LIFE EXPERIENCES AND PERFORMANCE PREDICTION:

TOWARD A THEORY OF BIODATA

Michelle A. Dean
Craig J. Russell
Division of Management
Michael F. Price College of Business
University of Oklahoma
Norman, OK 73019-4006

and

Paul M. Muchinsky

Department of Business Administration
Joseph M. Bryan School of Business and Economics
University of North Carolina-Greensboro
Greensboro, NC 27412-5001

In G. Ferris (Ed.), *Research in Personnel/Human Resources*, published by JAI Press, 1999.

Craig J. Russell, Ph.D.
Chair

Timothy D. Rasnic
Director

J.C. Penney Leadership Center
Research Series 98-001

ABSTRACT

Fleishman (1988) described biodata selection technology as among the most promising avenues for generation of new knowledge in personnel selection. However, researchers generally hold biodata selection systems in low regard due to its perceived atheoretical nature. Further, surveys indicate biodata is used in less than 5% of personnel selection decisions. We argue that biodata systems are no more atheoretical than other popular selection technologies. We review aspects of biodata instruments that make them unique among selection devices and biodata theory as embodied in the ecology model (Mumford, Stokes, & Owens, 1990) before offering two extensions. First, we propose to extend the ecology model by focusing on negative life events, reviewing diverse literatures addressing affective and cognitive reactions to these events. Second, an individual difference variable labeled "moxie" is proposed as a key mediator and/or moderator of latent negative life event-job performance relationships. Specific directions for needed research are presented.

INTRODUCTION

Performance prediction is one of the fundamental tasks of management. Specifically, managers align raw materials, capital, and human resources with market opportunities in ways that yield targeted performance outcomes. Performance predictions made in the context of personnel selection systems thus become a primary contribution of human resource managers. Similarly, development of theories or models that provide insight into why and when applicants are able to perform on the job becomes a primary objective of human resource management research.

The literature on performance prediction suggests three selection technologies tend to achieve the highest predictive power or criterion validity across situations: cognitive ability tests, work sample tests (e.g., assessment centers), and biographical information inventories (hereafter referred to as biodata). Ample meta-analytic evidence exists to support use of these technologies in a wide variety of personnel selection settings (Schmitt, Gooding, Noe, & Kirsch, 1984; Reilly & Wareck, 1990; Russell & Dean, 1994b). However, only an extremely small portion of the research literature has been devoted to explaining why these technologies demonstrate criterion validity.

Hunter (1986) speculated that paper and pencil tests of general cognitive ability (*g*) demonstrate criterion validity due to controlled cognitive processing requirements found on the job, consistent with his finding that *g* criterion validities increase with job complexity. To be sure, an immense literature yields insight into the nature and development of cognitive ability (e.g., Kanfer, 1990; Kanfer &

Ackerman, 1989). The same cannot be said for insight into latent causal processes between an individual's cognitive ability and subsequent work performance. It remains somewhat humbling to inform a lay audience that "we just recently have been able to conclude that smart people do better on the job, though we aren't really sure why."

Work sample and simulation criterion validities have led investigators to generate competing explanations for latent causal processes. For example, Klimoski and Strickland (1976) identified numerous competing explanations for assessment center criterion validity. Unfortunately, Klimoski and Brickner (1987) were able to generate an even larger set of competing explanations in the virtual absence of any systematic tests of the issues during the preceding 10 years (see Russell & Domm, 1995, for a test of two competing explanations). The absence of compelling theory becomes most troublesome when the focus of work sample information shifts from performance prediction to training needs assessment; that is, if we do not know what we are measuring we cannot infer what interventions (e.g., training or otherwise) might be appropriate. Note that a discussion of a common explanation of assessment center criterion validity, i.e., the "consistency principle" (Wernimont & Campbell, 1968), and its shortcomings appears somewhat later in the paper.

It should not be surprising then, that the third selection technology (i.e., biographical information inventories) can also be described as lacking strong theory. What might be surprising is the verve with which labels such as "dustbowl empiricism" have been selectively applied to this technology when other technologies provide equally viable targets (Childs & Klimoski, 1986; Dunnette, 1962; Owens, 1976; Nickels, 1994). Common use of pejorative labels in the biodata literature (e.g., "atheoretical" and "dustbowl empiricism") may explain its low usage rates among practitioners; of 348 firms surveyed, Hammer and Kleinman (1988) found only 6.8% had ever used biodata in employment decisions and only 0.4% currently used biodata. Notably, 40% of Hammer and Kleinman's respondents indicated invasion of applicant privacy was a primary concern in biodata use. Similarly, a Bureau of National Affairs survey of human resource specialists reported only 4% used biodata. The human resource specialists indicated perceived invasiveness of applicant privacy as a major reason to avoid biodata selection systems (as reported in Mael, Connerley, & Morath, 1996). However, while most consider an individual's performance on standardized tests to be "personal" information (Note 1), at the item level few would consider the question "What does 2 + 2 sum to?" as invasive as "How many magazines did your parents subscribe to while you were in high school?" (Note 2).

We suspect a number of unique characteristics contribute to how often authors reference biodata's "atheoretical" nature. For example, investigators have only recently started to employ traditional psychometric construct validity techniques in biodata item development (Mumford, Costanza, Connelly, & Johnson, 1996; Russell, 1994). An early emphasis on sorting biodata items into taxonomies did not evolve into a traditional psychometric assessment of construct validity. Further, empirical keys are common in biodata applications, generally devoid of theoretical rationale, and virtually unheard of in other selection technology arenas.

Regardless, Fleishman (1988) argued that the greatest opportunities for advancing our understanding of human performance in organizations lie in examining relationships between life experiences and subsequent job performance. Fleishman was not alone in his observations, they have been echoed over the years by many others (Dunnette, 1962, 1966; Guion, 1965; Owens, 1976). Biodata has likely been viewed as so promising due to the (as yet unrealized) potential of a biodata theory to guide life experience interventions (i.e., the purposeful, theory-guided exposure of individuals to certain quantities and qualities of life experiences).

The primary purpose of this paper is to review existing biodata theory development and propose a partial extension. Our approach builds on prior efforts as a first step toward specification of life experience interventions hypothesized to causally influence subsequent performance. Toward these ends, the remainder of this paper is structured in three sections. First, we review characteristics of biodata unique to this performance prediction technology, including efforts at developing biodata item taxonomies and empirical keying procedures. Fleishman's (1988) vision of biodata's "promise" will only be realized with an increase in research activity, which will require an understanding of unique biodata characteristics by a broader research audience. Second, we describe development of existing biodata theory. Third, we present an extension to this theory which specifies an additional life history construct domain and a relatively unexamined personality characteristic. A secondary purpose in reviewing biodata characteristics and theory is to provide a single source of such information for those interested in contributing to the body of research.

UNIQUE ASPECTS OF BIODATA PERFORMANCE PREDICTION TECHNOLOGY

Biodata Item Types

A number of authors have taken it upon themselves to develop taxonomies of biodata item "types" or characteristics. Perhaps most well known is Asher's (1972) taxonomy, initially offered as a means of "improving" biodata items. "Improvement" was defined in terms of eliminating potential sources of systematic error rather than increasing scale convergence with some latent construct domain, an important but unfortunate distinction that appears throughout the biodata literature. While other performance prediction procedures focus on latent constructs underlying the predictor measures (e.g., cognitive ability and Big-5 personality measures), biodata item taxonomies tend to focus on minimizing sources of prediction error.

Biodata item characteristics discussed by Asher (1972) included:

- Verifiable vs. unverifiable } Hard vs. soft (joint verifiable & factual vs. unverifiable & interpretive)
- Subjective vs. objective
- Historical vs. futuristic
- Actual behavior vs. hypothetical behavior
- Memory based vs. conjecture based
- Specific vs. general
- Response vs. response tendency
- External vs. internal event

While Asher's (1972) taxonomy was originally presented as a description of biodata item characteristics, subsequent authors have used these and other characteristics in a prescriptive manner (e.g., Gandy, Outerbridge, Sharf, & Dye, 1989).

A number of research efforts have examined taxonomic characteristics. For example, Owens, Glennon, and Albright (1966) found that consistency in subjects' responses to biodata items was related to item brevity, the presence of an escape option, and phrasing in a neutral or pleasant connotation. Mael (1991) recently echoed Asher's (1972) earlier suggestions that some biodata item attributes aid in obtaining accurate responses. Mael suggested biodata items should ask for respondent's first-hand knowledge, avoiding asking individuals about how others would evaluate the respondent. For example, asking, ?How did your parents evaluate your academic achievement?? would be second-hand information in which the respondent is asked to speculate. Verifiable items have been found to reduce the likelihood that applicants will provide bogus answers in hopes of achieving higher scores (Atwater, 1980; Cascio, 1975; Mosel & Cozan, 1952). However, Hough, Eaton, Dunnette, Kamp, and McCloy (1990) found that simply warning respondents that answers can be verified may act as a faking deterrent regardless of item content.

Mael (1991) also considered legal and moral issues surrounding the use of biodata. Items may vary in terms of controllability, accessibility, and visible job relevance. Controllability refers to the degree to which a person chose to perform or not to perform an action. This label parallels Owens and Schoenfeldt?s (1979) distinction between prior behaviors (i.e., behaviors that were chosen by the applicant) and input variables (i.e., things that happened to the applicant). Mael suggested all life events, whether consciously chosen or not, have the ability to shape a person's future behavior, and should be included on a biodata instrument.

Opposition to including noncontrollable items such as parental behavior and socioeconomic status typically speaks to the "fairness" of these items, due to applicants lack of control over their early environment. This leads to a potential legal question regarding whether items should tap skills and experiences that are equally accessible to all applicants (Stricker, 1987; 1988). Stricker (1988) suggested items asking about experience as a football team captain would be unfair because individuals of a particular gender, size, or size of school may not have the opportunity to engage in this role. In contrast, Mael (1991) suggested the concept of equal access is irrelevant to prediction or test fairness as operationalized by the Uniform Guidelines on Employee Selection Procedures (Equal Employment Opportunity Commission, 1978). Rather, what is relevant is that the person who had access to the role was changed in some meaningful way by the role while others who were not in-role received no benefit nor harm (i.e., individuals who were not football captains are not penalized for non-exposure).

Finally, biodata items vary in visible job relevance. Many view all life experiences as potentially developmental. However, Gandy et al. (1989) used only items with a point-to-point relationship between item content and job content on their public sector biodata instrument in order to avoid the accessibility issue and maximize job relevance.

Research has examined the influence of item attributes on biodata results. Barge and Hough (1988) analyzed biodata items in terms of item heterogeneity (i.e., degree to which items measure more than one construct), behavioral discreteness (i.e., degree to which items address ?a single, perhaps verifiable behavior rather than a more abstract or summary characteristic?? pp. 3-4), and behavioral consistency (i.e., degree of congruency between the content domain of the biodata item and the content domain of the target job, or the degree to which an item is a sign v. sample of behavior, Wernimont & Campbell, 1968). Barge (1988) analyzed 103 items taken from Owens and Schoenfeldts? (1979) Biographical Questionnaire (BQ) and found more homogenous items (i.e., in terms of consistency and discreteness) demonstrated higher criterion validity. In a different vein, Shaffer, Saunders, and Owens (1986) found ?soft,? or subjective, non-verifiable items, are nearly as predictive and reliable as ?hard,? or verifiable, factual items. Average test-retest reliability five years after initial administration was higher for objective than subjective items. The Barge and Shaffer et al. efforts are noteworthy because they are among the first to evaluate the influence of biodata item characteristics on criterion validity (Stokes & Reddy, 1992).

