May 1983

Legislative Evaluation of Social Welfare Programs: The Process and Consequences

Patricia K. Freeman
University of Tennessee

William Lyons
University of Tennessee

Follow this and additional works at: https://scholarworks.wmich.edu/jssw

Part of the Social Welfare Commons, and the Social Work Commons

Recommended Citation
Available at: https://scholarworks.wmich.edu/jssw/vol10/iss2/8
Program review has increased substantially in recent years. There has been concern that the upsurge in policy evaluation would be detrimental to social welfare programs, where data is often unavailable, little agreement exists concerning specific goals and documenting success is often virtually impossible.

In this analysis we compared the program evaluation of social welfare programs with non-social welfare programs. The review process differed substantially as did the recommendation made as a result of the review. Social welfare programs faced more criticisms than non-social welfare programs. However, the criticisms were less serious. Changes prior to the publication of the evaluation report were more common for social welfare programs. They were also considerably more likely to make significant administrative changes—action that helps them avoid legislative involvement. Non-social welfare programs were more than twice as likely to face changes mandated by the legislature. Those findings indicate that social welfare programs are not experiencing adverse consequences as a result of mandatory reviews.

The evaluation of public policies has increased substantially in the past decade. In the early 1970s, about two hundred new evaluation studies were begun each year with average budgets of about $100,000 each. By now the number of policy and program evaluation studies started each year has probably doubled, and costs risen substantially. (Freeman, 1977: 19) Another indication of the growing interest in program evaluation is the fact that federal and state legislation involving public programs increasingly includes formal evaluation of the policy's impact as part of the bill. (Nachmias, 1979: 2) Finally, and having the greatest impact, is the dramatic increase taking place in the states in the exercise of legislative oversight and the expertise with which it is conducted. Program review and policy evaluation have become a part of the legislature's workload. (Keefe and Ogul, 1981: 385-408)

The greater involvement with program evaluation among legislators can be attributed to several factors. There have been considerable improvements in staff, facilities and technical hardware in most state legislatures in the past two decades which have facilitated the conduct of policy evaluation. (Hamm and Robertson, 1981) Also important is the increased emphasis on legislative independence and legislative activity that accompanied the move toward professionalization of state legislatures.

* Both authors are affiliated with the Bureau of Public Administration, Knoxville.
Thus the mechanisms have been established for policy evaluation and program review and legislators are committed to increased activity in this area; from all indications legislatures will become increasingly involved in policy evaluation.

Significant questions that have received little or no attention are one, how these evaluations are to be conducted and secondly, the types of changes that are recommended and enacted into law as a result of these program evaluations.

These questions are of particular relevance for those interested in the future of social welfare programs. For several reasons, the problems inherent in the evaluation of any program are magnified when the policy is in this area. First, most programs are adopted without a clear statement of purpose. Indeed, ambiguity was necessary to build a coalition large enough to win policy adoption. Ambiguity is particularly a characteristic of social welfare policies, which involve a redistribution of income and are therefore highly controversial. (Ripley and Franklin, 1982: 158) Second, the art of program evaluation is not well developed. It is difficult enough to determine the effectiveness of a new weapon, much more problematic is determining whether the goals of a social welfare program are being met. How does one determine the extent to which a juvenile correction program has reduced recidivism, or whether a program for teenagers has raised self esteem? So many factors are influential in these areas it is extremely difficult to establish cause and effect relationships. (Dubnick and Bardes, 1983: 227)

This paper examines the program evaluation process through an analysis of the reviews conducted under sunset, a law mandating periodic review of agencies by the legislature. We will focus on one state, Tennessee, to determine the impact of comprehensive legislative review on social welfare programs.

SUNSET LEGISLATION

Sunset legislation, which is directed towards improving fiscal control and accountability through a periodic review of agencies by the legislature has received considerable attention since 1976, when Colorado became the first state to adopt a sunset law; thirty-four other states have also enacted sunset legislation. Sunset has three goals: to force accountability to the legislature, to curtail unnecessary bureaucratic activity and to formalize long neglected oversight activity. These goals have garnered support from a wide variety of sources, with different expectations. However, the distinguishing feature of sunset, and the factor which has received the most attention, is the termination of agencies on established dates unless recreated by law.

