validity designs, the type of biodata scale used, and the lack of information provided on the items used) that merit attention.

### 2.1. The heavy reliance of past studies on a concurrent validity design

Given the focus of many studies has been on estimating the validity of a biodata scale for predicting the behavior of newly hired employees, an important issue is whether the results of studies that involved gathering biodata from current employees are comparable to those that involved job applicants. This is a particularly important issue given how many biodata studies have involved the use of a concurrent validity design (the heavy reliance on a concurrent validity design is not unique to studies on biodata; McDaniel, Hartman, Whetzel, & Grubb, 2007; noted that the great majority of studies on situational judgment tests have involved current employees).

A study by Bliesener (1996) is the only one located that categorized biodata studies in terms of the type of validation design used (he did so as part of a meta-analysis he was conducting). Studies were categorized as involving a concurrent validity design (i.e., biodata was gathered from current employees), a predictive validity design (i.e., biodata was gathered from job applicants but it was not used for making selection decisions), or a predictive validity design with selection (i.e., biodata was gathered from applicants and was used for making selection decisions). Bliesener located 165 biodata studies; 135 of these (82%) involved a concurrent validity design Although the studies on which Bliesener's results are based are somewhat dated, more recent studies (e.g., Allworth & Hesketh, 2000; Mount et al., 2000; Schoenfeldt, 1999) suggest researchers continue to rely heavily on designs involving current employees (for an exception see Dean & Russell, 2005).

Given the prevalence of studies that have used a concurrent validity design, it is important to know whether the results of such studies are comparable to studies in which biodata was gathered from job applicants. The meta-analysis conducted by Bliesener (1996) is directly relevant to this issue. He found the mean validity coefficients for studies using a concurrent validity design, a predictive validity design, and predictive validity design with selection were respectively .35, .29, and .21. Such differences in magnitude suggest that attention should be given to the type of research design used.

Stokes, Hogan, and Snell (1993) conducted the most relevant individual study with regard to the comparability of results for job incumbents and job applicants. They utilized a sample of incumbents who were working in a sales position with an equipment company and a sample of applicants who had applied for this position. Both samples completed a 171-item biodata scale that addressed work experiences and preferences, educational background, personal history, and leisure activities. For purposes of their study, they dropped the three demographic items from their biodata scale. The remaining 168 items were used to predict turnover (i.e., left within 1 year of hiring or service exceeding 1 year).

Stokes et al. (1993) used the same strategy for developing a final biodata scale for job incumbents and job applicants (i.e., separate scales were developed). First, from the initial pool of biodata items administered, they selected those items that were significantly correlated with employee turnover. For the job incumbent group, 28 items were selected. Next, they utilized regression analysis to select a subset of these 28 items that contributed unique variance in predicting turnover. This regression analysis was run on a random sample (approximately 50%) of the incumbent group. For the job incumbent group, the final biodata scale consisted of 13 items. Stokes et al. assessed the validity of this 13-item scale for their development and cross-validation samples. The observed validities for these two groups were .40 and .22 (both $p<.01$). For their job applicant sample, Stokes et al. followed the same procedure. At the first step, they found 27 items were significantly correlated with turnover. Based on their regression analysis, 10 items were selected for inclusion in the final applicant biodata scale. The estimated validities of this scale for their development and cross-validation samples were .36 and .23 (both $p<.01$).

Given biodata scales are commonly developed with current employees for use with job applicants, Stokes et al. (1993) applied the 13-item job incumbent biodata scale to their applicant sample. The resulting validity coefficient was .08. This value was not statistically significant and was much smaller than the validity coefficient (.22) found for the job incumbent cross-validation sample. Perhaps most troubling is the fact that the final biodata scales developed for the job incumbent and job applicant groups had no items in common!

A study by Harold et al. (2006) also merits consideration with regard to generalizing from the results of concurrent validity studies. These researchers had 425 call center employees and 410 applicants for call center positions respond to 20 biodata items which were written to assess educational history, work history, and leadership. They found the validity coefficient for their biodata scale in predicting ratings of job performance was higher for job incumbents (.27) than for job applicants (.18). These results parallel those reported by Bliesener (1996).

In summary, the studies by Bliesener (1996) and Harold et al. (2006) suggest the results for a biodata scale used with current employees may considerably overestimate the validity coefficient an employer might find for job applicants. For example, in Harold et al.'s study, biodata accounted for over twice as much of the variance in job performance for incumbents as it did for job applicants (i.e., 7% vs. 3%). Furthermore, the validity coefficient of .08 reported by Stokes et al. (1993) suggests that researchers should be cautious in assuming that a biodata scale developed based on data from current employees will be useful for deciding which job applicants to hire. Adding to this cautiousness should be the fact that the final biodata scales developed by Stokes et al., for their job incumbent and job applicant groups had no items in common.

