DOCUMENT RESUME

ED 435 731 TM 030 363

AUTHOR	Lane, Eliesh O'Neil; Dietz, James S.; Chompalov, Ivan; Bozeman, Barry; Park, Jongwon
TITLE	Using the Curriculum Vita To Study the Career Paths of Scientists and Engineers: An Assessment.
SPONS AGENCY	Department of Energy, Washington, DC.; National Science Foundation, Arlington, VA.
PUB DATE	1999-11-03
NOTE	59p.; Paper presented at the Annual Meeting of the American Evaluation Association (Orlando, FL, November 3-6, 1999).
CONTRACT	DE-FG02-96ER45562; SBR-98-18229
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Career Development; *Careers; Coding; *Engineers; *Internet; *Research Methodology; *Resumes (Personal); *Scientists

ABSTRACT

The usefulness of the curriculum vita (CV) as a data source for examining the career paths of scientists and engineers was studied. CVs were obtained in response to an e-mail message sent to researchers working in the area of biotechnology who were funded by the National Science Foundation (55 responses) or listed as authors (industry only) in the "Science Citation Index" (19 responses). In addition, CVs were obtained passively from a search of the Internet (30 CVs). Methodological issues and problems of this data collection strategy are discussed, along with the results of an exploratory analysis. In sum, despite difficulties with coding and variation in CV formats, this collection strategy seems to hold much promise for examining career paths. An appendix describes a coding experiment performed to investigate ways to code CVs for data collection. A second appendix lists coding items that were not found to be reliable. (Author/SLD)

Using the Curriculum Vita to Study the Career Paths of Scientists and Engineers: An Assessment[1]

Eliesh O'Neil Lane

James S. Dietz

Ivan Chompalov

Barry Bozeman

and Jongwon Park

Research Value Mapping Program
School of Public Policy
Georgia Institute of Technology
Atlanta, Georgia 30332
http://rvm.pp.gatech.edu

November 3, 1999

This paper was prepared for the American Evaluation Association meeting in Orlando, FL, November 4, 1999.

[1] The authors gratefully acknowledge the support of the U.S. Department of Energy (DE-FG02-96ER45562) and the National Science Foundation (SBR 98-18229). The opinions expressed in the paper are the authors' and do not necessarily reflect the views of the Department of Energy or the National Science Foundation. This work was performed under the *Research Value Mapping Program* [http://rvm.pp.gatech.edu] of the Georgia Institute of Technology. The authors wish to thank Jeff Bournes, Marie Chesnut, Jungki Kim, Andy McNeil, Seth Sobel, Ryoung Song, and Larry Wilson for their assistance. The authors also thank Juan Rogers for his helpful advice.

Abstract:

In this paper, we assess the utility of the curriculum vita (CV) as a data source for examining the career paths of scientists and engineers. CVs were obtained in response to an email message sent to researchers working in the area of biotechnology who were funded by the National Science Foundation or listed as authors (industry only) in the Science Citation Index. In addition, a number of CVs were obtained "passively" from a search of the Internet. We discuss the methodological issues and problems of this data collection strategy and the results from our exploratory analysis. In sum, despite difficulties with coding and variation in CV formats, this collection strategy seems to us to hold much promise.

Using the Curriculum Vita to Study the Career Paths of Scientists and Engineers: An Assessment

1.0 Introduction

Scientists' and engineers' career trajectories have much in common with other professional paths. Motivational factors are not so very different, including income, need for achievement and recognition, and desire for "interesting" work. Scientists and engineers face many of the same constraints as others, choosing jobs because of a spouse's opportunities, the quality of schools available to children, distance from family, and so forth. Thus, the standard models available to labor economists can tell us much about scientists and engineers.

But there are some respects in which scientists and engineers differ dramatically from dentists or attorneys or airline pilots. Some are obvious like the peculiar formal assets, such as patents and publications, which scientists and engineers bring with them. Other assets are less obvious, less formal, but perhaps even more important. Each scientist and engineer can be thought of as a unique embodiment of "scientific and technical human capital" (S&T human capital), a walking set of knowledge, skills, technical know-how and—just as important—a set of sustained network communications, often dense in pattern and international in scope. In previous work (Bozeman, Dietz, and Gaughan, forthcoming), we outlined S&T human capital as an alternative model for research evaluation, originating in response to the limitations of traditional economic and state-of-the-art models.

The S&T human capital model puts more weight on the sustained ability of scientists and engineers to enhance their own capabilities and those with whom they work than do traditional models. S&T human capital includes not only the researcher's human capital but the social capital he or she draws upon in creating knowledge and interacting in various social and professional contexts. It includes not just the educational credentials normally recognized in traditional human capital models (Becker, 1962; Schultz, 1963) but the researchers' tacit knowledge (Polanyi, 1967; Polanyi, 1969), craft knowledge, and know-how. And, essential to the effective exploitation of all of these human capital endowments is the social capital (Bourdieu, 1986; Bourdieu & Wacquant, 1992; Coleman, 1988; Coleman, 1990) that scientists continually exercise in engaging their interests.

These endowments not only make the study of scientists' and engineers' career trajectories more difficult (e.g., less amenable to standard labor models) but more challenging. When a dentist changes jobs it is of interest chiefly to old and new clients. When a scientist or engineer changes jobs the implications are often profound: the movement of S&T human capital is, arguably, a vital element of scientific discovery, technological innovation, and even economic development. For S&T human capital transcends the intellect of any one individual. Thus, individual migration patterns of scientists and engineers can be likened more appropriately to the movement of the web of S&T human capital they possess—a web that continually manifests new shapes and patterns. Thus, for example, if Northern California was known as Dental Floss Valley there might well be a concentration of healthy teeth. But, for scientists and engineers network dependencies imply something altogether different for migration patterns, which rightly command researchers' and policymakers' attention.

If the career trajectories of scientists and engineers are often a bit more complicated and less predictable than dentists (or airline pilots, or attorneys), they also leave more marks along the trail. One of the great, albeit largely unexploited, advantages of studying the careers of scientists' and engineers' is the near universal reliance on the curriculum vita (CV). The utility of CV data for study of S&T human capital is striking, at least at first blush. The CV provides not only a clear-cut indicator of movement from one work setting to the next but is, in a sense, a representation of certain aspects of S&T human capital. Not only does it indicate the skills and knowledge embodied in scientists and engineers (through publications and other technical activities) but the professional association memberships, consulting, and co-authorship patterns serve as a crude index of social capital.

The CV, unlike other data sources, often recounts the entire career of the scholar in some detail. Thus, it is not simply a list of credentials, but an historical document that evolves over time capturing changes in interests, jobs, and collaborations. Whether viewed as historical record, marketing tool, or scientific resource, it is a potentially valuable datum for persons interested in career trajectories or, more generally, science and technology studies. Not only is the CV nearly universal, it is in some respect standard, and it is relatively easily obtained (sometimes even from the public domain). Most important, the CV contains useful, concrete information on the timing, sequence, and duration of jobs, work products (e.g., articles, patents, papers), collaborative patterns, and scholarly lineage. The CV is, indeed, a rich source of longitudinal data, which lends itself especially well to the study of phenomena associated with careers and labor flows—precisely the target of S&T human capital.

On the other hand, this proposed method is not without its limitations or problems. In fact, several of the advantages to using the CV as a data source can also be viewed as disadvantages. First, because the information is self-reported, its accuracy requires verification by the researcher. Second, the semi-structured format falls short of a purely standardized template, thus risking the elimination of valuable information or the inclusion of extraneous non-relevant data. Perhaps most significant, however, is the enormous work involved in coding the CV for subsequent data analysis. Not only is the coding time consuming, it is also tedious and runs the risk of introducing error due to coder fatigue. In some cases it is possible to have as many as 1,200 variables for one CV.

Despite its limitations, the potential of the CV as a research tool is enormous. Yet it has been used only sparingly—and sometimes incidentally—as a research device. We seek to address this neglect, to explain it, and to assess the promise and obstacles to a research agenda employing the CV as primary data. The development of such a methodology undoubtedly would provide a unique and potentially useful alternative for evaluating scientists' and engineers' careers.

1.1 Organization of Paper

A major objective of this paper is simply to determine the extent to which it is possible to obtain useful CV data and to assess the utility of various approaches to collecting CVs. In section two, we present a review of the literature on scientific careers.

In section three, methodological issues are presented and discussed with specific attention to several coding and data-related issues. In section four, we address issues of validity and reliability, data consistency and quality, and CV accessibility. After examining the descriptive findings (in section five), we reflect more broadly (section six) on our assessment of the utility of a CV-based methodology, including possible strategies for improving the quality and consistency of data. In section seven, we present our conclusions.

