
What Policymakers Need and Must Demand From Research Regarding the Employment Rate of Persons with Disabilities

Robert Silverstein, B.S., J.D.,^{*,†} George Julnes, B.S., M.S., Ph.D., M.P.P., M.B.A.,[‡] and Renee Nolan, B.S.W.[§]

Enabling persons with disabilities to prepare for and engage in gainful employment has become a priority concern in U.S. society. As a result, federal and state policymakers are adopting employment-related initiatives designed to enhance the employment rate of persons with disabilities. Policymakers need credible evidence in order to assess and reform these initiatives. This recognized need, however, begs the question of what constitutes “credible evidence.” Of particular concern in the disability policy arena is the debate over the types of conclusion about employment rates that can and cannot be drawn from analyses of national survey data sets. This article connects standard research methodology concepts with the complexities of evaluating disability policy to help stakeholders appreciate the issues involved in this debate. This appreciation can help policymakers (1) recognize unwarranted cause-and-effect conclusions based solely on existing national survey data and (2) demand better data and stronger research designs to complement the potential over-reliance on

*Correspondence to: Robert Silverstein, CSADP. E-mail: Bobby@csadp.org

†Director, Center for the Study and Advancement of Disability Policy (CSADP); B.S. in Economics, The Wharton School, University of Pennsylvania; J.D., Georgetown University Law Center. This article was funded in part by a grant from the National Institute on Disability and Rehabilitation Research, U.S. Department of Education supporting the Rehabilitation Research and Training Center on Workforce Investment and Employment Policy for Persons with Disabilities (Grant No. H133B9800420) and the Rehabilitation Research and Training Center on State Systems and Employment (Grant No. H133B30067). This article was also funded in part by a grant from the Robert Wood Johnson Foundation. The opinions contained in this article are those of the authors and do not necessarily reflect those of the U.S. Department of Education, the Robert Wood Johnson Foundation, or Utah State University. For related projects see the CSADP website at <http://www.disabilitypolicycenter.org>

‡Associate Professor of Psychology, Utah State University; B.S., Psychology and Philosophy, University of Wisconsin, Madison; M.P.P., Institute of Public Policy Studies, University of Michigan; M.B.A., School of Business Administration, University of Michigan; and Ph.D. Psychology, University of Hawaii—Manoa.

§Doctoral student in Psychology, Research and Evaluation Methodology Program, Utah State University.

Contract/grant sponsor: National Institute on Disability and Rehabilitation Research, U.S. Department of Education; contract/grant numbers: H133B9800420; H133B30067.

Contract/grant sponsor: Robert Wood Johnson Foundation.

correlational studies using problematic survey data to estimate policy impacts. To this end, the article concludes with a practical framework with a checklist for assessing the adequacy of research regarding the employment rate of persons with disabilities. Copyright © 2005 John Wiley & Sons, Ltd.

INTRODUCTION

Enabling persons with disabilities to prepare for and engage in gainful employment commensurate with their strengths, abilities, and capabilities is a priority concern to federal and state policymakers. As a result, policymakers are adopting initiatives designed to improve the employment rate of working-aged persons with disabilities, while at the same time maintaining appropriate income protection and health care programs. Policy initiatives adopted in recent years include laws relating to civil rights protections, income protection (cash assistance), health coverage, employment-related services and supports, asset development and retention, and tax policy. Policymakers need to know whether these initiatives are succeeding in enhancing economic self-sufficiency for persons with disabilities and achieving other goals of disability policy, such as full participation (self-determination) and independent living, and if not why not and what alternative approaches might be more successful.

The value placed on well informed policymakers is embodied in the current emphasis in government on evidence-based decision making. This emphasis, however, highlights the importance of developing and applying a consensus on what constitutes adequate, or actionable, evidence for a given policy area. The challenges and opportunities for policy analysis in providing evidence adequate for informing disability policy are on display in the 2003 volume edited by David C. Stapleton and Richard V. Burkhauser entitled *The Decline in Employment of People with Disabilities—A Policy Puzzle* (hereinafter referred to as *Stapleton and Burkhauser*). Because of the centrality of employment as an outcome and the divergent conclusions reached by different researchers and others regarding the appropriate use of data in national surveys, the book has sparked an important debate over what policymakers and other stakeholders involved in the policymaking process need and must demand from research regarding the employment rate of working-aged people with disabilities.

The debate regarding the appropriate use of data in national surveys is occurring in four related domains:

- the quality of the data (i.e. reliability and validity),
- the validity of the comparisons being used to identify employment trends for persons with disabilities,
- the adequacy of the analyses being used to support cause-and-effect conclusions, and
- the utility of the research in informing policymakers and other stakeholders involved in the policymaking process about the effectiveness or value of recent policy changes.

Our concern is that problems in each of these four domains may be combining to result in misleading conclusions being derived from a fairly common policy research paradigm. The purpose of this article, therefore, is to help policymakers, researchers, and others involved in the policymaking process appreciate the issues involved in this debate so that they are better able to (1) make sense of available research to guide policy reform and (2) promote, argue for, and demand research approaches that would yield even more informative findings for guiding policy reform.

To serve this purpose, we first identify and examine (from a policymaker's perspective) the key factors that should be considered in evaluating current public policy initiatives designed to enhance the employment rate of persons with disabilities while maintaining appropriate income protection and health care programs. This examination involves presenting a standard set of criteria for judging the adequacy of research-based evidence and an overview of the complexities that challenge our efforts to conduct quality research. We do not presume that researchers in this area are unaware of the criteria for good research nor of the specific challenges in evaluating disability-related policies. Our sense, however, is that these issues need to be raised together to provide policymakers with the perspective needed to judge the adequacy of policy research.

With this perspective in mind, we then analyze the extent to which current research efforts relying on national data sets have properly and appropriately considered these issues. For this analysis, we review the evidence regarding employment-related trends, including trends in the overall employment rate of persons with disabilities, the employment rate of persons with disabilities who are able and available to work, and the proportion of persons reporting work limitations and an inability to work. Much of the evidence reviewed has been developed by researchers who contributed chapters to *Stapleton and Burkhauser*, but we have brought in other evidence as well. Much of what we cover from these researchers and scholars is presented with extensive quoting of their statements or light paraphrasing. We do this because of the sensitivity of the subject and our desire to represent the differing points of view correctly. Further, because of the extensive use of national data sets in *Stapleton and Burkhauser* and in the broader policy debate, we devote most of our attention to quantitative research. We hope, however, that our review encourages policymakers to appreciate the importance of using multiple research approaches, including mixed methods that blend experimental, quasi-experimental, and correlational designs with qualitative analyses, to guide policy decisions.

Finally, we organize the lessons learned from a review of the policy-related research and offer a practical framework for promoting a critical appreciation of the evidence available and attainable in the arena of public policy initiatives designed to enhance the employment rate of persons with disabilities. The appendix makes this framework concrete with a suggested checklist for assessing the adequacy of research regarding the employment rate of persons with disabilities.

Because the development of our research critique and associated suggestions covers a wide range of material, we preface our presentation with an overview that summarizes our main points. Our goal in this is that readers might better identify where they agree or disagree with our conclusions and thus be more reflective about the evidence and logic we use to support those conclusions.

SUMMARY OF MAJOR FINDINGS AND CONCLUSIONS

Social science research plays a critical role in informing policymakers and others regarding efforts to design, implement, and evaluate policy options for improving the employment rate of persons with disabilities and determining appropriate income protections and health care programs. Burkhauser and Stapleton (2003b, p. 372) made the following statement in regard to the role of social science in informing the policymaking process:

[O]bjective evidence is always controversial to those unwilling to allow their policy hopes to be tempered by reality. When public policy is subject to the scrutiny of social science, the outcome of the research is, by its nature, uncertain. Yet, an unflinchingly objective examination of the evidence is critical if we are to learn from the past and improve the employment opportunities for working-aged people with disabilities while maintaining appropriate income protection for those unable to work.

From our perspectives as a former staffer to a Federal policymaker¹ and as a contributor to evaluation theory,² we agree with this statement by Burkhauser and Stapleton. While evidence is rarely as “objective” or unambiguous as one would like, the Burkhauser and Stapleton goal is appropriate. To the extent possible, policymakers do need and, in fact, should demand “unflinchingly objective examination of the evidence” to aid them to carry out their public policy responsibilities.

Our overarching concern is that the desire to be “objective” may lead to an unwarranted over-reliance on one source of information that might bias our understanding and mislead policymakers (see Kirchner, 1996). This potential for bias might seem benign if it were only a matter of academic debates and the associated prestige of competing research paradigms. The policy community, however, is influenced by the outcomes of these debates and, as a result, real people may either benefit or suffer accordingly.

To highlight what we see as important for informed policy debates, we summarize five points regarding the conclusions of this article about what policymakers and others involved in the policymaking process need and must demand from policy-related research regarding the employment rate of persons with disabilities.

First, policymakers must understand the degree of controversy regarding the appropriate use of so-called “work limitation” questions in the national data sets. The three primary sources of national survey data on trends in the employment rate of persons with disabilities are the Current Population Survey (CPS), the National Health Interview Survey (NHIS), and the Survey of Income and Program Participation (SIPP). Each survey has its strengths and limitations; all use different questions to identify the population of persons with disabilities. A major component of the definition in all three surveys is one or two questions on limitations in work.

¹Robert Silverstein worked for Senator Tom Harkin (D. Iowa) in various capacities, including staff director and chief counsel for the Subcommittee on Disability Policy, Committee on Labor and Human Resources, U. S. Senate.

²George Julnes has directed many policy evaluations, including current work in Utah on the Medicaid Infrastructure Grant and the Social Security Administration (SSA) §1 for §2 Benefit Offset Demonstration Pilot. His scholarly contributions include Julnes and Mark (1998), Julnes (2004), and Mark, Henry, and Julnes (2000).

The CPS includes the questions “Does anyone in this household have a health problem or disability which prevents him or her from working or which limits the kind or amount of work they can do? If so, who is that?” The NHIS asks “Does any impairment or health problem now keep [respondent] from working at a job or business? Is [respondent] limited in the kind or amount of work [respondent] can do because of any impairment?”. Finally, the SIPP addresses work limitations with “Does [respondent] have a physical or mental or other health condition which limits the kind or amount of work [respondent] can do?”.

The meaning that we may derive from these limited questions is controversial. Several individuals and organizations believe that it is inappropriate to use work limitation questions to identify people with disabilities when studying employment trends and question the extent to which the associated findings may be generalized (see Hale, 2001). Indeed, the National Council on Disability (2002, p. 27) expressed concern that CPS data could lead to “ineffective or even dangerous public policy decisions” based on a “Federal consensus that certain CPS items are not adequately designed to elicit accurate and reliable information from people with disabilities.”

Some researchers, however, have chosen to study trends in the employment rate of persons with disabilities by focusing on a population identified by these work disability questions, despite their own recognition of the limitations in the use of the data. For example, Burkhauser and Stapleton draw near-definitive conclusions from these data despite statements acknowledging that the work-limitation question over-represents those with impairments who are not employed and misses those with impairments who are sufficiently integrated into the workforce that they do not report being work limited (Burkhauser & Stapleton, 2003a, p. 7, Burkhauser, Houtenville, & Wittenburg, 2003, p. 36). These limitations are exacerbated in the typical use of these survey data in trend analyses that examine change over time. Examples include studying trends in the overall employment rate of persons with disabilities, the employment rate of persons with disabilities able and available to work, and the proportion of persons reporting work limitations. Others using these national data sets did not rely on the work limitation questions, but rather used other data in the national surveys and, as a result, reached different conclusions about trends in the employment rate of persons with disabilities (Kruse & Schur, 2003a).

Second, researchers relying only on national data sets are unable to take into consideration critical factors such as the social, political, and economic context in which the policy is implemented. These factors are important because policy-related research is dynamic and complex. That is, the impact of public policy may vary greatly in light of the interplay among specific factors such as the characteristics of the subpopulations of persons with disabilities, the significant policy differences/inconsistencies among policy initiatives and among the states, and the degree of compliance and implementation. An understanding of these factors requires information beyond the national data sets. As a result, regardless of the sophisticated analyses conducted, there are serious limitations to the ability of researchers and others in the policy community to “tease out” or isolate the effects of a federal policy on the employment rate of persons with disabilities using only employment data from existing national surveys.

Third, researchers conducting policy-related research that relies on national data sets should be held to the Government Auditing Standards applicable to

research conducted by the General Accounting Office (GAO). The GAO was established by Congress to provide research to members of Congress and their staff making policy-related decisions. According to the Government Auditing Standards (U.S. General Accounting Office, 2003, § 7.61; 2003 Revisions), findings, conclusions, and recommendations must be based on reliable and valid data. As elaborated below, reliability refers to the consistency of the measured results (e.g. across raters or across time); measurement validity refers to whether researchers are measuring the constructs that they intended to measure (e.g. whether self-report questions on a survey are distinguishing those with a disability and those without). When auditors are unable to obtain sufficient, competent, and relevant evidence about the validity and reliability of the data, they may find it necessary to use the data, but they must indicate in their report the data's limitations and "refrain from making unwarranted conclusions or recommendations." Accordingly, we argue that there are enough sources of potential bias (including questionable measurement and threats to internal validity) that it is unwarranted to use data from the work limitation questions in these national data sets as the sole basis for drawing cause and effect policy conclusions. For example, analyses of these national data sets should not be presented as having demonstrated that a particular federal policy *caused* the decline in the employment rate of persons with disabilities.

Fourth, the definition of disability used by researchers should comport with the scope and purpose of the policy question addressed. It is not valid from a research point of view to draw conclusions about the effect of a program that has a limited purpose and addresses a limited population by using data sources that are not limited to the specified purpose or population. For example, given that the limited purpose of Title I of the Americans with Disabilities Act (ADA) is to protect the subpopulation of qualified persons with disabilities (i.e. persons who can perform the essential functions of the job, with or without a reasonable accommodation) from discrimination by employers, it is not valid to draw policy-related conclusions about the effect of the program by using data sources that focus on a different subpopulation (e.g. using the CPS work limitation data that focuses on persons with impairments who report that they are unable to work or otherwise are limited in work; for further discussion, see Blanck, Schur, Kruse, Schwochau, & Song, 2003).

Fifth, there is consensus that the quality of existing disability-related data is inadequate to support the needs of policymakers, researchers, and other stakeholders involved in the policymaking process. Policymakers, researchers, and others need better data sources to ascertain the overall employment rate of persons with disabilities that is based on the new paradigm of disability policy, includes the full spectrum of persons with physical or mental impairments, and allows researchers to disaggregate the data based on subpopulations of persons with disabilities (i.e. given the heterogeneity of the population containing persons able to work, persons able and available to work, and persons unable to work).

These five points support our pragmatic conclusion—to improve the quality, utility, and relevance of policy-related research regarding the employment of persons with disabilities, researchers should not rely solely on the use of national data sets. Rather, to provide evidence worthy of informing the policy debates, we must promote the use of multiple research strategies and data collection efforts, including the use of mixed research methods.