Sunset statutes grant no additional authority to legislatures for performing their oversight task. Many of the checks on bureaucratic growth can be accomplished by existing review mechanisms. However, the wide-spread adoption of sunset reflected a recognition among lawmakers that traditional oversight procedures were not working. In most states legislative review is largely discretionary. Since little personal pay-off is perceived for stringent oversight activities, the process, at best, is characterized as hit or miss supervision. Many hoped that sunset would provide the tools and create the incentive for periodic and comprehensive evaluation.

Despite the popularity of the sunset concept, not everyone was an enthusiastic supporter. Several reasons for this skepticism of this highly publicized "reform" have been cited. First, many questioned whether states would be willing to invest
the time and staff necessary to perform an adequate evaluation. Critics were con-
vinced that there would be very little review. (Behn, 1977) Some voiced concern
that the difficulties in establishing uniform evaluation criteria would create sub-
stantial problems for agencies which have difficulty documenting output. There were
fears that the threat of termination would lead to an increased bureaucracy as agen-
cies fought to prove themselves productive. Finally, there was apprehension, which
has proven well founded in at least some states (e.g., Texas), that sunset review
would serve as a "marvelous generator of legislative campaign funds" as interest
groups sought a favorable evaluation. (Azama, 1977) Those fears were particularly
prevalent with regard to social welfare issues, where improvements are very difficult
to document. Because of the problems in documenting output, there was concern that
more time would be spent justifying the agency, taking away from service to clients.
Social welfare programs usually lack the number of lobbyists that regulated agencies
depend on. In short, many felt that social welfare programs would be highly vulner-
able to substantial criticisms and some of the smaller programs would face termina-
tion under the sunset law.

Because this study involves only one state, as with any case study caution must
be exercised in generalizing from this study to the experiences that other states
have had with sunset review. Tennessee has a comprehensive rather than selective
sunset law. Under a comprehensive law, all agencies are evaluated while under selec-
tive review only certain agencies (primarily regulatory bodies and occupational
licensing boards) come under review. However, because it has more adequately co-
ordinated sunset activities (review is performed in stages rather than simulta-
neously), Tennessee has not been in the position in which so many entities were to be
reviewed there was little more than routine re-authorization.¹ As in most states,
Tennessee was confronted with a significant, but not overwhelming number of entities
to evaluate. Thus, its experience is compatible with that of most other states. A
discussion of the data used in this study follows a brief description of the law.

The Tennessee Experience

The Tennessee sunset statute, known as the 1977 Government Entity Review Law, is
a comprehensive law mandating review of all agencies, departments and compacts to
which the state is a party. The responsibility for conducting reviews of each entity
was assigned to the Division of State Audit under the office of the Comptroller of
the Treasury. The division was already involved in program reviews and key officials
were consulted in regard to provisions of the Tennessee law. Termination dates for
every agency of state government were established over a six-year cycle with over 220
agencies being cited. A sunset review staff in the Division of State Audit was es-

dtablished who reports to the joint committee on Government Operations. After a
public hearing, the committee forwards legislation to the entire general assembly.
This legislation can mandate the continued existence of an entity, its alteration or
its cancellation. Without any action, the entity will automatically terminate after a
year.

Because the amount of time and money that can be spent on sunset review is
limited, more resources are allocated to major agencies, or programs where the

¹This was the situation in Alabama, Arkansas, and Louisiana.
potential savings are greater, and the services provided are of greater significance. For example, less than 350 hours were charged to the Board of Cosmetology, while almost 4,000 hours were spent on the evaluation of the Department of Insurance. For each review, the Joint Government Operations Committee is presented with a written report of the evaluation. This is followed by a public hearing after which the committee makes a decision regarding the action that will be recommended to the legislature. The sunset staff is to be present throughout the deliberations to explain the options available with regard to the entity under review.

