Using the Delphi process to analyze social policy implementation: A post hoc case from vocational rehabilitation

ANDREW J. BUCK,[1] MEIR GROSS,[2] SIMON HAKIM[3] &
J. WEINBLATT[4]
[1]*Department of Economics, Temple University, Philadelphia, PA 19122;* [2]*Department of
Landscape Architecture and Regional Planning, University of Massachusetts, Amherst, MA
01003;* [3]*Department of Economics, Temple University, Philadelphia, PA 19122;* [4]*Department of
Economics, Ben Gurion University, Beer Sheva 84105, Israel*

Abstract. This study uses a Policy Delphi to discern differences in perspective among and
within groups responsible for formulating and implementing vocational rehabilitation policy.
Four groups of players were chosen for our analysis: government officials, academics, directors
of rehabilitation centers, and the staff who interface with program participants. Significant dif-
ferences were found between the groups regarding the relative importance of possible legislative
goals. This suggests that the failure of vocational rehabilitation policy to promote a work agenda
may be attributed to a lack of consensus among policy implementors. The Delphi technique
could help policy planners understand the different perspectives within the implementation
community, and hence craft more realistic legislation.

1. Introduction

It has been 20 years since policy analysts began to recognize the importance
of the implementation process in achieving legislative goals. As idealistic laws
concerning model cities, civil rights, education and employment failed to pro-
duce desired results, scholars turned from a comparatively simple study of
how laws and regulations are produced to a more complex analysis of how
legislative intent is thwarted as it filters through bureaucratic structures. How,
these scholars asked, can the implementation process be made to succeed?

The literature on implementation indicates that a major problem lies in the
way the members of a policy community diverge in terms of perspectives,
goals, and priorities. This suggests that a precise understanding of these diver-
gences would help us to gain a firmer grasp of why legislation goes awry and
how the difficulty might be alleviated. To approach such an understanding, we
use a variant of the Delphi method to analyze the lines of divergence within
the policy community concerned with the implementation of the Federal
Rehabilitation Act of 1973.

The Policy Delphi is a reiterative statistical technique used to discern dif-
ferences in perspective among and within groups responsible for formulating
policy. In our analysis of vocational rehabilitation policy, four groups of
players were chosen: government officials, academics, directors of rehabilita-
tion centers, and the staff who interface with program participants. Each
group in the policy community has a different operational perspective and

hence a potentially different perception of the overall goals of the rehabilitation process. Where and to what extent do differences exist on the perceived relative importance of rehabilitation goals?

The formal purpose of the Federal Rehabilitation Act of 1973 (as amended by Public Law 99–506, Section 2) is to

> ...guarantee equal opportunity, comprehensive and coordinated programs of vocational rehabilitation and independent living for individuals with handicaps in order to maximize their employability, independence, and integration into the work place and the community.

Some state acts are more specific. The Pennsylvania Vocational and Rehabilitation Act of 1987 states that the disabled should be able to 'engage in a gainful occupation,' or 'engage in competitive work,' and 'achieve such ability as independent living.' In both cases, lawmakers emphasized competitive employment, independent living, and earning ability as major goals of the rehabilitation process.

In practice, legislation emphasizing a work agenda for the handicapped has met with little success. Bryen et al. (1987) tracked the 18 months rehabilitation programs of 83 people. The clients were being served at two different centers in New York and Philadelphia. By the end of the programs, only 13 clients had found competitive employment. Within six months, when the financial incentives to the employers were exhausted, only five clients remained in employment. Vachon (1990) relates the results of a 1986 Harris survey and a 1987 General Accounting Office audit of the effectiveness of vocational rehabilitation. The Harris survey found that although 60 percent of disabled adults knew about vocational rehabilitation, only 10 percent had used it, and about half of those reported that it provided little or no help in finding a job. The GAO audit found that less than one percent of Social Security Disability Insurance (SSDI) recipients left the rolls after receiving Vocational Rehabilitation services.

Several studies have explored reasons for failure in the implementation of vocational rehabilitation policy. Rogan and Murphy (1991) identify four principles of supported employment as identified in the 1986 Rehabilitation Act amendments: integrated work settings, paid employment, ongoing support, and priority service provision to people with the most severe disabilities. The authors explore the failure of service providers to implement vocational rehabilitation programs according to these four legislative principles and conclude that the principles were redefined in order to make them more compatible with the existing service system.

Burton (1982) reports on a Classical Delphy study that was used to identify internal and external variables relevant to a rehabilitation agency's ability to provide civil rights protections to its clients. Study participants were asked to rank the desirability, feasibility, and effectiveness of service delivery models designed to protect these civil rights. The Delphi panel concluded

that the most important factor in successful implementation is the attitude of elected and appointed government officials, of rehabilitation agency personnel, and of the voting public toward the goal of civil rights enforcement.

Berkowitz (1987: pp. 178–180) cites a number of practical reasons for the failure of vocational rehabilitation programs to meet employment goals. For instance, the 1973 act mandated that the severely disabled be given preference over those with mild impairments in rehabilitation programs. This eliminated the practice of 'creaming,' where participants likely to succeed were hand-picked by program administrators. He also identifies rational disincentives to finding work, such as the threat of losing medical insurance.