CRITERIA FOR ACTIONABLE POLICY EVIDENCE

There is no shortage of “evidence” being offered to support or to challenge a particular policy such as the ADA. On what basis are we to judge some evidence as adequate for guiding action, or actionable, and other evidence as inadequate? This section lays out the criteria developed in field of research methodology that are used routinely for judging the adequacy of research findings (see Cook & Campbell, 1979; Cronbach, 1989; Shadish, Cook, & Campbell, 2002). Keeping these criteria in mind helps clarify our critique of current research and promote efforts to develop a more appropriate evidence base for policymakers in the disability policy arena.

The central goal of most policy research is to understand the effects of identified policies of interest (Mark et al., 2000). In the disability area, primary questions include the following:

- What are the effects of the antidiscrimination protections embodied in the ADA on the employment rate of qualified persons with disabilities?
- What are the effects of programs providing employment-related services and supports on the employment rate of persons with disabilities targeted by the programs?
- What impact do work disincentives have on the willingness of persons with disabilities receiving cash benefits to risk working, and how may we modify current policies to reduce work disincentives and make work pay?

Related questions consider whether the impacts of particular policies vary across persons with different capacities and what aspects of the policies produce these effects (e.g. policy changes that reduce barriers or increase incentives; see Julnes & Mark, 1998). It is important for policymakers to appreciate that it is challenging to design studies that yield straightforward answers to these and similar policy questions. Some of the challenges involved are inherent in social science research; others are complicated by the context of public policy (Weiss, 1998). In this section, we summarize lessons learned by policy researchers in terms of four criteria that need to be met for the results from policy research to be considered adequate evidence for policy debates.

- Are we measuring what we claim to be measuring?
- Are we making the right comparisons to help us understand policy impacts?
- Are we using appropriate analyses to reveal important relationships?
- Are we using indicators of success that embody the values that we want to promote through public policy?

Note that the adequacy of evidence is often a relative matter (what is judged adequate in a new and uncharted policy arena may be dismissed in a more developed arena), but accepting evidence that misleads us is always a problem. Keeping this in mind, the following section reviews the specific parameters of the disability policy context that need to be considered in assessing the value of particular research approaches in this field.

Are We Measuring What We Claim to Be Measuring?

One of the most basic criteria for policy analysis is for it to be based on valid measurement. The short definition of valid measurement is that you are measuring

what you intend to measure. A prerequisite for valid measurement is that the measure employed is reliable. This means that the measure yields consistent results when what is being measured is itself consistent. For example, if different reviewers arrive at different assessments of whether an individual is considered an “individual with a disability,” or whether the individual is precluded or limited in his or her ability to work, the measurement is not reliable and one has to question the validity of those assessments.

A second requirement of valid measurement is for the measures used to measure what are called the “constructs” of interest (construct validity; see Cronbach, 1989). Consider, for example, the discussion section of a journal article that claims to have shown that counseling focused on improving self-esteem increases work motivation among people with disabilities. You might be hopeful that their intervention is effective in increasing work motivation and yet be skeptical based on the information provided in the methodology section of the article. Did they really measure “work motivation” (with, perhaps, a scale formed from three survey questions)? Perhaps their survey questions measured, instead, overcoming “work anxiety” or even the desire to please the counselors who had been nice to them. Similarly, you might question whether the counseling should be described as focusing on self-esteem rather than depicted as focusing on reducing anxiety. To complete this scenario, you might also question whether the people receiving the counseling were representative of the people with work-relevant disabilities that you are most concerned about. In each of these cases, the concern is whether the interpretation offered elsewhere in the article matches the activities and measures described in the methodology section. This emphasis on proper interpretation highlights an important point about construct validity. We do not judge a measure, be it a survey or an ability test, as being valid or not in itself. Instead, we examine the validity of a particular interpretation that resulted from using the measure in a specific manner with an identified population.

In addition, there are operational issues that must be addressed to have valid measures of the constructs of interest. For example, even if the quarterly wages of consumers matches the construct of interest in theory, we are aware of limitations of available measures of quarterly wages. The Unemployment Insurance file is reliable, but it does not include employment with the federal government or with churches, nor does it include informal jobs not reported. Self-reported data may include these neglected sources of earned income, but these data may not be valid or reliable for other reasons. The lesson generally drawn from the acknowledged inadequacies of specific measures is the importance of employing multiple measures of the major research constructs, such as employment as an outcome (Shadish *et al.*, 2002).

Similar concerns about measurement (and similar lessons about the need for multiple measures) arise with regard to having adequate measures of the implementation of the intervention (e.g. was the policy applied appropriately?; is everyone eligible aware of the program?; are the people who are signed up for the program aware of which program they are part of and the benefits involved?) and about the population served (e.g. are the characteristics of the program participants in accord with the population discussed in your conclusions?). For example, if a state’s Medicaid Buy-In program (which allows workers with disabilities to buy into Medicaid even if their earnings exceed eligibility criteria for regular Medicaid) does not lead to an overall increase in the employment rate of SSDI beneficiaries, it

would be important to have data on the degree to which those eligible knew of the program (even in the group of those who sign up for Medicaid coverage, we would want measures of what they understood of the program benefits).

Are We Making the Right Comparisons to Help Us Understand Policy Impacts?

Even if one has adequate measures of the constructs of interest, there are other considerations in making the best use of evidence to guide policy. One issue relates to our desire to make cause-and-effect statements about policies. For example, as noted at the beginning of this article, the edited volume by Stapleton and Burkhauser is concerned with understanding what caused the alleged decline in the employment rate among people with disabilities. Similarly, one might want to understand the degree to which the 1996 Temporary Assistance for Needy Families (TANF) welfare reform caused an increase in employment, or an increase in hardships, among former welfare recipients with disabilities.

Supporting our ability to reach appropriate cause-and-effect conclusions is the primary focus of research design in social science. Many of the developments in this field are sophisticated to the point of being arcane, but it is useful to remember that quantitative techniques for supporting causal conclusions are based on making a comparison of one form or another (Mohr, 1995). Ideally, we would like to know how people responded to a policy reform and compare it to how they would have responded had there been no policy reform (this hypothetical “what would have happened” is referred to as the counterfactual).

In that it is impossible to know what would have happened to a group of policy-affected people if the policy change had instead never occurred, we try to approximate this ideal comparison. The two major alternative design approximations are (1) to compare people who were subject to a policy reform with those not subject to the specific reform (e.g. Medicaid Buy-in recipients in different states who are subject to different fee structures), and (2) to compare outcomes (e.g. quarterly earnings) of the same people before the policy change and their outcomes after the reform is initiated. These two alternatives are sometimes combined (pre-post measures of both intervention and comparison groups), and statistical analyses are often used to enhance the comparisons between groups and across time.

Choosing among these options for strengthening causal conclusions is a matter of choosing how to respond to several threats to valid conclusions—primarily selection bias, history, and time order confusions. In the research methodology field, these are referred to as threats to internal validity.³

Selection Bias

The concern here is when the comparison of a policy of interest (or, generally, the treatment intervention) with another policy involves different types of person. The

³See Shadish et al. (2002) for a comprehensive treatment of the range of threats, and Mohr (1995) and Julnes (2004) for summaries of these three general threats.

exemplar of this is comparing outcomes (e.g. employment) of people who receive a service (e.g. counseling) with outcomes of those not receiving the service. If the people were assigned randomly to the different policy or program conditions (e.g. flip a coin to determine assignment), there would be no reason to believe that an observed difference in outcomes for the two groups is biased by the pre-existing differences of the people in the two groups (except the “bad luck” of randomly assigning more of the motivated individuals to the treatment group; as Mohr, 1990, points out, minimizing, or at least estimating, this problem is the point of statistical inference).

On the other hand, if people self-select themselves into groups, as when the more motivated or more employable among those with disabilities seek employment counseling and benefits planning, there is reason to suspect that the impact of the program or policy, whether measured by wages or some other outcome of interest, is co-mingled with the effect of the pre-program differences in motivation or degree of disability. This is a source of self-selection bias in that it may appear that a program is having a considerable impact when in fact the program participants would have succeeded almost as well without the program (Mohr, 1995).

The most common scenario for selection bias is when you are comparing outcomes of two or three defined groups (e.g. people with disabilities in two or three cities with different policies). An interesting variation is when comparing average outcomes based on surveys conducted at different times and where the composition of the groups change over time based on individuals being self-identified in one group or another. If the definition of the term “person with a disability” were to change over time (which it has through the courts’ interpretations), then comparing the employment outcomes of people defining themselves as having disabilities today with the outcomes of those defining themselves as having disabilities a decade ago would be vulnerable to the same selection bias.

History

Because of the difficulties posed by selection bias, many studies compare outcomes for the same people at two or more points in time, such as prior to a policy change and also after the change. Trend analyses use many data points over time to offer a finer-grained depiction of policy-related changes, and so represent a major improvement over simple pre- and post-intervention measures. Even for trend analyses, however, the obvious threat to valid conclusions is that many things might have occurred over time besides the change in a particular policy or program participation. For example, the past decade has witnessed changes in attitudes about individuals with disabilities in addition to significant policy changes at the federal and state levels. Estimates of policy impacts based on comparisons over time, therefore, are vulnerable to bias from these other sources of changes in outcomes. Taken together, these sources of changes over time are referred to as history.

Time Order Confusion

Another problem, common when using survey information gathered at a single point in time, is distinguishing the directionality of a causal relationship. For

example, if a survey of people with disabilities who have left welfare revealed a negative correlation between self-reported barriers to employment (e.g. inability to find day care for a child or lack of public transportation) and employment outcomes (e.g. weekly wages or consistency of employment), it might be natural to conclude that the increased barriers existed prior to, and thus are responsible for, the lower outcomes. But this might be an incorrect assumption. It might be that those unemployed are likely to report barriers that would account for, or justify, their unemployment. Because this involves a misattribution of what was prior, it is called a time order confusion. Individuals with disabilities might be similarly likely to report significant disabilities if they are not employed. While this would overstate the impact of disability on employment, it also would overstate the unemployment rate of persons with disabilities.

Research methodologists recognize that the best way to counter these threats is to make use of research designs with appropriate comparisons. Random assignment experiments are especially effective for supporting causal conclusions in most circumstances, and a set of research designs referred to as quasi-experiments (named for being “almost” as adequate as a randomized experiment because of the attention given to ruling out threats to internal validity) approach the strength of the random experiment (Shadish et al., 2002). An example of a quasi-experiment is when you select a comparison group not by random assignment but because they are similar in important ways. One example of this would be to compare persons with disabilities who participate in a Medicaid Buy-In program with non-participants who are similar in terms of education, work history, and job skills. Bell, Orr, Blomquist, and Cain (1995) discuss how people who apply for such a program but end up not participating can be a more appropriate comparison than most for estimating program impacts on participants. This is better than using a comparison group that is known not to be similar, but it still begs the question of how similar is similar enough.

The degree to which random assignment experiments should be preferred to others is a matter of controversy beyond our scope. Our brief advice is: if you are not versed in strengths of random assignment experiments, you will be impressed when you realize how effective they are in estimating the causal impacts of policies (see Orr, 1998); on the other hand, if you believe that random experiments should always be conducted, you probably need to learn more about their limitations in many contexts (Cronbach, 1986). The point for our review is that statistical analyses of national data sets are correlational studies that face greater challenges than randomized experiments or some other research designs in supporting causal conclusions. The implication for policymakers is that they need to be aware of these threats to valid interpretation of research findings so that they can exercise appropriate caution in judging the causal conclusions of survey researchers and argue for policies that encourage stronger research approaches in contexts where valid causal conclusions are essential for effective policies (Julnes, 2004). One way to strengthen our confidence in causal conclusions is to promote multiple studies (indeed, a ‘fleet of studies’) that use different methods (and, hence, are vulnerable to different types of bias). To the extent that the different studies, for example some based on interviews, others on correlations, and still others on experimental manipulations, show consistent results, our conclusions on policy impacts will be better grounded.

One other aspect of using proper comparisons is relevant to the disability policy debate. In any research study there is the concern over the degree to which the results (based on the specific features of one's study) can be generalized to the population(s) of concern (this issue is referred to as generalizability or external validity). There are several dimensions that we typically seek to generalize across in policy research, including different people, different versions of a program, different ways of measuring key outcomes, and different policy contexts. As discussed elsewhere in this article, this is particularly problematic when conducting research on disability policy in that there are often disagreements about how the population of concern should be defined. There is also uncertainty about the degree to which the impacts of a policy in one decade can be expected to generalize to another decade (particularly with the changes in our cultural views of disabilities).

Are We Using Appropriate Quantitative Analyses to Reveal Important Relationships?

A strong design is sometimes betrayed by the analyses used to identify relationships. Flaws result from using one analysis when another is better suited for the task at hand. (The reasons for the superiority of one analysis over others are generally technical and are not important for this review; one example would be using ordinary least squares regression when the dichotomous dependent variable being predicted, such as being employed vs. not employed, calls for logistic regression.) Often, however, the problem comes from making improper assumptions about the nature (i.e. the functional form) of the relationship being studied. For example, if researchers presume that willingness to hire individuals with disabilities decreases proportionately (i.e. is a linear function) to the actual health insurance costs associated with their disabilities, but the real relationship has a different shape (e.g., employers erroneously assume that any and all individuals with disabilities would increase their health care obligations to an untenable level), then a typical analysis (e.g. regression analysis with no transformation of variables) would suggest that there is minimal relationship between these factors. Shadish *et al.* (2002) refer to problems with quantitative analyses that compromise valid conclusions as threats to statistical conclusion validity.

Additional problems with conducting appropriate quantitative analyses concern going beyond the traditional aggregate conclusions. For example, an overall analysis may show no relationship between a policy change and desired outcomes, but this overall lack of relationship might mask a moderated relationship in which some individuals do much better after the policy change while many others do not (a moderated relationship means that the effect of something, such as a policy change, is influenced, or moderated, by the value of another variable, such as the severity of a disability or the level of prior work experience). Failure to include the appropriate moderated relationships in your analyses (often with the use of interaction terms in multiple regression analysis) would lead in such cases to misleading conclusions. As another example of a moderated relationship, a policy change may be effective, but only when other supportive policies are in place (e.g., a Medicaid Buy-In program might be more effective in promoting employment if participants did not face an

abrupt discontinuation, the so-called “cash cliff,” of SSDI benefits for earnings beyond the SGA). Conducting these moderated analyses is particularly important given the diversity of the population of persons with disabilities and the diverse assortment of federal and state policies on disabilities and employment. Further, this is in line with the trend in all policy research to move from simple conclusions of overall impact to more nuanced conclusions of “for whom is the policy effective and under what circumstances?”

Another way that proper analysis helps is in countering threats to valid causal inference (internal validity). This general use of analysis includes many of the statistical approaches reported in the various chapters of *Stapleton and Burkhauser*. For example, if, even with no impact from policy reform, people in one group were more likely than those in another to be employed, it is possible to model the dynamics that led to their greater predisposition for employment (e.g. higher average education or lower severity of disabilities). If done correctly, we control for these biasing factors and obtain the “real” impact of the policy reform. One approach to this desired statistical correction is the use of propensity scores (Rosenbaum & Rubin, 1984). Another approach is the selection bias modeling associated with Heckman (Heckman & Hotz, 1989) and other economists. The actual effectiveness of applying these statistical adjustments is controversial. For example, LaLonde and Maynard (1987) demonstrated that applying this selection bias analysis to a non-experimental study yielded different results about program impact than the estimate from a random assignment experiment. Work on improving selection bias modeling continues, but there is still a fundamental uncertainty in knowing how effective selection bias modeling is in most applications. As such, the position offered by Shadish et al. (2002) is prudent, endorsing the use of these techniques, but not as replacements for solid research design comparisons: they (2002, p. 161) offer the motto “statistical adjustment only after the best possible design controls have been used.”