However, Asher's (1972), Mael's (1991), and others' item taxonomies, like all taxonomies, can be evaluated in terms of 1) ability to reduce variance in measurement of the latent construct domain, and 2) prediction of some criterion (e.g., job performance) of interest (Fleishman & Quaintance, 1984; McCall & Bobko, 1992). Biodata item taxonomies tend to fall short in terms of the former due to emphasis on the latter; indeed, most biodata research has focused on enhancement of criterion validity with a distinct lack of attention to latent construct domains. Decreasing the presence of random and systematic measurement error in subjects' responses will certainly contribute to biodata criterion validity and ease of use. Further, taxonomic efforts pinpoint the types of items typically found on a biodata instrument and differentiate biodata items from other closely related, but conceptually distinct, measures such as personality scales. However, there is no theoretical basis for ?correctness? of any of these attributes. The attributes need to be understood in terms of their contribution to measurement of some latent construct domain. While biodata taxonomies may prove useful for generating some theory or model of inter-related construct domains, item taxonomies cannot be considered theory itself (Bacharach, 1989). Unfortunately, investigations conducted in the 26 years since Asher's (1972) study suggested taxonomic efforts have not lead to an understanding of latent predictor constructs and causal processes. It would seem that such insight will not be forthcoming without a shift in discussion of "biodata item types" to a discussion of latent biodata construct domains.

Empirical Keying Procedures

Empirical keys constitute a second unique aspect of biodata performance prediction systems. Empirical keying procedures assign "optimal" weights to items or response options for purpose of performance prediction. Subject responses are typically multiplied by the respective weight associated with each item/response option and summed into a "biodata score." Weights are derived in ways that reflect 1) some empirical relationship between the item/response option and the criterion of interest (externally referenced) or 2) empirical relationships among the item/response options (internally referenced). Empirical keys in many ways might be considered a "defining characteristic" of typical biodata systems because no other behavioral science applications in organizational settings use similar scoring procedures. For example, no published research breaks down Likert scale job satisfaction items into dummy coded (0 = did not select, 1 = did select) response options before determining the optimal weighting required to predict job performance. Indeed, the statistical sophistication associated with empirical keying procedures may serve as a "barrier to entry" for many investigators and practitioners. Empirical keying is one of the primary reasons for criterion validities summarized in the meta-analytic estimates reported in Table 1 (Mitchell & Klimoski, 1982).

Insert Table 1 about here

A wide variety of empirical keying methods exist. Unfortunately no single study has compared all methods in a single data set. Further, even if such a study existed, its findings would not generalize beyond the biodata items and latent criterion construct domains (whatever they might be) captured in that study. A brief review of the most frequently used externally referenced methods (i.e., where weights are a function of biodata-criterion relationships) is presented below, including the vertical percent difference, horizontal percent difference, mean criterion, phi coefficient, and regression methods. One internally referenced method (i.e., where weights are a function of relationships among responses to biodata items) is also discussed, as are a few emerging methods.

The vertical percent difference method (Stead & Shartle, 1940) historically appears to be most frequently used in empirical biodata key construction (Mumford & Owens, 1987). This method calculates differences in response percentages for high- and low-criterion groups in a three-step process applied to a "key development" portion of the original sample. First, high- and low-criterion groups are established using individuals who are most or least successful on the criterion of interest. Second, the percentage within each of the extreme criterion groups who chose each item response option is computed. Third, a difference in percentages for each alternative is calculated by subtracting the percentage for the low-criterion group from the percentage for the high criterion group. Strong?s (1926) tables (see Stead & Shartle, 1940, p. 255) can then be used to transform percentage point differences to net weights, though alternate weighting procedures are common (e.g., simple unit weights when some arbitrarily large percentage point difference is present). Once derived in the key development sample, subject responses in the remaining portion of the original sample, called the cross validation sample, are multiplied by the respective weights, summed, and correlated with the target criterion.

Perhaps the most flexible application of the vertical percent difference method is captured by a custom keying procedure developed originally by Abrahams (1965) and subsequently crafted into a computer program labeled "KEYCON" at the U.S. Navy Personnel Research and Development Center

(Abrahams, 1998). The KEYCON program permits the investigator to systematically vary the following parameters:

1. The number of subjects randomly selected for the key development versus cross validation samples.
2. The cut points for determining "high" and "low" performing criterion groups in the key development sample.
3. The percent difference in response frequency between high and low performing criterion groups required for a response option to enter the key.
4. The weight associated with each response option entering the key (e.g., use of unit weights for all percentage differences larger some specified level, use of percentage difference as the weight, etc.).

Thus, an investigator can explore the tradeoff between increasing key development sample size (increasing expected reliability of the key) while decreasing cross validation sample size (increasing expected sampling error in estimating criterion validity), and vice versa. Programs such as KEYCON greatly facilitate investigator judgment calls in construction of vertical percent difference empirical scoring keys.

The horizontal percent difference method (Guion, 1965; Stead & Shartle, 1940) also arrives at weights for each response option by identifying subjects in a key development sample who were high on the performance criterion. Response option weights are derived by dividing the number of subjects selecting a particular response option from the high criterion group by the total number of subjects. Weights for response options selected by each subject in the cross validation sample are summed to form a biodata score. A criterion validity estimate is derived by correlating the key-based biodata score with the criterion in the cross validation sample.

The mean criterion technique (Devlin, Abrahams, & Edwards, 1992) does not require use of high- and low-criterion groups, rather weighing each response option by the average criterion score of respondents in the key-development sample who chose that response option. The key is then applied to a cross validation sample. The correlation between the biodata score and criterion in the cross-validation sample estimates criterion validity in the population.

Several correlation/regression methods of empirical keying are also available, including the phi coefficient method, regression technique, rare response method (Neidt & Malloy, 1954), and multiple regression method (Malone, 1977). All typically develop empirical weights in a key development sample before estimating criterion validities in a cross validation sample.

The phi coefficient method (Lecznar, 1951; Lecznar & Dailey, 1950) uses correlations between each item response option (response versus no response) and a binomial criterion as response option weights. A variation of this method uses the point biserial correlation coefficient when the criterion is continuous. Regression techniques are similar in that standard regression analysis coefficients weight items or response options on the basis of their ability to add to the prediction obtained from items already in the regression equation (Malone, 1977). Interpretation of regression weights derived for each item is subject to typical restrictions associated with common least square linear regression procedures. However, when response options are used as independent variables, the parametric assumption of independence among the predictors is violated (e.g., subjects selecting response option "1" on a question by definition did not select response option "2"), and interpretation of regression weights becomes impossible.

Russell and Dean (1994a) used a variation of the phi coefficient method in which point biserial coefficients between response options and the criterion were derived. However, they noted when distributional characteristics of two variables do not conform with parametric assumptions, distributional characteristics of phi, point biserial, or Pearson product moment correlations between the two variables are unknown and no probablistic inferences can be drawn. As a result, two response options with equal point biserial correlations with a criterion might have different standard errors (Dean, Russell, & Broach, 1998). For example, $s_{r_{pb}}$ is expected to be the same for two response options with equal r_{pb} <u>if</u> assumptions of bivariate normality hold (e.g., Fischer, 1970 found the best estimate of $s_r = s_r = \frac{\sqrt{(1-r^2)}}{N-2}$). However, r_{pb} between response options and performance criteria cannot exhibit bivariate normality because one of the two variables (i.e., the response option) is discrete and usually highly skewed. In this instance, true values of $s_{r_{pb}}$ might vary across response options yielding the same estimate of r_{pb}. Russell and Dean (1994a) argued the estimated criterion-related validity of a response option with a smaller $s_{r_{pb}}$ is less subject to random sampling error and should receive more weight in the key (e.g., by dividing r_{pb} by a bootstrap estimate of $s_{r_{pb}}$).

Consequently, Russell and Dean (1994a) developed an empirical key based on both estimates of effect size (r_{pb}) and accuracy of those estimates ($s_{r_{pb}}$), using the bootstrap technique to estimate the population value of $s_{r_{pb}}$. They found marginal increases in criterion validity using response options weighted by r_{pb} and $s_{r_{pb}}$. Further, examination of bootstrapped frequency distributions of r_{pb} indicated that a single population value of r_{pb} was unlikely for a number of response options. Subsequent analyses by Brown (1995) indicated meaningful differences in r_{pb} between applicants from urban versus rural areas. Addition of this configural relationship (i.e., moderation of r_{pb} by urban versus rural applicant point of origin) incrementally increased criterion validity.

Telenson, Alexander, and Barrett (1983) proposed the rare response method in which item responses are weighted based on how few individuals select a given response option. This method differs from others discussed in this section in that it does not weigh response options by their criterion relationships. Instead, this method derives weights using all candidates who completed the biodata inventory regardless of whether subsequent criterion measures were available. Infrequently endorsed responses are assigned larger weights than frequently chosen responses because infrequent responses were presumed to convey more information regarding individual standing on a particular characteristic (Telenson et al., 1983). Webb (1960) and Malloy (1955) both found rare response procedures added more to prediction than a criterion-based empirical key, though Webb (1960) suggested the predictive power of deviant response keys may not hold up under cross-validation. This procedure is rarely used in biodata practice, though Smith and McDaniel (1998) recently used it to infer which subjects were giving nonsense answers.

Recently, Brown and McManus (1994) investigated an application of classification-regression tree procedures in key development. The statistical technique has its roots in Automatic Interaction Detection (AID) procedures developed by Morgan and Sonquist (1963) and is described in detail by Breiman, Friedman, Olshen, and Stone (1984). The readers are referred to Breiman et al. (1984) for a thorough discussion of this technique, though for brevity we provide a simplified description that permits comparison to other keying methods. Specifically, regression tree applications are sensitive to the existence of linear, nonlinear, and configural (interactive) relationships between biodata items/response options and the criterion. All other empirical keying procedures at best only capture linear and nonlinear predictor-criterion relationships. Brown and McManus (1994) reported a comparison of this procedure to more traditional empirical keying procedures in a large sample of applicants for life insurance sales positions. Findings failed to demonstrate a significant or meaningful difference in estimates of population criterion validity. Again, recall that any such comparisons cannot be generalized beyond the specific biodata items and criterion construct domain used in that comparison. It remains to be seen whether procedures using configural predictor combinations incrementally increase prediction or understanding at a level that justifies their complexity and cost.

Empirical Key Comparisons.