DATA AND METHODS

The sunset evaluation process was examined through an analysis of the written reports conducted between 1978-1980. For the content analysis, information was collected regarding the number and the type of evaluation methods used, the objectives of the evaluation, the number and type of criticisms made in the evaluation reports, and managements’ response to the criticisms. An analysis of legislative records and interviews with staff provided information regarding the action taken by the legislature and the agency staff in response to the report. The responses were categorized as follows: (a) no action, (b) administrative change, (c) legislative change, and (d) termination. An administrative response consisted of a significant change in the program’s administration, either a change in structure or in its operation. Any change mandated by the legislature constituted a legislative response. Forty-seven reports were analyzed, seventeen of these involved social welfare issues, thirty were non-social welfare programs. Social welfare policies were defined as those which transfer benefits or social resources to individuals or groups. The definition was not restricted to programs assisting the financially needy (seventeen percent); criminal justice (thirty-seven percent), mental health (seventeen percent), alcohol and drug abuse programs (twelve percent), and policies providing assistance for select groups in the state (e.g., the elderly—seventeen percent) were also included.2

To insure that any differences between the reviews of social welfare programs and non-social welfare programs were not due to factors such as size or revenue, a matching procedure was used. The two groups of programs were compared on the basis of expenditures, size (number of staff), and type of entity. Non-social welfare programs were selected that matched the characteristics of the social welfare programs. The percentage of each group falling into each category was as follows:

<table>
<thead>
<tr>
<th>SIZE</th>
<th>Non Social Welfare (n = 30)</th>
<th>Social Welfare (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of staff</td>
<td>Percentages*</td>
<td></td>
</tr>
<tr>
<td>over 100</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>51-100</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>under 50</td>
<td>43</td>
<td>47</td>
</tr>
</tbody>
</table>

*figures are rounded to nearest tenth

2A list of the programs evaluated is available from the authors upon request.
The fact that the two groups of programs are highly similar to each other on these characteristics increases our confidence that any difference found between the social welfare and non-social welfare programs can be attributed to the different kinds of services provided.

There were fifteen review methods used in the sunset evaluations. To simplify comparisons of the methods a varimax rotated factor analysis was used. Three factors emerged. One encompassed methods involving staff and other "advocates" of the entity under evaluation. Advocates were defined as those benefiting from the entity’s existence—staff, interest groups and those regulated. We expect staff to present as favorable a view as possible to evaluators. Among the most intense criticisms of...
government agencies, regulatory agencies in particular, is that they are unduly influenced by the industries they are supposed to regulate or monitor. Thus, we also expect interest groups and those regulated to be positive towards the entity. Of course, not all individuals falling into these categories will be highly supportive. It is likely that some will raise significant issues to the reviewers with a few being highly critical. Nevertheless, the literature on interest groups and bureaucracy indicate that the staff, the clientele, and the interest groups involved with a program generally serve as advocates for that program (McCurdy, 1977: 118-121). The second factor consisted of more "objective" evaluation methods—contact with non-advocates of the entity. The third factor was comprised of review methods involving the examination of records. The following items fell into each factor.

Evaluation Methods Using Advocates: (1) interviews with entity staff, (2) surveys of entity staff, (3) interviews with clientele, (4) surveys of clientele, (5) surveys of interest groups.

Evaluation Methods Using Non-Advocates: the legislative staff conducting the review (1) corresponded with other states, (2) corresponded with professional organizations, (3) observed administration and operation, (4) analyzed existing data.

Examination of Records: the examination of (1) minutes, records, documents, (2) material related to the entities' operation—publications, articles, reports, (3) minutes, files, employee qualifications.

All items had loadings of at least .40 on the factor on which it was listed. There was no significant overlap across factors as no item had a loading greater than .15 on any of the other factors.