2.0 CVs, Scientific and Technical Human Capital, and Research Value

2.1 The Research Value Mapping Program

Our interest in S&T human capital, and the potential of CVs as a research tool for mapping flows of this capital, stems from a general interest in assessing the impacts of government-financed research projects. *The Research Value Mapping (RVM) Program* within the School of Public Policy at the Georgia Institute of Technology began in 1996, using 30 intensive case studies of research projects as sources of both qualitative and quantitative information about the nature and intensity of the projects' scientific and socioeconomic impacts (see Bozeman, et al., 1999; http://rvm.pp.gatech.edu).

The Phase I work, sponsored by the Department of Energy's (DOE) Office of Science, focused entirely on DOE-sponsored projects in government and university labs. We are beginning Phase II based on continued funding from DOE and with new funding from the National Science Foundation. The mission of Phase II is to compare research impacts in multiple fields and in the U.S. and France. Whereas Phase I focused on information developed in the case studies, Phase II will focus on S&T human capital impacts, using the CV as one research tool to examine labor flows and career trajectories. The core hypothesis of Phase II is that many of the impacts of projects are not easily confined within normal project boundaries but occur over considerable time as S&T human capital diffuses into other settings.

We will ultimately test several hypotheses about the connection between the characteristics of team-oriented R&D projects and the diffusion of S&T human capital via the projects' "graduates." A preliminary study (just begun for this paper) of scientists and engineers in the area of biotechnology provides the opportunity to explore the use of the curriculum vita as a methodological tool for garnering such information.

Studies of innovations have already established the importance of close coupling for knowledge transfer and the diffusion of innovations in the economy (Rogers, 1995). The flow of people from one organization, firm, or group to another is key in the process of knowledge exchange. But, despite some good attempts (e.g., Stephan and Levin, 1997; Simonton, 1997), the extant literature has not managed to fully capture the dynamic nature of these flows over time and across research contexts. Careers are inherently dynamic—evolving and intersecting in planned and unplanned ways, but traditional research evaluation models view them as static or at best, additive and cumulative over time. We hope, in the next round of the RVM Program, to address this need.

2.2 Curriculum Vitae and Credit Allocation in Science

Despite the potential value of CVs as both data collection instruments and sources of data on productivity, recognition, career trajectories, and mobility in R&D, there is a paucity of theoretical and empirical investigations. One of the few studies that shed some light on the importance of CVs is Latour and Woolgar's anthropological account of the social production of scientific knowledge in a neuroendocrinological laboratory (Latour and Woolgar, 1986). From their point of view the CV is considered a "balance sheet" of a scientist's past investments and a testament to his or her credibility. Latour and Woolgar claim that, apart from accreditation, awards, collaborations, and publications, there is an element of the CV that plays a crucial role in estimating a researcher's total value. In their view, value is a three-part notion that incorporates academic rank, situation in the field, and geographical location (Latour and Woolgar, 1986). CVs and interviews could serve not only as valid sources of information to reconstruct individual career trajectories, but also group dynamics and the accumulation of social capital in the form of credit.

Surprisingly, CVs have been used only infrequently to illuminate well-studied processes of how the social system of science operates. Given the tradition in sociology of science to focus on rewards and credit allocation, for example, the dearth of studies employing CVs as rich data sources to trace the award of credit is quite noteworthy. Most of the past research on recognition in science has been carried out within some economic model of knowledge production.

The psychosocial mechanisms of reward in science have been, if anything, well known and investigated since at least the 1960s. The best known thesis regarding scientific credit is the "Matthew effect" described by Robert Merton (1973a). He defines this phenomenon as reflecting an "accumulation of advantages," so that already outstanding scientists receive disproportionately more credit for their contributions than younger researchers who are perhaps less visible in the field. Merton notes that while he coined the term to refer to the greater recognition scientists of higher rank receive for their discoveries, it undoubtedly has implications for the communication system (e.g. visibility), as well as for resource distribution (e.g., grant funds). The Matthew thesis fits well within the framework of the normative structure of science (Merton, 1973b) and its system of social stratification (Cole and Cole, 1973). And, although the effect may be pervasive in all fields, Zuckerman and Merton point out that, in all likelihood, it operates more strongly in less codified fields such as the social sciences and humanities. In these disciplines, "the personal and social attributes of scientists are more likely to influence the visibility of their ideas and the reception accorded them" (Zuckerman and Merton, 1973, p. 516).

Overall, the Mertonian treatment of credit allocation is consistent with the neoclassical economical view of early "entrepreneurial capitalism," where scientists operate on a free market and try to maximize their utility. Another economics model—the exchange system of gift-giving in primitive societies and other settings (Hyatt and Hopkins, 1998)—dominates Hagstrom's conception of scientific recognition. In a nutshell, scientists give away information expecting to be rewarded by field recognition for their contributions: "social control in science is exercised in an exchange system, a system wherein gifts of information are exchanged for recognition from scientific colleagues" (Hagstrom, 1965, p. 52).

Unlike Hagstrom who postulates a mechanism analogous to primitive gift exchanges, Latour (1986) argues that the control system of science operates similar to, and apparently derived from, the Marxist political economy of capitalism. In Latour's view, scientists are interested in gaining credit or credibility mainly because it gives them access to other resources, which, in turn, can be translated into further credit. The resulting image is that of a perpetual cycle where the accumulation of credit and faster rates of turnover become ends in and of themselves. This "cycle of credibility" involves conversions between different forms of capital (e.g., money, equipment, data).

Latour's model strives to overcome a weakness in Hagstrom's and Bourdieu's (1975) theories, namely their failure to consider demand (of scientists for each other's work). In Latour's view, the researcher as *homo economicus* gets caught up in the objective of market activity "to extend and speed up the credibility cycle as a whole" (Latour and Woolgar, 1986, p. 207).

2.3 Scientific Productivity and the Life Course

There is a strong support for the thesis that academic career trajectories and especially promotion are significantly affected by productivity in terms of both quantity and quality of publications. Such a relationship is perceived as a confirmation of the Mertonian normative model of scientific knowledge production and particularly the operation of the "universalism" norm. Empirically, it has been demonstrated that allocation of citations follows a "repayment of intellectual debt" mode, rather than a social constructivist "network" perspective (Baldi, 1998). Promotion or the achievement of a higher rank in science is explained by institutional prestige, at least so is true in the case of academic psychologists (Hurlbert and Rosenfeld, 1992). Rank advancement has been proven to depend more on the sheer quantity than quality of publications for university departments in biochemistry (Long et al., 1993). The results from event history analysis of promotional patterns that Long and associates report also indicates that the likelihood for promotion is lower for women than for men for a job change from assistant to associate professorship. However, when a battery of control variables was added to the model, the gender difference was cut in half and was no longer statistically significant. Of course, although gender differences in research productivity have been well documented, there have been big controversies regarding the explanation of these differentials. This has led some authors to label the phenomenon as 'the productivity puzzle' (Cole and Zuckerman, 1984; Xie and Shauman, 1998). Field effects also account for a significant amount of variation in productivity (Bonzi, 1992) and, consequently, on academic promotion.

Life cycle models view the careers of scientists as a longitudinal function of the individual's skill levels and his or her incentives to act productively (Diamond, 1984; 1986). At earlier stages of career building, productivity incentives are strong while skills are growing. At the middle stages (and sometimes even earlier), both incentives and skills are strong as productivity peaks. And at later stages, both begin to wane, as does productivity. The concept of a career life cycle originated in human capital theory from an economics tradition (Becker, 1963). Human capital theory sought to relate investments in human beings (education, training, job and life experiences, and personal health) to an individual's earnings trajectory.

In the scientific life cycle model, Levin and Stephan (1991) report that scientific productivity follows one of two general patterns (depending on scientific discipline): one where productivity simply declines with age, the other where it increases at first but then declines with age. Although there is plenty of empirical evidence to support this notion of diminishing marginal rates of productivity, such models fail to explain much variation in productivity (Stephan, 1996). Moreover, as Stephan and Levin have pointed out, many of these life-cycle models lack sufficient attention to the research process and the institutional setting of the process (Stephan and Levin, 1997).