### 2.2. The type of biodata scale used

A key decision in deciding to use a biodata scale is whether items should be closely tailored to fit the unique aspects of the position being filled or should be more generic in nature. Either option has potential disadvantages. For example, using a position-specific measure can require considerable development costs (e.g., writing items, pilot testing). In contrast, using a more generic biodata measure (e.g., one offered by a vendor) may cut costs but do so by sacrificing validity. In past studies, generic biodata scales have frequently been utilized (Bliesener, 1996). A few studies suggest that using such scales may result in sacrificing validity.

Mount et al. (2000) conducted a concurrent validity study that involved the use of biodata for selecting clerical employees. Their developmental sample ($N=222$) responded to 196 items that were written to predict four aspects of performance (i.e., task performance, interpersonal facilitation, problem solving, and retention probability). Mount et al. developed a separate biodata scale to predict each dimension of performance. Items that were significantly related to the corresponding aspects of performance for their development sample were retained for cross-validation. The "work habits" biodata scale consisted of 28 items, the "problem-solving abilities" scale had 40 items, the "interpersonal relations skills" scale had 38 items, and the "situation perseverance" scale had 40 items. The respective coefficient alphas for these scales were .54, .81, .64, and .69.

For their cross-validation sample ($N=154$), Mount et al. (2000) found their targeted biodata scales correlated significantly with their corresponding performance measures ($r$'s = .37, .33, .28, and .34). In contrast to these sizable validity coefficients, consider if Mount et al. had used a general biodata scale (i.e., one that was not targeted at predicting a specific performance dimension). By computing the median correlation of the four biodata scales Mount et al. used, a rough estimate of the validity of this general scale for a given performance dimension can be derived. These values for the four performance dimensions were .18, .10, .13, and .14. In summary, Mount et al.'s study suggests that developing situation-specific biodata scales can result in substantial validities in an absolute sense (median $r=.33$) and considerably higher validities (accounting for 4–10 times as much performance variance) than if a more generic scale (reflected by the median correlation of the four biodata scales) had been used (median $r=.13$).

Although one should be cautious in drawing conclusions from one study, the results of Bliesener's (1996) meta-analysis also suggest that a situation-specific biodata scale may have higher validity than a more generic scale. He categorized studies as having used either a scoring key that was develop for the specific setting in which the study was conducted or a standard scoring key that was used across organizations. Bliesener reported an average validity coefficient of .33 for the situation-specific key and an average validity coefficient of .22 for the standard key.

In summary, there is some evidence that a situation-specific biodata scale will have higher validity than a more generic scale. However, in those situations in which an employer is unable or unwilling to create a situation-specific biodata scale, as shown by the studies by Rothstein et al. (1990), and Carlson et al. (1999), the use of a more generic scale can have value.

### 2.3. The lack of information provided on biodata items

A third concern with past studies is the lack of information that has been provided on the biodata items used. For example, even a cursory review of the biodata literature makes apparent that most researchers (e.g., Harold et al., 2006) have not reported the actual items they used. This lack of reporting is likely due to two factors. First, researchers have frequently used very lengthy biodata measures (frequently more than one hundred biodata items were used). Therefore, journal space is an issue. Second, many researchers have used biodata scales sold by vendors, who do not allow their items to be published. Given the length of the biodata scales commonly used and concerns about the proprietary nature of items, it is not surprising that most studies have not reported: (a) the correlation between each biodata item and the criterion used in the study, (b) how each item was weighted in creating a biodata scale, (c) whether an item provided unique variance in predicting a criterion variable, (d) whether an item had adverse impact, and (e) the correlations among the biodata items.

Although information concerning the items used, whether an item accounted for unique criterion variance, etc. is not typically reported, such information would be valuable for both practitioners and researchers. For example, consider an employer that is going to develop a biodata scale. Knowing which items have proven to be valid predictors in past studies might be invaluable. Similarly, knowing whether a given item is likely to have adverse impact would be important. Having access to such information would allow an organization to select biodata items that are likely to be of maximum value while limiting the number of items that are used. For a researcher, having a sense of the types of items used and which were predictive can be essential for understanding the results reported (including non-significant findings).

Although other concerns with past biodata studies could be raised, given the three concerns addressed have received little attention to date, they were selected for emphasis. Conceivably, the failure of biodata researchers to reach consensus on two other issue (i.e., What is biodata? Why does it predict employee behavior?) should be of greater concern. Given the importance of these issues, they are addressed in the next section of this paper.

## 3. What is biodata? And why does it predict employee behavior?

In her introduction to a special issue on biodata that appeared in this journal a decade ago, Stokes (1999) stressed that "the role of life history in determining the success of individuals in different jobs has not been adequately understood" (p. 112). Although a decade has passed, Stokes' concern remains. That is, at present, we still do not have a sound understanding of why biodata predicts employee behavior.