These practical concerns are symptomatic of a disability policy which represents diverse and often conflicting ideologies. While civil rights laws and training programs were enacted to open the doors to productive employment, advocates for the handicapped have argued for entitlements and release from requirements to look for work. 'Independent living' provisions in the 1973 bill were battled by the Nixon administration for cost reasons (Berkowitz, 1987). Vachon (1990: p. 46) states that a fundamental problem with Vocational Rehabilitation is that 'it is not clear if it is an employment program, a social welfare program, or something else.' There is apparently much disagreement within the disability policy community concerning preferred goals.

This study uses a Policy Delphi to discern differences in perspective among and within groups responsible for formulating and implementating vocational rehabilitation policy. Four groups of players were chosen for our analysis: government officials, academics, directors of rehabilitation centers, and the staff who interface with program participants. Each group in the policy process has a different operational perspective and hence a potentially different perception of the overall goals of the rehabilitation process. Where and to what extent do differences exist on the relative importance of rehabilitation goals?

Section 2 presents a brief review of the relevant implementation literature and a description of the Policy Delphi method. The empirical results of the Delphi study and their interpretation are given in Section 3. In Section 4, we summarize our major findings and conclusion.

2. Policy implementation and the Delphi method

Scholars of public policy have long theorized on the sources of general failure in the implementation of legislative directives. Edwards (1984), Hogwood and Peters (1985), and Jenkins (1978) attribute failure in policy implementation to lack of consensus, coordination, and vertical communication of intentions. Levin and Ferman (1986) believe that the heterogeneous nature of American society and its fragmented political system are a source of difficulty in policy implementation, and show that in the field of youth employment, executives have improved implementation by building consensus. Schulman

(1980) adopts a biological-organis view of failure in the policy implementation process. He claims that all policy implementation components (e.g., legislation, budget appropriations, personnel-client interaction) should be in accord with the desired result. Otherwise, like organs in living bodies, they will try to reject each other. In the lack of such accord, successful goal achievement cannot be expected.

Pressman and Wildavsky (1973) present a case study of the failure of the Economic Development Administration to successfully administer a jobs creation program targeted to low-income populations in Oakland, California. They conclude in part that administrtion is hampered by contradictory criteria and by antagonistic relationships among participants in the implementation process. Bardach (1977) notes that policy initiatives are typically sponsored not by individuals, but by coalitions encompassing diverse and incongruent goals. The resulting legislation is vaguely worded, general, and subject to a multitude of interpretations. The implementation process is marked by gamesmanship, where each group of actors interprets the legislation in a way which best serves its own interests.

While these studies provide useful theoretical structures for the analysis of policy failure, they do not attempt to quantify the degree of conflict among players regarding specific policy objectives. While Bardach's qualitative analysis explores this area most incisively, a quantitative reading of the differing programmatic objectives brought to the implementation process can lead to a clear understanding of exactly where difficulties lie and aid lawmakers in their attempts to formulate successful legislation.

The Delphi statistical method has traditionally been used to generate a group opinion or 'statistical consensus' from a panel of experts on technical or prognostic issues. The primary components of the Classical Delphi are the systematic generation of options, the opportunity for panel members to assign weights to each option (and sometimes attach a written rationale), anonymous feedback of responses within the panel in a multi-round process (which gives members a chance to modify their views in accordance with group opinion), and a final statistical group consensus (Linstone and Turoff, 1975). The underlying assumption of the Delphi technique is that the statistical information provided to each panelist in successive rounds brings about a reconsideration of previous estimates and hence generates a higher degree of consensus.

The original technique was formulated by the Rand Corporation in the early 1950's to estimate the probable effects of an atomic bomb attack upon the United States (Helmer and Rescher, 1959). In the late 1950's and early 1960's, Delphi was applied to technological forecasting, management science, and operations research (Helmer, 1966). In the 1970's, the method was used in the fields of planning, project evaluation, and cost-benefit analysis (Shefer and Strousma, 1982).

Since its inception, the technique has been modified to meet diverse needs. The Policy Delphi was introduced in 1970 as a means of defining and dif-

ferentiating the viewpoints of advocate groups rather than establishing the consensus of unbiased experts. It was recognized that the idea of unbiased technical expertise does not apply to policy issues, which are by their nature defined by conflicting special interests. It is unlikely, too, that consensus could be generated through the process within a group of individuals who are each strongly committed to their ideological stance. The Policy Delphi is used as a means of analysis instead of a tool for decision-making or prognostication (Turoff, 1975; Rauch, 1979).

In this study, the Policy Delphi method is used to test for differences and consistency in the perception of legislative goals and objectives among and within the various participant groups who create and implement vocational rehabilitation policy.

A list of potential outcomes and suboutcomes (goals and objectives) of the rehabilitation process is systematically generated by a panel of experts. Representatives from various groups within the policy community assign numeric weights to these outcomes, indicating relative preferences, in a multi-round process. Between rounds, statistical information regarding the numeric responses of the group is fed back to all participants, who change their responses as they see fit. The result is a snapshot of the fissures within the policy community and the strength of conviction with which views are held.