Researchers have also called attention to the role of early career collaboration and mentoring as spurs to longer-term scientific productivity. Long and McGinnis (1984) found significant and lasting effects of predoctoral collaboration with mentors on the careers of biochemists. The productivity of the mentor was positively and strongly related to the biochemists' own publication productivity six years later. For students who had not collaborated with their mentor, there was no relationship. Similarly, Reskin (1977), studying chemists who obtained their Ph.D. in the late 1950s, found graduates from higher "caliber" departments were more likely to have collaborated with their doctoral mentor and showed higher productivity after their first postdoctoral decade than graduates from lesser-prestige departments. Zuckerman (1977) found that Nobel prize winners viewed their doctoral apprenticeship as crucial to their later success and, specifically, in building broad skills such as proper standards of achievement, tastes in choice of research problems, and confidence in their work and abilities.

Life *course* models can be thought of as an enhancement or conceptual expansion of life *cycle* models. Elaborated by Elder (1994), the life course paradigm views individual lives as affected by the historical period in which events occur, the developmental timing and sequence of events, and the involvement of the individual in relevant social relationships. Elder refers to the concept of human agency, which—as applied to science—can be thought of as the unique set of abilities that each scientist uses to translate his or her training and skills into scientific outputs. All individuals have "human agency," although in different mixes. In a sense, human agency is a recognition that individuals vary in the predispositions (both strengths and weaknesses) they bring to the construction of a life course. Elder warns, however, that life course is more than just human agency. It is human agency constrained by developmental timing and history effects.

The most important contribution that life course models have made to the understanding of the scientific careers and to S&T human capital, for that matter, is the notion that human lives are linked, or interdependent with other each other, and—not just statically—but dynamically over time. Merton ([1965] 1993) recognized this in titling his book, *On The Shoulders of Giants*, in which he illustrates how Newton made his intellectual advances using the contributions of his scientific peers and forefathers. The life course concept illustrates the dynamic form of learning and communication among individual scientists and the meso and macro social contexts in which they are engaged. It is not completely socially deterministic, but nor does it rely strictly on individual reductionism.

3.0 Framing the Research Issues: Some Practical Concerns

Very few studies have employed CVs as data sources about trends in job mobility in science. Typically, CVs are used as a supplemental source of information that serves to fill in the gaps from other documents (Long et al., 1993; Gomez-Mejia and Balkin, 1992). Even when CVs[1] are utilized as the primary or only data source, their advantages or disadvantages are rarely discussed (Bonzi, 1992). The notion of using CVs as a research tool is hardly a novel idea. But the actual utility of CVs lies in answers to some quite practical questions and resolution of some fundamental methodological issues.

3.1 What are the labor issues?

A wealth of information is provided in most CVs but the coding of the information and its entry into a database is not at all straightforward. When one considers that some CVs include hundreds of publications and conference papers, many with multiple authors, the costs of labor become apparent. The options are few. If one wishes to capture almost all the information in a CV into standard databases, the enterprise likely involves thousands of observations per case. This requires a small army of labor, well trained and perhaps not all "low end" inexperienced data entry personnel. A second option is to mechanize as much as possible, through scanners, but this too has substantial labor and set-up cost. The final option, almost inescapably, must be pursued: limiting data capture. Absent prodigious data entry resources, the only option is to forgo much data or to categorize data at a relatively high level of abstraction (e.g., count

[1] Methodologically, CVs have been found to closely match information from other secondary sources such as the American Psychological Association's directory. Nevertheless, these other secondary sources have been shown to undercount the number of published journal articles as compared to CVs (Heinsler and Rosenfeld, 1987).

articles). The trick of the trade, then, is to optimize time, data capture, and labor. The hazards include insufficient culling, poorly predicted labor requirements, or settling on data at so high a level of aggregation that inferences are obscured. Can one develop heuristics or some empirical base for making such decisions? That is one of our concerns in the paper and the overall project.

3.2 How to Operationalize the CV data?

With anything less than complete data capture, the particular operationalization of CV data becomes vital. Even after whole sections of data are dismissed (in our case, such sections as conference papers, courses taught, internal working papers), one still must grapple with measuring the remainder. Is it important, for example, to capture not only article publications but also author numbers and author order? How does one represent data in more economical indices and, more to the point, how does one know which indices are most useful without sufficient original data to employ in indices? Are data best represented in arrays, across time, or in cross-sectional detail? To be sure, many of these answers depend upon specific hypotheses of specific studies, but if the CV database is to serve as a general resource for multiple research objectives, specific hypotheses provide little relief.

3.3 Where and how does one obtain CVs?

Virtually every scientist and engineer has a CV. But how does one obtain it? One way is simply to write a letter or an email message requesting one. Least obtrusive is simply going to public domain websites and downloading the publicly available data.

How does one know the yield and peculiarities of return from each of these various approaches?

3.4 Are CV's consistent?

An interesting problem is "which CV does one obtain?" First there is the time issue. If one wishes to examine CVs over time, one finds that most people do not keep old CVs, only recent ones. This is a problem when one considers that a great many CVs get truncated (e.g., "publications for past ten years") and that the information that is important in an early career (e.g. all conference presentations) may be unimportant to the scientists later and may disappear from the CV. The results are possible differences in time, periodicity, and cohort. Interestingly, the availability of CVs on the web has been helpful and hurtful to those interested in the CV as data. The popularity of the web has meant that a great many more CVs are accessible, but the institutionalization of websites has led to a stylistic conformance of CVs, which is not itself a problem, and, typically, significant abridgment, which can be a great problem. If the CV on the web is typically an institutional rather than individual marketing resource, the rational marketing approach is succinct information about more people, rather than detailed information about particular people.

3.5 How to link to benchmarks and secondary data?

A great advantage of CV data is that it is so easy (conceptually) and useful to link to cognate data. The availability of a wide array of citation data through the Science Citation Index is extremely valuable. These same databases also include information on the "power index" (i.e., the likelihood of citation) of journals. Similarly, the aggregate data provided in the SESTAT database of the National Science Foundation also serves as a potentially fruitful linkage. The problem, of course, is that these activities multiply data entry and manipulation costs another order of magnitude. Furthermore, the decision to use such benchmarks and cognate data requires making significant "up front" decisions on data collection strategies.

3.6 What is the coding validity and reliability?

By most any standard, the coding of more than a few CVs is a daunting task. We know that coding error rates from relatively tractable survey data range from about 5-10 percent (Fowler, 1988). What about more difficult, less obvious CV data? While coding errors can at least be determined with some ease, it is not clear even which standard is best for coding reliability. Moreover, good measures of coder reliability require a good number of coders, again accelerating costs. Most important of all, however, is coding validity. Except for the most straightforward issues, CV coding is almost always sure to cause problems for any but the best-trained eye. Explaining to a coder how to deal with visiting professors working at (apparently) three different places, in two sites, with three ambiguous titles requires time, patience, and imagination. For example, the difference between a postdoc and a fellow may be vital in some instances, not others. And how does one determine if a proceedings publication is consequential when working in a number of very different fields. Is it possible to conveniently detail such matters for coders in anything less than a 50-page codebook?

4.0 Research Methodology

4.1 CV Selection Strategy

Three approaches were used to obtain an expected sample of 350 CVs: a "targeted agency" search, a "targeted industry" search, and a "passive Internet" search. For both targeted searches, a direct email message was sent to potential respondents who had either recently conducted funded research or published in the area of biotechnology.[1] Respondents were asked to submit a full CV via email or fax, although a few respondents actually preferred to mail a hard copy because they felt there may be security issues in sending us an electronic copy. For the passive search, various Internet search engines and search phrases were used to identify a subgroup of web-posted CVs. Of the sample group, 50 CVs were solicited from industry scientists and engineers, 200 from academic researchers, and 100 from the web.

4.2 Collection Procedures for Targeted Agency Search.

A sample of 200 researchers funded by the National Science Foundation's (NSF) Biotechnology program and working at US institutions was obtained from NSF's awards database. This strategy has the main advantage of identifying a group of active, biotechnology researchers whose email addresses are provided by NSF. An email was sent inviting the researchers to submit their full CV via email. Approximately 20 percent of the email addresses taken from the database were erroneous or obsolete and were returned. The researchers attempted to obtain a current address for all of undelivered emails using institutional directories via the web. Ten could not be located. No follow-up was done on the nonrespondents. Fifty-five resumes resulted. Four respondents formally refused and one claimed not have been funded by NSF. The effective response rate was approximately 28 percent.