While inferences are necessarily limited to the biodata items, population, and criterion domains examined, a handful of studies have directly compared alternate keying procedures. Malone (1977) compared two regression methods and the vertical percent difference keying method, finding significantly higher cross-validities for the regression methods in predicting absenteeism and no difference between methods when predicting tenure. Aamodt and Pierce (1987) found the vertical percent method superior to the rare response method, though this study failed to report cross-validities. However, in analyses of multiple biodata selection applications in the U.S. Navy, Devlin et al. (1992) found different levels of shrinkage across empirical procedures. Hence, no firm conclusion can be drawn from Aamodt and Pierce?s results. Telenson et al. (1983) compared vertical percent, horizontal percent, and rare response methods, finding only one rare response cross-validity to be significant, and none of the horizontal percent or vertical percent methods yielding significant results. Telenson et al. (1983) remains one of the few studies reporting evidence of criterion validity using the rare response method.

Devlin et al. (1992) compared nine empirical keying procedures in terms of shrinkage and cross-validity. The mean criterion method yielded the highest validities in keying samples but also the greatest amount of shrinkage. Devlin et al. found vertical percent methods to generally yield the highest cross-validities. The lowest cross-validities were obtained from the rare response and mean criterion techniques. Devlin et al. (1992) suggested non-criterion-based procedures, such as the rare response method, are unlikely to be useful.

In sum, methods that directly estimate strength of relationships between response options and criteria consistently yield higher criterion validities in cross validation samples. Mitchell and Klimoski (1982) noted that theory-based (Note 3) keys will necessarily yield lower expected criterion validities unless the theory or model fully specifies all causal influences. When performance prediction is the primary goal, the vertical percent difference method seems to capture latent linear and nonlinear predictor-criterion relationships that maximize criterion validity (Devlin et al., 1992). However, advanced configural keying procedures involving classification/regression trees and neural networking hold promise for incremental increases in predictive power (Brown, 1994).

Implications for theory have rarely been explored in the application of empirical keying procedures. In one example, Russell and Domm (1990) reported items loading on a "negative life event" factor tended to have middle-range response options (i.e., 2 through 4 on a 1 to 5 point Likert scale) enter a vertical percent difference-based empirical key. They noted this was consistent with findings from the goal setting literature, i.e., that moderately difficult goals tend to yield higher performance than too easy or difficult goals. Unfortunately, empirical keying remains an under-traveled bridge between the classic prediction versus explanation issues. A strong theory linking life experiences to future job performance should result from a combination of 1) grounded theory building (Glaser & Strauss, 1967) from inspection of empirical key derived weights, and 2) examination of theory-based predictions regarding which response options are likely to enter empirical keys are needed. We now turn to a brief review of theories and models put forth to explain biodata criterion validities before presenting our extension.

BIODATA THEORY

Whereas labels such as "dustbowl empiricism" and "atheoretical" have followed biodata since its inception, meaningful efforts have been mounted to develop explanatory models. Our extension is best understood after tracing the development of these efforts from their roots in the "consistency principle" through more recent nomological propositions.

The Consistency Principle Revisited

The consistency principle has been routinely called upon as an explanation of biodata, work sample, and assessment center criterion validities (Klimoski & Brickner, 1987; Russell & Domm, 1995; Wernimont & Campbell, 1968). Owens (1976) stated ?one of our most basic measurement axioms holds that the best predictor of what a man will do in the future is what he has done in the past? (p. 625). It can be expressed as a simple time series forecast:

$$Y_t = f(Y_{t-1})$$

where human performance (Y) at time t is a function of performance at time t-1.

The consistency principle provides an elegant and quick answer to the question "Why do biodata selection systems work?" Unfortunately, this is where the discussion often ends even though the consistency principle does not explain anything. A closer look at exactly what the consistency principle implies should make this more clear.

Specifically, the state of economic and psychometric assessment indicates we cannot measure human performance (Y) without error, hence:

$$\hat{Y}_t = f(Y_{t-1} + e_1)\text{, and}$$

$$\hat{Y}_{t+1} = f[(Y_{t-1} + e_{t-1}) + e_t].$$

Importantly, measurement error cumulates as the time between performance measures increases; that is, Y_t predictions about Y_{t+50} are much less accurate than predictions about $Y_{t+}1$. Both predictions are attenuated by the sum $\sum_{t=1}^{k} e_t$, though k = 1 for the former prediction, while k = 50 for the latter. As measurement error occurs at each point in time, the further apart in time two performance measures occur the larger the cumulative error component and the smaller the correlation between them. This is a very common phenomena observed in the operations research and production planning literatures, yielding what has become known as a “simplex” matrix of correlations among performance measures taken at adjacent periods over time (Hulin, Henry, & Noon, 1990). In this matrix, time lag between performance measures increases as one moves off of the main diagonal and correlations systematically decrease.

Importantly, the simplex matrix suggests past performance will, as error accumulates over time, become an increasingly poorer predictor of future performance. In a recent example of this phenomena, Deadrick and colleagues (Deadrick & Madigan, 1990; Deadrick, Bennett, & Russell, 1997) reported a decrease in correlations between weekly performance measures among N = 82 sewing machine operators from r = .92 with a one week time lag to r = .55 and r =

.25 with 26 and 51 week time lags, respectively. The simplex pattern has been a major issue since the early 1960s in debates over how to interpret changing criterion validities with criterion measures taken over long periods of time (Hulin, 1962).

Hence, the simplex phenomena would seem to be at odds with predictions made by the consistency principle. How can 1) performance or behavior in the distant past as captured by biodata responses accurately predict future performance or behavior <u>and</u> 2) performance measures obtained on the exact same job exhibit decreasing correlations as the time lag between measures increases? How can both the simplex pattern of successively less correlated performance measures and biodata criterion-related validity co-occur?

The answer would appear to be in the operations research distinction between forecasting and time series. Time series equations assume no causal link between the independent variable (e.g., prior performance measures) and dependent variable (e.g., future performance). Instead, time series predictions assume whatever (generally unknown) causal processes in effect at time t-1 will tend to operate the same way at time t. No insight into latent causal processes is needed to explain predictive power for equations 1 or 2 above, one only need assume causal processes change slowly over time. In the absence of strong theory, forecasting involves harvesting causal “candidate” variables for use in powerful statistical optimization algorithms in hopes of capturing an isomorphic (exact) or paramorphic (parallel but not identical to) model of latent causal influences. For example, for an Iowa pig farmer, a time series predicts change in pork belly prices from trends in past pork belly prices, while forecasting involves predicting pork belly prices from the cost of feed corn, subsoil moisture levels (which causally influence feed corn crop yields and subsequent feed supply and price), natural disasters in other large pork or pork substitute (meat) producing markets (e.g., mad cow disease), and so forth.

Those familiar with biodata selection systems should immediately see the parallel; biodata generally does not literally involve the prediction of future performance (or behavior) from past measures of the same. Instead, biodata is hypothesized to capture causal events or correlates of causal events that influence job candidates' future behavior (Note 4). The most well developed explanation of biodata criterion-related validities, i.e., Mumford, Stokes, and Owens' (1990) ecology model, describes causally recursive sequences of life events as a learning process. Specifically, Mumford et al. (1990) described a causally related sequence of "resources" (e.g., knowledge, skills, and abilities) and "affordances" (e.g., values, expectancies, or what Kanfer & Ackerman, 1989, labeled "distal motivation") whereby individuals act in a way consistent with their environment, resources, and affordances at t_1. Individual actions at t_1 yield a new environment (including performance outcomes resulting from t_1 activities), and (possibly) a new profile of knowledge, skills, and abilities (resources) and modified preferences or desires for outcomes resulting from application of knowledge, skills, and abilities (affordances) at t_2.

The inability of a strict interpretation of the consistency principle to explain biodata criterion validity becomes clear when one considers the prediction problem faced by Russell, Mattson, Devlin, and Atwater (1990). Russell et al. faced predicting performance of Naval officers five to six years after graduation from the U.S. Naval Academy for candidates who are generally still in high school. High school seniors tend not to have prior performance records as Naval officers, so Russell et al. developed measures of past behaviors and performance in domains in which high school seniors had been exposed. As described by Binning and Barrett (1989), a strong theory of performance prediction must explain latent nomological relationships between prior life events and future job performance. The latent nomological network clearly must relate more than prior performance to future performance, given no prior performance as Naval officers existed. Russell and Domm (1990) presented the model portrayed in Figure 1 as a way of conceptualizing what such a nomological net must minimally look like. We now review efforts aimed at developing such a theory and present an extension.

Insert Figure 1 about here

The Ecology Model

Efforts have been mounted to develop theories to explain biodata criterion validity that reach beyond the consistency principle (Mael, 1991; Mumford & Owens, 1987; Mumford & Stokes, 1992; Mumford, Stokes, & Owens, 1990; Owens 1968, 1971, 1976; Owens & Schoenfeldt, 1979). One such theoretical rationale for biodata is the ecology model (Mumford et al., 1990). The ecology model acknowledges individuals have their own unique hereditary characteristics and exposures to environmental circumstances that determine initial individual differences, focusing specifically on how individual difference characteristics shape the subsequent choices individuals make. The ecology model grew from Owens' developmental-integrative (D-I) model (Owens, 1968, 1971, 1976; Owens & Schoenfeldt, 1979), which initially proposed biodata items need to capture prior behaviors and experiences affecting personal development on individual difference characteristics (e.g., knowledge, skills, and abilities). These individual differences were hypothesized to subsequently affect a person?s performance on organizational criteria of interest.

Mumford, Stokes, and Owens (Mumford & Owens, 1987; Mumford & Stokes, 1992; Mumford, et al., 1990) continued to refine the ecology model framework. The ecology model not only considered individual differences as predictors of future performance, but also considered the processes that motivate and influence an individual's choices. Specifically, the model suggested people select themselves into situations based on the value of expected outcomes as well as pre-existing individual difference characteristics. Each choice requires adaptation to new situations and represents a developmental experience. The model represents an iterative process of choice, development, and adaptation. People are constantly faced with making choices and over time will tend to develop characteristic patterns of choices and behaviors.

The ecology model posits that ?life events indicating successful engagement in activities requiring the application of KSAOs similar to those required on-the-job might prove to be useful predictors? (Mumford & Stokes, 1992, p. 81), as well as those events that play a role in developing knowledge, skills, abilities, and other personal characteristics (KSAOs). Note, the former capture consistency principle or time series-based performance prediction, while the latter tap latent causal processes captured in forecasting applications.

In an early attempt to understand dimensions underlying the ecology model, Nickels (1990) identified a framework of characteristics and individual differences posited to influence performance later in life. Nickels reviewed over one hundred and fifty citations of individual differences and known predictive relationships between past behavior/experience and later performance, yielding a preliminary list of 500 possible dimensions. The 500 dimensions were subjected to a series of reviews by subject matter experts to obtain a manageable and interpretable number of dimensions. Dimensions were excluded from further investigation based on a consensus decisions that a dimension...

> a) demonstrated an obvious content overlap with another dimension (e.g., gregariousness and sociability); b) could not feasibly be rated given the information provided by background data items (e.g., attractiveness); was inappropriate with respect to the population (e.g., paranoia in a normal population); or d) seemed unlikely to influence the life history of individuals in adolescence and young adulthood (Nickels, 1990, pp. 28-29).

