

Child Welfare Privatization Initiatives—

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

Topical Paper #4

**Evaluating Privatized
Child Welfare Programs:
A Guide for Program
Managers**

August 2008



**U.S. Department of Health and Human Services
Office of the Assistant Secretary for Planning and Evaluation**

Acknowledgements

This project builds on the resources available at the Quality Improvement Center on the Privatization of Child Welfare Services (QIC PCW), funded by the Children's Bureau. We want to acknowledge all of the state and county child welfare administrators and private providers that shared their experiences with us and the QIC PCW. Additional information on child welfare privatization issues is available through the QIC PCW Website: <http://www.uky.edu/SocialWork/qicpcw/> .

This paper was prepared by Planning and Learning Technologies, Inc. in partnership with The Urban Institute for the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, under contract HHSP233200600242U. The opinions expressed in this paper are those of the authors and do not necessarily represent positions of the U.S. Department of Health and Human Services.

This issue paper was written by Dr. Jacqueline Smollar. Paper review and comments were provided by Elizabeth Lee and Karl Ensign of Planning and Learning Technologies, Inc.

This document is available online at:
<http://aspe.hhs.gov>

Table of Contents

| | |
|---|-----------|
| I. Introduction | 1 |
| A. <i>About this Paper.....</i> | 1 |
| B. <i>Why Privatization Initiatives Should be Evaluated</i> | 2 |
| C. <i>Using a Professional Evaluator</i> | 3 |
| D. <i>The Role of the Program Manager in Program Evaluation</i> | 3 |
| II. Preparing for a Program Evaluation: Developing the Conceptual Framework. 4 | 4 |
| A. <i>Developing a Logic Model.....</i> | 4 |
| B. <i>Establishing Evaluation Questions.....</i> | 8 |
| C. <i>Establishing Evaluation Measures</i> | 10 |
| D. <i>Determining the Evaluation Design.....</i> | 12 |
| III. Implementing the Evaluation | 16 |
| A. <i>Making Sure Key Stakeholders Are Adequately Informed about the Evaluation</i> | 16 |
| B. <i>Making Sure that Program Staff and Evaluation Team Members are Well Informed about Their Roles and Responsibilities in the Evaluation</i> | 17 |
| C. <i>Establishing the Target Population and Evaluation Timeframes.....</i> | 19 |
| IV. Using Evaluation Information Effectively | 21 |
| V. Cost Evaluations — A Special Consideration..... | 22 |
| VI. Conclusion | 23 |
| Appendix A Evaluation Resource Guides | |
| Appendix B Examples of Program Evaluations with Cost Components as well as Guides to Developing Cost Estimates | |
| Appendix C Federally Established Child Welfare Outcomes and Measures | |

I. Introduction

This is the fourth paper in a series of topical papers on the privatization of child welfare services. In 2006, the Office of the Assistant Secretary for Planning and Evaluation funded the **Child Welfare Privatization Initiatives Project** to provide information to State and local child welfare agency administrators who have implemented, or are considering implementing, a privatization initiative. The project includes six technical assistance papers covering key topic areas pertaining to child welfare privatization initiatives. This technical assistance paper addresses the topic of evaluating privatization initiatives. The following topics are covered in the other technical assistance papers. (These are available online as they are completed at <http://aspe.hhs.gov/hsp/07/CWPI/>.)

- **Assessing Site Readiness: Considerations about Transitioning to a Privatized Child Welfare System**
- **Program and Fiscal Design Elements of Child Welfare Privatization Initiatives**
- **Evolving Roles of Public and Private Agencies in Privatized Child Welfare Systems**
- **Developing Effective Contracts for Child Welfare Services**
- **Contract Monitoring and Accountability in Child Welfare Privatization Initiatives**

A. About this Paper

The term “privatization” is used in child welfare to describe a variety of situations in which a public agency contracts with a private agency to perform functions that were at one time performed by the public agency. For the *Child Welfare Privatization Initiatives Project*, privatization is defined as the contracting out of the case management function with the result that contractors make the day-to-day decisions regarding the child and family’s case. Typically, such decisions are subject to public agency and court review and approval, either at periodic intervals or at key points during the case. A privatization initiative of this type may be implemented statewide or at a county or regional level.

Although there are a variety of privatization approaches, for this paper, privatization initiatives are described in this paper as either “new initiatives” or “revised initiatives” according to the following criteria:

- *New initiatives*—Initiatives in which the public agency decides to transfer some or most case management and decision-making responsibilities for particular children and families to one or more private agencies.
- *Revised initiatives*—Initiatives in which the public agency already has contracted with private agencies to assume some or most of the case-management and/or decision-making responsibilities for particular child welfare cases, but plans to change a particular aspect of the contract to improve attainment of desired outcomes. One example of this is when the public agency decides to replace a case-rate payment system with a performance-based payment system.