Are We Using Indicators of Success that Embody the Values We Want to Promote Through Public Policy?

This last question may seem ambiguous, particularly as different interest groups in the American polity espouse different values, but the need to consider the differing values of different stakeholders is precisely the point (Julnes & Foster, 2001).

- Should disability policy be judged as successful exclusively in terms of the economic benefits that resulted from the policy change and the economic costs of the activities required by the policy change?
- If other values are to be included, which values should we include—such as self-determination, the opportunity to contribute to society, and generally the opportunity to be a full member of society—and what indicators may we measure to represent these values?
- If multiple values are to be embodied in the research, should we attempt to balance the relative importance of these diverse values (e.g., is promoting self-determination as important as economic efficiency)?

The point here is that policymakers are generally aware of multiple competing values in policy decisions. This has long been recognized by scholars of policy analysis but is often neglected by individual researchers. If policy research and evaluation are conducted to be relevant to only one of these competing values, this limitation should be noted and the results interpreted accordingly (Henry & Julnes, 1998). Otherwise, there is risk that the research findings will misrepresent the true value of the policies under study. As in other aspects of policy evaluation, the danger is that efforts to be objective result in some values (in particular, those most easily expressed in monetary terms) being emphasized at the expense of others that are, ultimately, also important to society.

In sum, there are several criteria that social scientists routinely consider in assessing the adequacy of research findings. Our concern is that researchers in the disability policy arena conduct research consistent with these criteria and policymakers become more aware of them, particularly with research that is limited to analyses of national data sets. One of the underlying themes of our analysis is that policy debates are too important to be based on evidence from a single methodological approach, that we benefit from promoting multiple research methodologies (see Caracelli & Greene, 1997; Cook, 1985). We highlight the challenges confronting those who rely on the major national survey data sets by turning next to the complexities of conducting useful research on policies related to the employment of persons with disabilities.

CONTEXT OF DISABILITY POLICY ANALYSIS

The previous section provided an analytical framework to call attention to the challenges in supporting causal conclusions when researchers use national data sets as the basis for their research. In this section, we detail the specific complexities that arise in conducting research regarding the employment rate of individuals with disabilities. As noted, researching the impact of public policy is complex because the political, social, and economic factors affecting public policy are dynamic in nature. Nonetheless, policymakers expect that researchers understand and appropriately consider factors, or sample parameters, that impact the implementation of public policy. For example, it is important for researchers to recognize that the impact of disability policy may vary based on the characteristics of the subpopulations of person with disabilities, the significant policy differences (e.g. inconsistencies) among policy initiatives and among the states, and the degree of compliance and implementation. The National Institute on Disability and Rehabilitation Research (NIDRR), in its long-range plan, recognized this point:

Researchers must develop an understanding of the public policy context in which disability is addressed, ignored, or exacerbated. General fiscal and economic policies, as well as more specific policies on employment, delivery and financing of services, telecommunications, institutionalization, education, and long-term care are critical factors influencing disability and disabled persons. Their frequent inconsistencies, contradictions, and oversights can inhibit the attainment of personal and societal goals of persons with disabilities (National Institute on Disability and Rehabilitation Research, 1999, p. 68,581).

This section describes the sample parameters that researchers must consider in conducting policy research relating to improving the employment rate of persons with disabilities and determining the appropriate income protections and health care programs. For a comprehensive review of relevant sample parameters, see Blanck, Hill, Siegal, and Waterstone (2003).

The Population of Persons with Disabilities Is Heterogeneous

While researchers in the field understand that the population of persons with disabilities is a heterogeneous group, individual studies may be compromised by failing to address this diversity (see Blanck et al., 2003b; see also Burkhauser et al., 2003, pp. 33–34). The points below highlight the sources of error when characterizing “persons with disabilities” (a problem with construct validity). As noted elsewhere in this article, such potential errors could lead to invalid conclusions when comparing the employment outcomes of “persons with disabilities” to those of other groups.

- *Variations in type, severity, and onset of impairment and age of the individual.* The variations in type and severity of disability and the age of the individual may affect the individual’s ability to work and the level of work effort. Another variable is the time of onset of disability, i.e. birth, during teens, after years of employment, which affect a person’s capacity to work. The category of individuals considered, “persons with disabilities,” includes persons with physical or mental impairments (that affect physical or mental processes of the human body such as walking, hearing, seeing, talking, and understanding) that do not limit one’s capacity to function in a work activity; limit one’s capacity to function in a work activity in varying degrees; or prevent individuals from functioning in a work activity at all.
- *Variations in capacity and predisposition.* Variations in levels of skills, education, and experience may affect an individual’s ability to work and the level of work. These standard employment-related factors, however, may be only a portion of the important capacities for successful employment. It may be, for example, that strong social support networks represent a different type of capacity that supports work effort (Julnes, Fan, & Hayashi, 2001). Likely interacting with these capacities are individual predispositions, such as willingness to take risks and achievement motivation. As a result, it is easy to imagine that 10 persons with disabilities having, hypothetically, equal capacities for work would end up with 10 very different employment patterns.
- *Different self-perceptions.* The perception of the individual as to whether he or she is a person with a disability may vary depending on the interaction with the environment, e.g. availability of reasonable accommodations, auxiliary aids and services, long-term services and supports such as assistive technology and personal assistance services, architectural and communication accessibility; personal and family attitudes; attitudes of others; and employment status (Burkhauser et al., 2003).
- *Multiplicity of barriers.* Not all persons with disabilities face the same barriers to employment. The range of barriers identified by Federal policymakers (see The

Rehabilitation Act of 1973 and the Ticket to Work and Work Incentives Improvement Act) include

- discrimination (e.g. disparate treatment, disparate impact, perceived, associational, and retaliatory),
 - lack of adequate education and employment-related services and supports (including vocational rehabilitation),
 - lack of job-ready employment experience,
 - lack of or fear of losing necessary health care (including long-term services and supports such as personal assistance services and assistive technology),
 - fear of losing cash assistance, financial disincentives to work and retain assets, the complexity of work incentives, and lack of benefits counseling,
 - lack of accessible housing and other community living options and transportation, and
 - difficulty navigating the separate programs and policy silos (fragmentation).
- *Divergent range of employment potential.* Persons with disabilities enjoy a range of employment potential (with or without reasonable accommodation) that may vary significantly from full-time to part-time employment as well as entry level and more advanced employment options. Some individuals may lack the capacity to work at all. The ability to work may vary over time for certain individuals (intermittent) because of the episodic nature of particular disabilities.

Over Time, the Precept of Public Policy for Addressing the Needs of Persons with Disabilities Has Changed

Historically, people with disabilities were treated as “defective” and in need of “fixing.” The term “disability” was equated with inability to work and low expectations. This treatment is commonly referred to as the old paradigm of disability policy. The old paradigm of disability policy has been replaced by the new paradigm, which provides that disability is a natural and normal part of the human experience that in no way diminishes a person’s right to fully participate in all aspects of society, including employment (Silverstein, 2000, p. 1695).

The four goals of our nation’s disability policy are enunciated in the ADA (Silverstein, 2000, pp. 1712–1713). The first goal is equality of opportunity. This has three components—(1) treat people as individuals based on facts, objective evidence, and science and not on the basis of fear, ignorance, prejudice, stereotypes, or administrative convenience; (2) provide meaningful and effective opportunity (e.g. reasonable accommodation, reasonable modification to policies, practices, and procedures, and auxiliary aids and services); and (3) administer programs in the most integrated setting appropriate.

The second goal of disability policy is full participation in the decisionmaking process at the individual and systems levels. This includes the concepts of self-determination, empowerment, real informed choice, and self-advocacy. The third goal of disability policy is independent living, which includes the provision of independent living skills, health care (including long-term services and supports such as personal assistance services and assistive technology devices and services), and cash assistance programs. The fourth goal of disability policy is economic

self-sufficiency, including the provision of appropriate employment-related services and supports, and income protections with work incentives.

Federal Policymakers Have Enacted a Range of Public Policies to Address the Needs of Persons with Disabilities

To address the multiplicity of barriers faced by and needs of persons with disabilities, federal policymakers have enacted a range of policies, including civil rights laws, cash assistance programs, health care programs, employment-related services and supports, and tax policy (Silverstein, 2000, pp. 1699–1708). These policies include a statement of policy and purpose, identify the target population, and specify proscribed and prescribed behaviors to achieve congressional intent. One practical implication of these policies is that it is difficult to separate the impacts of one policy from the others that often have overlapping objectives.

Civil Rights

The first category of laws includes federal civil right statutes that prohibit covered entities (such as employers, businesses, and state and local governments) from discriminating against persons on the basis of or by reason of disability. Examples of civil right statutes include the Americans with Disabilities Act (ADA), which prohibits discrimination on the basis of disability in employment, public services, public accommodations, and telecommunications; Section 504 of the Rehabilitation Act, which prohibits discrimination by recipients of federal aid, such as hospitals, universities, and public schools; and the Fair Housing Act. In addition, Section 508 of the Rehabilitation Act specifies requirements for federal departments and agencies regarding the development, procurement, maintenance, or use of accessible electronic and information technology.

Cash Assistance Programs

The second category includes cash assistance programs. The Social Security Disability Insurance program (SSDI) is a program of federal disability insurance primarily for workers who have contributed to the Social Security Trust Fund and became disabled or blind before retirement age. Note that there are also people with disabilities who receive benefits based on disability through the Social Security Retirement and Survivors portions of the program, as disabled adult children or disabled widows or widowers. For the purposes of this article, the reader should assume that these groups are included in the references to the SSDI program.

A second cash assistance program is the Supplemental Security Income program (SSI). The SSI program is a means-tested program providing monthly cash income to low-income persons with limited resources on the basis of age and on the basis of disability and blindness. The SSI program is funded from general revenues of the Treasury.

The definition of disability used for initial eligibility for both the SSDI program and the SSI program is identical. Disability is defined as the inability to engage in

any substantial gainful activity (SGA) by reason of a medically determinable physical or mental impairment that is expected to last for a continuous period of not less than 12 months, or result in death. SGA is defined in the federal regulations as earnings of \$810 per month for 2004.

The SSDI and SSI programs include a number of so-called “work incentives” designed to encourage earnings and work, although each program has different incentives. A significant difference between the two programs is the time-limited nature of the work incentives under the SSDI program. Once an SSDI recipient has completed a trial work period, and it is determined that he/she is able to engage in SGA and, after a three month grace period, he/she will lose cash benefits. This is commonly referred to as the “cash cliff.” Under the SSI program, an individual can earn more than SGA and still remain eligible for the program. Under Section 1619(a) of the Social Security Act, the SSI recipient faces a gradual rather than a precipitous loss of benefits and may receive no cash payments (non-payment status) and remain eligible under the program.

Health Coverage

The Medicare and Medicaid programs are particularly relevant to persons with disabilities. The Medicare program authorizes health insurance benefits to specified elderly persons and certain persons with disabilities (e.g. disabled workers receiving SSDI benefits). The Medicare program is divided into four parts. Part A authorizes hospital insurance benefits; Part B provides supplemental medical insurance benefits; Part C contains miscellaneous provisions, including coverage for end-stage renal disease; and Part D authorizes the new prescription drug benefit.

Medicaid is the nation’s major public financing program for providing health and long-term services and supports to low-income persons. Medicaid is a means tested entitlement program financed by state and federal government out of general revenues. Under Medicaid, states are required to serve some population groups and are permitted to serve others. Most states provide automatic eligibility for Medicaid to SSI recipients. Under Section 1619(b) of the Social Security Act, individuals can continue to be eligible for Medicaid even if their earned and unearned incomes make them no longer eligible for SSI cash benefits.

The Ticket to Work and Work Incentives Improvement Act, enacted in 1999, makes several changes to the Social Security Act designed to make work pay for SSDI and SSI recipients and reduce or eliminate their dependency on cash benefits. Changes include a Medicaid Buy-In program (encourages states to adopt the option of allowing disabled workers to purchase Medicaid coverage necessary to maintain employment); providing individuals with disabilities the option of maintaining Medicare coverage while working; improved work incentives; benefits planning and outreach; and specific demonstration authority vested in the Social Security Administration to test the implementation of a \$1 for \$2 gradual rather than precipitous loss of benefits under the SSDI program to replace the cash cliff.

Employment-Related Services and Supports and Other Programs

Federal policymakers have enacted a number of employment-related programs providing services and supports to persons with disabilities. Examples of federal

programs are described below. Part B of The Individuals with Disabilities Education Act provides financial assistance to state and local educational agencies to help them meet their constitutional responsibilities to provide a free and appropriate public education (including transition services) to children with disabilities, regardless of the nature and extent of their impairments. The vocational rehabilitation program established by Title I of the Rehabilitation Act assists states in operating a comprehensive, coordinated program of vocational rehabilitation for persons with disabilities. Title I of the Workforce Investment Act (WIA) supports workforce investment activities, such as job training, through state and local workforce investment systems. Communities, through their Local Workforce Investment Boards, establish “One-Stop” career centers that under Section 188 must be programmatically and physically accessible to persons with disabilities. The Temporary Assistance for Needy Families program authorizes communities to provide cash assistance to needy families and provides flexibility to provide employment-related services and supports to accommodate the needs of persons with disabilities in these families. The Ticket to Work and Self-Sufficiency program provides SSDI and SSI recipients with additional choices of providers of employment-related services and supports.

Tax Policy

The tax code provides incentives for covered entities to implement existing responsibilities. For example, the Disabled Access Credit provides tax credits to small businesses for expenses incurred in becoming compliant with the ADA. In addition, the Architectural/Transportation Tax Deduction permits businesses to take a specified annual deduction for expenses incurred to remove physical, structural, and transportation barriers for persons with disabilities at the workplace. Furthermore, the Work Opportunity Tax Credit (which replaces the Targeted Jobs Tax Credit program) provides a tax credit for employers who hire certain targeted low-income groups, including SSI recipients and vocational rehabilitation referrals.

States and Localities Have Enacted a Range of Public Policies to Address the Employment Needs of Persons with Disabilities

Twenty-five states have enacted civil rights law prohibiting discrimination on the basis of disability by private entities (Hotchkiss, 2003, pp. 126–128). Thirty-one states have authorized and twenty-eight states have implemented Medicaid Buy-In programs and there are 116 Benefits Planning, Assistance and Outreach programs operating in all states and the District of Columbia (see agency website, <http://www.cms.gov/twiiia/ede.asp>). Several states (e.g. Wisconsin and Massachusetts) are developing comprehensive, person-centered employment initiatives designed to create a seamless system that addresses the multiplicity of needs facing persons with significant disabilities who want to work. These state and local variations hold promise for making meaningful comparisons of where a policy has and has not been implemented. Usually, however, there are so many specific differences across localities and states that comparisons are not realized as planned.