We utilize the following Delphi features: (1) *Anonymity:* Panel members are not told who provided specific responses, although they may know the other members of the panel. This enables a panelist to express an opinion and to change it without undue influence from other panel members. (2) *Controlled feedback:* The feedback procedure assures that only directly relevant information is asked of, and provided to, the panel. This differs from the committee method, which tends to produce discussions that obscure issues. (3) *Statistical measurement of agreement and dissent:* The information fed to panelists after successive rounds includes a statistical analysis of the panel's previous answers to a questionnaire. For each question, this analysis includes measures of central tendency for the panel's responses to each question and at least one measure of dispersion. Such measures determine whether more rounds are necessary.

The criteria used to decide when to terminate the Delphi process are *stability* and *convergence* (Nelson, 1978). *Stability*, or consistency, measures the similarity or central tendency of the panel's response to each question across rounds. If the difference in, say, the response frequencies from one round to another is below a given threshold, the particular question is not included in the next round's questionnaire. *Convergence* is based on the degree of agreement, or consensus, achieved by the panel in its response to a given question. A significant increase in the dispersion of responses for a question indicates a decreasing level of agreement, whereas a decrease in dispersion indicates a move toward statistical consensus.

We established two panels of participants. The first group approximated an 'expert panel' in the Classical Delphi sense. This panel consisted of the 20

members of the Vocational Rehabilitation Advisory Council. The panel was asked to generate a list of desired outcomes of the vocational rehabilitation process. Each member was requested by mail to list the outcomes that they expected to emerge from vocational training.

We classified and generalized responses into two lists, the first consisting of six general and non-quantifiable goals of the rehabilitation process. These six 'outcomes' were labeled as financial status, employment opportunities, job satisfaction, working skills, functional independence and social integration.

A special effort was made to limit the number of outcomes to six. The purpose was to make the rating of their relative importance an easy task with reliable results. To make the outcomes as clear as possible to the respondents, a brief definition and description was attached to each outcome. (The reader will note that the names of the categories have been abbreviated in the tables. Responses which addressed specific programmatic objectives were categorized into a related list of discrete 'sub-outcomes.')

Each general outcome was divided into three more specific sub-outcomes. (In table 1, the sub-outcomes have been shortened somewhat for space considerations.) Each sub-outcome or objective had to include a basic characteristic of its corresponding general outcome. In addition, each set of three sub-outcomes was to be treated as categorically discrete. This list was returned with the original lists to each member of the advisory council. They were requested to add, modify, or delete from the list, as well as to revise their definitions and explanations. From this, the final Delphi questionnaire was created and used as the main instrument for allowing the second panel to rate the importance.

The second panel consisted of position advocates in the implementation process. This group will be referred to as the 'policy panel.' A comprehensive list of interested parties was compiled, including academicians, federal and state government officials involved in rehabilitation issues, and directors of rehabilitation agencies. This policy panel was given the task of assigning relative weights to each of the general outcomes and sub-outcomes on the list generated by the first panel. Policy panelists were requested to allocate weights to the six general outcomes so that they totalled 100, and to do the same for each group of three specific outcomes.

To complete the testing of policy implementation, a third group was created. We used a sample of staff workers in two rehabilitation centers located respectively in Philadelphia and New York City. They were given the final questionnaire and were asked to rank the general outcomes of the rehabilitation process. Staff workers were tested as a group so that their responses could be compared directly with those of higher echelon implementors and policy makers. The purpose of this was to test the degree of consensus between the two ends of the policy implementation process. Government agency officials represent one end, mainly responsible for budget appropriation and the determination of criteria, but without direct contact with the clients. At the other end, rehabilitation center staff members are closely involved with the clients and are in charge of the execution of the policy.

Table 1. Average weights given to outcomes by the policy panel.

	Round 1 (n = 111)			Round 2 (n = 89)		
	Mean weight	Standard deviation	Coeff. of variation	Mean weight	Standard deviation	Coeff. of variation
General outcomes						
1. Financial	22.3	14.9	0.67	21.4	11.1	0.52
2. Employment	23.2	14.5	0.62	24.0	13.1	0.54
3. Job satis.	13.1	8.4	0.64	12.7	6.3	0.50
4. Work Skills	15.0	8.4	0.56	15.4	7.9	0.51
5. Independence	15.8	9.2	0.58	15.6	7.8	0.50
6. Integration	10.6	9.0	0.86	10.9	9.1	0.83
Specific outcomes						
1.1 Wages	34.1	17.8	0.52	34.1	15.1	0.44
1.2 Income	34.9	18.5	0.53	37.5	16.5	0.44
1.3 Less public support	31.0	16.5	0.53	28.4	12.9	0.45
2.1 Stable employment	36.0	16.7	0.46	34.8	13.7	0.39
2.2 Competitive employment	41.5	16.6	0.40	42.8	14.1	0.33
2.3 Advancement	22.5	12.7	0.56	22.4	11.7	0.52
3.1 Satis. work	39.7	12.1	0.30	38.5	12.9	0.33
3.2 Satis. wage	34.9	12.8	0.37	34.6	13.0	0.38
3.3 Colleagues	25.4	8.2	0.32	26.9	7.6	0.30
4.1 Productivity	29.5	12.5	0.43	29.5	11.0	0.37
4.2 Work adjust.	36.9	14.1	0.38	37.8	11.9	0.32
4.3 Spec. skills	33.6	14.2	0.43	32.7	12.9	0.40
5.1 Living	48.9	18.0	0.37	49.9	14.1	0.28
5.2 Transit	23.5	10.7	0.45	22.9	10.3	0.45
5.3 Appearance	27.5	13.4	0.49	27.2	9.9	0.36
6.1 Family	34.8	12.7	0.36	35.7	9.4	0.26
6.2 Friends	30.6	9.3	0.31	31.5	5.6	0.18
6.3 Community	34.6	14.9	0.43	32.8	10.5	0.32