The primary purpose of this paper is to provide child welfare agency administrators and program managers with guidance on evaluating privatization initiatives. The paper highlights the key features of program evaluation and describes the tasks that program managers can perform

to ensure a successful and effective evaluation. This paper also provides a brief discussion about the value of cost-effectiveness analysis and the kinds of information that cost analyses can generate. Although this paper is not a program evaluation manual, it may be used in conjunction with other resources providing more detailed information about program evaluation, such as *The Program Manager's Guide to Evaluation*, which is available from the Office of Planning, Research and Evaluation of the Administration for Children and Families, U.S. Department of Health and Human Services. **Appendix A** provides information about additional resources for program managers pertaining to program evaluation. **Appendix B** provides resources for conducting cost analyses as well as examples of program evaluations with cost components.

B. Why Privatization Initiatives Should be Evaluated

The most important benefit of an evaluation is that it provides objective evidence as to whether or not a program is effective in attaining desired outcomes. Program evaluation is different from outcome monitoring. Outcome monitoring identifies particular outcome measures or indicators and examines performance on those measures over time to assess whether performance is improving, declining, or remaining the same. Outcome monitoring does not link changes in performance to particular programmatic features or activities. Outcome monitoring provides information about whether performance on a particular outcome measure improved, but not why it improved. Both the Federal Child and Family Services Review (CFSR) and the Federal Annual Report to Congress on Child Welfare Outcomes are examples of outcome monitoring.

Program evaluation also differs from implementation monitoring (sometimes called “process” evaluation). Implementation monitoring examines whether an initiative is being (or was) implemented as intended. That is, were the required practices, policies, and procedures appropriately implemented, and if not, why? Implementation monitoring does not provide information about what happened as a result of the new initiative or program—that is, the outcomes for the individuals served. Studies of “fidelity” to a particular service model or program approach are examples of implementation monitoring.

Program evaluation is a systematic approach that combines aspects of both outcome and implementation monitoring to examine the linkages between the programmatic aspects of an initiative and the outcomes observed. As a result, a program evaluation can provide the following basic information:

- Whether the initiative was implemented as intended
- Whether it resulted in improvements in achieving particular outcomes
- The relationships between programmatic aspects of the initiative and the observed outcomes; that is, what did and did not “work”

Program evaluations use various research designs and methods to link program elements to results. Different designs yield different levels of confidence about causal relationships. As will be noted in later sections, a professional evaluator can help the program manager determine what evaluation design best suits the agency's needs.

Although all new programs or initiatives should be evaluated, it is particularly important to evaluate privatization initiatives. These initiatives represent considerable changes in child welfare system operations, and their initial implementation can be expensive. Also, because privatization can be controversial, evaluating the initiative may help allay concerns because it ensures that attention will be given to whether the initiative is benefiting children and families.

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

Some program managers may be hesitant to evaluate privatization initiatives because they believe that the complexity and scope of the initiative make it too difficult to evaluate. While many privatization initiatives involve complex models, the basic principles of evaluation are still relevant. The evaluation process itself is not necessarily affected by the complexity of the initiative; rather it is the interpretation of evaluation findings that is affected. This issue can be addressed to a large extent by ensuring that information is collected during the evaluation regarding the broader context in which both the program and the evaluation are implemented. This is discussed later in the section on data collection.

Because privatization initiatives reflect a whole new way of “doing business” for a child welfare system, it is important not only to evaluate them, but also, when possible, to pilot test them. A pilot test refers to the practice of implementing an initiative on a relatively small scale, evaluating the initiative, and then using information from the evaluation to guide decisions about whether to expand the implementation of the initiative or make changes in the initiative to improve outcomes. A pilot test is recommended even if the initiative being tested involves only a change in the existing privatization initiative, such as implementing a new payment structure in the contractual agreements.

*A **pilot test** refers to the practice of implementing an initiative on a relatively small scale, evaluating the initiative, and then using information from the evaluation to guide decisions about whether to expand the implementation of the initiative or make changes in the initiative to improve outcomes.*

Pilot testing an initiative also allows a child welfare agency to garner support from the public, the legislature, and public and private agency child welfare staff for a full-scale implementation at either a county or statewide level. If a methodologically rigorous evaluation of a pilot test indicates that the new initiative is more effective in achieving desired outcomes than the existing approach, stakeholders may be more willing to support a large-scale implementation.