This process resulted in the dimensions being reduced in number from 500 to 44. Five general dimensional categories emerged from the reduced set of 44. Three captured general categories of individual differences posited to influence subsequent performance: personality resources, interpersonal (social) resources, and intellectual resources. Two other categories covered motivation and beliefs/attitudes, labeled choice and filter processes, respectively.

The Nickels (1990) study was one of the first attempts to operationalize the ecology model. Mumford, Stokes, and Owens (1990), the primary ecology model architects, subsequently elaborated this framework as described in Figure 2, changing some of the labels though not the substance of Nickels dimensions (Mumford & Stokes, 1992). The ecology model suggested individual difference constructs ?facilitate the attainment of desired outcomes while conditioning future situational choice by increasing the likelihood of reward in certain kinds of situations? (Mumford & Stokes, 1992, p. 81). The first three categories (i.e., personality, social, and intellectual resources) are personal characteristics posited to influence future behavior and decisions. The remaining two categories are motivational variables that might affect situational selection and resource application (i.e., choice and filter processes; Mumford & Stokes, 1992; Nickels, 1990). These five constructs are discussed below.

Insert Figure 2 about here

Personality Resources

Nickels suggested this category represented ?stylistic or emotional attributes thought to impact effective environmental interactions? (1990, p. 29) such as adaptability, emotional stability, and persistence. These resources closely resemble the ?Big Five? personality constructs (cf., Barrick & Mount, 1991). The five factors include: extraversion (e.g., sociable, assertive, ambitious), emotional stability (e.g., secure, anxious, well-adjusted), conscientiousness (e.g., dependable, efficient, achievement oriented), and openness to experience (e.g., cultured, curious, broad-minded; Barrick & Mount, 1991; Mount, 1997). Construct validity evidence supportive of these five personality factors has been found consistently across longitudinal studies, raters, personality inventories, and protected subgroups (Digman, 1990; Mount, 1997). Results reported by Caspi and associates (e.g., Newman, Caspi, Moffitt, & Silva, 1997) have demonstrated differences in personality temperament as early as age three predict meaningful differences in interpersonal functioning in work contexts at age 21.

Social Resources

Nickels (1990) posited that social resources influence effectiveness of interpersonal relations and therefore play a role in situation selection and subsequent behavior/performance. Some examples of constructs in this category include self-monitoring, dominance, and empathy. Mael?s integration of social identity theory with the ecology model also speaks to the influence of group membership on one?s own personal identity (Mael & Ashforth, 1989; Mael & Hirsch, 1992; Mael, 1991). It could be argued that the more group memberships held, the greater one?s interpersonal adeptness. Large numbers of group memberships may suggest individuals are high in self monitoring and behavioral flexibility, i.e., able to adjust behavior to match expectations generated from a variety of groups.

Intellectual Resources

Intellectual resources represent attributes that enable one to assimilate and retain knowledge affecting one?s ability to make choices and perform efficiently and effectively. An example of an underlying construct that biodata items might capture is general cognitive ability, or ?*g*.? The ecology model explicitly includes *g* in intellectual resources, but also speaks to the iterative, bi-directional relationship of *g* to life events and performance over time. Many extant biodata items seem to tap *g* (Dean, Russell, & Broach, 1998). For example, questions on Owens' (1971) Biographical Questionnaire asked individuals about academic achievement, academic attitude, and intellectualism. Biodata instruments may also tap *g* by asking about past experiences requiring general cognitive ability or associated with its development. In contrast, paper and pencil cognitive ability tests infer *g* from the number of correct answers selected by an applicant to questions tapping various knowledge content domains. Answers deemed "correct" reflect some universal agreement as to what is truth within each item content domain (e.g., ?4" is the correct answer to the arithmetic knowledge question, ?What is 2 plus 2?).??? The "correct" answer to a *g*-loaded biodata item is any response option that 1) is related to the latent life history construct of interest (i.e., exhibits construct validity relative to *g*), and 2) contributes to explanation of subsequent performance differences.

<u>Choice Processes</u>

The choice processes domain represents the ?differential motivational influences with respect to individual differences in performance? (Nickels, 1990, p. 43). Example components of the choice processes domain include goal orientation, personal performance standards, and desirability of the reward (e.g., ?valence? in expectancy theory terminology). Research on performance prediction suggests individuals must have motivation and ability to generate performance outcomes (Campbell, Dunnette, Lawler, & Weick, 1970). Gottfredson (1997) suggested biodata items capture motivational components of task performance better than paper and pencil mental ability tests, speculating that cognitive ability measures may best estimate what applicants ?can do,? but measures not specifically targeting cognitive ability (i.e., ?non-cognitive? measures) such as biodata may best estimate what applicants ?will do.? Some investigators suggested high criterion-related validities may be partially due to biodata?s ability to tap ability constructs, motivational constructs, and past examples of the interaction between the two domains (Mael, 1991; Mitchell, 1996).

Filter Processes

Nickels (1990) suggested this category represents values, beliefs, and attitudes which may influence self-perception and, consequently, decisions an individual makes. Constructs included in this category are self-esteem, self-efficacy, and locus of control.

Model Deficiencies

After reviewing biodata theory development, one might ask "What latent causal process(es) explain biodata predictive power?" The D-I and ecology models hypothesize coarse groupings of construct domains deemed relevant, on the basis of theory and research in non-biodata venues, to the development of important individual differences. This coarseness becomes clear when one attempts to generate items unique to each domain. For example, while "dominance" is clearly an individual difference characteristic associated with social interaction (and hence part of the "social resources" category), it is equally clearly a personality characteristic that might be found in the "personality resources" category. The possibility that a single life event can be causally linked to more than one outcome (i.e., what von Eye & Brandtstädter, 1998, referred to as a "fork" dependency) makes traditional convergent and discriminant assessments of biodata construct validity difficult.

Unfortunately, the models portrayed in Figures 1 and 2 also exhibit a shortcoming common to many path model conceptualizations; that is, the absence of any explanation of latent processes underlying the paths themselves (cf. Russell & Van Sell, 1986). At least one explanation of processes linking prior life experiences reflected in the ecology model involves learning or knowledge acquisition. Prior life experiences that lend themselves to learning <u>and</u> biodata items demonstrating criterion validity could be defined as situations where 1) people create meaning out of patterns of information where no meaning and/or different meanings had been previously applied, and 2) newly created meaning causally influences subsequent knowledge acquisition, behaviors, and performance. To be sure, any learning could be ecologically valid (i.e., subjects learned the correct lessons from experience) or demonstrate any one of a number of common errors. Hence, the forecasting model becomes:

$$\hat{Y}_t = f(X_{1_t}, X_{2_t}, X_{3_t}, \ldots, X_{k_t}, X_{k+1_{t-1}}, \ldots) + e$$

where the various Xs are causal influences on performance Y. Note, Y occurs at time t, while life event X_k can be chronologically coincident (time t) or earlier (t-1). Integrating this model with a consistency principle-based time series forecast, performance Y_t is generated from past causal influences (captured by Y_{t-1}), any systematic changes in those causal influences (e.g., learning driven by X_t and X_{t-1} variables), and random error.

Models of how and when people assign meaning to life experiences are needed to provide guidance for biodata item development. Such a model would likely describe multiple alternative sequences of experiences that may or may not yield similar meaning across individuals. If experience sequences and attendant patterns of meaning are consistent within or across labor markets, industries, jobs, and so forth, a model of performance prediction will result linking profiles of "lessons learned" predictor constructs to subsequent job performance. In turn, systematic examination of "lesson acquisition event" domains would address latent processes within the ecology model and serve to guide both biodata item development and keying.

TOWARD A MODEL OF LIFE EXPERIENCE LEARNING

The training and development literature suggests at least three characteristics of life experiences influence learning: learning aids in the environment, time needed to reflect on experience, and failures (Goldstein, 1986). Learning aids might include sources of passive stimulation (e.g., number of magazines subscribed to by parents) and active intervention (e.g., task versus consideration-oriented leadership style of an adult role model). Quantity and quality of reflection time has been shown to influence learning in educational settings, where overstimulation decreases reflection time in preventing pattern recognition and meaning making. The literature suggests individual differences in stimulus seeking behavior are likely moderators of reflection time influence (Cacioppo, Petty, Feinstein, & Jarvis, 1996). Environmental support for learning and time needed to reflect on experience would seem to have fairly straightforward implications for theory-based biodata item development and application (see Russell, 1986, for a description of how biodata items were generated from life history essays to reflect environmental support).

Negative Life Experiences

What we find to be most interesting, however, are implications drawn from the third facet: failures, or negative life experiences. Negative life events are imbedded in theory and empirical findings in a wide array of literatures, though there appears to be few bridges between these concurrent research streams. We present a sampling of these research streams as well as literature bearing on a relevant individual difference construct. Implications for extending our understanding of processes underlying the ecology model are drawn.

Before doing so, it seems prudent to ask whether any evidence exists suggesting negative life event-oriented biodata items exhibit criterion validity. Two biodata criterion-related validity studies suggest negative life events are both recalled particularly well and demonstrate strong criterion-related validity when compiled in a biodata inventory. First, Russell et al.'s (1990) factor analysis of biodata items created for selection of midshipmen into the U.S. Naval Academy yielded a dominant initial factor containing unpleasant life events. Example items included "How often have you worked really hard to achieve something and still came up short?", "How often have you wondered whether people like you because of something you've done (e.g., sports, good grades, cheerleading, etc.) rather than for who you are?", and "How often has someone fooled you so that afterwards you felt stupid or naive?" Russell and Domm (1990) found a similar dominant factor in a separate study using a different labor pool, target job, and biodata items. Negative life event items generated from life history essays written by retail store managers included "How often have you found yourself in positions where you were forced to make frequent decisions even though you felt you needed more time or information?", "How often have you followed a friend's recommendation only to find out things did not turn out well?", and "How often have you lost sales or had to order items for a customer because the items were not in inventory?"

Results obtained from qualitative research methods also point to negative life events. Specifically, Lindsey, Homes, and McCall (1987) interviewed over 190 highly successful top-level executives to assess "key events" in executive lives. A frequently cited event across a heterogeneous sample involved having worked for a very demanding, almost abusive boss early in one's career. They also found 17.4% of all "key developmental events" reported by successful top-level managers were severe enough to be deemed "hardships" (p. 87). Cattell (1989) reported longitudinal evidence suggesting a substantial number of people exposed to severely abusive early life experiences survive and, indeed, thrive later in life (Note 5). Holman and Silver (1998) recently reported evidence suggesting individual differences in temporal orientation moderate individuals' abilities to cope with traumatic events.

Perhaps the most significant body of findings and theory suggesting the importance of negative life events is found in Lewin's (1951) field theory. The first step of the classic sequence of unfreezing, change, and refreezing generally involves receipt of some negative piece of evaluative information. Descriptions of the unfreezing stage in field theory-based interventions such as t-groups (Lewin, 1951) suggest this life event can be particularly unpleasant.