After the first round, each member of the policy panel received a copy of his or her assigned weights as well as the entire panel's mean weights for each outcome. The participants were requested to review the material and, if they chose, to change their weightings in light of the responses of the other participants. Ideally, the iterations continue as long as significant changes occur. Student t statistics were calculated to identify if there were significant differences between the means for each of the general and specific outcomes for successive rounds. As in previous studies, only slight changes occurred after the second round (e.g., Gross et al., 1983). Thus, only two rounds were conducted (table 1).

In total, 189 questionnaires were sent, of which 111 were returned with the requested information. These were used to determine the weights shown in Table 1, round 1. For the entire panel and each subgroup of the panel we computed means, standard deviations, and coefficients of variation. These measures were used to test for stability and convergence of responses. Such measures are meaningful only if the responses for each of the goals and objectives approximate the normal distribution. The larger the size of the group, the better the chances of approaching a normal distribution. By the end of the second round, the number of respondents had diminished to 89.[1]

A practical barrier to the pursuit of several rounds in the Delphi procedure is the decline in the response rate. This problem is particularly acute when Delphi is conducted by mail and does not generally exist when a panel is gathered in situ for the Delphi procedure. On the other hand, Delphy by mail has the major advantage of incorporating a larger panel. We chose to pursue Delphi by mail in order to exploit his latter attribute, despite potential shortcomings in the response rate.[2]

3. Results and analysis

In examining the implementation of a legislatively mandated policy, we first consider the intent of the law. Bardach (1977: p. 313) supports the adoption of the legislative perspective for similar reasons: while taking on the perspective of any advocate group may lead to a facile understanding of how to subvert implementation, it is more challenging to think of ways to make our current implementation process succeed.

The following section presents the views of government officials, agency heads, academicians, and staff regarding the relative importance of each outcome of the rehabilitation process. We test for consensus among and within the various groups that comprise the panel and between the staff members and the panel. Because the legislation does not provide a ranking of the relative importance of each outcome we are not able to statistically compare the priorities of the panel with those of legislators. We do, however, make a qualitative comparison based upon the wording of the legislation.

Five tables of statistical information are constructed to illustrate similarities and differences in perspective among and within participant groups. Tables 1 and 2 look at the average weights given to each outcome and suboutcome by the policy panel – that is, the three groups of government officials, agency heads, and academicians combined. Table 3 tests for consistency in that it asks whether the second round of the Delphi process produces a significant difference in average statistical opinion within each group – for instance, did academics as a group change their mind about any specific outcome after becoming aware of the opinions of other panel members? Table 4 looks at differences not among but between groups – after the second round, are there significant differences in opinion between government officials,

Table 2. Average outcome weights given by experts: Only those participating in both rounds.

General outcomes	Round 1 (n = 89)			Round 2 (n = 89)		
	Mean rank (%)	Standard deviation	Coeff. of variation	Mean rank (%)	Standard deviation	Coeff. of variation
Financial	21.6	14.3	0.66	21.4	11.1	0.52
Employment	23.4	15.1	0.65	24.0	13.1	0.55
Job satis.	12.9	7.4	0.57	12.7	6.3	0.50
Work skills	15.1	8.5	0.56	15.5	7.9	0.51
Independence	16.0	9.2	0.57	15.6	7.8	0.50
Integration	10.7	9.5	0.89	10.9	9.1	0.83

agency heads, and academicians concerning the relative importance of the different policy outcomes? Table 5 tests for differences in perspective between the policy panel and the staff at rehabilitation centers – those most directly responsible for carrying out policy directives.

3.1 General results of the Delphi process

Table 1 shows the average views of the entire policy panel for both rounds one and two. The information in the upper part of the table includes the mean weights given by the panel as a whole to the six general outcomes. The sum of the weights assigned to the six categories should be 100.[3] Standard deviations and coefficients of variation are included for the range of responses to each outcome.