C. Using a Professional Evaluator

Although there are a number of useful evaluation-related manuals and guides for program managers, the most effective evaluations will involve a professional program evaluator as part of the evaluation team. A professional evaluator can either lead the evaluation effort or provide consultation on such critical evaluation tasks as identifying evaluation measures, determining the evaluation design, guiding the data collection and analyses, and interpreting findings. Involving a professional evaluator is particularly important for evaluations of child welfare privatization initiatives. These initiatives involve multiple, interrelated systemic components in the way services are designed, delivered, monitored, and funded. Program evaluations of privatization initiatives should consider how the various components of an initiative may affect both the implementation of the initiative and the outcomes attained. A program evaluator will have the necessary skills to design an evaluation to address these issues and obtain the information needed to determine if the initiative is effective in attaining desired results.

D. The Role of the Program Manager in Program Evaluation

A central message of this paper is that program managers have several important roles in a program evaluation, even if a professional evaluator is used. An effective evaluation is a team effort, and program managers must be key players on the team. For example, a program manager could take a lead role in the following tasks:

- Preparing for the program evaluation by supporting the development of a clear

conceptual framework for the evaluation that describes the intervention, its objectives, and how the intervention is expected to achieve those objectives

- Ensuring that all relevant parties, both in and outside the child welfare agency, are well informed about the evaluation and about their roles and responsibilities during implementation of the evaluation; and
- Ensuring that data analyses and interpretations of findings are relevant for future program decision making.

This paper is designed to help program managers accomplish these tasks.

II. Preparing for a Program Evaluation: Developing the Conceptual Framework

The conceptual framework is the most critical piece of an evaluation. It ensures that the evaluation is connected to the initiative being implemented and that the information produced by the evaluation will be useful to the agency implementing the initiative. The conceptual framework includes the following components:

- A logic model that describes your theory of change. The logic model links the program initiative, the programmatic features of the initiative, and the desired outcomes;
- A set of evaluation questions that specify what you want to learn from the evaluation;
- Evaluation measures that transform the evaluation questions into measurable events; and
- An evaluation design that delineates the methodology to be used to answer key evaluation questions.

A. Developing a Logic Model

A logic model is the most basic structure for an evaluation. Although there are a variety of logic models, the most basic logic model specifies the following:

- *The theory of change underlying the initiative.* These are beliefs or underlying assumptions about what is necessary to bring about the desired change(s), that is, why or how X change is expected to lead to Y result.
- *The implementation objectives.* These are the programmatic features to be implemented in the initiative that are expected to bring about the desired change(s).
- *The outcomes objectives.* These are the desired changes that are expected to occur as a result of implementing the programmatic features.

A **theory of change** is a general statement about why and/or how X change is expected to lead to Y result.

An **implementation objective** refers to what is expected to be done in a program or initiative to achieve the desired outcome, for example, what actions are to be taken, what services are to be implemented, and what policies are to be changed or developed.

An **outcome objective** is a change expected to happen as a result of a particular "action." For a privatization initiative, an outcome objective may refer to an improvement in children's safety, permanency, or well-being as a result of implementing the privatization initiative.

The model is called a *logic* model because there must be logical connections among the parts of the model; that is, the outcome objectives must be logically related to both the theory of change and the implementation objectives. The basic goal of an evaluation is to determine whether the causal relationships expressed in a particular theory of

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

change are valid, that is, was the intervention successful in producing the desired outcome. Therefore it is critical that the logic model depicts the causal relationship between a particular implementation objective (for example, child specific recruitment) and a particular outcome objective (for example, more timely adoptions).

A logic model is not intended to describe every aspect or feature of a program initiative, but only to capture the general framework of the initiative to ensure that the evaluation corresponds to what is being implemented. Exhibits 1 and 2 on the following pages present a sample logic model for the two types of privatization initiatives described in the introductory section—a new initiative (Exhibit 1) and a revised initiative (Exhibit 2). Both examples pertain to privatizing child welfare functions for children in foster care (out-of-home care). Exhibit 1 focuses specifically on privatizing child welfare functions pertaining to adoption. These examples are intended for use by a program manager as references in completing the following tasks.