### 3.1. What is biodata?

An important first step in increasing our understanding of why biodata is able to predict employee behavior would be reaching a consensus with regard what constitutes biodata. In this regard, most researchers (e.g., Allworth & Hesketh, 2000; Mumford & Owens, 1987) agree that biodata involves the gathering of self-report data. However, beyond that, opinions differ markedly (Gatewood et al., 2008). For example, as noted earlier, Mael (1991) has argued that biodata "pertains to historical events that may have shaped the person's behavior and identity" (p. 763). In contrast to this focus on an applicant's past behavior and experiences, several researchers have taken a much broader view. Among the factors that biodata has been viewed as encompassing are: (a) personality traits, (b) attitudes, (c) preferences, (d) future expectations, (e) self-assessed skills, (f) values, and (g) interests (e.g., Lefkowitz et al., 1999; Oviedo-Garcia, 2007; Sisco & Reilly, 2007; Van Iddekinge et al., 2003). The difficulty with viewing biodata so broadly is that it becomes difficult to define biodata as being anything more than self-report data.

Although there is general agreement that biodata is a method of data gathering, whether it represents a class of self-report data gathering methods (e.g., paper-and-pencil, telephone, web-based) or a specific method is an open question for future debate. What is more important for advancing the literature is a discussion of whether biodata represents only a method (i.e., it can measure a wide range of variables) or a method that is defined in terms of its measuring a limited range of variables. In keeping with the definition of biodata offered by Mael (1991) and others (e.g., Allworth & Hesketh, 2000), the position taken in this paper is that biodata should be defined only in terms of an applicant's past behavior and experiences. These past behaviors and experiences can reflect events that occurred in a work context (e.g., quit a job without giving notice), an educational setting (e.g., graduated from college), a family environment (e.g., traveled considerably growing up), community activities (e.g., led a cub scout troop), or other domains (e.g., active in local politics, religious activities).

Defining biodata in terms of historical events does not mean that a job applicant's past experiences are unrelated to such variables as interests, personality, values, knowledge, and skills. Rather, as described by Schmitt et al. (1999), it is likely that an individual who possesses certain interests, personality traits, values, and/or KSAs will be more likely to seek out certain situations that are captured by biodata (e.g., the type of job held, getting a college degree) than an individual who may not fit the situations as well. Similarly, having had certain experiences, an applicant may be more likely to have developed certain interests, values, and KSAs. In summary, it is suggested that many of the variables (e.g., personality traits) that have commonly been confounded with biodata are actually antecedents or consequences of the personal experiences that biodata taps. In the next section, this theme is more fully developed.

### 3.2. Why does biodata predict employee behavior?

In contrast to the number of empirical studies that have focused on the criterion-related validity of biodata, few models have been offered to explain "Why" biodata is able to predict employee behavior. The most programmatic model development effort has involved research by Mumford, Owens, and Stokes (e.g., Mumford & Owens, 1987; Mumford, Stokes, & Owens, 1990). Given the complexity of the model development work of these researchers and its evolution over time (see Dean & Russell, 2005), it is not possible to provide a detailed description of their efforts. However, an overview of their "ecological model" provides an important lens for viewing past research and for design future studies.

According to the ecological model, a person begins life with a certain environmental and hereditary resources (e.g., a nurturing mother, excellent eyesight) and certain limitations (e.g., substandard nutrition, poor coordination) which determine individual differences early in life. Given these individual differences (e.g., high cognitive ability, poor health, self-confidence), an individual attempts to maximize his/her adaptation to his/her environment. The ecological model presumes that an individual makes decisions about what situations to enter (e.g., "What college to attend?" Whether to accept a job offer?) based upon the perceived value of the outcomes (e.g., social status, financial rewards, intrinsic satisfaction) that are likely to be derived from the situations. The valences attached to these outcomes are based on the person's needs, interests, and values. The ecological model is interactive in that an individual's choice at a given point in time about what situation to enter affects his/her subsequent development, which influences his/her future choices of situations, which affect future development, etc. Thus, over time, an individual may develop new skills, satisfy existing needs, increase academic goals, or decrease his/her work ethic.

Much of the research on the ecological model conducted by Mumford, Owens, Stokes, and others has involved undergraduate students in a college environment (for an exception, see Dean & Russell, 2005). However, this model clearly has implications for understanding why biodata may predict individual behavior in a work context. Fig. 1 presents a simplified model of the ecological process (i.e., only three time periods of an ever-unfolding process are portrayed).

As portrayed in Fig. 1, at a given point in time (e.g., Time 1), an individual possesses several attributes (e.g., knowledge, values, goals). Based on these attributes (e.g., a consideration of one's abilities and goals), the ecological model suggests (i.e., the arrow labeled "Choice") the individual will actively choose to enter a new situation/environment (e.g., take a new job, begin an MBA program, start volunteering with a community organization) that is perceived as likely to aid his/her development (e.g., satisfy goals, increase important skills). Having experienced the new environment at Time 2, some attributes of the individual are likely to change (the arrow labeled "Environmental Effects"). Therefore, at Time 3, the person is likely to be different than at Time 1 on one or more attributes (e.g., increased interpersonal skills, different career goals, a diminished work ethic).