4.3 Collection Procedures for Targeted Industry Search.

Prior to conducting the targeted industry search, the Science Citation Index (SCI) was explored for the years 1999-95 using the key word "biotechnology." Over 3,100 titles were returned. Because searching each title for industry affiliation would have required extensive time and would not necessarily yielded useful results, we opted to identify five journals[2] that were likely to draw readers and authors from industry. Two[3] of these were not available on the SCI and were substituted by two additional journals.[4] Fifty-nine email notes were mailed and 19 responses were received—a response rate of approximately 34 percent. Based on their email addresses, thirteen of the industry recipients did not reside in the U.S., seven appeared to work in government agencies (which we considered non-academic for our purposes), and two worked with university-affiliated hospitals.

---

[1] Biotechnology was chosen because it represents an interdisciplinary field where research is conducted both in academia and industry. This was expected to result in high variance among CV contents and formats which was ideal for this exploratory phase of the research.

[2] *Journal of Heart Valve Disease, Annals of Biomedical Engineering, Journal of Biomechanics, Echocardiography,* and *American Society of Echocardiography*

[3] *Journal of Echocardiography* and *Echocardiography*

[4] *Biotechnology and Applied Biochemistry* and *Biotechnology and Bioengineering* (an electronic journal).

4.4 Collection Procedures for Passive Internet Search

Our third group of CVs was acquired via a passive search strategy over the Internet using popular web search engines. The target sample for this group was individuals with a Ph.D. in either biology, chemistry, life sciences, or a related engineering field (with at least some remote connection to the area of biotechnology). This approach has several advantages: it is non-intrusive, it utilizes CVs already available in the public domain, it is cost effective, and there is virtually no wait time. On the other hand, a major disadvantage is that the sample includes only people who posted their CV on the web.

Web searches were conducted using the keywords "curriculum AND vita*[1] AND biotechnology." We tested ten search engines: Alta Vista, Excite, Google, Goto, Hotbot, Infoseek, Look smart, Netscape, Snap, and Yahoo. The initial results using standard keywords were disappointing since these search engines turned up too many irrelevant web pages such as job search lists, academic newsletters, job announcements or résumé writing guide pages—"noise."

Too much noise makes the process more time consuming since the researcher has to go through longer list of irrelevant web pages. Two different tactics were used to remedy noise problems. First, we employed various keywords in different combinations. For example, using biochemistry instead of biotechnology produced more applicable pages. In particular, using Yahoo, the combination of "curriculum AND vitae AND biotechnology" produced over a thousand hits, about forty percent of which were resume pages broadly. Using "curriculum vitae" turned up more European researchers' web pages than US researchers' resumes. Using Yahoo, the string "curriculum AND vita* AND biotechnology AND PhD" turned up about 800 hits and approximately half of them were applicable. Second, we used various advanced search features to reduce the noise level. These advanced search methods are variations of Boolean expressions in most cases. For example, for the Alta Vista search we added "NOT job NOT jobs" since most noise was related to job-related websites. Refining the key words dramatically reduced noise level (see Table 1).

Recently developed search engines such as GOTO and HOTBOT seemed to employ different search algorithms and their noise levels were extremely low while they turned up relatively small numbers of hits. For example, GOTO produced 240 hits with noise level less than 20 percent using the simplest keyword combination, "curriculum vita* biotechnology." With refined keyword combinations, "curriculum vita* + biotechnology - job - jobs," GOTO produced 197 hits with noise level around 10 percent.[2]

It took about five to seven hours to collect and print out 30 resumes. If individual scientists or engineers update their own web page, it was more likely that the information was up to date. However, we also encountered many summary biographies in cases where some organizational unit such as academic department managed the website.

[1] We used the asterisk as wildcard since both curriculum vita and curriculum vitae were used. Some web search engines did not allow using asterisk as wildcard and turned up zero hits. We used Boolean expression "vita or vitae" in those cases.

[2] Another search engine HOTBOT produced similar results with similar level of noise.

4.5 Coding Methodology

To develop a useful and meaningful codebook, we reviewed a subset of the CVs to identify problems with the information included or eliminated in each of our three core groups. Over 300 potentially useful variables were identified. However, many of these were "variable sets"—multiple degrees received, multiple publications, multiple patents, and so forth. Several practice coding exercises[1] were conducted to determine the expected time required to effectively code the "typical" CV. Times were recorded to determine the range required to code a CV. Based on the information garnered from this experience, a training protocol was developed for future CV coding. The protocol was designed so that a work-study student could be trained to code the typical resume with minimal reliability problems within 30 minutes.

5.0 Data Analysis

5.1 Intercoder Reliability

Before we focus on some trends that emerged from the preliminary analysis of the CV data, we will briefly examine reliability issues associated with how consistently the information could be coded by independent coders. Intercoder reliability refers to the extent to which different coders using the same instrument to measure the same set of observations and their attributes achieve the same results (Singleton et al., 1993). In other words, it denotes the level of agreement among coders when they code one and the same thing using equivalent coding schemes. The higher the agreement is among coders, the higher the reliability. Discrepancy among coders can stem from two main sources: the code (when it is poorly designed and creates ambiguities, for instance) and the coders (when they disagree because of differences in experience, training, or personal style). Various measures of intercoder reliability have been proposed (see Scott, 1955; Craig, 1981; Funkhauser and Parker, 1968; Fleiss, 1971; Landis and Koch, 1977; Crittenden and Hill, 1971; Montgomery and Crittenden, 1977), but for our purposes we found Crittenden and Hill's measure of intercoder reliability (Rs) (Crittenden and Hill, 1971) the most useful, widely applicable, intuitively appealing, and simple.

This statistic denotes the level of intercoder agreement that is not based on chance. The formula for this measure is fairly simple and compares observed differences among n number of coders across interview items to the maximum possible differences among the same coders:

$$Rs = \frac{\sum D_{max} - \sum D_o}{\sum D_{max}} = 1 - \frac{\sum D_o}{\sum D_{max}}$$

[1] A hard-copy data entry table was developed for test-run on ten resumes from each target group. A group of faculty, postdoctoral fellows, doctoral students, masters students, and work-study students totaling eight carried out the test coding.

where D_o refers to the number of observed differences between pairs of coded responses, and D_{max} to the number of maximum possible differences. Thus, Rs can range in value from 1 (perfect reliability) to 0 (perfect unreliability). Although Crittenden and Hill do not talk about how they derive ΣD_{max}, we found that it could be done in the same way as the calculation of the number of possible reciprocal ties in a network (Scott, 1991). Thus, borrowing from social network analysis, this number can easily be calculated as n(n-1)/2, or, alternatively $(n^2 - n)/2$. This is the general case, when the number of response categories k is equal to, or larger than the number of coders n. We developed an algorithm that handles cases where k is equal to 1/2n or larger, but less than n, such that $\Sigma D_{max} = n(n-1)/2 - (n - k)$ for $n > k \geq 1/2n$. Since we used five coders to test for intercoder reliability, the only other instance to be accounted for is when k=2. Then ΣD_{max} becomes 6.

This technique is helpful in both estimating the overall intercoder reliability for a given interview and for item-by-item analysis. We conducted a test with five coders who independently coded ten resumes in order to check for reliability. One of the coders was experienced and had been involved in all the stages of preparation of the pilot study (sampling, development and modification of the codebook, test-coding to pinpoint potential problems with the coding schemes). The other four had no prior knowledge of the development of the instrument or the discussions that had led to its present form. They were briefly instructed on the use of the codebook and were given an overview of the standards pertaining to the coding protocol.

6.0 Findings and Discussion

The use of CVs as a data source as well as a collection instrument, while appealing and potentially fruitful, has made us aware of a variety of problems that can be summarized in three categories:

6.1 Sampling

There was some difficulty, in this exploratory phase of the research, in how to adequately specify the sampling framework. It was often problematic to identify the population of biotechnology researchers given the interdisciplinary nature of that field. Does it include biochemists, for example? Should national and international scientists be included? Are biotechnologists without a Ph.D. relevant for our purposes? In cases where multiple subdisciplines are represented as one field, how do subdisciplinary differences in field norms affect the quality of our data? To compound the problem, what happens when multiple fields are represented (e.g., biotechnology, physics, mechanical engineering, etc.)? How can we account for differences in the reward and incentive structure in how fields are compared in any number of ways?

There were also problems with sampling methodologies (e.g., with nonprobability sampling we will not be able to use inferential statistics, strictly speaking), and, more fundamentally, how do we sample if we do not know the population? From our experience in this early exercise, CVs from industrial scientists and engineers were difficult to identify as such and obtain. A non-statistical comparison of industry CVs with academic CVs did not indicate any glaring dissimilarity, with one exception. Many of the CVs of industrial scientists and engineers contained a section indicating current research interests. Most of the academic CVs did not spell these out in separate narrative form. Surprisingly, only one industry respondent specifically mentioned his management activities and related training—a section not included on any of the academic CVs. Perhaps, the passive Internet search yielded the highest variance among CVs, in terms of scientific fields, levels of experience and backgrounds of the scientists, and institutional affiliation.