In the first round, the general outcome of improved employment status (outcome 2) is ranked the highest in importance by the entire panel (weight = 23.2 percent). This is closely followed by financial independence (22.3 percent). Next in the ranking are functional independence (15.8 percent) and improved work skills (15 percent). Integration into the client's family and community was considered the least important by the panel as a whole. The policy panel assigned the greatest weight to those outcomes that are related to employ-ability.

The lower part of Table 1 shows the weights assigned to the suboutcomes or specific objectives. The three suboutcomes within each outcome category must also sum to 100. Among the suboutcomes, competitive employment, increased total income and independent living are ranked the highest in their respective groups. This ranking is consistent with the purpose of legislation both at the federal and state levels, where the emphasis in rehabilitation is on 'competitive' employment, 'gainful occupation' (income) and 'independent living.' In other words, the Delphi process reveals a general consistency between the policy panel and the goals of the legislators as articulated in the 1973 legislation.

Table 3. Significant differences between rounds by group and issue[1].

Issue	Group	Round	Mean	Standard deviation	t-Statistic	F-Statistic
Financial	Entire panel	1	22.1	15.1	0.52	1.93[2]
		2	21.1	10.8		
	Academicians	1	9.4	7.2	−0.60	1.15
		2	11.2	7.7		
	Gov't officials	1	23.9	15.1	0.31	1.98[3]
		2	23.3	10.7		
	Agency heads	1	25.0	15.4	0.82	3.89[4]
		2	20.8	7.7		
Employment	Entire panel	1	23.2	14.5	−0.33	1.23
		2	24.0	13.1		
	Academicians	1	27.0	21.6	−0.03	1.28
		2	27.2	19.1		
	Gov't officials	1	22.8	14.3	−0.42	1.24
		2	23.8	12.8		
	Agency heads	1	23.3	8.3	0.36	1.57
		2	22.1	6.7		
Job satisfaction	Entire panel	1	13.1	8.4	0.34	1.76[3]
		2	12.8	6.4		
	Academicians	1	11.9	7.1	0.17	1.02
		2	11.5	7.1		
	Gov't officials	1	13.9	9.1	0.38	1.95[3]
		2	13.4	6.5		
	Agency heads	1	9.8	3.5	−0.55	1.39
		2	10.6	3.0		
Work skills	Entire panel	1	15.0	8.4	−0.51	1.06
		2	15.4	7.9		
	Academicians	1	14.7	9.3	0.35	1.02
		2	13.5	9.2		
	Gov't officials	1	14.9	8.2	−0.64	1.05
		2	15.1	8.0		
	Agency heads	1	18.8	8.8	−0.50	1.47
		2	20.6	2.4		
Independence	Entire panel	1	16.0	9.1	0.47	1.51[5]
		2	15.4	7.4		
	Academicians	1	18.5	12.0	0.06	1.01
		2	18.2	11.9		
	Gov't officials	1	16.0	8.9	0.82	2.05[2]
		2	14.9	6.2		
	Agency heads	1	13.3	5.4	−0.66	1.30
		2	15.0	6.1		
Integration	Entire panel	1	10.6	9.0	−0.47	1.07
		2	10.9	9.1		
	Academicians	1	18.0	17.9	0.00	1.07
		2	18.5	18.5		
	Gov't officials	1	9.4	5.9	−0.43	1.06
		2	9.6	5.8		

Table 3 (continued).

Issue	Group	Round	Mean	Standard deviation	t-Statistic	F-Statistic
	Agency heads	1	9.8	6.4	−0.46	1.90
		2	11.0	4.7		

[1] Where variances were not homogeneous the t-test was based on separate variance estimates. Both t and F tests are two tailed.
[2] Significant at the .005 level.
[3] Significant at the .01 level.
[4] Significant at the .1 level.
[5] Significant at the .05 level.

Table 4. Test statistics for between group differences by issue second round.

		Academicians		Agency heads	
		Different mean: t	Different variance: F	Different mean: t	Different variance: F
Gov't officials	Financial	3.87[1]	1.93	0.67	1.93
	Employment	−0.63	2.23[3]	0.59	3.71[3]
	Job satisfaction	0.95	1.21	2.17[3]	4.69[3]
	Work skills	0.65	1.34	−1.94[3]	1.20
	Independence	−0.97	3.70[4]	−0.04	1.03
	Integration	−1.71	10.17[1]	−0.70	1.04
Agency heads	Financial	2.88[2]	1.00		
	Employment	−0.89	8.27[2]		
	Job satisfaction	−0.41	5.66[3]		
	Work skills	−1.93[4]	1.61		
	Independence	−0.83	3.80[4]		
	Integration	−1.39	15.78[1]		

[1] Significant at the .005 level.
[2] Significant at the .01 level.
[3] Significant at the .05 level.
[4] Significant at the .1 level.
All test are two tailed.

When we look at the F-Statistics, we see that the entire policy panel moved toward consensus on the relative importance of issues of financial independence, job satisfaction, and functional independence. Both government officials and agency heads moved toward agreement on the outcome of financial independence by the second round. Variance of opinion was most markedly reduced within the group of government officials – they significantly moved toward inter-group consensus on the issues of job satisfaction and functional independence as well as financial independence.