Step 1: Specify the “theory of change”

New or revised privatization initiatives are implemented because there is a belief among administrators, managers, and/or legislatures that the initiative will improve attainment of desired outcomes. A theory of change explains how an activity or approach will result in desired outcomes. It articulates the assumptions about the process and may be derived from a variety of sources, including personal experience or beliefs, research findings, or anecdotal information.¹

Articulating a theory of change can be difficult. If you are planning the evaluation at the same time you are designing the new contracts, it may be useful for program managers, agency administrators, and other members of the program and evaluation teams to complete the following statement. “I think we will better achieve the outcomes we want for children in out-of-home care if we . . .”. By completing this statement, the program manager and other stakeholders will be delineating their theories about what is needed to bring about desired change. The people involved in developing and implementing an initiative may complete this statement in different ways. One person may say something like the following: “I think we will better achieve the outcomes (or a specific outcome) we want for children if monetary incentives are attached to attaining the outcomes.” Another person may suggest: “I think we will better achieve the outcomes (or a specific outcome) we want for children if we involve families and the community in meeting the needs of the children.” Differences in statements can serve as areas for discussion among stakeholders when planning the initiative.

Different theories of change may be combined in one initiative as long as they do not contradict one another and are used to determine implementation objectives. An example of how this might work is shown in the exhibits. In Exhibits 1 and 2, the theory of change is provided in the first column of the logic model. In Exhibit 2, the theory of change is that improved outcomes will occur if monetary benefits are attached to attainment of those outcomes. This assumes that the monetary benefits will be sufficient to bring about change regardless of the services offered. In Exhibit 1, the theory of change is that improved outcomes will occur if there are monetary incentives **and** if services are provided that are consistent with what is known about best practices in the field. According to the theory of change for this initiative, monetary incentives by themselves will not be sufficient to bring about desired change.

¹ Theories of change derived from methodologically sound research findings are likely to be the most useful in generating desired outcomes.

As can be seen in the models, the broader theory of change for this initiative is consistent with a larger number of implementation objectives.

Exhibit 1
Logic model for a new privatization initiative that targets children with a case goal of adoption.

| Theory of Change | Implementation Objectives | Outcome Objectives |
|--|---|---|
| Desired outcomes will be achieved if there is a monetary incentive for attaining outcomes. | <ul style="list-style-type: none"> • Establish a payment structure that is based on attainment of outcome objectives. • Provide rewards for outcomes achieved beyond what is expected. | <ul style="list-style-type: none"> • Adoptions will increase. • Timeliness of adoptions will improve. |
| Desired outcomes will be achieved if services are implemented that have been found to be effective in achieving specific outcomes. | <ul style="list-style-type: none"> • Clinical social workers trained in the adoption field will provide at least 24 hours of adoption preparation services to children. • Case managers will have at least monthly face-to-face contact with each child in their caseload until the child is in an adoptive placement. • Case managers will have at least weekly contact with pre-adoptive parents for the first 3 months after placement. | <ul style="list-style-type: none"> • Disruptions of adoptive placements will decrease. |
| | <ul style="list-style-type: none"> • Relative searches will be conducted within 2 weeks after goal change to adoption, and then on an ongoing basis every 3 months, until the child is placed in a pre-adoptive placement. | <ul style="list-style-type: none"> • Relative adoptions will increase. |
| | <ul style="list-style-type: none"> • A private agency will collaborate with the courts with regard to establishing timely judicial reviews and decision making. | <ul style="list-style-type: none"> • Timeliness of adoptions will improve. |
| | <ul style="list-style-type: none"> • Specialized, child-specific recruitment will be conducted for children with special needs at least once a month. | <ul style="list-style-type: none"> • Timeliness of adoptions of children with special needs will improve. |

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

Exhibit 2 Logic model for a revised privatization initiative that has implemented a new payment structure for agency contracts.

| Theory of Change | Implementation Objectives | Outcome Objectives |
|---|--|--|
| Timeliness of permanency for children will be improved if there is a monetary incentive for attaining outcomes and the monetary incentive incorporates a reward system as well as a penalty system. | <ul style="list-style-type: none">• Establish a payment structure that is based on attainment of outcome objectives.• Provide rewards for outcomes achieved beyond what is expected for each of the outcomes addressed. | <ul style="list-style-type: none">• Timeliness to permanency will improve.• Timeliness of adoptions will improve.• Timeliness of reunifications will improve without increasing re-entries.• Re-entries into foster care will decrease. |

Step 2: Specify the implementation objectives

As defined previously, implementation objectives specify the programmatic features of an initiative. For privatization initiatives, the implementation objectives included in the logic model are likely to reflect some of the key specifications of the contractual agreement between the public and private agencies. These specifications may include the payment structure and may also include particular services or service approaches, reporting requirements, and monitoring. Not all contractual specifications need to be included in the implementation objectives of the logic model. The ones that must be included, however, are those that are essential to the theory of change, for example, the specific policies, practices, and activities that are expected to result in attainment of desired outcomes.