6.2 Methodology

The principle methodological problems stem from the limited standardization in CV formats (possible international effects compound this problem). While CVs are as a matter of course semi-standardized, it seems that even limited amount of variance in contents and formats poses problems for coding and item response. One major problem encountered involves selection effects: the problems of CV "embellishment" vs. "truncation." That is, some researchers included everything *ad nauseum*, while others deleted relevant information (usually in the form of reporting only recent materials or deleting earlier work that the researcher now felt was less relevant). It is the latter problem that is most limiting. We are certainly able to delete extraneous information, but information not included is information lost. Related, were the problems that reflect formatting standardization—some CVs were "coder friendly" in terms of their formatting while others were not. In addition, there were definite cohort effects in CV construction. Thus, older biotechnologists were more likely to include personal information and to exclude conference presentations than younger researchers.

Tables 2 and 3 summarize the results from our preliminary coding test (statistics for only 37 out of 115 variables originally used are reported here). Overall, the average reliability coefficient value of .766 shows that we would be wise to heed the advice to further refine the instrument (created ad hoc on the basis of the actual information contained in the curriculum vita) and use several coders for check-ups, at least at the beginning. While there is no widely accepted "threshold level" of intercoder reliability for this particular coefficient, anything below .850 should probably be considered problematic, and anything below .600, outright unacceptable. Only 16 out of 37 items satisfy the .850 requirement. Moreover, 7 out of 37 items fall below .600. A closer inspection demonstrates that those variables concern awards, publications, and jobs. The difficulty stemmed from the unrefined coding scheme for these items as well as the time sequences presented in the CVs themselves. This feedback was valuable for improving the codebook.

The average time to code a resume was 23.6 minutes (it ranged from a minimum of 15 minutes to a maximum of 31 minutes), or about what we had anticipated it to be (between 20 and 30 minutes). In examining the actual Rs obtained, the most problematic CVs are clearly #3 and #6. In both of those instances coder disagreement seemed to have been caused by incomplete or ambiguous data. What stands out here is that coding time is unrelated to the error rate (resume #6 took least time and resume #3 took more than the average, yet both vitas had the lowest reliability).

6.3 Technical Problems

One major technical problem involves the fact that the number of variables for each scientist depends on the length of the CV of the respondent. Junior researchers could have as few as 25 variables per CV; seasoned veterans could have thousands. Throw into the mix the problems discussed above with CV truncation and embellishment, making it difficult to discern just what career record we have in hand. Aside from the database problems such variation can cause, there are significant validity issues and issues regarding coding accuracy, coding labor/duration, and coder fatigue.

7.0 Conclusions

Results from this preliminary assessment of the use of CVs as data for research on scientists' and engineers' career trajectories shows that the potential of this approach is mitigated by several practical problems, some easily remedied, some not. Our study seems at least to provide some notion of the likely magnitude of problems.

One of the most basic issues is the availability of CVs and best approaches to obtaining them. We found that obtaining CVs was more difficult than expected. Since CVs are routinely requested for all sorts of purposes, we expected the routine nature of our request would yield considerable returns, at least in the case of the National Science Foundation database. We expected that our sponsorship by the NSF would, if not result in eager participation, at least provide some needed and useful rationale. The 28 percent response from the NSF database sample is lower than we expected. To be sure, there were inaccuracies in the database and, just as important, we did not (for this preliminary assessment) implement a follow-up request. The results show that developing an adequate response likely requires some use of incentives (e.g., a promise of information about the results of the study) and, probably more important, some considerable effort to ensure the accuracy of the address data. Allowing more time for response, in connection with a follow up, will likely yield better results as scientists are highly mobile and sometimes do not receive email quickly. Further, those receiving large quantities of email in all likelihood have some heuristic for triage and, quite likely, our request would not survive the triage.

The results for the web search for CVs were interesting and, despite disappointing returns, the learning curve was such that greater familiarity with particular search engines is likely to yield improvement. At this point, the returns for the web search entail considerable selection effects since there is no reason to believe that the persons with readily available web-based CVs are representative of the entire population of scientists and engineers. While our analysis does not, at this point, permit us to draw inferences based on systematic data, a casual analysis indicated that persons with CVs on the web tend to be younger (in age and career stage), located in universities, and scientists seem better represented than engineers. It seems likely that the number of CVs on the web will continue to increase for some time and, along with the increase, will tend to be more representative. However, Lawrence and Giles (1999) report that search engines typically return hits from only about 16 percent of available websites on the Internet. Since most search engines actually perform their search in advance (this explains the speed of the return hits) and many matches result from links on other sites, a majority of sites may not even turn up in the search output. Worse still is the result of useless matches that bare no relation to the intended search string. However, if the web search has many flaws, the nobtrusiveness of the web CVs has great appeal and, at least, this approach seems useful from preliminary studies, pretests, methodological tests, and, especially, hypothesis development. It provides an expensive means of getting started on CV-based research.

If the acquisition of the CVs was somewhat more difficult than we expected, coding was somewhat less of a burden than expected. With little experience and only limited training, the coders were able to code CVs in less than 30 minutes each, despite collecting a considerable volume of data in a relatively sophisticated codebook (see Appendix C). The coding scheme we began with was, however, simplified a good deal. Developing even more detailed information would certainly break the 30-minute barrier. But we expect that with more experience and training, coding for most tasks can be reduced considerably, within a range acceptable for a large-scale project and a range not exceeding most similar questionnaire-based studies.

The intercoder reliabilty levels were not acceptable in many cases. The coding challenge is considerable, especially if one anticipates using student coders. However, the reliability levels were such that one can expect respectable results after the coders are exposed to a strong coding regime and more of the task becomes mechanized. We also found it useful to take great care determining which observations could be derived from others, thereby reducing both motor coding and interpretation requirements.

What of the validity of the data obtained on the CVs? This preliminary test was not sufficient to provide formal results on accuracy and validity of the CV data. However, it is clearly possible to perform spot checks to determine whether part, and an important part, of the information is complete. With publications (as well as several other output variables) it is a straightforward, though time consuming, matter to check data against actual publications. Similarly, occupational information is subject to audit. In all probability, a more difficult problem is the incompleteness of information on CVs. We found that a great many CVs entailed some sort of truncation. For example, a proportion of the CVs we analyzed mentioned specifically that some earlier years' publications were not reported on the current CV.

A less obvious, but potentially significant, issue in assessing the quality and utility of CV data is phenomenological in nature. Arguably, the CV means very different things to different scientists and engineers and the respective constructions of the CV may have implications for study objectives. This is redolent of the point made by Latour and Woolgar (1986). For example, the CV of the recent graduate is perhaps best thought of as a marketing tool whereby the CV's author seeks to maximize credibility (Bozeman, 1986) to potential employers, using such artifices as seem likely to achieve employment objectives. Similarly, late career scientists may well view the CV more as an historical record, focusing chiefly on the chronicle of output and activities. These are just two of the constructions one might envision and each construction may well embody different motives, different communication strategies (including variance in communicated content), and different CV revision strategies. In all likelihood, purveying a CV on the web can be understood in part as a reflection of a particular set of constructions (at least when the placement of the CV is dictated by the individual and not the institution). The fact that so many scientists have multiple versions of CVs (some more suitable for obtaining grants, others more suitable for employment, still others more suitable for consulting or service) seems to reinforce the notion of multiple roles and constructs for CVs.

In sum, the use of CVs as serious data for social inquiry seems to us to have much potential. But despite the familiarity of this everyday artifact, knowledge of its social meaning, its research utility and the attendant practical problems in its research use is just beginning to accumulate. Using CVs for research is not exactly a brave new world, but an old world seen a new way.

8.0 Appendices.

8.1 Appendix A: Coding Experiment Involving the Use of Curriculum Vitas

To detect problems and improve procedures, we performed an experiment on coding of CVs. We hoped this would accomplish four objectives: (1) to test the codebook to find out what refinements were necessary or desirable; (2) to check intercoder reliability; (3) to identify sources of error; (4) to obtain an estimate of the average length of time needed to code a typical CV (and, subsequently, to adjust accordingly the number of items to be coded.)