Table 3 gives an interesting overall perspective. While all three groups consider employment opportunities to be the most important outcome, government officials and heads of agencies regard financial status as being almost as

Table 5. Differences in outcome weights between expert groups and rehabilitation center staff members second round.

Group		General outcome	Finan-cial	Employ-ment	Job satis-faction	Work skills	Inde-pen-dence	Integra-tion
	Mean staff rank, round 2	14.85	18.67	13.55	24.04	15.63	13.07
Total expert panel	t-Statistic	3.42	1.93	−0.58	−3.36	−0.11	−0.99
	2 Tail prob.	0.001	0.056	0.562	0.002	0.914	0.325
Government officials	t-Statistic	−4.32	−1.82	0.13	3.44	0.39	2.39
	2 Tail prob.	0.000	0.072	0.900	0.002	0.699	0.019
Agency heads	t-Statistic	−2.10	−0.94	1.65	0.79	0.20	0.79
	2 Tail prob.	0.043	0.355	0.108	0.434	0.840	0.438
Acade-micians	t-Statistic	−0.74	−1.51	1.06	2.72	−0.79	−1.01
	2 Tail prob.	0.468	0.151	0.295	0.010	0.434	0.330
Center 1 staff versus center 2 staff	t-Statistic	−0.74	−1.64	0.09	0.17	−0.30	3.76
	2 Tail prob.	0.468	0.114	0.929	0.869	0.765	0.001

important as employment, while academicians rank it as least important. Academicians believe that independence and integration of the disabled are outcomes of major importance, a view which is not in accord with the other two groups. Agency heads consider the improvement of working skills an important outcome, a view which is not shared by the government officials and the academicians.

Table 4 tests for differences in mean and variance between the different groups. When we look at the t-Statistics, we see that government officials and agency heads disagree on the relative importance of job satisfaction and work skills, that government officials and academicians disgree on the importance of financial independence, and that agency heads and academicians disagree on financial independence and work skills. Financial independence, the development of work skills, and the attainment of job satisfaction seem to be the most controversial policy goals.

A look back to Table 3 gives us the absolute weights assigned by each group. Here we see that government officials gave the outcome of financial independence a relative weight of 23.3, while academics assigned a mere 11.2 out of one hundred. Agency heads were far more closely aligned with government than academia here, giving financial independence a 20.8 relative weight.

Other significant differences between groups can be looked at in the same way – where, for instance, is the policy community aligned on the outcome of

enhanced work skills? Agency heads consider this to be the second most important outcome, weighting it at 20.6, while academicians assign a much lower 13.5, giving first precedence to employment, independence, and integration. Job satisfaction, while not highly valued by any group, is given the lowest rating of all by agency heads and is viewed most kindly by government officials.

An important arithmetic comment on the findings in Table 4 is that the number of significant differences between groups is five. This seems to be a relatively small number when compared with the possible eighteen cases of disagreement (six times three pairs of groups). Still the results indicate the existence of substantial differences in the perceived importance of major outcomes among the groups.

t-tests on differences in ratings were conducted for the eighteen specific outcomes in a similar fashion to the tests which are shown in Table 4. Due to space constraints we do not show the additional results. Nonetheless, we found that the number of significant differences on the specific outcomes is much smaller than on the general outcomes. In the second round of the Delphi process, significant differences were found only for two specific outcomes, both between government officials and academicians.

3.2 Consistency between rounds

By consistency we mean the degree of change in the average weight assigned by each participating group to each outcome between rounds. The consistency between the rounds for all groups reflects both clarity of understanding and the degree of commitment felt toward the goals (general outcomes) and objectives (specific outcomes). If the weights change significantly from one round to the next then it may imply a lack of basic understanding or the lack of a firm commitment to beliefs.

The t-Statistics in Table 3 test for consistency between rounds – that is, the significant differences in the mean weightings assigned to each outcome by each participant group between the first and second rounds of the Delphi process. Thus, this table reveals changes in perspective within each group after members became aware of larger group opinions. The t-Statistics from Table 3 show no significant differences in mean weights of policy outcomes by any of the groups after the second round, indicating a high level of consistency. This implies that the panel and each of its group is consistent and has its own clear ideas on the importance of goals and objectives of the rehabilitation process.

In the same vein, Table 1 shows us that the rank ordering of the major outcomes did not change for the entire panel between rounds one and two. Apparently neither the entire panel or its constituent groups found it necessary to revise their average opinions between rounds.[4]

3.3 Variance of opinion

In the previous section it was inferred from the lack of change in mean weights that the panel participants had clear prior values concerning the mission of rehabilitation. But how broad is the distribution of beliefs within the entire panel and its constituent groups? In considering variation within the panel as a whole we return to table 1 and consider the coefficient of variation. A greater coefficient of variation indicates a greater range of beliefs among the panel members concerning an issue.

In both rounds of the Delphi process the issue of greater integration provoked the greatest diversity of responses. In fact, in the second round the coefficient of variation for all issues but integration are in the neighborhood of 0.5; in some instances it fell greatly. The coefficient of variation for integration is more than 50 percent higher than the others. Some panel members, after completion of the second round, continued to hold extreme views about the importance of achieving integration through the rehabilitation process.