In the sample logic models, implementation objectives appear in the second column of the model. The number of implementation objectives included in the model will depend on the theory of change. For example, in the privatization initiative shown in Exhibit 2, there are very few implementation objectives; the theory of change for this model is that change will occur if there are monetary incentives paid to providers that directly reward or penalize providers for their performance on a designated outcome. In this type of privatization initiative, the focus is on purchasing results rather than purchasing particular types of services expected to produce specific results.

As shown in Exhibit 1, other privatization initiatives may be based on a theory of change that links particular types of services to desired outcomes in addition to the provision of monetary incentives. As a result, a number of implementation objectives are listed, with each reflecting a particular service that is believed to be, or has been shown to be, effective in achieving desired outcomes.

Step 3: Specify the outcome objectives

An outcome objective is the change expected to happen as a result of achieving implementation objectives. Usually, outcome objectives are stated in the form of an expected change in performance, such as an increase or decrease in performance with regard to a

particular outcome. The underlying hypothesis for a program evaluation is that if all implementation objectives are achieved, then desired outcomes will result. The validity of this hypothesis is what is being tested in an evaluation.

In identifying program objectives, it may be useful again to have key members of the program and evaluation teams complete the following statement: “We are implementing this initiative because we want children in out-of-home care (foster care) to . . .”. Completing this statement will ensure that the goals for your initiative pertain to achieving desired outcomes for children.²

In the sample logic models provided in Exhibits 1 and 2, outcome objectives appear in the third column. They may be associated with particular types of implementation objectives, as shown in Exhibit 1, or they may be connected to the theory of change and therefore, linked in a general way to a single implementation objective, as shown in Exhibit 2.³ With clear guidance from the Federal Children’s Bureau, most states and communities are focusing their efforts on bringing about improvements in outcomes related to child safety, permanency, or well-being. Information about the impact of program and fiscal components of a privatization initiative on these central outcome objectives is particularly important for decision-making purposes.

Some outcome objectives may not be achievable until other outcome objectives are attained. In these situations, the logic model might identify short- and long-term outcome objectives. For example, a common outcome is to reduce the time a child remains in foster care before reunifying with their families. However, to reduce time to reunification, improving parenting skills may be necessary to help ensure child safety. Therefore, a short-term outcome may focus on changes in a caretaker’s knowledge and skills in effective parenting.

B. Establishing Evaluation Questions

The logic model creates the structure of an evaluation and the logical relationships between the implementation objectives and the outcome objectives. The logic model also serves as a basis for generating evaluation questions. Evaluation questions specify the information desired from the evaluation. The most basic evaluation questions pertain to the overarching goals of the evaluation, which, in the case of privatization initiatives, are likely to be one or both of the following:

Evaluation questions specify the information that you want to collect in your evaluation and provide a basic structure for the analyses.

- Was the privatization initiative more effective than the public agency in achieving desired outcomes?
- Was privatization model X more effective than privatization model Y in achieving desired outcomes?

² Many communities that launch privatization initiatives also are interested in how the costs of care compare to the former system. It should be remembered that costs are best considered in relationship to whether or not the initiative achieved better outcomes for the children and families served. In other words, information gained from a cost analysis will not be meaningful unless the desired client-level outcomes are attained. However, a reduction in costs of attaining particular outcomes could be incorporated into a logic model as a desired outcome if it is linked to a particular implementation objective.

³ In establishing outcome objectives, it is important to be aware that meaningful data on some outcome objectives may not be available for at least 2 to 3 years after program implementation, while meaningful data for other measures may be available within 1 year of program implementation. This aspect of the evaluation will be addressed further in the section on establishing evaluation timeframes.

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

However, there are other evaluation questions to consider in a privatization effort that will enhance the understanding of the questions presented above. For example, program staff may want to know more about the implementation of the privatization initiative and will want to answer the following questions:

- Was the privatization initiative implemented as intended (that is, were all implementation objectives attained, including adherence to contractual requirements)? If not, what differences were observed in services, staffing structure, service intensity, etc.? When did changes occur?
- If the initiative was implemented in more than one agency, how did implementation vary across those agencies with regard to services, staffing, etc.?
- If there were differences in implementation across agencies, were there also differences in attainment of outcome objectives?