Although most of the discussions of the ecological model have stressed the active choices an individual makes, as reflected in the bottom portion of Fig. 1 (i.e., the arrow labeled "Non-Choice"), entering into a new environment is not always volitional. For example, a person may become unemployed due to an economic downturn, a health problem may arise, or a family crisis may

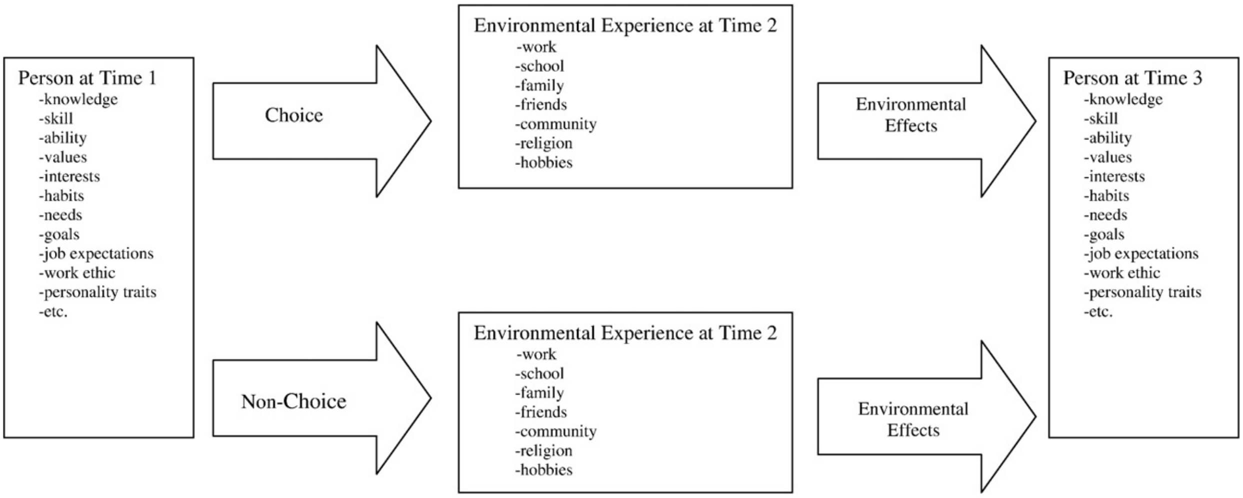


**Fig. 1.** A simplified representation of the ecological model.

occur. As reflected in Fig. 1, these and other uncontrolled events can result in an individual experiencing a new situation that was not sought out. In turn, such an experience can influence a person's attributes at Time 3 (e.g., having spent considerable time in a hospital, an individual's values may change).

This treatment of the ecological model was of necessity superficial (readers interested in a more detailed treatment of how past experiences may fit into a nomological network with other individual difference variables are referred to Mael, 1991; Mumford, 1999; Mumford, Reiter-Palmon, & Snell, 1994; Schmitt et al., 1999). For example, Mael (1991) discussed how social identity theory explains how interacting with valued others can affect a person's goals, interests, and beliefs. Despite this cursory treatment, Fig. 1 demonstrates how by limiting the definition of biodata to past behaviors and experiences, the focus of data gathering is on the events reflected in the boxes labeled "Environmental Experience at Time 2". Such events could include having graduated from college, how long one worked in one's previous job, the percentage of college expenses paid, whether one has had a leadership position in a community group, or whether one has management experience. In addition, the model makes clear that, if one defines biodata in terms of past behaviors and experiences, this does not mean a biodata score is unrelated to such variables as conscientiousness, ability, interests, knowledge, or work ethic. Rather, Fig. 1 shows how such variables are likely to be antecedents and/or consequences of an individual's behaviors and experiences.

Given the complexity involved, it is unlikely that the discussion in this section will bring resolution to the issue of how broadly or narrowly researchers define biodata. However, it is hoped that the preceding discussion will make more salient a number of important issues that have received insufficient attention in the past. For example, it should be apparent that a researcher could try to measure certain person attributes such as personality, interests, values, or job knowledge directly (e.g., via a test or an interview). Alternatively, a researcher could measure past behaviors and experiences (e.g., time on prior job, quitting without giving notice, past sales experience) with the assumption that they may serve as proxy measures for person attributes. For example, a researcher might assume that, having spent considerable time in a prior position that involved external sales, the person is dependable, knows what selling involves (e.g., has realistic job expectations), has good communication skills, and is extraverted. Such an indirect measurement strategy may be appropriate when the person variables of interest are not easily measured (e.g., applicants may be motivated to distort their responses). Given biodata has been shown to be one of the best predictors of new employee performance and retention, such an indirect approach to person measurement seems to be well-grounded.

## 4. Biodata: future research directions

Throughout this paper, a number of fundamental issues (e.g., Can one be confident generalizing from the results of a concurrent validity study to a selection context?) have been raised about which there is currently no consensus. In this section, seven important areas meriting additional research are highlighted.