Federal and State Policy are Based on an Inconsistent Mix of the Old and New Paradigms About Persons with Disabilities

Some policies enacted by federal and state policymakers are based on the new paradigm of disability policy, some on the old paradigm, and some on both the old and new paradigm. For example, the ADA, IDEA, and Rehabilitation Acts are all based on the new paradigm of disability policy; i.e., disability is a natural and normal part of the human experience that in no way diminishes a person's right to fully participate in all aspects of society, including employment. The definition of disability used for determining eligibility for the SSI and SSDI programs is based on the old paradigm (disability means inability to work). Many of the work incentive provisions in the SSI program are based on the new paradigm. The work incentive provisions in the SSDI program send mixed messages in light of the continued existence of the cash cliff.

It is Important to Distinguish Disability-Specific and Generic Programs

Sometimes public policies addressing the needs of persons with disabilities are disability specific and sometimes they are generic. Disability-specific initiatives include the ADA, Rehabilitation Act, Developmental Disabilities Assistance, and Bill of Rights Act. Generic programs that impact directly on persons with disabilities include Title I of WIA, TANF, and Medicaid (Silverstein, 2000, pp. 1704–1705). To the extent that people with disabilities historically have not been part of the generic support system, there may be limited experience on the part of those working within that system on how to provide effective services to such individuals. In these cases, there is a critical need to build the capacity of the generic support system (through disability-specific policies, practices, and procedures) to provide meaningful opportunities for persons with disabilities (see U.S. Department of Labor, 2002).

Public Policies Enacted by Federal and State Policymakers Often Have Specific Goals and Objectives

Public policies enacted by federal and state policymakers often have specific goals and objectives. Sometimes the goals and objectives of a specific policy initiative address the needs of a limited population. For example, the specific employment-related provisions of the ADA (Title I) are not designed to address the overall employment rate of persons with disabilities; rather, these provisions are designed to enhance equality of opportunity (prohibit discrimination) for the subset of the population of persons with disabilities who are currently qualified, i.e. have the requisite skills, experience, and expertise to perform the essential functions of a job with or without reasonable accommodations.

Sometimes the objectives of particular policy initiatives overlap with the objectives of other policy initiatives. For example, the non-discrimination objectives of

the ADA and Section 504 overlap with Section 188 of WIA, which requires One-Stop career centers receiving WIA funds not to discriminate against, among others, persons with disabilities. Title I of the Rehabilitation Act provides vocational rehabilitation services to persons with disabilities (including SSI and SSDI recipients) in need of such services. Similarly, the Ticket to Work and Self-Sufficiency Program pays employment networks for achieving specified employment outcomes for SSI and SSDI recipients.

Sometimes the goals and objectives of policy initiatives are in conflict and send mixed messages. SSI is designed to provide cash assistance to low-income persons with impairments who are “unable to work.” SSI work incentives are designed to make work pay but are conceptually inconsistent with the definition of disability used to determine eligibility (inability to work). SSDI uses the same definition of disability as used for SSI (unable to work). SSDI requirements (e.g. a recipient loses eligibility if he or she earns more than SGA) after the Trial Work Period, however, are time limited and result in precipitous rather than gradual loss of benefits (cash cliff) and hence may not make work pay. In addition, in some cases a person receives both SSDI and SSI (because SSDI benefits are low) and the confusion that results from two sets of rules may not make work pay.

The Public Policies Enacted by Congress and the States Are Dynamic, Not Static—They Change Over Time in Response to New Realities

Congress and the Executive Branch are constantly changing policy to respond to new realities. For example, in 1984, Congress amended the rules governing the SSDI and SSI programs, particularly for those with mental illness, to ensure that persons with disabilities who were eligible for benefits actually received them. According to the Disability Policy Panel of the National Academy of Social Insurance, “it is plausible that the increase in the number of persons with mental illness who qualify for work-based SSDI benefits reflects positive consequences of community integration (in lieu of institutionalization) as well as improvements in prescription drug therapy that help stabilize disabling conditions” (Mashaw & Reno, 1996, p. 96).

During the 1990s, federal policymakers recognized that existing work incentives were not working. GAO found that less than one-half of one percent of SSDI and SSI recipients left the rolls due to work and earnings (Ticket to Work and Self-Sufficiency Program Regulations, 2001, p. 67,370). In 1999, The Ticket to Work and Work Incentives Improvement Act was enacted to enhance employment of SSDI and SSI recipients and to reduce or eliminate their dependency on cash assistance by including, among other things, a Medicaid Buy-In option for states (to enable recipients to buy into Medicaid if they increased their earnings), improved work incentives, benefits planning and outreach (to address the complexity of existing programs), the Ticket to Work program (to increase choice of providers of employment-related services and supports for recipients), and authorized SSA to conduct demonstrations to test alternative policies (\$1 for \$2) to address the cash cliff disincentive to work.

The Public Policies Enacted into Law at the Federal, State and Local Levels Are Subject to Varying Degrees of Compliance and Implementation

The fact that a bill is enacted into law creating a new public policy does not mean that covered entities are complying with and implementing the law. Compliance may vary from total non-compliance to full compliance, and compliance may be inconsistent across covered entities and across states and localities. Similarly, implementation may vary from adopting practices based on the old paradigm of disability policy to practices that reflect best and promising practices, i.e. are state of the art. The degree of continuous improvement, monitoring, and enforcement of policies, practices, and procedures may affect compliance and implementation. For example, the National Council on Disability (NCD) recently found significant deficiencies in efforts to monitor and enforce the ADA and IDEA. According to the NCD, enforcement of civil right laws prohibiting discrimination on the basis of disability presents “troubling and pervasive issues,” as documented in its report *Promises to Keep: A Decade of Federal Enforcement of the Americans with Disabilities Act* (2000b, p. 37). This report analyzed ADA enforcement efforts of the Department of Justice, the Equal Employment Opportunity Commission, and other Federal agencies (see also the NCD’s report *Back to School on Civil Rights*, 2000a). If these variations in implementation are not measured and understood, conclusions about effectiveness are vulnerable to bias. In sum, lack of compliance and implementation may adversely affect the extent to which the behavior changes and outcomes intended by the policymakers are realized.

CHARACTERISTICS OF NATIONAL DATA SETS

The previous sections summarized the criteria of valid research and the complexities in the policy environment that complicate research. This section describes the various federal data sources on disability, a framework for conceptualizing disability, and the limitations in the use of work disability definition of disability.

Federal Data Sources on Disability

According to the Congressional Research Service (CRS), there are four Federal data sources on disability and labor force status (Levine, 2000):

- Census of Population
- National Health Interview Survey (NHIS)
- Current Population Survey (CPS)
- Survey of Income and Program Participation (SIPP).

Three of these federal data sources (CPS, SIPP, and NHIS) are generally used by researchers to show trends in the employment rates of persons with disabilities. Hence, CPS, SIPP, and NHIS are frequently referred to throughout this article. The following descriptions are taken from the CRS Report to Congress.

Census of Population

Disability questions are included on the long-form of the decennial census of population. The number and nature of disability questions have differed over time. The 1970 form asked about work disability; the 1980 form about work disability and the ability to use public transportation; and the 1990 form about work disability, the ability to go outside the home alone to shop or visit a doctor's office, and about self-care (e.g. bathing or dressing without assistance). The work disability questions were fairly consistent from 1970 through 1990 (i.e. having a physical, mental or other health condition that prevents an individual from working, or that limits the amount/kind of work an individual can do).

The few disability questions were substantially revised on the 2000 long form. Households that received the revised questionnaire were asked whether a physical, mental, or emotional condition of at least 6 months duration creates difficulty in working at a job or business; going outside the home alone or visit a doctor's office; dressing, bathing, or getting around inside the home; or learning, remembering, or concentrating. In addition, a question was included on whether respondents have long-lasting sensory impairments (e.g. blindness) or substantial limitations on physical activities (e.g. reaching or climbing stairs).

The long form of the decennial census will eventually be replaced by the more frequently administered American Community Survey (ACS). The disability questions from the 2000 long form have been used in the ACS since 1999.

National Health Interview Survey (NHIS)

The National Health Interview Survey (NHIS) is a nationally representative household survey conducted by the Census Bureau for National Center for Health Statistics, part of the Centers for Disease Control and Prevention. The sample is limited to civilian members of U.S. households; not included are people living in institutions. The purpose of the annual survey is to provide information on chronic health conditions, health status, health care utilization, and disability. It asked consistent disability and labor force questions between 1982 and 1996. Although the survey underwent a major revision in 1997, some of the earlier disability questions were carried forward. However, there is no regularly published federal document in which NHIS's labor force and disability data appear.

In the annual NHIS prior to 1997, individuals were considered to have disabilities if they reported a chronic condition or an impairment that prevented them from or limited their ability to perform age-dependent major life activities (e.g. attending school in the case of children 5–17 or taking care of oneself in the case of someone at least 70 years old), or limited their ability in any way to perform any other activities (e.g. social or family pursuits). The NHIS also asked people if they needed assistance in performing ADLs or IADLs.

Respondents aged 18–69 who reported that an impairment or health problem completely prevented them from working or restricted the kind/amount of work they could perform were deemed to have work limitations. More specifically with respect to work limitations, NHIS includes the following questions: "Does any impairment or health problem NOW keep [person] from working at a job or business? Is

[person] limited in the kind OR amount of work [person] can do because of any impairment?”. Those who answer “yes” to either question are considered to report a work limitation.

Beginning with the redesigned 1997 NHIS questionnaire, some changes were made to the above-described items. For example, the questions that determine work limitation were reworded to ascertain whether a physical, mental, or emotional problem prevents or restricts the kind/amount of work a person can perform; to eliminate an upper age limit; and to allow acute in addition to chronic conditions as causes of work disability.

Current Population Survey (CPS)

The CPS is a monthly survey of the non-institutionalized population of the United States conducted by the Bureau of the Census on behalf of the Bureau of Labor Statistics. Information is collected on labor force characteristics (e.g. employment, earnings, hours of work). In March of each year, the CPS basic monthly survey is supplemented with the Annual Demographic Survey. The supplement focuses on sources of income, government program participation, previous employment, insurance, and a variety of demographic characteristics. The CPS began in the early 1940s; however, the work limitation variable was not asked until 1981. In 1994, major revisions were made to the basic monthly survey and the labor force questions.

The basic CPS questions are organized to first determine whether the respondent is employed or unemployed (i.e. actively looking and available for work). If the respondent does not meet the criteria for either category, he or she then is classified as “not in the labor force.” When individual respondents are asked whether they did any work for pay or profit, some instead respond that they have a disability, which is recorded. Following is the question on disability from the basic CPS: “LAST WEEK, did you do ANY work for (either) pay (or profit)? Yes, No, Retired, Disabled, Unable to Work, Don’t Know, Refused” (U.S. Census Bureau, 2004, p. 2). If the respondent meets the criteria for either employed or unemployed, the response of “disabled” is erased from the file because it is inconsistent with the labor force categories of employed and unemployed.

The March supplement, which is the primary source used by researchers, asks “Does anyone in this household have a health problem or disability which prevents them from working or which limits the kind or amount of work they can do? [If so,] who is that? (Anyone else?)” (U.S. Census Bureau, n.d., p. 2). Those who answer “yes” to this question are considered to report a work limitation.

Survey of Income and Program Participation (SIPP)

The SIPP is a longitudinal survey that contains detailed demographic, program, employment, and health characteristics for a nationally representative sample of the non-institutionalized resident population of the United States. The SIPP is conducted by the Bureau of the Census on behalf of the Bureau of Labor Statistics. The purpose of the SIPP is to provide comprehensive information regarding the income

and program characteristics from a representative sample of the United States population. In addition to basic information about work limitations, extensive data on functional limitations are collected periodically.

Respondents to the SIPP are considered to have a disability if they indicate that they experience difficulty performing certain sensory or physical functions (e.g. seeing words in newspapers, hearing normal conversations, or lifting/carrying something that weighs 10 pounds); activities of daily living, ADLs (e.g. bathing, eating, dressing, or getting around inside the home); or instrumental activities of daily living, IADLs (e.g. going outside the home, keeping track of money or bills, taking prescription medicines, or using the telephone). People who indicate that they have a physical, mental, or other health condition that prevents them from or restricts the kind/amount of work they can perform around the house or at a job also are deemed to have a disability. In addition, respondents who state that they use assistive devices (e.g. wheelchairs or canes), have specific conditions (e.g. developmental disorders or Alzheimer's disease), or are under age 65 and covered by certain benefit programs (e.g. SSI) are regarded as having a disability. More specifically, with respect to work limitations, the first core interview asks "Does [] have a physical, mental, or other health condition which limits the kind or amount of work [] can do?"

Definition of Disability: A Conceptual Framework

The varying disability-related questions used in the national data sets highlight the role of measurement and the importance of construct validity in policy research. To measure the employment of the working-aged population with disabilities, therefore, it is first necessary to define the population. Using the disability-related questions from NHIS, CPS, and SIPP, Burkhauser and Stapleton organize the available empirical evidence into a nested framework with three definitions of disability (Burkhauser et al., 2003, pp. 28, 32; see also Nagi, 1991).

The largest group, "**impairment**," is constituted of individuals who report having an impairment, that is, a physical or mental loss or abnormality that limits a person's capacity to function. According to Burkhauser et al. (2003, p. 32), "[t]his population could be considered to represent the potential population that many of the supporters of the ADA intended to protect." Burkhauser et al. (2003, p. 33) assert that NHIS functional limitation questions capture this population.

The next group, "**activity limitation**," involves a subsample of people with impairments, people who report some type of activity limitation because of their impairment. According to Burkhauser et al. (2003, pp. 33–34, emphasis added), "...those with a pathology that causes a physical or mental impairment that subsequently limits one or more life activities such as work but who, nevertheless, work would *not* be considered to have a work limitation For example, a person with deafness who is accommodated at the workplace with a TTY machine that permits him or her to use the telephone would *not* be considered work-limited despite his or her impairment." They report that the most commonly used activity-limited definition of disability includes those who report a work limitation, which is available in the NHIS, CPS, and SIPP.

The third category, “**longer-term activity limitation**,” represents persons with the most severe and long-term limitations. This population is most likely to be eligible for SSI and SSDI based on their inability to perform any substantial gainful employment. According to Burkhauser *et al.* (2003, p. 34), this category includes people who report a work limitation in both the CPS and CPS follow-up survey.

With these three groups in mind, Burkhauser *et al.* report (pp. 36–37) that the employment rate in 1996 of the “impairment” population of men was 77.3%. They also report that the employment rate for the “activity-limitation” population was 50.1% and the employment rate of the “longer-term activity limitation” population was 23.6%.

Limitations in the Use of Work Disability Definition of Disability

According to Burkhauser *et al.*, the “work disability” definitions of disability used in the various surveys have the following limitations (note, these limitations overlap with each other and are presented in the order discussed by the authors).