The first question examined was whether the evaluation methods used in the review of social welfare agencies differed from the methods used with non-social welfare entities. This issue was addressed by comparing the variety and type of evaluation methods used with each. The first step was to identify the sunset reports in which there was a high use of one of the three types of evaluation methods—use of advocates, non-advocates, and record examination. A score was assigned to each report for the number of research components used within each of the three dimensions produced by the factor analysis. A report was designed as having an "extensive" number of evaluation methods used in its compilation if the score was at least one standard deviation above the mean of all reports on that research dimension. Approximately one-third of the total reports fell into the high category for each dimension (use of advocates—28 percent; use of non-advocates—34 percent; examination of records—38 percent).

FINDINGS

The Review Process

Table 1 shows the percentage of reports in which there was "extensive" use of each type of evaluation method, comparing social welfare programs with non-social welfare programs. There are marked differences between the two types of programs in
the kind of evaluation process that predominates, sixty-five percent of the social welfare programs involved advocates compared to fifty-seven percent of the non-welfare programs. Although a majority of social welfare evaluations also relied heavily on methods involving non-advocates of the entity (fifty-two percent), somewhat more extensive use of non-advocates was made in the review of non-social welfare programs (seventy-three percent). Non-social welfare program reviews also made more extensive use of records.

The types of review methods were broken down through time to determine if the evaluation techniques have changed since sunset reviews first began. Table 1 also presents the figures for the three years (1978 to 1980) in the percentage of sunset reports containing "extensive" use of each of the evaluation methods. As the number of cases is not large when viewing each year separately, particularly for the social-welfare programs, any difference appearing from year to year should not be overemphasized. Nevertheless, the data provide information regarding trends. An examination of Table 1 indicates that similar types of changes are occurring in the review of social welfare and non-social welfare programs. Both types of programs show greater use of all three types of evaluation methods. However, the use of advocates show a particularly sharp increase. It appears that reviews of social welfare programs will continue to rely more heavily on "subjective" data—interviews with "advocates"—than more objective data. However, evaluations of non-social welfare programs are also increasingly involving agency staff and others who primarily serve as advocates of the program.

The second question examined was the objectives of the report, that is, what factor(s) did the evaluation center on? As Table 2 shows, the non-social welfare evaluations focused primarily on the effectiveness and efficiency of the operation. In contrast, most reviews of social welfare agencies centered (1) on the type of personnel used, the extent to which they were qualified and whether affirmative action requirements were met, (2) public disclosure questions, and (3) whether there were alternative ways of providing the service. Over seventy-five percent of the sunset evaluations of non-social welfare agencies made extensive use of review factors related to efficiency and effectiveness while less than thirty percent of the welfare

### Table 1

**"EXTENSIVE USE" OF EVALUATION METHODS IN SUNSET REPORTS**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Advocates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 17)</td>
<td>65</td>
<td>57</td>
<td>(n = 6)</td>
</tr>
<tr>
<td>Non-advocates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 30)</td>
<td>52</td>
<td>73</td>
<td>(n = 7)</td>
</tr>
<tr>
<td>Examination of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>records</td>
<td>24</td>
<td>37</td>
<td>(n = 9)</td>
</tr>
</tbody>
</table>

-265-
TABLE 2
OBJECTIVES OF SUNSET EVALUATION CITED IN REPORT (figures are %s)

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Social Welfare</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average during 3-year period (1978 - 80)</td>
<td>(n = 17)</td>
<td>(n = 30)</td>
</tr>
<tr>
<td>Efficiency/effectiveness</td>
<td>29</td>
<td>76</td>
</tr>
<tr>
<td>Personnel qualifications</td>
<td>53</td>
<td>30</td>
</tr>
<tr>
<td>Public disclosure</td>
<td>58</td>
<td>23</td>
</tr>
<tr>
<td>Alternative ways of</td>
<td>53</td>
<td>40</td>
</tr>
<tr>
<td>providing the service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Welfare</td>
<td>(n=4) (n=6) (n=7)</td>
<td>(n=9) (n=12) (n=9)</td>
</tr>
<tr>
<td>Other</td>
<td>77 75 77</td>
<td>77 75 77</td>
</tr>
</tbody>
</table>

programs were reviewed using that criteria. By contrast, in less than one-third of the non-welfare programs was it reported that extensive use was made of review methods involving personnel or public disclosure issues. The general pattern is that social welfare programs were reviewed with a focus on the extent to which procedures were followed correctly while the output of the program and the efficiency with which the program operated was the focus of the reviews of non-social welfare agencies. The pattern has changed little through time. Thus, in terms of the review process, a different type of review was made of social welfare programs. Probably because of the difficulties discussed above in evaluating program success, the reviews centered on procedural questions.