### *4.1. What is biodata?*

Given the attention given to this issue earlier in this paper, it will not be addressed in detail here. However, I would contend that, unless general agreement among researchers can be reached concerning how broadly or narrowly biodata should be defined, advances in biodata research will be limited. It seems unlikely that rigorously designed empirical biodata studies will result in a consensus with regard to the breadth of the domain that biodata encompasses. Rather, it is more likely that cogent arguments offered by experts on the topic will need to persuade the research community.

### *4.2. Do results for concurrent validity studies generalize to a selection context?*

As Bliesener (1996) documented, biodata researchers have relied heavily on the use of concurrent validation studies. Given the results of Bliesener's meta-analysis and the results of a study by Harold et al. (2006) suggest validity coefficients derived from studies that used current employees may overestimate the validity coefficients found for studies involving job applicants, there clearly is a need for more studies that utilize predictive validation designs. Particularly important would be studies such as that conducted by Stokes et al. (1993) which raised the question of whether the same items are likely to be predictive for current employees and applicants.

### *4.3. Increased research with an item-focus*

As should be apparent from the literature reviewed earlier, relatively little attention has been given to the specific items used in biodata studies. For example, although a few studies (see Gatewood et al., 2008) have suggested that biodata items that are historical, verifiable, and focused on discrete events have higher validity than items lacking these attributes, little data are available on whether the context of biodata items (e.g., education, work, family) makes a difference. Furthermore, there is little data to draw upon with regard to how best to operationalize a job applicant's experience. For example, in attempting to predict job performance, at present it is unclear whether an applicant should be asked the number of times he/she has done a task or how long he/she has been in a job.

The lack of lack of attention in the biodata literature to measuring work experience is surprising. Although data are lacking in the context of biodata, a study by Quiones, Ford, and Teachout (1995) suggests decisions that are made about how to measure experience can be important. For example, these researchers showed that, in terms of predicting job performance, the number of times a person has completed a task ($r = .36$) may be more important than how long a person has been on the job ($r = .22$). In the future, increased attention to item-level issues (e.g., work vs. education, amount vs. times) certainly is merited.

### *4.4. Greater focus on the use of technology*

With the exceptions of studies by Van Iddekinge et al. (2003) and Ployhart et al. (2003), most biodata studies have involved the use of paper-and-pencil instruments. From an informal investigation of vendor websites, it appears that an organization can contract for the rights to administer a biodata instrument via the internet. However, in most cases, it appears that doing so simply involves administering a copyrighted instrument via a different medium.

In the future, a more sophisticated use of technology (e.g., computer terminals at the host organization) offers considerable promise. For example, the use of computer technology allows for customizing the items administered to the job applicant (or category of applicant) based upon certain characteristics. It also affords the ability for a biodata item to be scored differently depending on applicant characteristics. Although a detailed discussion of the value of computer-administered testing is beyond the scope of this paper (see Bartram, 2005), two issues merit attention.

An examination of the biodata literature makes apparent that many items have focused on educational experiences (e.g., academic honors awarded, number of high school sports participated in) and work experiences (e.g., experience interacting with customers, number of jobs held in the past 5 years). Yet other biodata items have involved other domains (e.g., age at which a driver's license was obtained). One can imagine that items focusing on educational experiences that may have occurred 30 years ago may lack face validity to older job applicants. In fact, inaccurate reporting may occur because of the faulty recall of events that took place years or even decades earlier.

In addition to concerns about perceived job relatedness and accurate reporting, another factor that should be considered is whether events that occurred several years ago lose their ability to predict. In this regard, a study by Roth, BeVier, Switzer, and Schippman (1996) is informative. These authors examined the relationship between grade point average (GPA) and job performance using meta-analysis. They reported an overall relationship of .16. However, what is more interesting is how this relationship varied depending upon how long ago a person graduated. Specifically, it was found that the correlation between GPA and performance was .23 for individuals who graduated within the last year, .15 for those who graduated between 2 and 5 years ago, and .05 for those who graduated 6 or more years ago. Although one should not attach too much significance to these data, they suggest that with the passage of time GPA becomes a less valid predictor. Although data are lacking on other experience-related variables that happened several years before, logic suggests that a similar diminishing relationship might be found.

This same time frame issue may be relevant for younger applicants. For example, as has been noted, items that refer to an applicant's work history (e.g., time in previous position, number of job held in the past 5 years) have been shown to be good predictors of employee turnover. Such items seem quite reasonable in trying to predict future turnover for an experienced worker. However, they appear less relevant in making selection decisions from an applicant pool that is comprised young applicants. Many of these individuals may never have held a full-time job or may have held several part-time jobs over the course of several summers. This issue not only applies to work-related biodata items. Items concerning such topics as marital status or home ownership may seem nonsensical for a teenager who has not worked before.