Finally, recent research on performance prediction has re-examined the influence of work experience on job performance (Quinones, Ford, & Teachout, 1995; Tesluk & Jacobs, 1998). While tenure and seniority are the typical time-based surrogate measures of work experience, recent research indicates the concept is much more psychologically complex. Tesluk and Jacobs (1998) classified work experiences along a density continuum, with high density experiences providing greater developmental impact than low density experiences. Personal growth from work experience is thus related to the qualitative nature of the life event, providing a basis to differentiate "ten years of job experience from one year of job experience repeated ten times." We would argue that negative life events in work and non-work domains constitute some of the most developmentally "dense" experiences encountered over a lifetime. Research conducted in non-work arenas (e.g., Suedfeld & Bluck, 1993) has begun to also examine effects of stressful life events on integrative complexity. Integrative complexity is very similar to developmental density, defined as "a state-dependent, cognitive style variable derived from conceptual complexity and system theories . . . as an . . . ability to perceive and think about multiple dimensions or perspectives of a stimulus . . . and the ability to recognize connections between differentiated characteristics" (Pennell, 1996, p. 777).

Potential Causal Mechanisms

Hence, Lewin's (1951) model of individual change, recent research on work experience-job performance relationships, results reported from biodata criterion-related validity studies, and evidence from more qualitative, ethnographic approaches all suggest negative live events constitute key markers in adult development. While suggestive, these results do not explain why negative events are apparently recalled more frequently (i.e., yielding the dominant factor in each study), or the process by which these experiences come to be related to subsequent performance outcomes. A literature on different attributional styles and attentional resources allocated in the face of failure or disappointment provides an initial theoretical foundation. Taylor (1991) reviewed the literature on differences in attributional style as a function of failure versus success. She noted that "diverse literatures in psychology provide evidence that, other things being equal, negative events appear to elicit more physiological, affective, cognitive, and behavioral activity and prompt more cognitive analysis than neutral or positive events" (Taylor, 1991, p. 67).

Results suggest negatively valent events initially lead to mobilization of personal resources (e.g., physiological arousal, affect, attention, and differential weighing of negative cues in decision making). Peeters and Czapinski (1990) found negative events elicited more frequent and complex causal attributional activity than positive events, while others have shown negative events are considered longer (e.g., Abele, 1985) and elicit more extreme attributions (e.g., Birnbaum, 1972). Models of affiliation (Schachter, 1959) and social support (House, 1981) also suggest negative or threatening events cause people to seek companionship, support, and assistance from others. Consistent with criterion-related validities obtained for biodata items tapping negative life events (Russell et al., 1990, Russell & Domm, 1990), Brunstein and Gollwitzer (1996) reported experimental results suggesting task failure caused higher levels of subsequent task performance when the task was relevant to subjects' self definitions (i.e., self-defined goals).

Schulz and Heckhausen (1996) proposed several general principles regulating human development across the life span. One principle pertains to compensating for and coping with failure. Life presents a myriad of negative events which frustrate goal attainment and foster negative self-perceptions. Schulz and Heckhausen postulated the capacity to compensate for frustration is critical, for without it we could not persist when faced with adversity. At one level frustration can foster commitment to a goal. However, at another level failure experiences can undermine self-esteem and individuals' self-ascribed competencies. Thus, negative life events can increase our resolve for goal attainment, yet not compensated for can undermine our self-concept.

Taylor, Pham, Rivkin, and Armor (1998) recently described a program of research focusing on mental stimulation as a core construct driving 1) resolution of stressors (i.e., negative life events) and/or development of goals or visions of desirable future events, 2) anticipation and management of emotions, and 3) initiation and maintenance of problem-solving activities. Taylor et al.'s (1998) efforts target identification of effective and ineffective sources of mental stimulation in relation to the learning/self-regulation process. Initial findings suggest mental stimulation brought about by either process or outcome "mental simulation" help individuals manage affective reactions and subsequent problem solving activities (Rivkin & Taylor, in press) associated with negative life events.

As noted above, negative life events are associated with physiological arousal, affect, attention, differential weighing of negative cues in decision making, more frequent and complex causal attributional activity, are considered longer, elicit more extreme attributions, and cause people to seek companionship, support, and assistance from others. All of these responses are expected to help individuals take action to end or attenuate negative life events, and learn how to avoid similar negative events in the future. Snell (1988) described how exposure to negative events at work socialized managers to be "graduates of the school of hard knocks" (p. 27). The hard knocks consisted of situations and events in which the managers made embarrassing mistakes, or felt overextended in role responsibilities. The essence of the hard knocks was to produce distress, caused by having encountered impasses, having suffered defeats or injustices, and having come under attack.

The effect of these experiences was to teach the managers to be more aware of potential trouble, priming oneself for challenges and anticipating areas of doubt. Specific reactions acquired by the learner-managers included anticipating likely problems, preparing for unfavorable circumstances, obtaining early readings of differences of opinions, and being counseled by a trusted colleague in advance of predicaments. These learning practices can be emotionally unsettling, but they are seldom distressing. The explanatory basis of their effectiveness is that the discomfort they produce is felt currently in mild form because it is expected, rather than later, in a more severe form, when it is not expected.

There would seem to be times when people are particularly receptive to learning from negative life events. Rovee-Collier (1995) proposed the concept of - time-windows in cognitive development. A time-window is a critical period where information about a current event is integrated with previously acquired information. However, if the same information is encountered outside of the time-window, it will not be integrated. Time-windows are not restricted to a particular age or stage of development. Nonetheless, they are open for a limited duration before closing. Discrete events that occur outside of a time-window are treated as unique and thus are not assimilated into the reservoir of collective memory.

Rovee-Collier (1995) speculated that time-windows may be the cornerstone of individual differences in cognitive domains involving integration of successive experiences. She asserted that to the degree to which personal experiences of same-age individuals differ from moment to moment, so will their time-windows, what they remember from those time-windows, and whether the new information will be integrated with existing information in the future. Negative life events that occur when time-windows are closed may be more likely to produce intensified distress upon re-exposure. The concept of time-windows may provide an explanation for the adage that "the most painful lessons in life are often those that we once learned but forgot."

Differential Recall of Negative Life Events

Curiously, Baddeley (1982), Ehrlichman and Halpern (1988), Linton (1982, 1986), and others found that people tend to remember positive life events more frequently than negative events. Further, Taylor and Brown (1988) reviewed literature indicating people actively attempt to re-interpret negative life events as positive life events. These results are inconsistent with those reported by Russell et al. (1990), Russell and Domm (1990), and Lindsey et al. (1987). However, none of the autobiographical memory research has examined targeted autobiographical recall. To generate initial biodata items, Russell and Domm (1990) asked incumbents to describe prior life events which were related to performance on each dimension of target job performance. Two of the life history essays used by Russell et al. (1990) targeted task performance (i.e., individual and group accomplishments), with follow-up questions that paralleled aspects of the performance appraisal criterion applied to Naval Officer performance. Finally, Lindsey et al. (1987) asked subjects to recall events viewed as central to their personal development as managers. Hence, while general autobiographical memory seems to be biased toward positive life events (Burger, 1986; Isen, 1984), quantitative and qualitative evidence (Lindsey et al., 1987; Snell, 1988) suggests negative events within the confines of career-relevant experience are vividly recalled.

One interpretation of these seemingly conflicting findings involves the public versus private nature of autobiographical recall. Specifically, the act of asking individuals for rich, narrative descriptions of prior life events that participants know will be read by others in studies of autobiographical memory may evoke self-serving attributional biases. In fact, subjects' “private” recollections, attributions, and efforts to learn and grow from prior life events may be much less subject to such biases (Csikszentmihalyi, Rathunude, & Walen, 1993). Typical biodata inventories do not request rich, detailed descriptions of prior life events, and hence may be less likely to be influenced by such biases. Lavallee and Campbell (1995) reported findings supportive of this interpretation, finding daily goal-related negative events elicited higher levels of self-focused attention, self-concept confusion, and subsequent self-regulatory responses. Regardless, Kluger, Reilly, and Russell (1991) found the common vertical percent difference method of biodata key development effectively eliminated self-serving biases induced by motivation to obtain a job offer from biodata scores (though not necessarily from biodata responses).

Individual Differences in Moxie

It seems likely that virtually every construct valid individual difference measure examined by behavioral scientists is somehow related to negative life events. We focus our discussion on one we feel is particularly relevant, "moxie." Moxie is an old term from the 1920s and 30s meaning courage or nerve. Authors in the personality literature have labeled it "ego-resiliency" (Block & Block, 1980; Block, 1993), or an openness to experience combined with ability to adapt and balance accommodation and assimilation. Ego-resilient children are "well equipped to assimilate experience -- to understand and interpret events -- with existing structures of thought when these structures are appropriate" (Hart, Keller, Edelstein, & Hoffman, 1998, p. 1278). Hart et al. (1998) reported evidence that ego-resilience at age seven predicts social-cognitive development at ages nine, 12, 15, and 19 independent of general cognitive ability, attention, or social participation. A similar individual difference is labeled "hardiness" in the social psychology literature and has been shown to affect relationships between negative life events and symptoms of illness (Kobasa, Maddi, & Kahn, 1982).

Why dwell on ego-resilience or moxie? First, we appear to be in an era of motivation theory focused on how behavior is sustained over long periods of time (e.g., cybernetic control theory, Bozeman & Kacmar, 1997). Motivation theory seems less concerned with one-time choices or predicting constant high-levels of effort devoted to a single task. Focus has shifted toward persistence of behavior over prolonged time periods. Note, unless an individual has been uniformly successful at all endeavors, persistence of behavior over prolonged periods of time must occur with some degree of failure or negative life events. Moxie would appear to be a critical individual difference in predicting persistence in the face of inevitable negative feedback received as one learns.

Second, if there is any constant in life it seems to be adversity. Adversity may be operationalized similarly for an entire generation (e.g., survivors of the Great Depression, the "global economic change" of the 1990s, etc.), or it may be uniquely private. Some individuals find ways to rise above the adversity; others enter prolonged periods of "victimhood;" while still others become relatively immobilized in the web of consequences. What varies is how the adversity manifests itself, when it occurs in an individual's life span, and how individuals respond to it.

Last, given our prior focus on negative life events, any model that does not address the presence of relevant individual difference variables and the possibility of person-situation interaction is deficient. We cannot think of an individual difference characteristic more relevant in the presence of negative life events than "resilience to negative life events," or moxie. Its potential role as a moderator and/or mediator in negative life event-job performance relationships needs to be examined.

Given these reasons for attending to moxie, where does it come from and how can we measure it? What life experiences promote/inhibit development of moxie? We suggest a number of research directions in the conclusion below.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

We have reviewed the literature on one of the most powerful performance prediction technologies currently available, showing why traditional reliance on the consistency principle must give way to more theory-rich models of biodata. We summarized characteristics of this prediction technology that make it unique (e.g., empirical keying) and troublesome (e.g., absence of strong theory). Research focusing on the nature of latent theory-based relationships in the ecology model is needed to extend our understanding of why biodata exhibits criterion-related validity. We argued that, based on learning theory predictions, emphasis on examination of negative life events is likely to shed insight into these processes. Finally, a critical individual difference, moxie, was hypothesized to moderate how negative life experiences influence the five ecology model construct domains and subsequent job performance.