- “When one asks a person if he or she has a disability or more specifically a ‘work disability’ the answer might depend on the person’s current employment status” (Burkhauser & Stapleton, 2003a, p. 7).⁴
- “The work limitation definition used in CPS and NHIS significantly underestimates the numbers of persons in the broader population with impairments and over-represents those with impairments who are not employed” (Burkhauser & Stapleton, 2003a, p. 7).
- “Caution must be exercised in using global self-reported responses on whether a person’s health limits the kind or amount of work he or she can perform because they are subjective and can vary from individual to individual” (Burkhauser *et al.*, 2003, p. 31).
- “Health responses may not be independent of the economic variables being examined” (Burkhauser *et al.*, 2003, p. 31).
- “Self-reported disability questions must be used with caution, particularly if the answers are sensitive to the respondent’s socioeconomic environment” (Burkhauser *et al.*, 2003, p. 31).
- “Relying on a current work-limitation question to define the ‘true disability population’ misses those with impairments who are sufficiently integrated into the workforce so that they do not report being work limited” (Burkhauser *et al.*, 2003, p. 36).
- “A work-limited response can be influenced by the work environment, rehabilitation opportunities, or the inner capacity of the individuals to overcome both their impairments and the barriers to work they face” (Burkhauser *et al.*, 2003, p. 36).
- “A work limitation question overstates the size of the population with longer-term work limitations” (Burkhauser *et al.*, 2003, p. 36).

Consistent with these limitations, the National Council on Disability (NCD, 2002 and 2004) and Hale (2001) have concluded that the CPS work limitation

⁴This is an example of the time-order bias discussed above.

measure cannot be used to provide credible information to policymakers with respect to the employment of working-aged persons with disabilities (see Burkhauser & Stapleton, 2003a, p. 7). NCD, an independent agency established by Congress, expressed concern (2002, p. 27) that CPS data could lead to “ineffective or even dangerous public policy decisions” based on a “Federal consensus that certain CPS items are not adequately designed to elicit accurate and reliable information from people with disabilities.” NCD concluded (2002, p. 20) “The Federal government should not encourage or support the dissemination of employment data until a methodology for assessing employment rates among people with disabilities that is acceptable to leading researchers and demographers in the field and credible to persons with disabilities can be developed.”

Hale (2001) contends that the CPS work limitation question neither captures the larger population of persons reporting “impairments” nor provides a representative sample of that population with respect to employment. More specifically, Hale (2001, p. 38) asserts that “conclusions by researchers [using CPS data] about the [decline in the] employment rate trend for persons with disabilities and the underlying causes [of the decline] are not valid.” According to Philip Rones, Deputy Commissioner of the Bureau of Labor Statistics in the U.S. Department of Labor, “Tom Hale’s article is the place to go for a description of the limitations of the current CPS questions on disability” (personal communication).

With respect to the use of CPS “work disability” data, the official position of the U.S. Census Bureau (2004) includes the following statements.

- “Researchers and others often have used CPS to identify the population who are said to have a ‘work disability.’ The questions used for this purpose, however, were not designed or tested with the intent of measuring disability, and thus the reliability and validity of the estimates generated from these questions is unknown. The questions were not placed in the CPS to measure disability, but rather, to achieve other goals. For example, the March work limitation question is a screener question to identify persons who subsequently will be asked various income questions” (p. 1).
- “The Census Bureau has not specifically tested these questions to determine how well they identify those who actually have a work disability. As a result, the data generated from the questions can only be said to represent a count of those who said yes (or another designated response) to the questions in the criteria, and should be understood as such” (p. 1).
- “Data from the [SIPP] (U.S. Census Bureau, 1997) show that many people who do have a condition or impairment that could reasonably be viewed as limiting the types of work they can do, respond ‘No’ to [the CPS question in the March Supplement]” (U.S. Census Bureau, 2004, p. 4). For example, 38.5% of deaf persons (unable to hear at all) respond that they do not have a work limitation.
- “Individuals, who do in fact have a work disability, may be likely to say ‘no’ they do not have a work disability if they are working. In other words, individuals may respond to these questions independent of their disability status” (p. 5).
- “[T]here is an inherent problem in defining individuals into the work-limited population based largely on their stated work limitations, and then cross tabulating this with their employment status. Doing so will, as a rule, generate low

employment rates. This should not be seen as confirmation that the questions identified correct population, but rather that the employment rate for this group is low because individuals were defined into this group because they were not working" (p. 5).

Kruse and Schur (2003a) and also Blanck *et al.* (2003b) make related points in identifying potential difficulties in using the work limitation measure to study employment trends of persons with disabilities.

- "The work-limitation measure may be affected by the success of the ADA in making workplaces more accessible, as people who obtain jobs would no longer say they are limited in the ability to work.... This could cause measured employment of people with work limitations to decline as the ADA increases job opportunities among people with disabilities" (Kruse & Schur, 2003a, p. 284; Blanck, Schwochau, & Song, 2003, p. 321).
- "Many people reporting a work limitation may not be protected by Title I of the ADA... because they are not qualified" (Kruse & Schur, 2003a, p. 285; see also Blanck *et al.*, 2003b).
- "The likelihood of reporting a disability may be intertwined with employment status.... Among people with the same medical conditions, functional limitations, and other characteristics, those who are not employed may be more likely to report that they have a work limitation as a way of justifying their lack of employment.... Those who obtain jobs may become less likely to cite a work limitation even if they have the same impairments and medical conditions as before" (Kruse & Schur, 2003a, pp. 285–286).

Researchers note additional cautions regarding the use of existing data sources due to major revisions made to the surveys. In 1994, an extensive revised basic questionnaire was introduced in the CPS survey. Kaye (2003, p. 222) contends that it is not appropriate to compare pre- and post-1994 data. Major revisions were made to NHIS in 1997. According to Kaye (2003, p. 220), "because of substantial changes to the survey, data prior to 1997 are not directly comparable to data from later years."

TREND COMPARISONS RELATING TO THE EMPLOYMENT OF PERSONS WITH DISABILITIES

As explained, most authors agree that there are significant limitations in the use of work limitation definitions of disability. There is, however, disagreement as to whether such data should be used at all, and if so, under what circumstances and for what purposes. Some researchers have chosen to study trends in disability by focusing on a population identified by these work disability questions, despite these limitations. Several other individuals and organizations have argued that it is inappropriate to use work limitations questions to identify people with disabilities when studying employment trends and question the extent to which the associated findings may be generalized. This section re-examines research findings regarding trends in the overall employment rate of persons with disabilities, the employment rate of persons able to work, the employment rate of persons able and available to

work, and the proportion of persons reporting work limitations and inability to work.

Trends in the Overall Employment Rate of Persons with Disabilities

The overall employment rate of persons with disabilities should include the full range of persons with physical and mental impairments using the conceptualizations suggested by Burkhauser and Stapleton, that is, persons with functional impairments, persons with activity limitations and persons with longer-term activity limitations (Burkhauser et al., 2003, pp. 32–34). Kaye uses NHIS data (see Burkhauser et al., 2003, p. 32) in an attempt to report trends for the overall employment rate. According to Kaye (2003, p. 222), there is no statistically significant trend in the overall employment rate, with the 2000 value of 24.5% being about the same as the 1994 value of 24%.

According to Levine (2000) of the Congressional Research Service, with the NHIS trend data the employment rate of working-aged adults with any limitation on any activity (e.g. housework or social pursuits) appears to have been virtually unchanged during the decade of the 1990s. Within this broadly defined population with disabilities, the increased proportion over time of adults with severe work disabilities may partly account for the stubbornness of the overall employment rate. If the subset of working-aged adults who report in the survey that they have impairments so severe they are unable to work remains elevated, it could be difficult to narrow the large gap between the employment rates of the population with and without disabilities.

Trend analysis conducted by Burkhauser et al. (2003, pp. 36–41), based primarily on the use of work limitation data, shows a decline in the employment rate for persons with activity limitations and longer-term activity limitations. However, it is important to note that there is general consensus that the work limitation data reported in CPS, SIPP, and NHIS does not capture the broad category of persons with impairments, but does capture (albeit with significant limitations) persons with activity limitations and longer-term activity limitations (Burkhauser et al., 2003, pp. 32–34). According to Burkhauser et al. (2003, pp. 33–34), excluded from these groups are persons who have impairments that result in functional limitations such as the inability to work but who report they can work because of reasonable accommodations. They also point out (pp. 33–34) that the group of persons with longer-term limitations is likely to consist of persons eligible for SSI and SSDI because they are unable to work. According to Philip Rones (personal communication, March 23, 2004), the Deputy Commissioner of the Bureau of Labor Statistics, “It is nonsensical to use as a broad measure of disability the single question in the March CPS Supplement on work disability.”

Thus, based on current data sources there are divergent opinions regarding the overall employment rate of persons with disabilities—some individuals believe current work limitation data cannot be used to determine the overall employment rate for persons with disabilities, others believe the overall rate did not change during the 1990s, and still others believe that the overall employment rate of persons with disabilities decreased during this period.

Trends in the Employment Rate of Persons Able to Work

According to Kaye (2003, pp. 221–222), for people with disabilities who say they are able to work, there is evidence of substantial improvement in their employment rate following the 1990–1991 recession. In the NHIS data, there is a 3.1 percentage point increase in employment rate of persons with disabilities who can work; “from a 1992 low of 70.2% employed, the rate climbed to a high of 73.3% in 1995, (dropping slightly, but not significantly to 72.3% in 1996). The upward trend is statistically significant and is comparable to the gains experienced by people without disabilities during the same period” (p. 221). The results may be more positive if the comparison used is a subset of the general population that is demographically similar to the population of persons with disabilities. Using CPS data, “among people with disabilities who are able to work . . . there is an 8.3 percentage point increase . . . [T]he upward trend is highly statistically significant” (p. 222). Using SIPP data, Kruse and Schur (2003a, p. 292; see also Kruse & Schur, 2003b) found that during 1991–1993 there were non-significant relative employment increases among people reporting any or severe functional and ADL limitations and among people reporting work limitations but an ability to work. Strikingly, there was a significant relative employment increase of 5.9 percentage points in weeks worked for 1991–1993. Surveys of human resource personnel indicate positive, improved practices toward the employment of persons with disabilities, although notable problems remain (see, e.g., Dixon, Kruse, & Van Horn, 2003; Kaye, 2003).

Trends in the Employment Rate of Persons with Disabilities Able and Available to Work

According to Kaye (2003, p. 223), using CPS data, there is a striking increase in employment from 71.9% in 1994 to 80.5% in 2000 for persons with disabilities who are able to work and available to work (the group includes people who are either labor force participants, working, on layoff, or actually looking for work) or non-participants who consider themselves able to work and report they would like to have a job. This increase of 8.6 percentage points “is more than twice that for working-age adults without disabilities who are available for work” (3.6 percentage point increase).

Trends in the Proportion of Persons Reporting Work Limitations

According to Burkhauser *et al.* (2003, p. 44), “the size of the able-to-work subpopulation declined substantially as a share of the entire work-limitation population in the 1990s in all three surveys, particularly in the CPS . . . [T]he decline in the overall size of the able-to-work population more than offsets the gain in employment by this group.”

According to Kaye (2003, p. 218), “the proportion of persons with disabilities who consider themselves able to work has declined over the years. The overall

disability rate among working-aged adults rose dramatically during the early 1990s, with a disproportionate share of that increase occurring among people reporting an inability to work. [The change] in the composition of the disability population accounts for the difference between the bleak employment picture evident when everyone is included and the far brighter outlook when the analysis is limited to those oriented toward working.”

According to Kruse and Schur (2003a, p. 290), “the SIPP data show a significant increase in the percentage of people reporting work limitations between 1991 and 1993 More important, among those reporting a work limitation, there was a statistically significant 2.8 percentage point increase in those reporting any of the measured functional or ADL limitations . . . and a statistically significant 5.2 percentage point increase in those reporting severe functional or ADL limitations.”

According to Hotchkiss (2003, pp. 28–29), the observed decline in the labor force participation rate was not the result of persons with disabilities fleeing the labor force, but was most likely due to the re-identification of some non-participants from non-disabled to disabled. In other words, there is less movement across disability status among labor market participants than among non-participants. In addition, these results indicate that the observed decline in the labor force participation rates should not be considered as casting a shadow on the measured impacts of the ADA on employment.

CONCLUSIONS: LESSONS LEARNED FOR EVALUATING POLICY RESEARCH

This article has addressed challenges in conducting research that warrant use in guiding policy. This section summarizes the lessons learned from a review of the research relating to the employment of persons with disabilities (from a policymaker’s perspective) with respect to

- understanding how the complex policy context affects research,
- using appropriate measurement to yield meaningful data,
- planning adequate comparisons and analyses to avoid threats to valid cause-and-effect conclusions,
- representing desired values in the indicators of program success,
- considering the implications of our critique for research on the impact of the ADA and the SSDI program,
- recommending additional areas of and approaches for research, and
- using a checklist for assessing the adequacy of research.

Understanding Research Within the Policy Context

Conducting research on the impact of a policy requires an understanding of the intent of the policy and of the constellation of related federal and state policies. A review of the research literature on the employment rate of persons with disabilities highlights the following aspects of the policy environment.

- The population of persons with disabilities is heterogeneous (e.g. variations in type, severity, onset, capacity, age, predisposition, self-perceptions, range of barriers and employment potential).
- Over time, the precept of public policy for addressing the needs of persons with disabilities has changed.
- Federal and state policymakers have enacted a range of public policies to address specific needs of persons with disabilities.
- Some policies enacted by federal and state policymakers are based on the new paradigm of disability policy, some are based on the old paradigm, and some are based in part on both the old and new paradigm.
- Sometimes the public policies addressing the needs of persons with disabilities are disability specific; sometimes they are generic.
- Public policies enacted by federal and state policymakers often include specific goals and objectives. Sometimes the goals and objectives of a specific policy initiative address the needs of a limited population; sometimes the goals and objectives of particular policy initiatives overlap with other policy initiatives; and sometimes the goals and objectives are in conflict and are inconsistent.
- The public policies enacted by Congress and the states are dynamic, not static—they change over time in response to new realities.
- The behavior changes and outcomes intended by policymakers may not be realized (in whole or in part) because of the lack of compliance and implementation by the entities covered by the policy.

As an example of the impact of specific policy features, and of how policies interact, consider that the limited number of SSDI beneficiaries that leave the SSDI rolls due to earnings and employment may result in part from the continued existence of the “cash cliff” under the SSDI program (once a recipient earns SGA after a trial work period, he/she is no longer eligible). Because of the policy overlap, the success of a State’s Medicaid Buy-In program and the Ticket to Work program (designed to increase the employment rate of SSDI and SSI beneficiaries and reduce or eliminate their dependency on SSI and SSDI) may be adversely affected by the “cash cliff.”

Adequacy of Data

As explained in the section on the need for valid measurement, current national data sets used to determine the overall employment rate of persons with disabilities are inadequate to support the needs of policymakers, researchers, and other stakeholders. NCD expressed concern that the use of the CPS data could lead to “ineffective or even dangerous public policy decisions.” This concern was based on a “Federal consensus that certain CPS items are not adequately designed to elicit accurate and reliable information from people with disabilities,” with NCD concluding that the reliance on current data is so suspect that there should be a moratorium on its use (National Council on Disability, 2002, p. 27).