In summary, given the time frame of biodata items (e.g., was high school two or 40 years ago? has the person ever held a full-time position?) may differ greatly depending upon the age of an applicant, three different options appear to exist when an applicant pool is not homogeneous (the focus here has been on age, but other variables such as gender may also merit consideration—e.g., females, especially older females may have had less opportunity to participate in high school sports). First, an organization could use a different set of items depending upon the type of applicant (although this is difficult to do with paper-and-pencil instruments, it can be easily accomplished using computer technology). Second, an employer could develop multiple scoring keys that enable the criterion behavior of individuals from varying background to be predicted with maximum accuracy. A third option is to use more generic items that are likely to be predictive of everyone in the applicant pool (Mumford & Owens, 1987). Readers interested in a more detailed discussion of the relevance of time frame in developing biodata items are referred to Mumford et al. (1990) who presented an “ecology model” of human individuality based on a person's life history.

### *4.5. Ways to increase the accuracy of biodata information*

Given the concern that job applicants may not provide accurate biodata, a number of strategies have been proposed for addressing score inflation (e.g., stating that responses will be verified). In this context, recent research by Schmitt et al., 2005 is interesting. These researchers found that requiring individuals to provide written elaboration for biodata responses resulted in lower scores. Although requiring such elaboration appears to lower biodata scores, data are lacking with regard to whether such elaboration increases validity. Research addressing this issue is needed.

### *4.6. The value of a biodata clearinghouse*

A review of the literature makes apparent that most researchers have not reported the biodata items they used, nor statistical information about the items (e.g., the correlation between a biodata item and the criterion). The absence of such information is likely the result of many studies having used a large number of items. Although these types of information are typically not reported, this does not mean this information is unimportant. For example, consider an individual who is interested in creating a biodata scale to predict new employee turnover. Given the lack of information in the published literature, it is difficult to know what items to use. More specifically, although studies by Barrick and Zimmerman (2005) and Breaugh and Dossett (1989) support the use of a few items (e.g., number of jobs held in the prior 5 years), little information on other items that have predicted turnover are available. Putting together a scale to predict new employee performance is even more of a challenge. No items that consistently predict items have been highlighted in the research literature.

In order to address this problem, almost 2 decades ago, Mael (1991) called for a “clearinghouse for documentation of objective biodata items, complete with previous results and optimal scoring keys” (p. 786). He discussed the benefits of such a clearinghouse as well as noting some potential drawbacks (e.g., the scoring key for a copyrighted biodata scale could be compromised). Although it may be unrealistic to expect vendors to share information concerning a proprietary instrument, many researchers have developed their own scales. They may be willing to share such information, especially if access to it was restricted (e.g., a web site is only open to members of certain professional societies). In terms of running such a clearinghouse, this could be done by a professional organization or as a collaborative effort of professional organizations.

### *4.7. Rethinking the use of a factorial biodata development strategy*

A factorial strategy frequently has been used (e.g., Allworth & Hesketh, 2000; Stokes & Searcy, 1999) to investigate what constructs underlie the biodata items used in a study. Typically, this strategy involves either the use of principal components analysis (PCA) or principal axis factor analysis (PFA) with a varimax rotation. In the future, there are four reasons why researchers should think carefully about the use of a factorial approach. First, in most cases, the use of a varimax rotation is questionable (it is likely that the constructs underlying the biodata items are correlated). This may explain why many factor analytically-derived solutions are hard to interpret and/or account for little variance in the biodata items used (e.g., Schoenfeldt, 1999; reported that only 19% of the variance in his biodata items was accounted for). Second, given in many cases the biodata scales used involve a large number of items, frequently researchers lack the sample size needed to justify the use of PCA or PFA (e.g., Lefkowitz et al., 1999; conducted a PFA with a varimax rotation on the responses of 528 current employees to 160 items). Third, assuming some thought has been given to the selection of biodata items (e.g., what underlying variables they tap), confirmatory factor analysis likely represents a more appropriate analytic technique. Finally, as noted by Schmitt et al. (1999), the use of a factorial strategy can result in valid biodata items being dropped from the final biodata scale.

## 5. Concluding remarks

Considerable research has documented the value of biodata as a predictor of such important work outcomes as employee turnover and performance. Despite this fact, biodata is not commonly used by most organizations (Gatewood et al., 2008). This is both surprising and disappointing. It appears this lack of use is due to several factors. For example, concerns have been raised about biodata validity, practicality, and legality (Furnham, 2008). In addition to these issues, concerns have been raised about a lack of face validity to job applicants and others who are unfamiliar with the biodata literature, applicant faking, and applicant time demands (i.e., some of the scales used have involved over a 100 items).

From the research reviewed in this paper, it appears that some of these concerns are unfounded (e.g., biodata has been shown to be an excellent predictor), and other concerns are only true for certain biodata scales. For example, adverse impact and a lack of face validity may be minimized by the careful selection items (e.g., avoiding the use of demographic variables or variables that refer to childhood events). In a similar vein, with regard to practicality and applicant time demands, as shown by Barrick and Zimmerman (2005) and O'Connell et al. (2002), a biodata scale does not need to involve a large number of items (both studies involved fewer than 10 items) or a complex scoring key.