After several iterations we developed a reasonable draft of the codebook that was used in the experiment (see Appendix C). We limited the number of variables that appeared in series (e.g., job1, job2, job3...job_n) to not more than 10 for sake of experimental convenience. We also streamlined the codebook by combining several variables into a single variable. Finally, we made the codebook tidier by relegating to the end section all items that asked coders to verbally describe problems they encountered while coding. Altogether, we ended up with eight codebook sections (general demographic and coding information, degree variables, employment items, characteristics of publications, information on external awards, disciplinary memberships, patents and consulting activities, and problems encountered during the coding). The goal was to trim the coding time down to 20 to 30 minutes per CV.

Five coders participated in the experiment. One of them was experienced in the sense that he had been involved in all the stages of data planning, data collection, and codebook development, as well as in trial coding of several CVs beforehand. Three master's students and one undergraduate work-study student—who were all unfamiliar with the procedures, research design, and coding— were also used as coders. The coders were given fifteen minutes of instruction on the structure of the codebook and how some potentially problematic pieces of data should be coded. They were also informed of some peculiarities of the coding protocol (e.g., all the time-dependent variables were to be coded in chronological order <u>except</u> the publication series of variables which were to be coded in reverse chronological order because they share the common reference point of the present).

Ten CVs were selected for coding by all participants in the experiment to assure roughly equal representation of the three sources of data (three CVs each from the targeted industry search and the passive Internet search and four from the targeted agency search). The ten were chosen to represent the range of CVs available—we strove to have as broad a representation as possible (long and short, structured and semi-structured, vitas of young and old scientists and engineers, etc.). At this stage, we also decided to use the CV as it was submitted to us rather than eliminate items—such as teaching innovations, presentations at meetings, committee service—that, while perhaps important, were not relevant for our purposes. There were four main lessons learned from our experiment:

1. Overall, the codebook worked out fairly well in providing a satisfactory framework for coding of CVs. It was not overly complex and managed to capture the most important and most frequently listed variables. Coders were able to code a CV reasonably quickly and could maintain coding for several hours at a time without

much fatigue. Coders identified problems in the codebook that were both obvious and easy to eliminate. Some of them concerned outer its appearance and organization. It was pointed out by coders that there should be a better separation of the sections, we should consider doing all time variables in reverse chronological order, we should add hard page breaks between sections and add section headings. Other issues were more substantive. Examples include the need for an additional code for the Environmental Protection Agency as one of the external funding sources; the need to include a code for two or more funding sources for a single award; the duplication of two codes (702 and 703) for one and the same field (chemical engineering); lack of a field code for "medical sciences, general" or "biomedical sciences;" and need to move the engineering specialties codes up front where they would be more accessible.

2. The summary statistics and the conclusions from the test for intercoder reliability have already been presented in the body of the paper. There are two additional points that warrant consideration. The first is that coder fatigue was not a factor in this experiment, but it might be when we begin coding a large number of CVs in a comparatively short time frame, especially so because we plan to code the whole curriculum vita (and not limit ourselves to 10 jobs, 10 publications, and 10 awards, as in this coding experiment). The second point concerns the possibility of a falsely low coder reliability owing to the compounding of error in the coding of time-dependent series of variables (e.g., if one coder uses a different starting point in time for "first job" than another coder, this will offset the whole sequence of employment history variables).

3. We identified three sources of errors: the codebook, the coders, and the CVs. The first source does not seem to present a serious obstacle since it can easily be revised. For example, although we tried to come up with an exhaustive coding scheme for fields and external funding agencies, we were aware that we would need some postcoding of additional entries that would appear in the vitas. A more worrisome source of errors is the different backgrounds and levels of experience of the coders. What was not a severe problem, in this respect, was a difference in personal preferences, attention spans, or other personal traits. However, given the large number of novice coders that will take part in the full-fledged study, such a source of variation should not be overlooked. The CVs themselves presented quite a few coding problems. Some of them were missing crucial chunks of information (e.g. external funding agencies, dollar amounts of awards). Others were poorly organized (e.g., publications were given in three different sections; employment data were dispersed among several sections). Still others used abbreviations extensively without spelling them out. Convertibility of foreign scientific degrees was also something of a problem. Finally, there was the "resume writer selection bias," namely that some scientists and engineers padded their vitas by including every possible appointment, award, distinction (even those that were temporary or trivial), while others (especially older researchers) often presented only distilled information that they considered essential (e.g., only data on the last 10 years) and omitted relevant and useful data.

4. We have already discussed the average length of time that it takes to code a CV in the main text of the paper. It is probably worth reemphasizing that coding time was

unrelated to the error rate. The slowest (surprisingly, the only experienced participant in the experiment) and the fastest coders turned out to be the most accurate.

8.2 Appendix B: Seven Items That Did Not Pass the Reliability Test

Item #1: Type (title) of Job #1

Academic Track

100. Graduate Teaching Assistant
101. Graduate assistant other (including Graduate Research Assistant)
102. Postdoctoral position
103. Lecturer/Instructor
104. Assistant Professor
105. Associate Professor
106. [Full] Professor
107. Visiting position
108. Administrator (research)
109. Administrator (other)
110. Research position
111. Laboratory technician
112. Medical intern
113. Medical resident
114. Academic other

Industry

200. Graduate assistant/Intern (non medical)
201. Postdoctoral position
202. Visiting position
203. Chief Exec. Officer (CEO)/President
204. Administrator (research)
205. Administrator (other)
206. Research position
207. Laboratory technician
208. Medical intern
209. Medical resident

210.Industry other

Government

300.Graduate Assistant/Intern

301.Postdoctoral position

302.Visiting position

303.Administrator (research) (includes program officer/director)

304.Administrator (other)

305.Research position

306.Laboratory technician

307.Medical intern

308.Medical resident

309.Government other

Other

400.Other

Item #2: Type (title) of Job #2

(See Code List for Job #1)

Item #3: Publication #4 type

1. Only author, journal article
2. First author, journal article
3. Multiple (not first) author, journal article
4. Only author, book or book chapter
5. First author, book or book chapter
6. Multiple (not first) author, book or chapter

(ENTER IN REVERSE CHRONOLOGICAL ORDER)

Item #4: External (not university) funding source for award #1 <u>Federal Government</u> 100.Department of Energy (DOE) Office of Science (OS) 101.DOE other 102.National Institutes of Health (NIH) 103.National Science Foundation (NSF) centers programs 104.NSF other 105.National Aeronautics and Space Administration (NASA) 106.Defense Applied Research Projects Agency (DARPA) or Office of Naval Research (ONR) Department of Defense (DOD) other 107.Department of Agriculture (USDA) <u>State (other) Government (Not University)</u> 200.State government agency (any state, any agency) 201.Other governmental unit <u>Industry</u> 300.Company 301.Company Foundation <u>Foundations and Other</u> 400.Private or independent foundation (Not National Science Foundation) 401.Other
Item #5: External (not university) funding source for award #2 (See List of Codes for Award #1'
Item #6: External (not university) funding source for award #3 (See List of Codes for Award #1'
Item #7 External (not university) funding source for award #5 (See List of Codes for Award #1'

BEST COPY AVAILABLE

8.3 Appendix C: Codebook for RVM Pilot Study of CVs—10/18/99

QUESTION	CODES AND NOTES
Coder Last Name	Key in your last name
Date RVM requested resume	Printed on resume
Date resume was received	Printed on resume
ID number	Printed on resume
Response wave number	Printed on resume
Is resume full version or partial version?	1. Full 2. Partial (Leave blank if you are not sure)
Sex of respondent	1. Male 0. Female (Leave blank if you are not sure)
Year of birth	YYYY (if indicated on CV)
National origin	Key in as indicated on CV (otherwise assume US)
Citizenship	Key in as indicated on CV (otherwise assume US)

Degree #1 Type	1. Associate's 2. Bachelor's 3. Master's 4. Engineering Master's + 5. Doctoral 6. MD 7. Law 8. Other (Record in chronological order)
Degree #1 field	
Degree #1 year	YYYY
Degree #1 institution	Institution where earned (See abbreviation list)
Degree #2 Type	(See List for Degree #1)