Additional questions may involve the relationship between outcomes and various aspects of the initiative and/or the relationship between outcomes and characteristics of the children. Sample questions in this arena might be the following:

- Was the initiative more effective for some children than for others? For example, did the initiative result in improved outcomes for children who entered foster care due to neglect but not for children who entered foster care due to sexual abuse? Did the initiative result in more improved outcomes for older children than for younger children, or for children of a particular race or ethnicity?
- Which aspects of the initiative appeared to contribute most to its effectiveness? For example, what was the contribution of the payment structure alone to the effectiveness of the implementation? Or, can outcomes be attributed to the lower caseloads of case managers in the private (or public) agency?⁴

It also may be important to explore questions that provide more in-depth information about the process of implementing both the initiative and the evaluation and explore the broader context in which both efforts were conducted. This information can be collected through interviews with various stakeholders involved in the initiative and can be used to provide a context for understanding and interpreting evaluation findings. The following are examples of these types of evaluation questions:

- What were the barriers to implementing the initiative and were these barriers perceived as affecting the attainment of outcome objectives? (Potential barriers might include: characteristics of the service provider community such as use of new providers to the child welfare field; variations in resistance to and/or support of privatization in the public agency; lack of sufficient competition for the contracts; and ineffective specifications in contractual agreements).
- What happened during the implementation of the privatization initiative that may have affected the implementation and evaluation of the initiative and attainment of outcome objectives? For example, during the implementation of the initiative, was there a considerable loss in staff in the public agency due to a migration of staff to the private agencies implementing the initiative? Or, during the implementation of the initiative, was

⁴ This “teasing out” of the effectiveness of various aspects of the initiative often can be done using statistical analyses that identify the strength of the relationships between or among variables.

there a change in administration of the public agency, with the new administrator having a less favorable view of privatization initiatives than the prior administrator?

- What factors facilitated implementation of the initiative? These might include the decision to pilot test the initiative before full implementation; support of agency staff, the courts, and the community for the initiative; the availability of a sufficient service array in the community; and a high level of collaborative planning among service providers, the public agency, and the courts.

C. Establishing Evaluation Measures

Both the logic model and the set of evaluation questions should be used to guide decisions about evaluation measures. Evaluation measures determine the type of data or information to be collected during the evaluation that is needed to answer the evaluation questions.

Evaluation measures determine the type of data or information to be collected during the evaluation that is necessary to answer the evaluation questions.

Generally, there are three types of evaluation measures: implementation, outcome, and contextual measures. Exhibit 3 provides examples of each. As shown in the exhibit, implementation measures, which are in the first column of the exhibit, *assess whether implementation objectives were attained*. They reflect the observable or countable aspects of the implementation process, such as the number of times a service was provided, the length of time a service was provided, and/or the number of children or families that received a particular service. When evaluating privatization initiatives, you will likely want to monitor several other key features of the revised contract; for instance, whether the provider received the expected types of payments as specified in the contract (that is, while a contract may discuss certain penalties and rewards, did the providers actually perform well enough or poorly enough to be impacted by performance contracts). Program staff may want to measure whether the public and private agency monitored services as discussed in the contract. Again, because privatization initiatives are inherently systemic reforms, it is important to measure all key program and administrative features of the initiative as part of your implementation study.

Outcome measures, which are presented in the second column of the exhibit, also are observable events that can be counted or quantified in some manner to *establish a level of performance*.

To evaluate a program effectively, measures must be established for all implementation objectives and outcome objectives stated in the logic model. Performance on these measures will be used to determine whether implementation or outcome objectives were achieved and to answer the basic evaluation questions.

Contextual measures are displayed in the third column of the exhibit and pertain to information needed to enhance interpretation of evaluation findings. For example, if a difference is found in performance on a particular outcome measure either over time or between the privatization initiative and a comparison program, it cannot be automatically assumed that the difference is “caused by” the programmatic features of the initiative.

*Program evaluations should consider whether there are other events, policies and practices that are also impacting the outcomes of interest. **Contextual measures** provide information about these factors.*

Contextual measures provide information about possible factors that *also* may contribute to differences in performance on outcome measures. Examples of contextual measures include the number of children represented in different age categories or ethnic/racial categories, the number of children who entered foster care because of neglect and the number that entered

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

because of physical or sexual abuse, the number of agency staff who held Master's degrees in social work, and the case load sizes of case managers. Some of these measures may be implementation objectives if they are specified in the contractual agreement. However, if they are not specified in the contractual agreement, it is important to include them as contextual measures so that their potential affect on outcome attainment can be assessed.