Another interesting difference that appeared when comparing the sunset reviews of social welfare programs with other types of programs is that evaluations of welfare programs were more apt to involve advocates of the agency—those benefitting from the program and staff while the evaluations of non-welfare programs relied more heavily on information gained from more "objective" sources—observation, interviews with those not directly benefiting and the evaluation of records. The obvious question is whether the recommendations made from the evaluation staff and the subsequent action taken in response to the review differ for social welfare and other types of programs.

THE SUNSET REPORT:

Criticisms and Recommendations

Table 3 shows the frequency with which social welfare and non-social welfare programs received various types of criticisms from the sunset review team. Social welfare agencies were most likely to be criticized for their administration and for
the type of personnel employed. The types of criticisms leveled against non-social welfare programs relate to their purpose—whether the function they serve is still necessary, the extent to which the legislature’s mandate is met, and the entity’s financial management. This finding is what we would expect, given the fact that the evaluation of social welfare agencies centered on procedural questions rather than an analysis of output. With regard to the changes occurring through time, social welfare agencies received more criticisms in later years questioning whether various programs were necessary. Non-social welfare programs have received slightly more criticisms for not following the legislative mandate. Criticisms regarding administration and financial management increased for both social welfare and non-social welfare programs.

The average number of criticisms per report and the severity of the criticisms are shown in Table 4 along with the response, by the management of the entity, to the review. We distinguished between major and non-major criticisms. A major criticism was defined as any criticism involving a significant change in administration, staff, hiring procedures (the change recommended requires hiring different types of people or involving different actors in the decision making) or any challenge made to the program’s purpose or effectiveness. On the average, social welfare agencies received twice as many criticisms per report (eight) as other types of programs (four). However, while seventy-seven percent of the criticisms made of non-social welfare programs were major, this was true of only fifty-two percent of the criticisms made of social welfare programs. Finally, the management of social welfare programs was at least twice as apt to accept the criticisms made in the report (sixty-five percent

<table>
<thead>
<tr>
<th>Criticisms</th>
<th>Social Welfare (n = 17)</th>
<th>Other (n = 30)</th>
<th>Social Welfare (n = 4)</th>
<th>Other (n = 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative treatment of clients, distribution criteria</td>
<td>58</td>
<td>27</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>Program purpose—not following mandate</td>
<td>12</td>
<td>40</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Program not necessary</td>
<td>35</td>
<td>53</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td>Financial management</td>
<td>6</td>
<td>40</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>Personnel qualifications</td>
<td>52</td>
<td>33</td>
<td>50</td>
<td>33</td>
</tr>
</tbody>
</table>

TABLE 3
TYPE OF CRITICISM IN SUNSET REPORT
(figures are %s)

Average during 3-year period (1978 - 80) Change from 1978 to 1980 (1978 - 80)

Program not necessary 35 53 25 33 42 55 50 55
Financial management 6 40 0 0 13 33 42 44
Personnel qualifications 52 33 50 50 57 33 33 33

programs were major, this was true of only fifty-two percent of the criticisms made of social welfare programs. Finally, the management of social welfare programs was at least twice as apt to accept the criticisms made in the report (sixty-five percent
to twenty-seven percent). There were no great differences between the two types of programs in the changes occurring through time. The management of both social welfare and non-social welfare programs were increasingly more apt to accept the recommendations of the sunset report.

The final and most important issue addressed is the final action taken in response to the sunset reports. Is the fact that different evaluation methods and objectives were used in the evaluation of social welfare and non-social welfare programs reflected in the changes made in the entities?