In 1998 President Clinton signed Executive Order No. 13,078 (1998) establishing the Presidential Task Force On Employment of Adults with Disabilities. Specifically, the Executive Order directed “The Bureau of Labor Statistics of the Department of Labor and the Census Bureau of the Department of Commerce, in

cooperation with the Department of Education and Health and Human Services, the National Council on Disability, and the President's Committee on Employment of Persons with Disabilities to design and implement a statistically reliable and accurate method to measure the employment rate of adults with disabilities as soon as possible."

Levine (2000, p. 4) of the Congressional Research Service noted that the accomplishment of this charge "is complicated by the fact that a universally agreed-upon definition of disability does not exist and a data source is lacking that incorporates both labor force and disability questions which have undergone testing to determine whether people understand what is being asked (one aspect of validity) and to which people respond in a consistent manner over time (i.e., reliability)."

To inform policy with any confidence, researchers need valid data sources to ascertain the overall employment rate of persons with disabilities that is based on the new paradigm of disability policy, includes the full spectrum of persons with physical or mental impairments, and yet allows us to disaggregate the data based on subgroups of persons with disabilities, e.g. persons able to work, persons able and available to work, persons unable to work. The use of multiple research methods would facilitate this. With respect to self-report data, Kruse and Schur (2003a, p. 296) suggest that "it would be valuable for researchers to closely examine what leads people to report work limitations In particular, . . . how does employment status affect reports of work limitations [the time order confusion noted above] as well as other self-report measures?". They also suggest that "an ideal research project would follow individuals over time, independently recording medical conditions and impairments, as well as self-reported work limitation status and ability to work" (p. 296).

Having more adequate data would avoid other problems. For example, the definition of disability inherent in the data used by researchers should comport with the scope and purpose of the policy question being addressed. It is not valid from a research point of view to draw conclusions about the effects of a program that has a limited purpose and addresses a limited subpopulation by using data sources that are not limited to such populations. For example, Title I of the ADA only protects a subgroup of the population of person with disabilities (i.e. qualified individuals with disabilities; see Blanck et al., 2003b). It would not be valid to use work limitation data from CPS, which focuses primarily on persons who report that they are unable to work or limited in their ability to work. Similarly, it is unwarranted to make conclusions about the overall employment rate of persons with disabilities by using data that focuses on a subpopulation (e.g. work limitation data from CPS) and excludes other subpopulations.

Drawing Policy-Related Impact Conclusions from the Research

We have discussed above that quantitative research on policy impacts depends on meaningful comparisons. An ideal comparison from a research point of view would follow from randomly assigning persons with disabilities to groups with different policy coverage (e.g. one group confronting the standard cash cliff with regard to excess earnings and the other group having their pay benefits decrease gradually as

they earn more money). With random assignment, there would be no reason to believe that the groups were substantially different at the beginning of the study in their propensity to be employed. The only difference should be that members of the intervention group are covered by the policy in question while members of the other group are not. As such, any significant difference in employment a year after assignment would reasonably be viewed as due to the differences between competing policies. If, in contrast, we did have reason to believe that people in the intervention group (e.g. the group able to earn above the SGA limit without losing all benefits) were more motivated, on average, than those in the comparison group, we would need to qualify our estimate of the impact of the innovative policy.

If random assignment is not an option (while it should be viewed as an option more often than it is, there are compelling arguments against it in many contexts), we need to be careful in selecting the appropriate comparison group. In the disability area, the question for researchers should be “what group is as similar as possible to person with disabilities while at the same time not being covered by the policy being evaluated?”. The research studies included in the Stapleton and Burkhauser book appear to compare persons with disabilities to the general non-disabled population despite the reasonable assertion by DeLeire (2003, p. 266) that “studies should use individuals without disabilities who have similar skill levels as a comparison group by which to measure the post-ADA experiences of those with disabilities.” An appropriate examination of the evidence might suggest that the population of persons with disabilities should be compared with similarly situated non-disabled persons using such factors as work history, skills levels, and past income (see Bartik, 2001). While matching people in research designs is generally problematic, considering these factors is important because, as Burkhauser and Stapleton (2003a, pp. 2–3) note, there is a significant differential between the median income of men and women with and without disabilities. For example, in 1989 (before the ADA was enacted), the median income for men without disabilities was \$31,899 and the median income for men with disabilities was \$16,905.

A second lesson about policy impact conclusions concerns the need for disaggregated analyses. Particularly in the domain of disability policy, the people targeted by policy reforms are very diverse (e.g., ADA focuses on with persons with disabilities who are currently qualified to perform the essential functions of a job, with or without a reasonable accommodation; in contrast to the SSI and SSDI programs that focus on persons with significant impairments that are unable to perform substantial gainful activity). Thus, it is essential to consider whether a particular program is designed to help some individuals with disabilities but not others. It is also essential to recognize that different people have different needs (Jules, Hayashi, & Anderson, 2001), and so a program that helps only a small portion of those eligible may still be viewed as an important success.

A third lesson on identifying policy impacts is that the types of quantitative analysis conducted do matter. On the one hand, failure to select techniques that match the characteristics of the data will result in bias. On the other hand, proper use of techniques to model selection bias has the potential to minimize that bias. Unresolved in most correlational studies, however, is the extent to which bias remains after employing state-of-the-art analyses. In sum, these three lessons about drawing policy-related impact conclusions provide additional arguments for

employing multiple methods, even a mixture of quantitative and qualitative methods, when seeking to inform policy debates.

Supporting Value Judgments with Indicators of Success

Estimating policy impacts, as important and challenging as that is, is only useful if the right impacts are studied. At issue are how we understand the goals of a policy, how we relate those goals to definitions of policy “success,” and how we select appropriate indicators of success. Research regarding the impact of federal and state employment-related initiatives on the employment rate of persons with disabilities should include a clear articulation of the definition of success used. What measures of “improved employment outcomes” should be used to ensure that research activities and evaluation of policy efforts reflect the scope and variety of employment outcomes desired by the heterogeneous population of working-age persons with disabilities?

For example, “improved employment outcomes” could include the following.

- Increased percentage of persons with disabilities who are employed.
- Increase in the number of hours worked in a week or month.
- Increase in the rate of pay, earnings, or net income from self-employment.
- Increase in disposable income from working.
- Improvement in the employment-related health or other benefits comparable to persons without disabilities.
- Improvement in the employment-related accommodations to enable a person with a disability to accomplish tasks at the workplace comparable to persons without such disabilities.
- Improvement in satisfaction by an individual related to what they do to earn income.
- Improved satisfaction by an individual in their interpersonal work environment, including the level of inclusiveness, normalization, and non-segregation related to disability.
- Improvement in flexibility in work schedule to fit needs related to an individual’s disability.
- Improvement in the ability to adjust the number of hours worked to fit needs and capabilities of individuals related to their disability.

What definition of success should be used for evaluating the SSI and SSDI programs? One former House Ways and Means staffer (Jensen, unpublished manuscript, p. 5) suggests the following four criteria.

- “Financial Security—Improve the financial security of persons with disabilities related to their immediate, emergency and long-term needs.”
- “Accessibility—Ensure access for persons with disabilities to financial assistance programs and social, rehabilitation, health and related services.”
- “Independent Living—Improve the opportunities of persons with disabilities to live as independently as possible in the least restrictive setting.”
- “Work Opportunities—Ensure persons with disabilities adequate opportunities to increase level of self-support by ensuring continued income supplementation and health care financing if they work in spite of their impairment.”

Critical Review of Policy Implications

The issues raised above are important because they qualify understanding of the impacts of the programs intended to support people with disabilities. This understanding is important because the policies that we adopt and refine have real impacts on real people. We summarize the implications of our review for the Americans with Disabilities Act and the Social Security Disability Insurance program.

The Americans with Disabilities Act (ADA)

Any analysis of the impact of the ADA must begin with an understanding of the twofold intent of the ADA (Silverstein, 2000). First, the ADA is a civil right statute specifying a protected class, covered entities, standards of behaviors constituting discrimination, and remedies. With respect to the protected class, Title I of the ADA (employment discrimination) specifies that only “a qualified individual with a disability” is protected, that is, an individual with a disability who can perform the essential functions of the job, with or without reasonable accommodation. Hence, Federal policymakers did expect that the employment rate would increase for this subgroup of persons with disabilities. The specific provisions of Title I of the ADA were not intended, however, to have a direct impact on the employment rate of persons with disabilities who are currently unqualified.

Second, Congress used the ADA as a vehicle for enunciating the overarching goals of disability policy (equality of opportunity, full participation, independent living, and economic self-sufficiency) with the hope that federal and state policymakers would modify policies to reflect the new paradigm of disability policy (disability is a normal part of the human experience that does not diminish a person’s right to fully participate in all aspects of society, including work) rather than the old paradigm of disability policy (disability means inability). Federal policymakers never intended the ADA to be the “be all and end all” of disability policy but rather intended the ADA to prescribe the precept for and goals of our nation’s disability policy and establish the antidiscrimination component of a larger disability policy framework.

With this context and intent as background, it is interesting to note the ambiguity of the conclusions about the impact of the ADA. According to Burkhauser and Stapleton (2003a, p. 5), it is difficult to draw the conclusion that the ADA caused the decline in the employment rate for people with disabilities. Goodman and Waidmann (2003, p. 363) reject the ADA explanation for the decline in the employment rate of persons with disabilities. Burkhauser and Stapleton (2003b, p. 394) seem to agree, stating that “[a]lthough we are not convinced that the ADA had a significantly negative effect on employment of working-aged people with disabilities, we also find no unambiguous evidence that it had a significant positive effect At best it may have increased the employment of the decreasing share of that population who report being able to work at all.” Yet, at an Urban Institute Forum, Stapleton (2003) made the following statement regarding the impact of the ADA: “we [Burkhauser and Stapleton] have to leave open the possibility that the ADA contributed to the decline.”

Also acknowledging the ambiguity, Kaye (2003) notes that it is difficult to determine whether the ADA and/or the booming economy of the 1990s was responsible for the significant increases in the employment rate for persons reporting that they are able and available to work. However, he believes that one can make a reasonable argument that economic and legal (ADA) factors played a role in improving the employment prospects of working-aged Americans with disabilities. Kruse and Schur (2003a, p. 295; see also Kruse & Schur, 2003b) found “increased employment among people reporting functional and ADL limitations who do not receive disability income or who report the ability to work. The greatest increase occurred among those who are arguably most likely to be covered by the ADA: people with severe functional or ADL limitations who report the ability to work.”

Further acknowledging the subtlety of the issue, Burkhauser and Stapleton (2003b, pp. 388–389) note that “Blanck and coauthors [Blanck et al., 2003c] also make a point about the the ADA that is worth repeating. The ADA might have had a significant impact on the culture of disability [e.g., corporate culture], which could have long-term positive effects on employment. It has probably increased the visibility of people with disabilities who do work [and encouraged many] to attempt to work... [and] to invest in their ability to work, ... [and it might have] encouraged educators, providers, and advocates to see work and independence as a desirable and achievable goal, and it might have encouraged employers and workers without disabilities to look more objectively at the capabilities of workers with disabilities” (Burkhauser & Stapleton, 2003b, pp. 388–389). By contrast, according to DeLeire (2003) and others, the ADA is “responsible” for the decline in the employment rate of persons with disabilities.

In sum, although it appears that during the 1990s there was a significant increase in the employment rate for persons with disabilities able and available to work, it is difficult to conclude with certainty (based on the limitations of the data) that the ADA was the cause of the increase (although Kaye believes a strong case can be made indicating it was at least partially responsible). What does appear clear is that the conclusions reached by DeLeire and others that the ADA was responsible for the decline are currently unwarranted, in large part because of their lack of understanding of the purposes and scope of the employment-related prohibitions in the ADA and because they rely on work limitation data that may exclude significant numbers of people with impairments who are working (the precise and only group of persons protected by the ADA).

The Social Security Disability Insurance Program (SSDI Program)

During the 1990s, the SSDI program grew rapidly. According to Goodman and Waidmann (2003, pp. 364–365), however, “if we look at the entire period [between the 1970s and 1990s] and adjust for increases in the number of people eligible for benefits, the growth does not seem so dramatic.” Further, Goodman and Waidmann (p. 339) “conclude that the growth in the number of recipients can largely be attributed to two program changes: a period of liberalization in eligibility criteria, beginning in 1984, and a gradual increase in program generosity for low-wage workers [replacement rate].”

The reasons for the growth may not be limited to those articulated by Goodman and Waidmann. A list of additional possible explanations for the increase in the proportion of persons reporting work limitations and the proportion of SSDI recipients is set out below.

- Lessened stigma associated with disability (Kruse & Schur, 2003a).
- Economic factors (some researchers, e.g. Kruse & Schur, 2003a, suggest that during the 1990s recession some people may have chosen to emphasize their inability to work in order to remain eligible for benefits after unemployment insurance ran out).
- Social factors (e.g. work disincentives such as the “cash cliff”, fragmentation in various disability programs; see Goodman & Waidmann, 2003).
- Demographic factors (baby boomers nearing retirement age and the effect that this has on increasing the numbers of older persons (50+) who are becoming disabled and need SSDI).
- Epidemiologic factors (i.e. increase in the prevalence of severe disabilities associated with inability to work e.g. obesity and mental stress; see Kaye, 2003).
- Pressures on state personnel to remove persons from state welfare rolls, resulting in referral for SSI of persons with disabilities and to a lesser extent SSDI (Reno, 2004).
- Pressure on private employers to remove persons for long-term disability by moving them onto SSDI.
- Workers’ compensation reform forcing people off state worker compensation rolls and onto SSDI (Kruse & Schur, 2003a, p. 291).
- Changes in the nature of work, changes in retirement options (e.g. a decline in good early retirement options for workers in arduous jobs), and changes in health care coverage by private employers (Stapleton, Goodman, & Houtenville, 2003).

The rise in the SSDI rolls, in and of itself, is not necessarily a positive or negative development (Goodman & Weidmann, 2003, p. 364). For example, the increase in the proportion of persons reporting work limitations and receiving SSDI benefits may have positive implications because they reflect congressional intent. Congress and SSA intended the changes to the SSDI eligibility criteria in the 1980s to achieve specific policy objectives, including restoring eligibility to persons who Congress intended to cover and enabling persons with mental illness and other impairments to have the cash assistance they needed to live in the community rather than in institutions (Mashaw & Reno, 1996, p. 96). Enhancing community living need not be seen as a “social disaster,” as Burkhauser and Stapleton (2003a, p. 5) would conclude.

As Goodman and Waidmann (2003, p. 364) note, “[i]f the growth [in SSDI rolls] indicates that individuals with disabilities are ‘trapped’ in a cycle of dependence that is antithetical to the goals of disability policy articulated in the ADA and the independent living movement, the growth is troubling. If, on the other hand, the growth is a result of providing an increasingly meaningful safety net to individuals who cannot fully participate in the labor market because of their impairment, the growth may be desirable.”