In conclusion, it is hoped that this paper results in biodata being more widely utilized by employers and more widely studied by researchers. For this to occur, it is argued that three major issues need to be addressed. First, there needs to be some agreement on what biodata is. Second, there needs to be greater reliance on predictive validity designs. Finally, there needs to be greater attention given to the specific biodata items that were used in a study.

## References


Allworth, E., & Hesketh, B. (2000). Job requirements biodata as a predictor of performance in customer service roles. *International Journal of Selection and Assessment*, *8*, 137−147.

Barrick, M. R., & Zimmerman, R. D. (2005). Reducing voluntary, avoidable turnover through selection. *Journal of Applied Psychology*, *90*, 159−166.

Bartram, D. (2005). Computer-based testing and the internet. In A. Evers, N. Anderson, & O. Voskuijl (Eds.), *Handbook of personnel selection* (pp. 399−418). Malden, MA: Blackwell.

Bliesener, T. (1996). Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational and Organizational Psychology*, *69*, 107−120.

Breaugh, J. A., & Dossett, D. L. (1989). Rethinking the use of personal history information: The value of theory-based biodata for predicting turnover. *Journal of Business and Psychology*, *4*, 371−385.

Brown, S. H. (1978). Long-term validity of a personal history item scoring procedure. *Journal of Applied Psychology*, *63*, 673−676.

Brown, S. H. (1981). Validity generalization and situational moderation in the life insurance industry. *Journal of Applied Psychology*, *66*, 664−670.

Carlson, K. D., Scullen, S. E., Schmidt, F. L., Rothstein, H., & Erwin, F. (1999). Generalizable biographical data validity can be achieved without multi-organizational development and keying. *Personnel Psychology*, *52*, 731−755.

Dalessio, A. T., Crosby, M., & McManus, M. A. (1996). Stability of biodata keys and dimensions across English-speaking countries: A test of the cross-situational hypothesis. *Journal of Business and Psychology*, *10*, 289−296.

Dean, M. A. (2004). An assessment of biodata predictive ability across multiple performance criteria. *Applied H.R.M.*, *9*, 1−12.

Dean, M. A., & Russell, C. J. (2005). An examination of biodata theory-based constructs in a field context. *International Journal of Selection and Assessment*, *13*, 139−149.

England, G. W. (1971). *Development and use of weighted application blanks.* Minneapolis: University of Minnesota, Industrial Relations Center (Bulletin No. 55).

Evers, A., Anderson, N., & Voskuijl, O. (2005). *Handbook of personnel selection.* Malden, MA: Blackwell.

Furnham, A. (2008). HR professionals' beliefs about, and knowledge of, assessment techniques and psychometric tests. *International Journal of Selection and Assessment*, *16*, 300−305.

Gatewood, R. D., Feild, H. S., & Barrick, M. (2008). *Human resource selection.* Mason, Ohio: Thompson South-Western.

Goldsmith, D. B. (1922). The use of the personal history blank as a salesmanship test. *Journal of Applied Psychology*, *6*, 149−155.

Graham, K. E., McDaniel, M. A., Douglas, E. F., & Snell, A. F. (2002). Biodata validity decay and score inflation with faking: Do item attributes explain variance across items? *Journal of Business and Psychology*, *16*, 573−592.

Harold, C. M., McFarland, L. A., & Weekley, J. A. (2006). The validity of verifiable and non-verifiable biodata items: An examination across applicants and incumbents. *International Journal of Selection and Assessment*, *14*, 336−346.

Hogan, J. B. (1994). Empirical keying of background data measures. In G. A. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp. 69−107). Palo Alto, CA: Consulting Psychologists Press.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment*, *9*, 152−194.

Hough, L., & Paullin, C. (1994). Construct-oriented scale construction. In G. A. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp. 109−145). Palo Alto, CA: Consulting Psychologists Press.

Laurent, H. (1970). Cross-cultural cross-validation of empirically validated tests. *Journal of Applied Psychology*, *54*, 417−423.

Lautenschlager, G. J. (1994). Accuracy and faking of background data. In G. A. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp. 391−419). Palo Alto, CA: Consulting Psychologists Press.

Lefkowitz, J., Gebbia, M. I., Balsam, T., & Dunn, L. (1999). Dimensions of biodata items and their relationship to item validity. *Journal of Occupational and Organizational Psychology*, *72*, 331−350.

Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology*, *44*, 763−792.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, *60*, 63−91.

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, *85*, 812−821.

Mount, M. K., Witt, L. A., & Barrick, M. R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, *53*, 299−323.

Mumford, M. D. (1999). Construct validity and background data: Issues, abuses, and future directions. *Human Resource Management Review*, *9*, 117−145.