We feel this evidence overwhelmingly suggests that learning how to respond effectively to negative life events constitutes key, often defining, developmental life experiences. Examination of the capacity to learn from "hard knocks" (i.e., moxie) may be a more valuable line of inquiry than examination of particular sequencing of these life events. As noted in our discussion of the consistency principle, classic biodata emphasis on "behavioral consistency" adds little to our understanding of human development. It is not necessarily the consistency of the behavior we seek to understand, but the underlying psychological constructs which ultimately produce the behavior in question. Perhaps it is useful to think of negative life events as "high density conditions" which fashion our proclivities to either rise above them through resilience and moxie or be diminished by their deleterious effects.

We believe an analysis of life events and our responses to them is, in essence, the chronicle of our individual existence. From a research perspective, such an analysis would provide the cognitive and developmental explanation for behavior we index through biodata inventories. Thus, positive and negative life experiences can be construed as providing an on-going series of arenas through which our identities become defined. Our position is best captured by the following quote:

> Every experience in life, everything which we have come in contact in life, is a chisel which has been cutting away at our life statue, molding, modifying, shaping it. We are part of all we have met. Everything we have seen, heard, felt, or thought has had its hand in molding us, shaping us.
>
> Orison Swett Marden

We conclude by suggesting lines of future research that might shed light on the nature of negative life events, moxie, and construct domains tapped by the ecology model.

Basic Construct Development

First, it would seem appropriate to create a taxonomy of negative life events. Such a taxonomy would differ meaningfully from biodata item taxonomies described above in that it would target a latent construct domain, specifically, aspects of negative life experiences. Some research in this area exists. The Social Readjustment Scale (Holmes & Rahe, 1967) contains a list of 43 stress-inducing life events, including death of a spouse, divorce, being fired from work, financial difficulties, and personal injuries and illnesses. Certainly other life events could be identified through open-ended questions or life history essays (cf. Russell, 1986). Lindsey et al.'s (1987) interviews with successful top-level executives suggested working for a nasty, almost abusive boss early in one's career constituted one such event. Regardless, we suspect residents of Bosnia would come up with additional adverse life experiences. Further, the term "downsizing" as an adverse life event had not been coined when the Social Readjustment Scale was developed. At the risk of sounding like cynics, we suspect a list of adverse life events that causally influence human development will be none too short.

Second, a taxonomy of human responses to adversity is needed. A list of such responses would target latent cognitive and behavioral reactions such as "freeze, flight, and fight" responses found in research on fear. While these categories are probably deficient in their sampling of potential human response to adversity, we would hope such an effort would yield identifiable profiles of response patterns. At least one such profile is likely to be dominated by high moxie and resilience. Taylor et al.'s (1998) approaches to "mental simulation" may also prove useful.

Third, a taxonomy of affective responses to negative life experiences is needed. Emotions associated with loss of a loved one would conceivably be different than emotion associated with the loss of a job. Both may involve grief, though they may differ in terms of remorse versus anger. How remorse or anger becomes channeled into future behavior could be the prime, if not primal, basis of resiliency. Resilience born of anger (or any other emotion) may yield grossly different subsequent behavioral manifestations relative to resiliency born of the will to survive.

Toward an Integrative Model

Simply put, negative life events appear to represent the "what," cognitive and behavioral responses to negative life events represent the "how," and emotional responses represent the "why" behind key development life events. Each of these questions could be conceivably asked in the context of the five construct domains captured by the ecology model. However, we feel the true value of biodata measurement technology is in 1) assessing these three phenomena, and 2) constructing an integrative model showing how reactions to negative life events drive change and growth in the five domains underlying the ecology model. Of course, these suggested lines of research could start with existing biodata instruments, sorting items into negative, neutral, and positive life event categories.

These life event categories could then be examined for differences in attributional styles and attentional resources. Other things being equal, negative life event items are expected to contribute to empirical scoring keys with greater frequency and incrementally increase biodata inventory criterion-related validities by a larger amount than neutral or positive life events. Russell et al. (1990) and Russell and Domm (1990) reported results consistent with this expectation, though replication is needed.

Given the apparent developmental centrality of negative life events, subsequent research efforts must identify the array of behavioral and affective reactions associated with negative life event categories that contribute most to performance prediction. Identification of profiles of negative life events and reactions that are differentially related to performance outcomes will yield strong implications for biodata theory development (i.e., the "what, how, and why" of latent causal processes underlying the ecology model). An extended model of such processes would permit strong tests of theory-based inferences, specifically, that interventions designed to inject an appropriate quantity/quality of adversity into an individual's life during an appropriate time-window <u>and</u> guide reactions to that adversity will yield higher performance than any other combination of adversity and adversity reactions (cf. Taylor et al, 1998). Again, the cynic in us suggests most will never need adversity injected into their lives; it will come in one form or another to everyone. We can influence reactions to adversity, though current biodata theory provides little guidance. Basic biodata research holds promise for providing such guidance, but richer and more refined theory development is needed.

REFERENCES

Aamodt, M. G., & Pierce, W. L., Jr. (1987). Comparison of the rare response and vertical percent methods for scoring the biographical information blank. Educational and Psychological Measurement, 47, 505-511.

Abele, A. (1985). Thinking about thinking: Causal, evaluative, and finalistic cognitions about social situations. European Journal of Social Psychology, 15, 315-322.

Abrahams, N. (1965). The effect of key length and item validity on overall validity, cross-validation shrinkage, and test-retest reliability of interest keys. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Abrahams, N. (1998, June 27). Personal communication.

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

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. Academy of Management Review, 14, 20-39.

Atwater, D. C. (1980). Faking of an empirically keyed biodata questionnaire. Paper presented at the annual meeting of the Western Psychological Association, Honolulu, HI.

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. Academy of Management Review, 14, 496-515.

Baddeley, A. D. (1982). Emotional factors in forgetting. In A. Cope (Ed.), Your memory: A user's guide (pp. 65-73). New York, NY: Macmillan.

Barge, B. N. (1988). Characteristics of biodata items and their relationship to validity. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Barge, B. N., & Hough, L. M. (1988). Utility of biographical data for the prediction of job performance. In L. M. Hough (Ed.), Literature review: Utility of temperament, biodata, and interest assessment for predicting job performance (ARI Research Note 88-020). Alexandria, VA: U. S. Army Research Institute.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1-26.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. Journal of Applied Psychology, 74, 478-494.

Birnbaum, M. H. (1972). Morality judgments: Test of an averaging model with differential weights. Journal of Experimental Psychology, 93, 35-42.

Block, J. (1993). Studying personality the long way. In D.C. Funder, R.D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), Studying lives through time: Personality and development (pp. 9-41). Washington, D.C.: American Psychological Association.

Block, J., & Block, J. H. (1980). The California Child Q-Set. Palo Alto, CA: Consulting Psychologist Press.

Bozeman, D.P. & Kacmar, K. M. (1997). A cybernetic model of impression management processes in organizations. Organizational Behavior and

Human Decision Processes, 69, 9-30.

Brieman, L. Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Belmont, CA: Wadsworth.

Brown, S. H. (1994). Validating biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction. (pp. 199-236). Palo Alto: Consulting Psychologists Press.

Brown, S. H (1995, April). Personal communication.

Brown, S. H., & McManus, M. (1994, April). Comparison of additive and configural biodata keying methods. Presented at the eighth annual meeting of the Society of Industrial and Organizational Psychology, Nashville, TN.

Brunstein, J. C., & Gollwitzer, P. M. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. Journal of Personality and Social Psychology, 70, 395-407.

Burger, J.M. (1986). Temporal effects on attributions: Actor and observer differences. Social Cognition, 4, 377-387

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Disposition differences in cognitive motivation: The life and times of individuals varying in need for cognition. Psychological Bulletin, 119, 197-253.

Campbell, J., Dunnette, M. D., Lawler, E. E., & Weick, K. (1970). Managerial behavior and performance effectiveness. New York: McGraw-Hill.

Cascio, W. F. (1975). Accuracy of verifiable biographical information blank responses. Journal of Applied Psychology, 60, 767-769.

Cattell, H. B. (1989). The 16PF: Personality in depth. Champaign, IL: Institute for Personality and Ability Testing, Inc.

Childs, A., & Klimoski, R. J. (1986). Successfully predicting career success: An application of the biographical inventory. Journal of Applied Psychology, 71, 3-8.

Csikszentmihalyi, M. Rathunude, K., & Walen, S. (1993). Talented teenagers: The roots of success and failure. Cambridge: Cambridge University Press.

Deadrick, D. L., Bennett, N., & Russell, C. J. (1997). Using hierarchical linear models to examine performance trends: Implications for selection utility. Journal of Management, 23, 745-757.

Deadrick, D. L., & Madigan, R.M. (1990). Dynamic criteria revisited: A longitudinal study of performance stability and predictive validity. Personnel Psychology, 43, 717-744.

Dean, M.A., Russell, C.J., & Broach, D. (1998, April). Sample size needed to establish criterion validity for air traffic controller selection: Comparing bootstrap vs. parametric estimation. Presented at the annual meetings of the Academy of Management, San Diego, CA.

Devlin, S. E., Abrahams, N. M., & Edwards, J. E. (1992). Empirical keying of biographical data: Cross-validity as a function of scaling procedure and sample size. Military Psychology, 4, 119-136.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. Annual Review of Psychology, 41, 417-440.

Dunnette, M. D. (1962). Personnel management. Annual Review of Psychology, 13, 285-313.

Dunnette, M. D. (1966). Personnel selection and placement. Belmont, CA: Wadsworth.

Ehrlichman, H. & Halpern, J. N. (1988). Affect and memory: Effects of pleasant and unpleasant odors on retrieval of happy or unhappy memories. Journal of Personality and Social Psychology, 55, 769-779.

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, and Department of Justice. (1978). Uniform guidelines on employee selection procedures. Federal Register, 43(166), 38295-38309.

Fischer, R. A. (1970). Statistical methods for research workers. (14th edition). New York: Hafner Press.

Fleishman, E. A. (1988). Some new frontiers in personnel selection research. Personnel Psychology, 41, 679-701.

Fleishman, E. A., & Quaintance, M. K. (1984). Taxonomies of human performance: The description of human tasks. Orlando, FL: Academic Press.

Gandy, J. A., Outerbridge, A. N., Sharf, J. C., & Dye, D. A. (1989). Development and initial validation of the Individual Achievement Record (IAR). Washington, D.C.: U.S. Office of Personnel Management.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine Publishing Co.

Goldstein, I.L. (1986). Training in organizations: Needs assessment, development, and evaluation. Monterey, CA: Brooks/Cole.

Gottfredson, L. (1997). Racially gerrymandering the content of police tests to satisfy U.S. Justice Department: A case study. (Working paper).

Guion, R. M. (1965). Personnel testing. New York: McGraw-Hill.

Hammer, E. G., & Kleiman, L. A. (1988). Getting to know you. Personnel Administrator, 34, 86-92.