In reviewing the specific outcomes, the degree of harmony within the panel is much greater. This may be because the larger ideological issues represented by the major outcomes are not at stake at the level of the sub-outcome – respondents are asked to rank objectives where the contextual goal is a given factor.

The range of disagreement within each group of the policy panel is represented by either its variance or coefficient of variation. We use the coefficient of variation since it deflates the standard deviation by the mean. Although they are not included in table 3, one can calculate the coefficients of variation for each group and each outcome for both rounds of the Delphi. The coefficients of variation indicate that agency heads exhibit the greatest internal consensus, while academicians have the least internal agreement. This is true in both rounds.

The F-statistics in Table 4 test for significant differences in variance between the different groups of the panel. Some of the F values in Table 4 are statistically significant, implying that the variance (or the lack of consensus) within one group is significantly different from the variance within the other group. Most of the significant F values are between academicians and the other two groups, indicating that academicians exhibit more internal disagreement than the other groups.

3.4 Convergence

Convergence indicates a move toward consensus within a group. The coefficients of variation show that in the first round of the Delphi process the diversity of opinions was greater than in the second round. Table 1 shows a consistent decrease of the coefficients of variation between the first and the second round, indicating an increase in consensus (greater movement toward the mean) for all general outcomes and for most of the specific outcomes.

The comparison of Tables 1 and 2 indicates that even when the panel is the same in both rounds, the degree of consensus rises significantly from the first to the second round. Thus, the respondents took into account the group's weights in reconsidering their own weights for the various outcomes in the second round.

The F-statistics in Table 3 show statistically significant decreases in the variance of opinions (measured in units of standard deviation) from the first to the second round for the entire panel and for all three groups. While consensus increased for government officials and agency directors, variance did not significantly change among the academicians. In our analysis, academicians exhibit the most diverse opinions and are the most persistent in their views.[5]

Agency heads are also relatively intransigent in their views. Only in the objective of financial status do we witness a significant improvement in intragroup consensus from the first to the second round of the Delphi procedure. Government officials reveal the greatest flexibility regarding the relative importance of each outcome.

3.5 Panel groups versus agency staff

We tested whether the weights given by staff members in the two rehabilitation centers coincide with those of the groups in the policy panel. The first row of Table 5 shows that staff workers in the centers value the goal of working skills most highly in the rehabilitation process. Another important result is that staff members do not value the goal of financial status very highly. In this regard, their views are similar to those of academicians but very different from the views of agency heads and government officials.

Table 5 tests for significant differences in the mean ranking of the staff members compared with those of the three other groups. The table shows that significant differences between staff members and the panel are found on the three issues of financial status, employment opportunities, and working skills. The difference in opinion is the greatest between staff members and government officials. They differ on four outcomes. The difference in opinion is much smaller when agency staff are compared with academicians and agency heads. It is interesting to note that staff members from the two centers value the controversial outcome of social integration differently (the final rows of Table 5). This indicates that rehabilitation staff cannot be perceived as a homogeneous group.

These findings show significant differences in the perception of the goals between the actual executors of the policy (heads of agencies and staff), and the academicians and government officials who are remote from daily contacts with the clients. The particular views of the agencies may result from existing programmatic structures which stress client work skills over financial independence. It may be possible that agencies are rationalizing their failure

to achieve those goals of financial independence and improved employment which are most important from the legislators' and government officials' point of view. Conversely, since agencies are in closer contact with clients, they may have a more realistic view of what is ultimately achievable or desirable and this cleft in opinion may indicate unrealistic expectations on the part of academics and government officials.

4. Conclusions

In this paper, we have used the Delphi method to understand the different perspectives held by various groups involved in setting and implementing vocational rehabilitation policy. We tested for preference and consistency in evaluating the relative importance of rehabilitation goals and objectives among and within groups of policy players. We found that significant differences exist among the various professional groups involved in the policy process. Competitive employment and improved financial independence were rated the highest among the goals by both the federal and state legislators in our case study. Government officials and heads of centers regard financial independence as important as employment, while academicians rank it as least important. Academicians regard personal independence and social integration as goals of major importance, a view not shared by others. Heads of centers and the staff value work skills more highly than do the other groups. Unlike all other groups, the staff considered financial status as unimportant. With the exception of government officials, a high degree of consistency prevails in each group regarding the relative importance of each outcome.

The fact that policy players were more likely to agree with each other at the level of the suboutcome indicates that consensus on specific programmatic objectives is achievable when there is assent on clearly articulated goals of a larger scope. While the various groups held divergent opinions on the relative importance of large goals, agreement on objectives prevailed when any one goal was taken as a given.