The feeling of being “trapped” (GAO found that less than one-half of one percent of SSDI and SSI recipients left the rolls because of earnings and work) was one of the reasons articulated by Congress for enacting the Ticket to Work and

Work Incentives Improvement Act (TWWIIA). Preliminary research regarding states' experiences implementing the Medicaid Buy-In (authorized by TWWIIA) indicates that there has been an increase in earnings; however, upwards of 90% of participants are still earning below SGA (see Jensen, Silverstein, & Folkemer, 2002). One explanation for the limited success of the Medicaid Buy-In programs may be the continued existence of the SSDI "cash cliff," which may be a significant impediment to substantial progress in increasing employment and earnings of SSDI recipients.

The Social Security Advisory Board in its October 2003 report, *The Social Security Definition of Disability*, concludes "while positive incentives can be added [to the SSDI program] as long as benefit receipt is conditional on demonstrating a lack of ability to work, disincentives will be inherent in the system" (p. 1). The Advisory Board concluded "The definition of disability in the Social Security Act often appears to undermine the goals of disability policy articulated in the ADA by providing incentives for impaired individuals to prove to the agency and, presumably to themselves, that they are incapable of any substantial work" (p. 26).

In sum, the increase in the number of SSDI recipients may have multiple causes, some positive, some negative, and the negative causes may include the failure to remove ongoing federal barriers such as the cash cliff.

Critical Review of Research Methodology

If we are to embrace the value of "unflinchingly objective examination of the evidence" for assessing policies, we need to apply that stance as well to assessing the value of research methods. Our predispositions and personal fondness regarding particular methods should not distort views of what those methods can and cannot deliver. Because the dangers of being misled by well intended policy research are so high, it is important for researchers to adopt a critical stance with regard to their own findings. For this purpose it is useful to consider other views regarding making warranted conclusions from policy research. According to standards used by the U.S. General Accounting Office (2003, § 7.61), "when auditors are unable to obtain sufficient, competent, and relevant evidence about the validity and reliability of the data, they may find it necessary to use the data but clearly indicate in their report the data's limitations and refrain from making unwarranted conclusions and recommendations."

Most researchers recognize that there are significant limitations in the use of work-limitation data in the CPS, SIPP, and NHIS surveys for purposes of studying the employment rate of persons with disabilities. There is divergent opinion among researchers and others, however, regarding the reliability and validity of the data and implications of using these data sources (ranging from acceptable to use (e.g. Burkhauser and Stapleton) to unacceptable (e.g. National Council on Disability and Hale)). We have also discussed the limitations of national surveys as a research method for supporting cause-and-effect conclusions, preferring at least quasi-experimental studies with more attention given to the appropriate comparison groups for estimating policy impacts on persons with disabilities.

Given all of this, what are we to make of the following policy-related conclusions promoted by Burkhauser and Stapleton in *Stapleton and Burkhauser* on the basis of these limited and suspect data?

- “Given the robust expansion of the 1990s and the promises of the ADA, the decline in the employment rate might reasonably be considered a social disaster” (Burkhauser & Stapleton, 2003a, p. 5).
- “The evidence indicates that the widespread employment rate decline for people with disabilities is a consequence of public policies that were implemented in the late 1980s and early 1990s” (Burkhauser & Stapleton, 2003b, p. 393).
- “The decline occurred because of changes in the social environment—reductions in individuals’ incentives to work and reductions in employer incentives to hire them” (Burkhauser & Stapleton, 2003b, p. 399).
- “There has been a real and substantial decline in the employment rate of people with disabilities and that it was caused by a change in public policy” (Burkhauser & Stapleton, 2003b, p. 393).
- “The bottom line of this book is that the unprecedented fall in the employment rate of working-aged people with disabilities in the 1990s was a direct effect of the unintended consequences of public policies” (Burkhauser & Stapleton, 2003b, p. 403).

By their own admission, Burkhauser and Stapleton recognize the limitations of the data on which they are basing these definitive conclusions (we quoted many of their measurement concerns about using the definition of work limitation and related concepts as operationalized in the national surveys). As such, in accordance with the GAO standards and other commonly accepted standards, it is unwarranted to rely the work limitation data to draw policy-related cause-and-effect conclusions regarding the relationship between the employment rate of persons with disabilities and existing public policies. While the interpretations offered by Stapleton and Burkhauser deserve to be considered in future research, we must continue to qualify the conclusions that we draw from analyses of the national survey data. We must do this regardless of our personal beliefs and preferred policy paradigms.

Recommendations for Future Research

We address next the research strategies that could build on the findings presented by the many contributors to *Stapleton and Burkhauser* and help us develop the necessary actionable evidence base for guiding future reform of disability policy.

Situating Research in the Policy Context

The provisions of the Title I of the ADA (employment discrimination) were not intended to be the “be all and end all” of disability policy in terms of employment. Research should focus on other aspects of disability policy related to employment at the federal and state levels. Researchers need to take into consideration the

impact of other programs constituting the emerging disability policy framework including

- The Rehabilitation Act
- The Individuals with Disabilities Education Act
- The Workforce Investment Act
- The Temporary Assistance for Needy Families Act
- Medicaid work incentive provisions for workers with disabilities (including Medicaid Buy-In programs) and Medicare
- SSI Work Incentive provisions, including Section 1619
- demonstrations mandated by Congress including the \$1 for \$2 SSDI demonstration using gradual rather than precipitous loss of benefits
- Ticket to Work and Self-Sufficiency Program
- state SSI Supplementation programs
- benefits counseling
- asset development and retention
- tax policy.

Researchers must also take into consideration the interaction among the various components of the emerging disability policy framework, including a review of efforts to assess the impact of federal and state comprehensive, person-centered return to work initiatives. This interaction has been emphasized in the Medicaid Infrastructure Grants sponsored by the Centers for Medicare and Medicaid Services (2003).

Complementary Methods of Research

The challenges that we have been describing for research based on the major national surveys should not be understood as arguments not to conduct this type of research. Rather, it is a common problem for a single approach to research to be inadequate as a base for guiding policy decisions. The answer, therefore, is to promote a research agenda that makes use of complementary and, as Blanck has stressed, cumulative methods. In the area of quantitative research, it is understood that random assignment experiments have a key role in establishing the impacts of public policies but that other approaches, such as quasi-experiments and correlational studies remain necessary to fill in our understanding. Many fine experimental and quasi-experimental studies are being conducted on the impacts of particular policies (such as the Medicaid Buy-In program and the recent SSA \$1 for \$2 benefit offset demonstration program) that can be integrated with the correlational findings described in the Stapleton and Burkhauser book. Similarly, as Kruse and Schur (2003a, p. 296) point out, panel studies that would follow a reasonably large group of individuals with disabilities over time, through economic recessions and expansions, would yield answers just not possible with the current national surveys that include different people each year (multiple cross-sectional studies).

In addition, however, it is important to benefit from the insights that qualitative research can provide. To clarify this potential contribution, it is useful to distinguish between two types of question. One is trying to estimate the impacts of a known cause (e.g., what did the known policy reform cause to happen?). The other is trying to understand the causes of known effects (e.g., what were the causes of the observed decline in employment, as measured by the national surveys?). The second question is akin to what detectives do when investigating a known crime, and it is especially suitable for qualitative research (see, e.g., Yin, 1994). For example, informal discussions with employers could reveal that they are concerned about the cost of providing health care coverage and quite convinced that hiring persons with disabilities will make the cost of such coverage prohibitive. One chapter in the Stapleton and Burkhauser text (Hill, Livermore, & Houtenville, 2003) examines this dynamic and finds no evidence to support such a mechanism for reduced employment (based on the reasonable but untested assumption that employers' resistance to hiring persons with disabilities would be a linear function of the actual health costs associated with the applicants' disabilities). Such correlational evidence should be supported by the triangulation that would come from systematic efforts to gather qualitative information on this.

Summary and Checklist for Assessing Adequacy of Research

The basic premise of this article is that social science research plays a critical role in informing policymakers and others regarding efforts to design, implement, and evaluate policy options for improving the employment rate of persons with disabilities and determining appropriate income protections and health care programs. Policymakers need and, in fact, must demand unflinchingly objective examination of the evidence to aid them to carry out their public policy responsibilities.

An unflinchingly objective examination of the evidence requires an understanding of the social, political, and economic context in which policy is implemented, the appropriate use of data, and a recognition of the limitations in the use of data, including the recognition of the limited circumstances under which it is appropriate to draw cause-and-effect conclusions regarding the impact of public policy based on the data. When researchers understand these parameters, they serve the public well. When researchers ignore or diminish the import of these parameters, they perform a disservice to policymakers, the general public, and, importantly, to persons with disabilities. If the current data sets are inadequate, it is critical that new and better data sets be devised. If the data have substantial limitations, researchers must recognize these limitations and refrain from drawing unwarranted summative conclusions. If a policy is designed to achieve a limited purpose for a target population, researchers should not draw conclusions about the effects of the policy if they are unable to identify the target population with the available data sources. If there are multiple factors that affect a policy outcome, researchers must refrain from isolating a particular factor and drawing unwarranted policy implications.

The following appendix includes a checklist for assessing the adequacy of research regarding the employment rate of persons with disabilities.

APPENDIX: A CHECKLIST FOR ASSESSING THE ADEQUACY OF RESEARCH REGARDING POLICY IMPACTS ON THE EMPLOYMENT RATE OF PERSONS WITH DISABILITIES

I. Awareness of Social, Political, and Economic Context in Which the Policy Is Implemented

A. Policy Development

1. Does the research demonstrate an understanding of the intended outcomes of the policy or policies being evaluated?
 - Are there multiple intended outcomes in addition to employment, such as independent living and self-determination?
 - Is the particular Federal/State policy initiative being evaluated based on the old paradigm of disability policy (disability equates with inability to work) or the new paradigm (disability is a natural aspect of the human experience that in no way diminishes a person's right to fully participate in mainstream activities) or both?
 - Does the federal/state policy initiative being evaluated have specific goals and objectives (i.e., are they targeted to address a particular subpopulation or a particular strategy for enhancing employment)?
 - Are the intended outcomes of the federal/state policy initiative being evaluated disability specific or are they part of a generic program?
2. Does the research consider the heterogeneity of the population of persons with disabilities that are targeted by the policy or policies in question?
 - Variation in type, severity, and onset of disability,
 - age of the individual,
 - capacity of the individual,
 - predisposition of the individual,
 - differing self-perceptions,
 - differing barriers,
 - divergent range of potential (full/part time; intermittent).

B. Policy Environment

1. Does the research reflect an understanding of the range of federal/state policies that overlap and interact with the policy or policies being evaluated in terms of impact on the employment rate of persons with disabilities?
2. Are the other federal/state policy initiatives complementary or are they inconsistent with the policy being evaluated?
3. Have the major precepts, goals, objectives, and policies included in relevant federal/state policy initiatives changed over time (dynamic or static)?

C. Policy Implementation

1. Is it possible to determine whether any lack of results is due to the policy or the lack of compliance/implementation by federal/state agencies (including failure to adopt best and promising practices)?
2. To what extent are the policies included in the federal/state initiatives being ignored/complied with by covered entities, e.g. employers?

II. Appropriate Measurement

A. Selection of Data Sources

1. What data sources are used in the research?
2. If existing data sources are used, for what purposes were they originally collected? Are there relevant limitations stemming from these original purposes?
3. Are the data sets designed to address the old or new paradigm of disability policy?

B. Quality of Data

1. Are the measures in the data sets reliable and valid for the uses made in the research?
2. Is there consensus/disagreement among researchers in the field regarding the reliability and validity of the data? Are the researchers aware of this consensus/disagreement?
3. Are the other logical limitations of the data? Are the limitations significant or de minimus?
4. Are there practical difficulties in ascertaining reliable information regarding the employment of persons with disabilities, e.g. pre-employment inquiry prohibition, disclosure of disability is voluntary?

III. Impact Analysis

A. Research Design

1. Is the research based on only quantitative methods or does it make use of mixed methods?
2. If quantitative methods are used, is the research based on an experimental, quasi-experimental, or correlational design?
3. What comparisons are made to estimate the impact of the policy being evaluated?
 - Comparisons over time, or trend comparisons
 - Comparisons across people, cities, or states where some are subject to the policy and some not, or group comparisons
 - Combination of trend analysis and comparison groups.

B. Trend Comparisons

1. If trend comparisons are used, what definition of disability is used to report trends? Are the data sets used consistent with the definition of disability?

2. Do the definition and data sets comport with the trends being described?
 - Overall employment rate of persons with disabilities
 - Employment rate of persons able to work
 - Employment rate of persons able and available to work
 - Persons with disabilities reporting inability to work.
3. When several data sets are used, are the trend patterns consistent? If not, what is the utility of trend analysis?
4. Are the researchers aware of the problem of “history” with this design (the problem of other events happening around the same time as the policy change of interest)? If so, how do they try to minimize this potential bias? Are their arguments against the potential for bias compelling?

C. Comparison Groups

1. If cities or states with different policies are being compared to each other, to what extent are the cities or states comparable in all ways except the policy differences?
2. Are the researchers aware of the problem of “selection bias” with this design (the problem of the people in one group just being different from those in the comparison groups)? If so, how do they try to minimize this potential bias? Are their arguments against the potential for bias compelling?
3. If using a comparison group, to what extent are the individuals in the policy-affected group comparable to those in the comparison group(s)? Is, for example, the employment rate for persons with disabilities being compared to the overall employment rate for the general nondisabled population or to the rate for similarly situated persons based on income, work history, and skills levels?

D. Aggregate Analysis

1. If quantitative analyses are used, do they seem appropriate for the types of data being analyzed?
2. Do the quantitative analyses require assumptions (e.g. linear relationship between severity of disability and willingness of employers to hire) that are questionable?

E. Disaggregated Analysis

1. How meaningful is reporting the overall employment rate of “persons with disabilities” given the heterogeneity of the population?
2. Is it important to disaggregate the data by subpopulations?
3. Which subpopulations are relevant for purposes of disaggregation? For example
 - persons with impairments, persons with activity limitations, and persons with longer-term activity limitations
 - persons who are currently qualified (with or without accommodations)
 - persons who need vocational rehabilitation or other employment-related services and supports in order to be qualified

- persons who are able to and available for work
- persons who are qualified to work but cannot work without ongoing health care, including medications and long-term services and supports
- persons who are unable to work because of the nature or severity of their disability.

IV. Support for Value Judgments

- A. What definitions of success are used? How is the definition determined? For example, is it possible to cite to specific statutory language or other statements of congressional intent?
- B. Does the criterion of success comport with the scope and purpose of the policy question addressed? For example, does the researcher draw conclusions about the effect of a program/policy initiative that has a limited purpose and addresses a limited subpopulation by using data sources that focus on the appropriate subpopulation and use the criteria of success that reflects congressional intent? Success?