Hart, D. Keller, M., Edelstein, W., & Hofmann, V. (1998). Childhood personality influences on social-cognitive development: A longitudinal study. Journal of Personality and Social Psychology, 74, 1278-1289.

Holman, E. A., & Silver, R. C. (1998). Getting "stuck" in the past: Temporal orientation and coping with trauma. Journal of Personality and Social Psycology, 74, 1146-1163.

Holmes, I. H., & Rahe, R. H. (1967). The social readjustment rating scale. Journal of Psychosomatic Research, 4, 189 - 194.

Hough, L., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. Journal of Applied Psychology, 75, 581-595.

House, J. A. (1981). Work stress and social support. Reading, MA: Addison-Wesley.

Hulin, C. L. (1962). The measurement of executive success. Journal of Applied Psychology, 46, 303-306.

Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in generalizability of predictive relationships. Psychological Bulletin, 107, 328-340.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitude, job knowledge, and job performance. Journal of Vocational Behavior, 29, 340-362.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

Isen, A.M. (1984). Toward understanding the role of affect in cognition. In R.S. Wyer Jr., & T.K. Srull (Eds.), Handbook of social cognition (Vol. 3, pp. 179-236). Hillsdale, NJ: Erlbaum.

Kanfer, R. (1990). Motivation theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (Vol 1, pp. 75-170). Palo Alto, CA: Consulting Psychologist Press.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. Journal of Applied Psychology, 74, 657-690.

Klimoski, R. J., & Brickner, M. (1987). Why do assessment centers work? The puzzle of assessment center validity. Personnel Psychology, 40, 243-260.

Klimoski, R. J., & Strictland, W. (1976). Assessment centers: Valid or merely prescient. Personnel Psychology, 30, 353-363.

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

Kobasa, S.C., Maddi, S.R., & Kahn, S. (1982). Hardiness and health: A prospective study. Journal of Personality and Social Psychology, 42, 168-177.

Lavallee, L. F., & Campbell, J. D. (1995). Impact of personal goals on self-regulation processes elicited by daily negative events. Journal of Personality and Social Psychology, 69, 341-352.

Lecznar, W. B. (1951). Evaluation of a new technique for keying biographical inventories empirically (Research Bulletin No. 51-2). San Antonio, TX: Lackland Air Force Base, Human Resources Research Center.

Lecznar, W. B., & Dailey, J. T. (1950). Keying biographical inventories in classification test batteries. American Psychologist, 5, 279.

Lewin, K. (1951). Field theory in social science. New York: Harper & Row.

Lindsey, E. H., Holmes, V., & McCall, M. W. (1987). Key events in executives' lives. Technical Report 32. Greensboro, NC: Center for Creative Leadership.

Linton, M. (1982). Transformations of memory in everyday life. In U. Neisser (Ed.), Memory observed: Remembering in natural contexts (pp. 77-91). San Francisco, CA: Freeman.

Linton, M. (1986). Ways of searching and the contents of memory. In D.C. Rubin (Ed.), Autobiographical memory (pp. 50-67). Cambridge, England: Cambridge University Press.

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

Mael, F. A., & Ashforth, B. E. (1995). Loyal from day one: Biodata, organizational identification, and turnover among newcomers. Personnel Psychology, 48, 309-333.

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

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

Malloy, J. (1955). The prediction of college achievement with the life experience inventory. Educational and Psychological Measurement, 15, 170-180.

Malone, M. P. (1977). Predictive efficiency and discriminatory impact of verifiable biographical data as a function of data analysis procedure. Unpublished doctoral dissertation, Illinois Institute of Technology, Chicago.

McAdams, D.P., Diamond, A., St. Aubin, E.D., Mansfield, E. (1997). Stories of commitment: The psychosocial construction of generative lives. Journal of Personality and Social Psychology, 72, 678-694.

McCall, M., & Bobko, P. (1990). Research methods in the service of discovery. In. M.D. Dunnette and L.M. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., Vol. 1, pp. 381-418). Palo Alto, CA: Consulting Psychologists Press.

Mitchell, T. W. (1996). Can do and will do criterion success: A practitioner?s theory of biodata. In R. B. Stennett, A. G. Parisi, & G. S. Stokes (Eds.), A compendium: Papers presented to the first biennial biodata conference (pp. 2-15), Applied Psychology Student Association, University of Georgia: Athens, GA.

Mitchell, T. W., & Klimoski, R. J. (1982). Is it rational to be empirical? A test of methods for scoring biographical data. Journal of Applied Psychology, 67, 411-418.

Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal. Journal of the American Statistical Association, 58, 415-438.

Mosel, J. L., & Cozan, L. W. (1952). The accuracy of application blank work histories. Journal of Applied Psychology, 36, 356-369.

Mount, M. K. (1997). Big five personality tests. In L. H. Peters, C. R. Greer, & S. A. Youngblood (Eds.), Encyclopedic dictionary of human resource management (p. 25). Blackwell: Oxford, England.

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., & Stokes, G. S. (1992). Developmental determinants of individual action: Theory and practice in applying background measures. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology (2nd ed., Vol. 3, pp. 61-138). 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.

Mumford, M. D., Costanza, D. P., Connelly, M. S., & Johnson, J. F. (1996). Item generation procdures and background data scales: Implications for construct and criterion-related validity. Personnel Psychology, 49, 361-398.

Neidt, C. O., & Malloy, J. P. (1954). A technique for keying items of an inventory to be added to an existing test battery. Journal of Applied Psychology, 38, 308-312.

Newman, D.L., Caspi, A., Moffitt, T.E., & Silva, P.A. (1997). Antecedents of adult interpersonal functioning: Effects of individual differences in age 3 temperament. Developmental Psychology, 33, 206-217.

Nickels, B. J. (1990). The construction of background data measures: Developing procedures which optimize construct, content, and criterion-related validities. Unpublished doctoral dissertation, Georgia Institute of Technology, Atlanta, GA.

Nickels, B. J. (1994). The nature of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens, (Eds.), The biodata handbook: Theory, research, and applications (pp. 1-16). Palo Alto, CA: Consulting Psychologists Press.

Owens, W. A. (1968). Toward one discipline of scientific psychology. American Psychologist, 23, 782-785.

Owens, W. A. (1971). A quasi-actuarial basis for individual assessment. American Psychologist, 26, 992-999.

Owens, W. A. (1976). Background data. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology (pp. 609-644). Chicago: Rand McNally.

Owens, W. A., Glennon, J. R., & Albright, L. E. and the American Psychological Association Scientific Affairs Committee, Division 14 (1966). A catalogue of life history items. Greensboro, N.C.: The Richardson Foundation.

Owens, W. A., & Schoenfeldt, L. F. (1979). Toward a classification of persons. Journal of Applied Psychology, 53, 569-607.

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. European Review of Social Psychology, 1, 33-60.

Pennell, G.E. (1996). Integrating complexity into the study of life events: Comment on Suedfeld and Bluck (1993). Journal of Personality and Social Psychology, 71, 777-780.

Quinones, M.A., 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.

Reilly, R. R., & Warech, M. A. (1990). The validity and fairness of alternative predictors of occupational performance. Paper invited by the National Commission on Testing and Public Policy, Washington, DC.

Rivkin, I.D. & Taylor, S.E. (in press). The effects of mental simulation on coping with controllable stressful events. Personality and Social Psychology Bulletin.

Rovee-Collier, C. (1995). Time windows in cognitive development. Developmental Psychology, 31, 147-169.

Russell, C. J. (1986). Review of the literature and development of a biodata instrument for prediction of naval officer performance at the U.S. Naval Academy. Army Research Institute Report No. DAAG29-81-D-0100, Arlington, VA.

Russell, C. J. (1994). Generation procedures for biodata items: A point of departure. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction (pp. 17-38). Palo Alto: Consulting Psychologists Press.

Russell, C. J., & Dean, M. A. (1994a, May). Exploration of a point biserial bootstrap method of empirical biodata keying. Presented at the ninth annual meeting of the Society of Industrial and Organizational Psychology, Orlando, FL.

Russell, C. J., & Dean, M. A. (1994b, August). The effect of history on meta-analytic results: An example from personnel selection research. Presented at the annual meetings of the Academy of Management, Dallas, TX.

Russell, C. J., & Domm, D. R. (1990). On the construct validity of biographical information: Evaluation of a theory based method of item generation. Presented at the fifth annual meetings of the Society of Industrial and Organizational Psychology, Miami Beach, FL.

Russell, C. J., & Domm, D. R. (1995). Two field tests of an explanation of assessment center validity. Journal of Occupational and Organizational Psychology, 68, 25-47.

Russell, C. J., Mattson, J., Devlin, S. E., & Atwater, D. (1990). Predictive validity of biodata items generated from retrospective life experience essays. Journal of Applied Psychology, 75, 569-580.

Russell, C. J., & Van Sell, M. (1986). Toward a model of turnover decision making. Winner of the Edwin E. Ghiselli Award for Research Design, presented by the Society for Industrial and Organizational Psychology.

Schachter, S. (1959). The physiology of affiliation. Stanford, CA: Stanford University Press.

Schmitt, N., Gooding, R. Z., Noe, R. D., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.

Schulz, R. & Heckhausen, J. (1996). A life span model of successful aging. American Psychologist, 51, 702-714.

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

Smith, K. C., & McDaniel, M. A. (1998, April). Background experience correlates of job performance: An expanded predictor space. Presented at the 13th annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

Snell, R. (1988). Graduating from the school of hard knocks? Journal of Management Development, 5, 23-30.

Stead, W. H., & Shartle, C. L. (1940). Occupational counseling techniques. New York: American Book Company.

Stokes, G. S., & Reddy, S. (1992). Use of background data in organizational decisions. In C. L. Cooper & I. T. Robertson (Eds.), International review of industrial and organizational psychology. (pp. 285-321). London: John Wiley & Sons. Ltd.

Stricker, L. J. (1987). Developing a biographical measure to assess leadership potential. Paper presented at the annual meeting of the Military Testing Association, Ottawa, Ontario.

Stricker, L. J. (1988). Assessing leadership potential at the Naval Academy with a biographical measure. Paper presented at the annual meeting of the Military Testing Association, San Antonio, TX.

Strong, E. K. (1926). An interest test for personnel managers. Journal of Personnel Research, 5, 194-203.

Suedfeld, P. & Bluck, S. (1993). Changes in integrative complexity accompanying significant life events: Historical evidence. Journal of Personality and Social Psychology, 64, 124-130.

Suedfeld, P., Tetlock, P.E., & Streufert, S. (1992). Changes in integrative complexity accompanying significant life events: Historical analysis. Journal of Personality and Social Psychology, 64, 124-130.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. Psychological Bulletin, 110, 67-85.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological adaptation. American Psychologist, 38, 1161-1173.

Taylor, S.E., Pham, L.B., Rivkin, I.D., & Armor, D.A. (1998) Harnessing the imagination: Mental stimulation, self-regulation, and coping. American Psychologist, 53, 429-439.

Telenson, P. A., Alexander, R. A., & Barrett, G. V. (1983). Scoring the biographical information blank: A comparison of three weighting techniques. Applied Psychological Measurement, 7, 73-80.