Degree #2 field	
Degree #2 year	YYYY
Degree #2 institution	Institution where earned (See abbreviation list)
Degree #3 Type	(See List for Degree #1)
Degree #3 field	
Degree #3year	YYYY
Degree #3 institution	Institution where earned (See abbreviation list)
Degree #4 Type	(See List for Degree #1)
Degree #4 field	
Degree #4 year	YYYY
Degree #4 institution	Institution where earned (See abbreviation list)
Degree #5 Type	(See List for Degree #1)
Degree #5 field	
Degree #5 year	YYYY
Degree #5 institution	Institution where earned (See abbreviation list)
Degree #6 Type	(See List for Degree #1)
Degree #6 field	
Degree #6 year	YYYY
Degree #6 institution	Institution where earned (See abbreviation list)
Degree #7 Type	(See List for Degree #1)
Degree #7 field	
Degree #7 year	YYYY
Degree #7 institution	Institution where earned (See abbreviation list)
Degree #8 Type	(See List for Degree #1)

Degree #8 field	
Degree #8 year	YYYY
Degree #8 institution	Institution where earned (See abbreviation list)

Degree #9 Type	(See List for Degree #1)
Degree #9 field	
Degree #9 year	YYYY
Degree #9 institution	Institution where earned (See abbreviation list)
Degree #10 Type	(See List for Degree #1)
Degree #10 field	
Degree #10 year	YYYY
Degree #10 institution	Institution where earned (See abbreviation list)

Type (title) of Job #1	Academic Track 115.Graduate Teaching Assistant 116.Graduate assistant other (including Graduate Research Assistant) 117.Postdoctoral position 118.Lecturer/Instructor 119.Assistant Professor 120.Associate Professor 121.[Full] Professor 122.Visiting position 123.Administrator (research) 124.Administrator (other) 125.Research position 126.Laboratory technician 127.Medical intern 128.Medical resident 129.Academic other Industry

	211.Graduate assistant/Intern (non medical) 212.Postdoctoral position 213.Visiting position 214.Chief Exec. Officer (CEO)/President 215.Administrator (research) 216.Administrator (other) 217.Research position 218.Laboratory technician 219.Medical intern 220.Medical resident 221.Industry other Government 310.Graduate Assistant/Intern 311.Postdoctoral position 312.Visiting position 313.Administrator (research) (includes program officer/director) 314.Administrator (other) 315.Research position 316.Laboratory technician 317.Medical intern 318.Medical resident 319.Government other Other 401.Other
Year JOB #1 began	YYYY
Year JOB #1 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #1 held	Key in actual (See abbreviation list)
Type (title) of Job #2	(See Code List for Job #1)
Year JOB #2 began	YYYY
Year JOB #2 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #2 held	Key in actual

	(See abbreviation list)
Type (title) of Job #3	(See Code List for Job #1)
Year JOB #3 began	YYYY
Year JOB #3 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #3 held	Key in actual (See abbreviation list)
Type (title) of Job #4	(See Code List for Job #1)
Year JOB #4 began	YYYY
Year JOB #4 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #4 held	Key in actual (See abbreviation list)
Type (title) of Job #5	(See Code List for Job #1)
Year JOB #5 began	YYYY
Year JOB #5 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #5 held	Key in actual (See abbreviation list)
Type (title) of Job #6	(See Code List for Job #1)
Year JOB #6 began	YYYY
Year JOB #6 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #6 held	Key in actual (See abbreviation list)
Type (title) of Job #7	(See Code List for Job #1)
Year JOB #7 began	YYYY
Year JOB #7 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #7 held	Key in actual (See abbreviation list)
Type (title) of Job #8	(See Code List for Job #1)

Year JOB #8 began	YYYY
Year JOB #8 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #8 held	Key in actual (See abbreviation list)
Type (title) of Job #9	(See Code List for Job #1)
Year JOB #9 began	YYYY
Year JOB #9 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #9 held	Key in actual (See abbreviation list)
Type (title) of Job #10	(See Code List for Job #1)
Year JOB #10 began	YYYY
Year JOB #10 ended	YYYY (USE 9999 FOR CURRENT POSITION)
Institution where JOB #10 held	Key in actual (See abbreviation list)

Publication #1 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #1	YYYY
Publication #2 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #2	YYYY
Publication #3 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter

	5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #3	YYYY
Publication #4 type	7. Only author, journal article 8. First author, journal article 9. Multiple (not first) author, journal article 10. Only author, book or book chapter 11. First author, book or book chapter 12. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #4	YYYY
Publication #5 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #5	YYYY
Publication #6 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #6	YYYY
Publication #7 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #7	YYYY
Publication #8 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter

	5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #8	YYYY
Publication #9 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #9	YYYY
Publication #10 type	1. Only author, journal article 2. First author, journal article 3. Multiple (not first) author, journal article 4. Only author, book or book chapter 5. First author, book or book chapter 6. Multiple (not first) author, book or chapter (ENTER IN REVERSE CHRONOLOGICAL ORDER)
Year of publication #10	YYYY

Dollar amount of grant or contract #1	Key in figure (no $ sign, no commas)
External (not university) funding source for award #1	<u>Federal Government</u> 108.Department of Energy (DOE) Office of Science (OS) 109.DOE other 110.National Institutes of Health (NIH) 111.National Science Foundation (NSF) centers programs 112.NSF other 113.National Aeronautics and Space Administration (NASA) 114.Defense Applied Research Projects Agency (DARPA) or Office of Naval Research (ONR) Department of Defense (DOD) other 115.Department of Agriculture (USDA) <u>State (other) Government (Not University)</u> 202.State government agency (any state, any agency) 203.Other governmental unit 300.Industry 301.Company 302.Company Foundation

	400.Foundations and Other 401.Private or independent foundation (Not National Science Foundation) 402.Other

Dollar amount of grant or contract #2	Key in figure (no $ sign, no commas)
External (not university) funding source for award #2	(See List of Codes for Award #1'
Dollar amount of grant or contract #3	Key in figure (no $ sign, no commas)
External (not university) funding source for award #3	(See List of Codes for Award #1'
Dollar amount of grant or contract #4	Key in figure (no $ sign, no commas)
External (not university) funding source for award #4	(See List of Codes for Award #1'
Dollar amount of grant or contract #5	Key in figure (no $ sign, no commas)
External (not university) funding source for award #5	(See List of Codes for Award #1'
Dollar amount of grant or contract #6	Key in figure (no $ sign, no commas)
External (not university) funding source for award #6	(See List of Codes for Award #1'
Dollar amount of grant or contract #7	Key in figure (no $ sign, no commas)
External (not university) funding source for award #7	(See List of Codes for Award #1'
Dollar amount of grant or contract #8	Key in figure (no $ sign, no commas)
External (not university) funding source for award #8	(See List of Codes for Award #1'
Dollar amount of grant or contract #9	Key in figure (no $ sign, no commas)
External (not university) funding source for award #9	(See List of Codes for Award #1'
Dollar amount of grant or contract #10	Key in figure (no $ sign, no commas)
External (not university) funding source for award #10	(See List of Codes for Award #1'

Disciplinary society membership #1	Key in actual (DO NOT INCLUDE AAAS,

	GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #2	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #3	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #4	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #5	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #6	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #7	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #8	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #9	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)
Disciplinary society membership #10	Key in actual (DO NOT INCLUDE AAAS, GREEK LETTER ORGANIZATIONS)

Year of patent #1	YYYY
Patent #1 licensed or sold?	1. Patent licensed or sold
Year of patent #2	YYYY
Patent #2 licensed or sold?	1. Patent licensed or sold
Year of patent #3	YYYY
Patent #3 licensed or sold?	1. Patent licensed or sold
Year of patent #4	YYYY
Patent #4 licensed or sold?	1. Patent licensed or sold
Year of patent #5	YYYY

Patent #5 licensed or sold?	1. Patent licensed or sold
Year of patent #6	YYYY
Patent #6 licensed or sold?	1. Patent licensed or sold
Year of patent #7	YYYY
Patent #7 licensed or sold?	1. Patent licensed or sold
Year of patent #8	YYYY
Patent #8 licensed or sold?	1. Patent licensed or sold
Year of patent #9	YYYY
Patent #9 licensed or sold?	1. Patent licensed or sold
Year of patent #10	YYYY
Patent #10 licensed or sold?	1. Patent licensed or sold
Did respondent perform consulting activity?	PLACEHOLDER
Did you encounter any problems in coding this CV?	Key in description of problem
Year of patent #1	YYYY

Table 1. Search Results from Alta Vista with Refined Key Words Combinations.