Policy outcome is ultimately determined at the point where the local rehabilitation center and the handicapped individual interact. The Delphi method could be used to construct a feedback system that channels the views of field workers and clients to academics and government officials in order to broaden the scope of agenda-setting with the goal of scripting more realistic and implementable legislation. The Delphi approach could clarify and depersonalize communication among the various groups. Its anonymous goal differentiation and feedback structure allows a focus upon issues that is often stymied by the dynamics of a committee or the influence of powerful or vocal special interests in the political process. In helping to clearly frame and articulate potential legislative goals, the Policy Delphi can set the stage for fruitful debate and discussion.

In this study we have used the Delphi method, post hoc, to evaluate consis-

tency and consensus regarding the goals and objectives of the groups involved in the vocational rehabilitation of handicapped people. *A Priori* uses of the Delphi method in policy planning and implementation might provide for a more inclusive and realistic process – a process that might ultimately allow for a greater likelihood of program success.

Acknowledgements

Notes

1. Bardecki (1984) contends that the Delphi encourages a false conformity in its promotion of minimal opinion dispersion. Those whose opinions fall in the tails of the response distribution tend to drop out of the process unless they are highly motivated and dogmatic.
2. Using the mails also meant that, unlike the participants in most Delphi applications, our panel participants remained anonymous to one another.
3. Rounding errors may result in the table not adding to 100.
4. One possible statistical problem in comparing the rankings of the first and second rounds is that the number of respondents was not the same in the two rounds. To test how this would affect the result, we checked the first and second round mean weights and their standard deviations just for the group of experts who responded in both rounds. The results are shown in Table 2. Adjusting the calculations to reflect dropouts does not change the conclusion that the mean rankings did not change between the rounds.
5. Bardecki (1984) might interpret these results as indicating that academicians are the most dogmatic members of the panel.

References

Arrow, Kenneth J. (1963). *Social Change and Individual Values.* New York: Wiley.

Bardach, Eugene (1978). *The Implementation Game: What Happens After a Bill Becomes Law.* Cambridge: MIT Press.

Bardecki, Michael (1984). 'Participants' response to the delphi method: An attitudinal perspective.' *Technological Forecasting and Social Change* 25: 281–292.

Berkowitz, Edward (1987). *Disabled Policy: America's Programs for the Handicapped* (A Twentieth Century Fund Report). New York: Cambridge University Press.

Bryen, Diane et al. (1987). 'Barriers to employability of persons with handicaps.' Philadelphia: Temple University Development Disability Center, submitted to the U.S. Department of Labor (grant number J-9E-3-0172).

Burton, Lloyd (1982). 'The rehabilitation agency and the civil rights of clients: An analysis of strategies for policy implementation.' *Journal of Rehabilitation* 48 (2): 46–52.

Edwards, George C. III, ed. (1894). *Public Policy Implementation.* JAI Press, Connecticut.

Gross, Meir et al. (1983). 'Evaluation of alternative arterial roads: A modified delphi approach.' *Environments* 15 (1): 27–38.

Helmer, O. (1966). *Social Technology.* New York: Basic Books.

Helmer O. and N. Rescher (1959). 'On the epistemology of the inexact science.' *Management Science* 6(1): 25–52.

Hogwood, Brian W. and R. Guy Peters (1985). *The Pathology of Public Policy.* Oxford: Clarendon Press.

Jenkins, W. J. (1978). *Policy Analysis: A Political and Organizational Perspective.* New York: St. Martin Press.

Levin, Martin and Barbara Ferman (1986). 'The political hand: Policy implementation and youth employment programs.' *Journal of Policy Analysis and Management* 5(2): 311–325.

Masser, Ian and Paul Foley (1987). 'Delphi revisited: Expert opinion in urban analysis.' *Urban Studies* 24(3): 217–225.

Nelson, Bradley W. (1978). 'Statistical manipulation of delphi statements: Its success and effects on convergence and stability.' *Technological Forecasting and Social Change* 12(1): 41–60.

Pennsylvania General Assembly (1987). *Vocational and Rehabilitation Act,* Senate Bill No. 1190. Harrisburg, PA.

Percy, Stephen L. (1989). *Disability, Civil Rights, and Public Policy: The Politics of Implementation.* Tuscaloosa: University of Alabama Press.

Pressman, Jeffery and Aaron Wildavsky (1973). *Implementation.* Berkeley: University of California Press.

Rauch, Wolf (1979). 'The decision delphi.' *Technological Forecasting and Social Change* 15: 159–169.

Rogan, Pat and Stephen Murphy (1991). 'Supported Employment and vocational rehabilitation: Merger or misadventure?' *Journal of Rehabilitation* 57(2): 39–45.

Shefer D. and J. Stroumsa (1982). 'Street lighting projects selection: A rational decision making approach.' *Socio-Economic Planning Sciences* 16(6): 245–259.

Schulman, Paul R. (1980). *Large-Scale Policy Making.* New York: Elsevier.

Turoff, Murray (1975). 'The policy delphi,' in Hal Lindstone and Murray Turoff, eds. *The Delphi Method: Techniques and Applications.* Reading, MA: Addison-Wesley.

U.S. Congress. *Rehabilitation Act of 1973.*

Vachon, R. Alexander (1989–90). 'Employing the disabled.' *Issues in Science and Technology* 15(3): 44–50.