Exhibit 3 Examples of Evaluation Measures

| Measures of Implementation Objectives | Measures of Outcome Objectives | Measures of Context |
|--|---|--|
| <ul style="list-style-type: none"> • Percentage of all children served in the target period who received at least 24 hours of adoption preparation services. • Percentage of all children served who received adoption preparation services from a licensed clinical social worker with adoption expertise. • Percentage of all children served who had face-to-face contact with case managers at least once a month prior to placement in a pre-adoptive home. • Percentage of pre-adoptive families in which parents had weekly contact with the case manager during the first 3 months of the child's placement in the home. • Percentage of cases in which relative searches were conducted within 2 weeks of establishing adoption as the child's permanency goal. • Mean number of child-specific adoptive family recruitment activities conducted for children with disabilities. • Similarity of contract specifications regarding payment structures and actual payments. | <ul style="list-style-type: none"> • Median/mean length of time between establishment of goal of adoption and finalized adoption -- for all children and for children with diagnosed disabilities. • Length of time between establishment of goal of adoption and placement in a pre-adoptive home -- for all children and for children with diagnosed disabilities. • Percentage of adoptions occurring in less than 24 months of the child's entry into foster care -- for all children and for children with diagnosed disabilities. • Percentage of all adoptions in which the adoptive parent was related to the child. • Percentage of all adoptive placements that resulted in a disruption prior to finalization. • Percentage of all finalized adoptions in which the child re-entered foster care in less than 12 months of finalization. | <ul style="list-style-type: none"> • Children's characteristics (age, gender, race/ethnicity, time in foster care prior to goal of adoption, etc..) • Staff characteristics (age, gender, race/ethnicity, qualifications, etc.). • Agency characteristics (case load size, staff turnover, etc.). • Other state policies/programs implemented simultaneously to privatization effort. • Extent of collaborative planning in the design of the initiative. • Extent of information exchange about new initiative with key community stakeholders prior to implementation. |

Professional evaluators are very helpful in establishing evaluation measures. Although program managers may have some general concepts about what to measure, a professional evaluator can operationalize those concepts to ensure they can be quantified and their meaning is clear and consistent. A professional evaluator also can assist in identifying the contextual measures necessary to interpret findings and answer evaluation questions.

Although it is important that concerted efforts be made to review existing administrative data to determine its utility with regard to measures of implementation and outcome objectives, it is also important that measures are not selected simply because administrative data are available for that measure. The evaluation measures must reflect the implementation and outcome objectives stated in the logic model and must be relevant to the evaluation questions. Sometimes, a program manager may decide to use measures that were used in evaluations of other privatization initiatives. Or, a program manager may decide to use measures for which data already exist. Both of these decisions are useful **only** if the measures are consistent with the program's logic model and evaluation questions. It will be difficult for the evaluation to report anything meaningful about the initiative if there are no logical connections between outcome measures and outcome objectives or if the data collected are not relevant to the key evaluation questions.

Once measures have been established, the availability and quality of data for any particular measure can be explored and decisions made about new data that must be collected or about improving the quality of existing data. The availability of reliable data for outcome measures may be a concern for some states and localities. However, in recent years, many states have made considerable improvements in the quality of data reported to the Federal Adoption and Foster Care Analysis and Reporting System (AFCARS). These improvements are due in some part to the outcome measures developed for both the annual report to Congress on child welfare outcomes and the second round of the CFSR and the need to provide accurate data to AFCARS to ensure correct calculation of these measures by the Federal Government.

The specific outcome measures used in the Report to Congress and the second round of the CFSR are provided in **Appendix C**. If one or more of these measures is consistent with the outcome objectives specified in the logic model, then the data to calculate the measures should be available from the state's management information system. In addition, AFCARS data are available at the county level. Again, however, in deciding on the outcome measures, the starting point should be the outcome objectives in the logic model, not the availability of data for a particular measure.

D. Determining the Evaluation Design⁵

The final task in preparing the conceptual framework for an evaluation is determining the evaluation design that will be used to assess attainment of desired outcomes. Outcome evaluation is about making comparisons. Does program X work better than program Y, or does program X work better at time 2 than it did at time 1. For example, evaluations of privatization initiatives are likely to involve comparisons between private agency (or agencies) and public

⁵ Evaluation designs are relevant to consider only if the number of children expected to be affected by the initiative is substantial enough to permit meaningful statistical comparisons. For example, if a small county decides to privatize adoption services but there are only, on average 10 finalized adoptions each year, then it would not be useful for the county to conduct an outcome evaluation. A professional evaluator can advise program managers regarding the number of children in the evaluation sample that would be necessary for meaningful analyses of differences in performance.