Key Word Combination	Number of Hits	Approximate Noise Level (%)
Curriculum and vitae and biotechnology	1,879,550	Very High 90 - 95
Curriculum vitae and biotechnology and not job and not jobs	2,629	High 70-80
(curriculum vitae or vita or resume) and biotechnology and not job and not jobs and personal	1,342	Medium 50-60

Table 2. Intercoder Reliability and Time of Coding for 10 Resumes and 37 Items

Curriculum Vita	Rs(r)	Time per resume (in minutes)
1	.897	23.0
2	.797	31.0
3	.651	27.8
4	.839	23.4
5	.868	24.0
6	.608	15.0
7	.756	30.4
8	.800	19.4
9	.728	19.8
10	.714	22.2
Mean	.766	23.6

Note: Rs(r) stands for resume intercoder reliability.

Table 3. Intercoder Reliability for Coding 37 Items

Item #	Rs	Item #	Rs(i)
1	.933	20	.600
2	.933	21	.740
3	.960	22	.790
4	.920	23	.630
5	.920	24	.550
6	.880	25	.880
7	.690	26	.620
8	.860	27	.570
9	.880	28	.630
10	.960	29	.580
11	.810	30	.690
12	1.000	31	.560
13	1.000	32	.690
14	1.000	33	.660
15	1.000	34	.660
16	.560	35	.580
17	.490	36	.893
18	.610	37	.933
19	.680		

Note: Rs(i) stands for item intercoder reliability.

9.0 Reference List

Babbie, E. (1992) The Practice of Social Research. Sixth Edition. Belmont, CA: Wadsworth Publishing Co.

Baldi, S. (1998) "Normative versus Social Constructivist Processes in the Allocation of Citations: A Network-Analytic Model." American Sociological Review, 63(6): 829-847.

Becker, G. (1962) "Investment in human capital: A theoretical analysis." Journal of Political Economics, 70(5): S9-S49

Bohrnstedt, G. and D. Knoke (1988) Statistics for Social Data Analysis, Second Edition. Itasca, IL: F.E. Peacock Publishers, Inc.

Bonzi, S. (1992) "Trends in Research Productivity among Senior Faculty." Information Processing & Management 28(1): 111-120.

Bourdieu, P. (1986) "The forms of social capital." In J. G. Richardson (Ed.), Handbook of theory and research for the sociology of education. New York: Greenwood.

Bourdieu, P. and L. Wacquant (1992) An invitation to reflexive sociology. Chicago: University of Chicago Press.

Bozeman, B. (1986) "The credibility of policy analysis: Between method and use." Policy Studies Journal, 14(4): 519-539.

Bozeman, B., J.S. Dietz, and M. Gaughan (forthcoming) "Scientific and technical human capital: An alternative model for research evaluation." International Journal of Technology Management.

Cole, J. and H. Zuckerman. (1984) "The Productivity Puzzle: Persistence and Change in Patterns of Publication of Men and Women Scientists." Advances in Motivation and Achievement, 2: 217-258.

Coleman, J.C. (1988) "Social capital in the creation human capital." American Journal of Sociology, 94: S95-S120

Coleman, J.C. (1990) Foundations of social theory. Cambridge: The Belknap Press of Harvard University Press.

Craig, R. (1981) "Generalization of Scott's Index of Intercoder Agreement." Public Opinion Quarterly, 45(2): 260-264.

Crittenden, K. and R. Hill. (1971) "Coding Reliability and Validity of Interview Data." American Sociological Review, 36(6): 1073-1080.

Diamond, A.M. (1984) "An economic model of the life-cycle research productivity of scientists." Scientometrics, 6(3): 189-196.

Diamond, A.M. (1986) "The life-cycle research productivity of mathematicians and scientists." Journal of Gerontology, 41(4): 520-525.

Elder, G.H. Jr. (1994) "Time, human agency, and social change: perspectives on the life course." Social Psychology Quarterly, 57: 4-15.

Elder, G.H. Jr. and E.K. Pavalko (1993) "Work Careers In Men's Later Years Transitions, Trajectories, And Historical Change." Journals of Gerontology, 48(4): S180-S191.

Fleiss, J. (1971) "Measuring Nominal Scale Agreement among Many Raters." Psychological Bulletin 76: 378-382.

Fowler, F.J. (1988) Survey Research Methods. revised edition, Newbury Park, CA: Sage.

Funkhouser, G. and E. Parker (1968) "Analyzing Coding Reliability: The Random-Systematic-Error Coefficient." Public Opinion Quarterly, 32(1): 122-128.

Gomez-Mejia, L. and D. Balkin (1992) "Determinants of Faculty Pay: An Agency Theory Perspective." Academy of Management Journal, 35(5): 921-955.

Heinsler, J. and R. Rosenfeld. (1987) "Charting Academic Careers." Social Studies of Science, 17: 135-144.

Hurlbert, J. and R. Rosenfeld. (1992) "Getting a Good Job: Rank and Institutional Prestige in Academic Psychologists' Careers." Sociology of Education, 65(3): 188-207.

Hyatt, C.W. and W.D. Hopkins (1998) "Interspecies object exchange: Bartering in apes?" Behavioural Processes, 42 (2-3): 177-187.

Landis, J. and G. Koch. (1977) "The Measurement of Observer Agreement for Categorical Data." Biometrics, 33: 159-174.

Lawrence, S. and Giles, C.L. (1999) "Accessibility of information on the web." Nature, 400 (6): p. 107.

Levin, S.G. and P.E. Stephan (1991) "Research productivity over the life cycle: Evidence for academic scientists." American Economic Review, 81(1): 114-132.

Long, J.S., P. Allison, and R. McGinnis. (1993) "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity." American Sociological Review, 58: 703-722.

Long, J.S. and R. McGinnis (1985) "The effects of the mentor on the academic career." Scientometrics, 7(3-6): 255-280.

Merton, R.K. (1973) "The Matthew Effect in Science." In R.K. Merton, The Sociology of Science: Theoretical and Empirical Investigations. Chicago: University of Chicago Press.

Merton, R.K. (1993) On The Shoulders Of Giants: A Shandean Postscript. Chicago: University of Chicago Press.

Mincer, J. (1958) "Investment in human capital and personal income distribution." Journal of Political Economy, (Aug)

Montgomery, A. and K. Crittenden (1977) "Improving Coding Reliability for Open-Ended Questions." Public Opinion Quarterly, 41(2): 235-243.

Polanyi, M. (1967) The tacit dimension. London: Cox & Wyman.

Polanyi, M. (1969) The logic of tacit inference. Knowing and being. London: Routledge & Kegan Paul.

Reskin, B.F. (1977) "Scientific productivity and the reward structure of science." American Sociological Review, 42(June): 491-504.

Rogers, E.M. (1995) Diffusion of Innovations. Fourth Edition, New York: The Free Press.

Schultz, T.W. (1963) The economic value of education. New York: Columbia University Press.

Scott, J. (1991) Social Network Analysis. London: Sage Publications.

Scott, W. (1955) "Reliability of Content Analysis: The Case of Nominal Scale Coding." Public Opinion Quarterly, 19: 321-325.

Simonton, D.K. (1997) "Creative productivity: A predictive and explanatory model of career trajectories and landmarks." Psychological Review, 104(1): 66-89.

Singleton, R.A. Jr., B. Straits, and M. Straits. (1993) Approaches to Social Research. Second Edition. New York: Oxford University Press.

Stephan, P.E. (1996) "The economics of science." Journal of Economic Literature, 34: 1199-1235.

Stephan, P.E. and S.G. Levin (1997) "The critical importance of careers in collaborative scientific research." Revue d'Economie Industrielle, 79: 45-61.

Xie, Y. and K. Shauman (1998) "Sex Differences in Research Productivity: New Evidence about an Old Puzzle." American Sociological Review, 63: 847-870.

Zuckerman, H.A. (1977) The Scientific Elite. New York: Free Press.

Zuckerman, H. and R.K. Merton (1973) "Age, Aging, and Age Structure in Science". In R.K. Merton, The Sociology of Science: Theoretical and Empirical Investigations. Chicago: University of Chicago Press.

TM030363

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Using the Curriculum Vita to Study the Career Paths of Scientists and Engineers: An Assessment

Author(s): E. Lane, J. Dietz, I. Chompalov, B. Bozeman, J. Park

Corporate Source: Georgia Institute of Technology

Publication Date: November 3, 1999

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☒	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here,→ please

Signature: Eliesh O'Neil Lane	Printed Name/Position/Title: Eliesh O'Neil Lane, Sr. Research Associate
Organization/Address: Georgia Institute of Technology, School of Public Policy, Atlanta, GA 30332-0345	Telephone: 404-894-9455 FAX: 404-894-9372 E-Mail Address: eliesh.lane@pubpolicy.gatech.edu Date: Nov. 15, 1999

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com