Tsluk, P. E. & Jacobs, R. R. (1998). Toward an integrated model of work experience. Personnel Psychology, 51, 321-355.

von Eye, A. & Brandtstädter, J. (1998). The wedge, the fork, and the chain: Modeling dependency concepts using manifest categorical variables. Psychological Methods, 3, 169-185.

Webb, S. C. (1960). The comparative validity of two biographical inventory keys. Journal of Applied Psychology, 44, 177-183.

Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. Journal of Applied Psychology, 52, 372-376.

NOTES

1. Note the Buckley amendment to the 1974 Privacy in Education Act makes it illegal for schools to share this information without written permission from students or their guardians.

2. This point was first made known to the authors in discussion comments made by Joel Lefkowitz at a symposium presented at the 1994 Society of Industrial and Organizational Psychology meetings.

3. We will not use the label "rational keys" to describe scoring systems derived from a priori theory-based expectations drawn from some theory or model. The label "rational" implicitly infers that alternative procedures (e.g., empirical keys) are "irrational," and nothing could be further from the truth. We are unaware of any biodata inventories whose architects did not have some reason for selecting weights found in a key.

4. Of course, some biodata applications do contain items that mimic traditional time series predictions, e.g., a biodata inventory common within the life insurance industry asks experienced candidates for sales positions "How much life insurance did you sell last year?" However, these items tend not to dominate inventory content or subsequent empirical keys.

5. We would like to thank Alan Mead for directing us to this research stream.

Table 1
Meta-analyses of Biodata Criterion-Related Validities

Study	Criterion	K	ΣN_i	$\bar{r}$	s_r^2	s_e^2	s_r^2	r
Hunter & Hunter (1984)								
	Supervisor ratings	12	4429	-	-	-	.10	.37
	Promotion	17	9024	-	-	-	.10	.26
	Training success	11	6139	-	-	-	.11	.30
	Tenure	2	2018	-	-	-	.00	.27
Re-analysis of Dunnette (1972)	-	115	-	-	-	-	-	.38
Re-analysis of Reilly & Chao (1982)	-	44	-	-	-	-	-	.34
Re-analysis of Vineberg & Joyner (1982)	Supervisor ratings Global	12	-	-	-	-	-	.20[a]
	Supervisor ratings Suitability	4	-	-	-	-	-	.29[a]
	Supervisor ratings All ratings	16	-	-	-	-	-	.24[a]
Reilly & Chao (1982)								
	Tenure	13	5721	-	-	-	-	.32
	Training	3	569	-	-	-	-	.39
	Ratings	15	4000	-	-	-	-	.36
	Productivity	6	661	-	-	-	-	.46
	Salary	7	680	-	-	-	-	.34

Study	Criterion	K	ΣN_i	$\bar{r}$	s_r^2	s_e^2	s_r^2	r
Schmitt, Gooding, Noe, & Kirsch (1984)	Across all criteria	**99**	**58107**	**.243**	**.0183**	**.0015**	**.0168**	**-**
	Performance ratings	**29**	**3998**	**.317**	**.0357**	**.0059**	**.0298**	**-**
	Turnover	**28**	**28862**	**.209**	**.0144**	**.0009**	**.0136**	**-**
	Achievement/grades	**9**	**1744**	**.226**	**.0784**	**.0047**	**.0738**	**-**
	Productivity	**19**	**13655**	**.203**	**.0036**	**.0013**	**.0023**	**-**
	Status change	**6**	**8008**	**.332**	**.0014**	**.0006**	**.0009**	**-**
	Wages	**7**	**1544**	**.525**	**.0157**	**.0024**	**.0133**	**-**
Russell & Dean (1994b) (update of Schmitt et al., 1984, all criterion validities reported in *JAP* and *Personnel Psychology* from 1965-1992)	Across all criteria	**107**	**59172**	**.291**	**.0151**	**.0015**	**.0136**	**-**

Note: Meta-analytic estimates of r_{xy} where corrected only for sampling error.

Figure 1

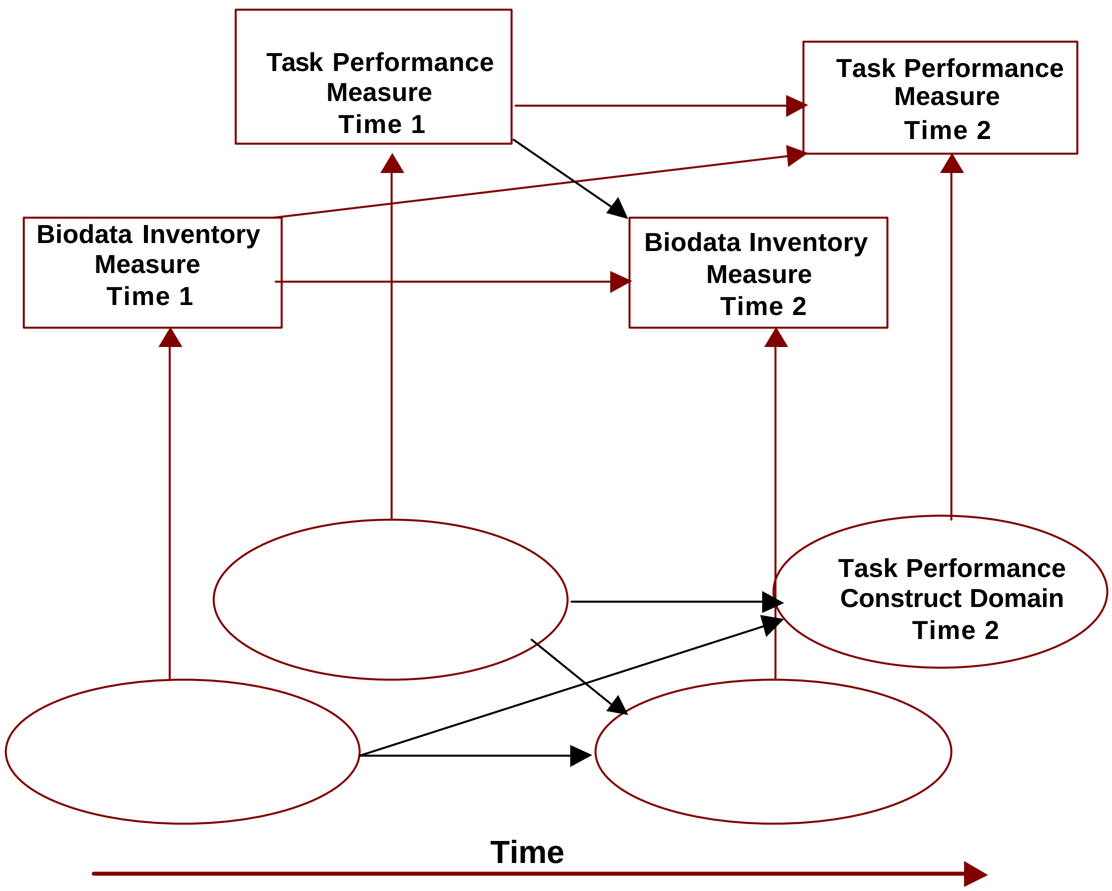
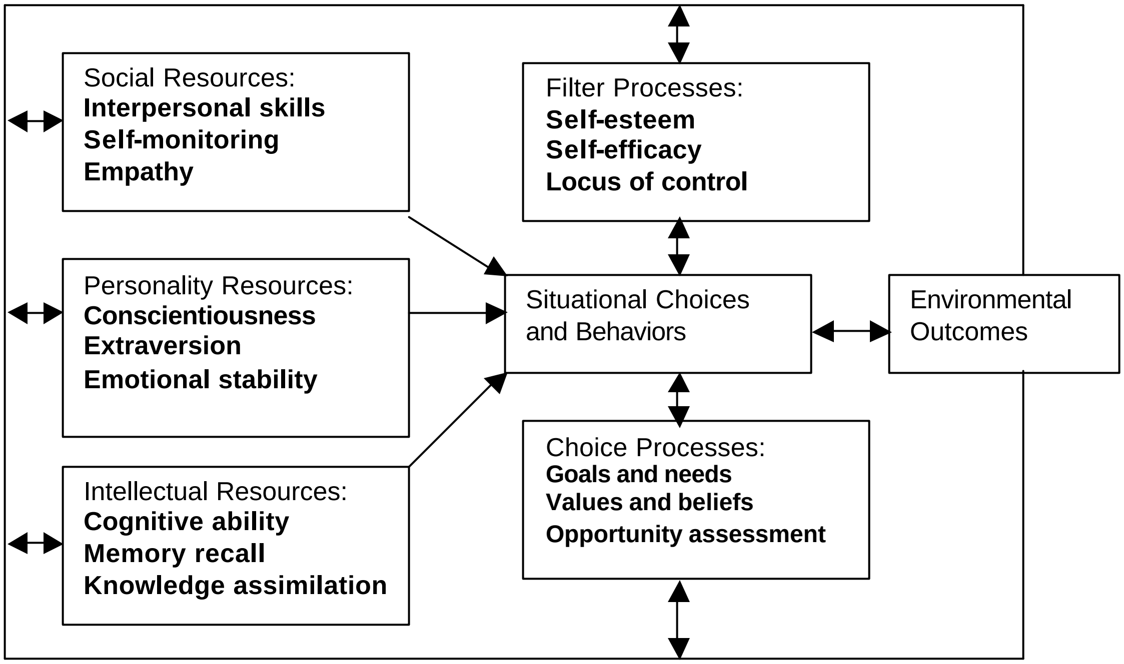


Figure 1: Model of biodata and performance construct domains and their relationship over time, adapted from Russell & Domm (1990).

Figure 2: The Ecology Model



Adapted from Mumford, M. D. & Stokes, G. S. (1992). Developmental determinants of individual action: Theory and practice in applying background measures. In M.D. Dunnette and L. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., vol. 3, pp. 61-138). Palo Alto, CA: Consulting Psychologists Press.

Authors

Michelle A. Dean is a doctoral candidate in human resource management and organizational behavior at Louisiana State University. She is a Visiting Assistant Professor in Management at the University of North Texas. Her current research interests are in personnel selection issues, specifically focusing on the prediction of performance using biodata. Her research has been published in the Academy of Management Journal and presented at the National Meetings of the Academy of Management and the Society for Industrial/Organizational Psychology.

Craig J. Russell is the J.C. Penney Chair of Business Leadership at the Michael F. Price College of Business, University of Oklahoma. He received his Ph.D. in business administration from the University of Iowa in 1982. His current research focuses on developing and testing models of leadership and management performance. He serves on the editorial boards of the Journal of Applied Psychology, Journal of Leadership Studies, Journal of Management, and Organizational Research Methods. His research has been published in the Academy of Management Journal, Journal of Applied Psychology, Journal of Management, and Personnel Psychology among others.

Paul M. Muchinsky is the Joseph M. Bryan Distinguished Professor of Business in the Joseph M. Bryan School of Business and Economics, University of North Carolina at Greensboro. He received his Ph.D. in industrial/organizational psychology from Purdue University in 1973. He has a long-standing interest in the scientific and practical value of biographical information.