Mumford, M. D., & Owens, W. A. (1987). Methodological review: Principles, procedures, and findings in the application of background data measures. *Applied Psychological Measurement*, *11*, 1−31.

Mumford, M. D., Reiter-Palmon, R., & Snell, A. F. (1994). Background data and development: Structural issues in the application of life history measures. In G. A. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp. 555−581). Palo Alto, CA: Consulting Psychologists Press.

Mumford, M. D., Stokes, G. S., & Owens, W. A. (1990). *Patterns of life history: The ecology of human individuality.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Mumford, M. D., Whetzel, D. L., Murphy, S. T., & Eubanks, D. L. (2007). Background data. In D. L. Whetzel & G.R. Wheaton (Eds.), *Applied measurement: Industrial psychology in human resource management* (pp. 201−233). New York: Lawrence Erlbaum Publishers.

Nickels, B. J. (1994). The nature of biodata. In G. A. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp. 1−16). Palo Alto, CA: Consulting Psychologists Press.

O'Connel, M. S., Hattrup, K., Doverspike, D., & Cober, A. (2002). The validity of "mini" simulations for Mexican salespeople. *Journal of Business and Psychology*, *16*, 593−599.

Oviedo-Garcia, M. A. (2007). Internal validation of a biodata extraversion scale. *Social behavior and personality*, *35*, 675−692.

Ployhart, R. E., Schneider, B., & Schmitt, N. (2006). *Staffing organizations: Contemporary practice and theory*, (3rd edition.). Mahwah, NJ: Lawrence Erlbaum.

Ployhart, R. E., Weekley, J. A., Holtz, B. C., & Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata, and situational judgment tests comparable? *Personnel Psychology*, *56*, 733−752.

Quiñones, M. G., Ford, J. K., & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, *48*, 887−910.

Roth, P. L., BeVier, C. A., Switzer, F. S., & Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology*, *81*, 548−556.

Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. A., & Sparks, C. P. (1990). Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology*, *75*, 175−184.

Ryan, A. M., & Huth, M. (2008). Not much more than platitudes? A critical look at the utility of applicant reactions research. *Human Resource Management Review*, *18*, 119−132.
Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2001). Predictors used for personnel selection: An overview of constructs, methods, and techniques. In N. Anderson, D. S. Ones, H. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology*, *Vol. 1* (pp. 165–199). Thousand Oaks, CA: Sage.
Schmidt, F. L. (1988). The problem of group differences in ability test scores in employment selection. *Journal of Vocational Behavior*, *33*, 272−292.
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.
Schmitt, N., Jennings, D., & Toney, R. (1999). Can we develop measures of hypothetical constructs? *Human Resource Management Review*, *9*, 169−183.
Schmitt, N., & Kunce, C. (2002). The effects of required elaboration of answers to biodata questions. *Personnel Psychology*, *55*, 569−587.
Schmitt, N., Oswald, F. L., Kim, B. K., Gillespie, M. A., Ramsay, L. J., & Yoo, T. (2005). Impact of elaboration on socially desirable responding and the validity of biodata measures. *Journal of Applied Psychology*, *88*, 979−988.
Schoenfeldt, L. F. (1999). From dust bowl empiricism to rational constructs in biographical data. *Human Resource Management Review*, *9*, 147−167.
Shaffer, G. S., Saunders, V., & Owens, W. A. (1986). Additional evidence for the accuracy of biographical data: Long-term retest and observer effects. *Personnel Psychology*, *39*, 791−809.
Sharf, J. C. (1994). The impact of legal and equal employment opportunity issues on personal history inquiries. In G. A. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook* (pp. 351−390). Palo Alto, CA: Consulting Psychologists Press.
Sisco, H., & Reilly, R. R. (2007). Development and validation of a Biodata Inventory as an alternative method to measurement of the five factor model of personality. *The Social Science Journal*, *44*, 383−389.
Steinhaus, S. D., & Waters, B. K. (1991). Biodata and the application of a psychometric perspective. *Military Psychology*, *3*, 1−23.
Stokes, G. S. (1999). Introduction to special issue: The next one hundred years of biodata. *Human Resource Management Review*, *9*, 111−116.
Stokes, G. S., Hogan, J. B., & Snell, A. F. (1993). Comparability of incumbent and applicant samples for the development of biodata keys: The influence of social desirability. *Personnel Psychology*, *46*, 739−762.
Stokes, G. S., & Searcy, C. A. (1999). Specification of scales in biodata form development: Rational vs. empirical and global vs. specific. *International Journal of Selection and Assessment*, *7*, 72−85.
Van Iddekinge, C. H., Eidson, C. E., Kudisch, J. D., & Goldblatt, A. M. (2003). A biodata inventory administered via interactive voice technology (IVR) technology: Predictive validity, utility, and subgroup differences. *Journal of Business and Psychology*, *18*, 145−156.
Vinchur, A. J., Shippmann, J. S., Switzer, F. S., & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, *83*, 586−597.