V. Critical Review of Findings, Conclusions, and Recommendations

A. Critical Review of Research Methods

1. What standards are used relating to the reporting of findings, conclusions and recommendations? For example, are the General Accounting Office Auditing Standards (7.61, 2003 Revisions) or an equivalent used (7.61 specifies that when auditors are unable to obtain sufficient, competent, and relevant evidence about the validity and reliability of the data, they may find it necessary to use the data but clearly indicate in their report the data's limitations and refrain from making unwarranted conclusions or recommendations)?
2. When there are substantial limitations in the existing data sets and disagreement among researchers regarding the reliability and validity of such data, does the research report the magnitude of findings from such data sets along with statements regarding the limitations of the data sets without drawing cause and effect conclusions?
3. When significant data limitations exist, do researchers use findings to identify additional areas of research and use additional research methods, e.g. qualitative as well as quantitative?

B. Critical Review of Conclusions and Recommendations

1. When reaching conclusions and making recommendations, are the researchers aware of existing barriers/constraints in current law recognized when the research is being conducted (for example, the possible impact of the SSDI cash cliff on the success of employment initiatives and work incentives as the Federal and State levels)?
2. Are the conclusions reached and recommendations made based on best practices and consistent with an "unflinching" acceptance of the data?

3. Is it possible to isolate reasons (causes) for the increase in the number of persons reporting inability to work and increases in the proportion of SSI and SSDI recipients based on current data set? What are some possible causes?

REFERENCES

- Americans with Disabilities Act. (1990). Definition of qualified individual with a disability, Section 101(8), 42 U.S.C. § 12111(8) (2000); Discrimination by employers prohibited, Section 102, 42 U.S.C. 12112.
- Architectural/Transportation Tax Deduction, Section 190 of the Internal Revenue Code, 26 U.S.C. § 190 (2000).
- Bartik, T. J. (2001). *Jobs for the poor: Can labor demand policies help?* Kalamazoo, MI: Sage-Upjohn Institute for Employment Research.
- Bell, S. H., Orr, L. L., Blomquist, J. D., & Cain, G. G. (1995). *Program applicants as a comparison group in evaluating training programs*. Kalamazoo, MI: Upjohn Institute for Employment Research.
- Blanck, P., Hill, E., Siegal, C., & Waterstone, M. (2003a). *Disability civil rights law and policy*. St. Paul, MN: Thomson-West.
- Blanck, P., Schur, L., Kruse, D., Schwochau, S., & Song, C. (2003b). Calibrating the impact of the ADA's employment provisions. *Stanford Law and Policy Review*, 14(2), 267-290.
- Blanck, P., Schwochau, S., & Song, C. (2003). Is it time to declare the ADA a failed law? In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 301-337). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Burkhauser, R. V., Houtenville, A. J., & Wittenburg, D. C. (2003). A user's guide to current statistics on the employment of people with disabilities. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 23-86). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Burkhauser, R. V., & Stapleton, D. C. (2003a). Introduction. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 1-22). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Burkhauser, R. V., & Stapleton, D. C. (2003b). A review of the evidence and its implications for policy change. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 369-406). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed-method evaluation designs. In J. C. Greene, & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. *New Directions for Evaluation*, No. 74 (pp.19-32). San Francisco: Jossey-Bass.
- Centers for Medicare and Medicaid Services. (2003). *Medicaid infrastructure grant to support the competitive employment of people with disabilities*. Department of Health and Human Services, CFDA No. 93.768.
- Cook, T. D. (1985). Postpositivist critical multiplism. In L. Shotland, & M. M. Mark (Eds.), *Social science and social policy* (pp. 21-62). Thousand Oaks, CA: Sage.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Skokie, IL: Rand McNally.
- Cronbach, L. J. (1986). Social inquiry by and for Earthlings. In D. W. Fiske, & R. A. Shweder (Eds.), *Metatheory in social science*. Chicago, IL: University of Chicago Press.
- Cronbach, L. J. (1989). Construct validation after thirty years. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy: Proceedings of a symposium in honor of Lloyd G. Humphreys* (pp. 147-171). Champaign, IL: University of Illinois Press.
- DeLeire, T. (2003). The Americans with Disabilities Act and the employment of people with disabilities. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 259-278). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Disabled Access Credit, Section 44 of the Internal Revenue Code, 26 U.S.C. § 44 (1994).
- Dixon, K. A., Kruse, D., & Van Horn, C. E. (2003). *A survey of employers about persons with disabilities and lowering barriers to work*. New Brunswick, NJ: Heldreich Center for Workforce Development, Rutgers, the State University of New Jersey.
- Exec. Order No. 13,078. (1998, Mar. 13). *Increasing employment of adults with disabilities*. 63 Fed. Reg. 13,111. Retrieved Aug. 24, 2004, from <http://www.jan.wvu.edu/cap/13078.htm>
- Fair Housing Act of 1969, as amended in 1988, 42 U.S.C. § 3604 (f) (1994).

- Goodman, N., & Waidmann, T. (2003). Social Security Disability Insurance and the recent decline in the employment rate of people with disabilities. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 339–368). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Hale, T. W. (2001). The lack of a disability measure in today's Current Population Survey. *Monthly Labor Review*, 124(6), 38–40. Retrieved September 3, 2004, from <http://stats.bls.gov/opub/mlr/2001/06/ressum1.pdf>
- Heckman, J. J., & Hotz, V. J. (1989). Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. *Journal of the American Statistical Association*, 84(408), 862–874.
- Henry, G., & Julnes, G. (1998). Values and realist valuation. In G. Henry, G. Julnes, & M. Mark (Eds.), *Realist evaluation: An emerging theory in support of practice. New directions for evaluation*, No. 78 (pp. 53–72). San Francisco: Jossey-Bass.
- Hill, S. C., Livermore, G. A., & Houtenville, A. J. (2003). Rising health care expenditures and the employment of people with high-cost chronic conditions. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 181–216). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Hotchkiss, J. L. (2003). *The labor market experience of workers with disabilities: The ADA and beyond*. Kalamazoo, MI: Upjohn Institute for Employment Research.
- Individuals with Disabilities Education Act, Part B, 20 U.S.C. §§ 1411–1419 (1994) (amended 1997, 1415 amended in 1999, 1414a omitted by Individuals with Disabilities Education Act Amendments of 1997, Pub. L. No. 105–17, 111 Stat. 37).
- Jensen, A., Silverstein, R., & Folkemer, D. (2002). *Policy frameworks for designing Medicaid Buy-In programs and related state work incentive initiatives*. Retrieved August 18, 2004, from <http://www.medicaidbuyin.org>
- Julnes, G. (2004). [Review of the book *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*]. *Evaluation and Program Planning*, 27(2), 173–185.
- Julnes, G. (2004). Experiment, overview. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement, Vol. 1* (pp. 885–899). San Diego, CA: Elsevier.
- Julnes, G., Fan, X., & Hayashi, K. (2001). Understanding self-sufficiency of welfare leavers in Illinois: Elaborating models with psychosocial factors. In G. Julnes, & E. M. Foster (Eds.), *Outcomes of welfare reform for families who leave TANF. New directions for evaluation*, No. 91 (pp. 33–44). San Francisco: Jossey-Bass.
- Julnes, G., & Foster, E. M. (2001). Crafting evaluation in support of welfare reform. In G. Julnes, & E. M. Foster (Eds.), *Outcomes of welfare reform for families who leave TANF. New directions for evaluation*, No. 91 (pp. 3–8). San Francisco: Jossey-Bass.
- Julnes, G., Hayashi, K., & Anderson, S. (2001). Acknowledging different needs: Developing a taxonomy of welfare leavers. In G. Julnes, & E. M. Foster (Eds.), *Outcomes of welfare reform for families who leave TANF. New directions for evaluation*, No. 91 (pp. 73–85). San Francisco: Jossey-Bass.
- Julnes, G., & Mark, M. (1998). Evaluation as sensemaking: Knowledge construction in a realist world. In G. Henry, G. Julnes, & M. Mark (Eds.), *Realist evaluation: An emerging theory in support of practice. New directions for evaluation*, No. 78 (pp. 33–52). San Francisco: Jossey-Bass.
- Kaye, H. S. (2003). Employment and the changing disability population. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 217–258). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Kirchner, C. (1996). Looking under the street lamp: Inappropriate uses of measures just because they are there. *Journal of Disability Policy Studies*, 7(1), 77–90.
- Kruse, D., & Schur, L. (2003a). The Americans with Disabilities Act and the employment of people with disabilities. In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 279–300). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Kruse, D., & Schur, L. (2003b). Employment of people with disabilities following the ADA. *Industrial Relations*, 42(1), 31–66.
- LaLonde, R., & Maynard, R. (1987). How precise are evaluations of employment and training experiments: Evidence from a field experiment. *Evaluation Review*, 11(4), 428–451.
- Levine, L. (2000). *The employment of people with disabilities in the 1990s*. Washington, DC: Congressional Research Service, The Library of Congress.
- Mark, M., Henry, G., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco: Jossey-Bass.
- Mashaw, J. L., & Reno, V. P. (Eds.). (1996). *The environment of disability income policy: Programs, people, history, and context*. Washington, DC: National Academy of Social Insurance. Retrieved September 3, 2004, from http://www.nasi.org/usr_doc/Environment_of_Disability_Income_Policy.pdf
- Medicare Program, Title XVIII of the Social Security Act, 42 U.S.C. §§ 1395c–1395ccc (1994).

- Medicaid Buy-In program, Section 201 of Pub. L. No. 106-170, which adds sections 1902(a)(10)(A)(ii)(XV) & (XVI) to the Social Security Act.
- Medicaid Program, Title XIX of the Social Security Act, 42 U.S.C. § 1396 (1994).
- Mohr, L. B. (1990). *Understanding significance testing*. Sage University paper series on quantitative applications in the social sciences, series No. 07-073. Thousand Oaks, CA: Sage.
- Mohr, L. B. (1995). *Impact analysis for program evaluation* (2nd ed.). Thousand Oaks, CA: Sage.
- Nagi, S. (1991). Disability concepts revisited: Implications to prevention. In A. M. Pope, & A. R. Torlove (Eds.), *Disability in America: Toward a national agenda for prevention* (pp. 309-327). Washington, DC: National Academy Press.
- National Council on Disability. (2000a). *Back to school on civil rights*. Washington, DC: NCD. Retrieved September 3, 2004, from <http://www.ncd.gov/newsroom/publications/2000/pdf/backtoschool.pdf>
- National Council on Disability. (2000b). *Promises to keep: A decade of federal enforcement of the Americans with Disabilities Act*. Washington, DC: NCD. Retrieved September 3, 2004, from <http://www.ncd.gov/newsroom/publications/2000/pdf/promises.pdf>
- National Council on Disability. (2002). *National disability policy: A progress report for December 2000-December 2001*. Washington, DC: NCD. Retrieved September 3, 2004, from http://www.ncd.gov/newsroom/publications/2002/pdf/progressreport_07-26-02.pdf
- National Council on Disability. (2004). *Improving federal disability data*. Washington, DC: NCD. Retrieved September 3, 2004, from http://www.ncd.gov/newsroom/publications/2004/pdf/improve_data.pdf
- National Institute on Disability and Rehabilitation Research. (1999, Dec. 7). *Correction notice for the final long-range plan for fiscal years 1999-2003*. 64 Fed. Reg. 68,576-68,614.
- Orr, L. L. (1998). *Social experiments: Evaluating public programs with experimental methods*. Thousand Oaks, CA: Sage.
- Rehabilitation Act of 1973, Title I, 29 U.S.C. §§ 720-751 (1994) (amended 1998, 740, 742-750 omitted by Workforce Investment Act of 1998, Pub. L. No. 105-220, 112 Stat. 936, 1116); Section 504, 29 U.S.C. 794 (1994) (amended 1998); and Section 508, 29 U.S.C. 794d). Section 100(a), as amended, 29 U.S.C. 720(a) (1994).
- Reno, V. (2004, April). *Social Security as part of integrated national disability policy—Is the Social Security definition out of sync?* Paper presented at Social Security Advisory Board Discussion Forum on the Definition of Disability, Washington, DC.
- Rosenbaum, P. R., & Rubin, D. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387), 516-524.
- Shadish, W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Silverstein, R. (2000). Emerging disability policy framework: A guidepost for analyzing public policy. *Iowa Law Review*, 85(5), 1691-1806.
- Social Security Act, Section 1619(a), 42 U.S.C. § 1382h(a) (2000); Section 1619(b), 42 U.S.C. § 1382h(b).
- Social Security Advisory Board (2003). *The social security definition of disability, October 2003*. Retrieved October 29, 2004, from <http://www/ssab.gov>
- Social Security Disability Insurance Program, Title II of the Social Security Act, 42 U.S.C. § 401 (1994) (amended 1999).
- Stapleton, D. C. (2003, November). *Safety net or tangle web?: Promoting the economic security of adults with disabilities*. Paper presented at the Urban Institute Policy Forum, Washington, DC.
- Stapleton, D. C., & Burkhauser, R. V. (Eds.). (2003). *The decline in employment of people with disabilities—A policy puzzle*. Kalamazoo, MI: Upjohn Institute for Employment Research.
- Stapleton, D. C., Goodman, N., & Houtenville, A. J. (2003). Have changes in the nature of work or the labor market reduced employment prospects of workers with disabilities? In D. C. Stapleton, & R. V. Burkhauser (Eds.), *The decline in employment of people with disabilities—A policy puzzle* (pp. 125-180). Kalamazoo, MI: Upjohn Institute for Employment Research.
- Supplemental Security Income Program, Title XVI of the Social Security Act, 42 U.S.C. §§ 1381-1383c (1994); 1382-1382b, 1382d, 1382e, 1383, 1383b, amended 1999; 1382c amended 1997; 1383c amended 1996.
- Temporary Assistance for Needy Families program, 42 U.S.C. §§ 601-619; See Preamble to TANF regulations, 64 Fed. Reg. 17720, 17722 (1999, Apr. 12).
- Ticket to Work and Self-Sufficiency program, Section 1148 of the Social Security Act, 42 U.S.C. § 1320b-19 (2000).
- Ticket to Work and Self-Sufficiency Program Regulations. Preamble, 66 Fed. Reg. 67,370 (Dec. 28, 2001).
- Ticket to Work and Work Incentives Improvement Act of 1999, Pub. L. No. 106-170, 113 Stat. 1860.
- U.S. Census Bureau. (n.d.). Survey of Income and Program Participation, 1996 Panel, Wave 5: August-November 1997.

- U.S. Census Bureau. (2004). *Uses and Limitations of CPS Data on Work Disability*. Retrieved August 24, 2004, from www.census.gov/hhes/www/disable/cps/cpstablexplanation.pdf
- U.S. Department of Labor. (2002). Delivering on the Promise (Report to the President on Executive Order 13217). In Department of Health & Human Services, *Compilation of individual federal agency reports of action to eliminate barriers and promote community integration*. Retrieved September 3, 2004, from <http://www.hhs.gov/newfreedom/final/pdf/dol.pdf>
- U.S. General Accounting Office. (2003). *Government auditing standards* (GAO-03-673G). Retrieved September 3, 2004, from <http://www.gao.gov/govaud/yb/2003/html/chap75.html#1035029>
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies*. Upper Saddle River, NJ: Prentice-Hall.
- Work Opportunity Tax Credit, 26 U.S.C. § 51 (1994).
- Workforce Investment Act (WIA), Title I, 29 U.S.C. §§ 2801–2945 (1998).
- Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.