The action taken as a result of the sunset evaluations is shown in Table 5. A comparison of the two types of programs shows considerable differences in the response made to the sunset evaluation of social welfare and non-social welfare programs. The evaluation of non-social welfare programs was more apt to invoke no response (twenty-three percent compared to twelve percent for social welfare programs). However, non-social welfare programs were more apt to be the focus of legislative action, and non-social welfare programs were more than twice as likely to be terminated (thirty percent of non-social welfare programs were terminated compared to twelve percent of welfare programs).

The changes made as a result of sunset evaluations of social welfare agencies were made by the entity’s management. The fact that fewer social welfare programs were terminated within this time period should not be construed as evidence that these programs are more "safe" than are other types of programs of comparable size. The higher termination rate of non-social welfare programs occurred because there were several programs falling into this category which no longer served any purpose.
and had little or no funding. We can offer a couple of explanations for the differences between the programs in the amount of changes made in administration. One, since administration is the focus of review for social welfare programs, the criticisms they receive are mostly in this area. It might also be the case that because social welfare administrators face more serious criticisms, they make the changes because of a fear of what the legislature will do. If this is what is happening, administrative action is taken with the purpose of pre-empting legislative involvement, it appears to be successful.

The changes through time vary little for social welfare and non-social welfare programs. Among non-social welfare programs there is a slight increase in the number of agencies making administrative changes, it is, of course, too early to determine whether the increase will be significant. The management of social welfare programs are increasingly making administrative changes and there has been somewhat of a drop in significant legislative action with regard to these programs. Given the small number of cases, however, the amount of change is not large and thus, its significance is questionable. Finally, the termination of social-welfare programs increased in the last two years of the years examined. This may indicate higher termination rates as sunset evaluations continue.

### Summary and Conclusions

The evaluation of any public policy is very complex. The problems are magnified when the programs under review are social welfare programs, where data is often unavailable, there is little agreement regarding the programs' goals and establishing cause and effect relationships is exceedingly difficult. For this reason, there has been concern that the recent upsurge in program evaluation would lead to significant reductions in social welfare programs since the policies' benefits could not be
easily documented. At the very least, many felt it would be necessary for staff to spend time trying to prove the significance of the policy taking time away from their service to clients.

We note that the significance of these findings is limited by the fact that the evaluations examined all occurred in one state and were all part of sunset review. This analysis does, however, offer some evidence regarding a question that has yet to receive much attention. Our findings show differences between the evaluation process of social welfare and non-social welfare programs. Although all evaluation methods were used for social welfare and non-welfare entities, staff and advocates of the agency were more likely to provide information for the evaluation of social welfare entities where more objective sources of information—observation, analysis of records—were the primary methods used for non-welfare programs. The goals of the evaluation also differed for the two types of programs. The extent to which administrative procedures were followed and the type of personnel employed constituted the focus of social welfare evaluations. Non-welfare programs, on the other hand, were primarily evaluated in terms of efficiency and effectiveness criteria. In other words, the focus of the review of non-social welfare programs was whether the program was accomplishing what it was supposed to and whether it was doing it as well as possible where the less challenging procedural questions dominated in the review of social welfare entities.

Social welfare programs faced more criticisms than non-welfare programs. However, the criticisms were less serious. There was also a difference in the response to the evaluation. Social welfare personnel were more apt to make changes prior to publication of the report of the sunset evaluation team, perhaps because they were fearful of action that the legislature would take if the report was highly critical. Additional evidence for the fact that social welfare personnel try to avoid legislative action comes from the fact that sixty-five percent of the sunset reports covering social welfare programs resulted in an administrative response by the entity's staff, only seventeen percent of the non-welfare programs experienced administrative change. Non-social welfare programs were almost twice as likely to face change mandated by the legislature.

Thus, the evaluation of social welfare programs has differed from the review that occurs of non-welfare programs. A number of changes have been made in social welfare programs as a result of the review. However, it does not seem to be the case the social welfare programs are faring badly in the review. Indeed, their administration appears better able to make changes in response to the evaluation that satisfy the legislature and thus, it would appear, probably exert slightly more control over the changes invoked as a result of sunset review.

REFERENCES