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

agency performance, or between a private agency's performances at different times. An evaluation design describes the nature of the comparisons that will be made. Evaluation design is one of the most important features of an evaluation because it determines the extent to which basic evaluation questions can be answered with a relatively high level of certainty. As noted in the section on evaluation questions, these basic questions are the following:

- Was the privatization initiative more effective than the public agency in achieving desired outcomes?
- Was privatization model X more effective than privatization model Y in achieving desired outcomes?

The level of certainty refers to the extent to which the differences found with respect to achieving desired outcomes can be attributed to differences in the type of agency providing the services (that is, the public or private agency) and not to some other variable.

The overarching goal of an evaluation of a privatization initiative is to determine whether there is a causal relationship between the privatization initiative and the attainment of desired outcomes. To achieve this goal, the evaluation must be able to answer one or both of these questions as either "yes" or "no" without a lot of "noise." Noise occurs when you cannot answer these questions with a high level of certainty because of the presence of other factors that may affect the results. The amount of noise in an evaluation determines the potential for incorrect interpretation of findings. The potential for noise in an evaluation is a primary reason for collecting data pertaining to contextual measures.

The key causes of noise in an evaluation are the characteristics of the population served and the conditions under which the initiative is implemented. For example, in the privatization initiative shown in Exhibit 1, evaluation findings may indicate that the private agency was more effective than the public agency in finalizing adoptions in a timely manner. However, suppose the data also show that the median age of children served by the private agency was 3 years younger than the median age of children served by the public agency. This would be a concern for interpreting evaluation findings because research has shown that adoptions of younger children tend to occur more quickly than adoptions of older children. Consequently, in this situation, it would not be possible to say with any level of certainty that the observed difference in performance with regard to timeliness of adoptions was due to the privatization initiative and not to differences in the ages of the children served. In these kinds of situations, the level of certainty may be increased somewhat by conducting particular statistical analyses to observe timeliness of adoptions as a function of the age of children in the two conditions. A variety of statistical techniques can assist with this, including multivariate analysis of variance and hierarchical linear modeling analyses. However, parsing out the age groups through statistical analyses sometimes results in considerable

Outcome Evaluations (sometimes referred to as **Impact Evaluations**) are done to assess whether the program attained its desired outcomes. This is done through making comparisons: does program X work better than program Y? Or, does program X work better in time 2 than it did in time 1?

An evaluation of a privatization initiative requires making a comparison between the outcomes achieved by a private agency and those achieved by a public agency. The **evaluation design** determines the nature of that comparison.

The overarching goal of an evaluation of a privatization initiative is to determine whether the privatization initiative was more effective in achieving desired outcomes than the public agency or than another privatization initiative. To achieve this goal, the evaluation must be able to answer this question as either "yes" or "no" without a lot of "noise." **Noise** occurs when the question cannot be answered with a high level of certainty because of the presence of other factors that may affect the results.

reductions in sample size. In addition, these types of analyses cannot control for the potential effects of an agency providing adoption services to a generally younger population of children. As a result, the evaluation findings in this situation will never reach the level of certainty that they would have if there had been no differences in the ages of children served by the public and private agencies.

When deciding on an evaluation design, it is important to select one that will reduce the level of noise to the extent possible by controlling for factors that might affect performance. Noise is always going to be present in an evaluation that takes place in real-world conditions. There simply are too many potential intervening factors to control for all noise. However, different evaluation designs have different potentials for minimizing noise. The three main types of evaluation designs are discussed below.

1. Comparisons of performance “before” and “after” implementing the initiative

This type of design often is referred to as a “pre-post” evaluation design. In this design, information is obtained regarding performance on the outcome measures prior to implementing the initiative or providing a specific service. Once the initiative is implemented or a particular service is provided, performance on the measures is obtained at specific follow-up intervals. In some instances, information is collected at multiple follow-up intervals.

This design is usually most appropriate for comparing performance on particular measures for the same population over a short time span. For example, when examining the affects of a particular type of educational or training curriculum on caseworkers’ knowledge in a particular area, one might measure the knowledge of the caseworker prior to receiving the curriculum and then the performance of that same population immediately after receiving the curriculum.

In a “pre-post” evaluation design, information is obtained regarding performance on the outcome measures prior to implementing the initiative or providing a specific service. Once the initiative is implemented or a particular service is provided, performance on the measures is obtained at specific follow up intervals.

For privatization initiatives, however, this design can result in a fairly high level of noise (with regard to interpretations of causality) because a privatization initiative usually does not focus on comparing performance of the *same population* over time. Instead, comparisons will almost always be made between different populations of children served over time. Two of the biggest causes for noise would be 1) differences in the children served before and after implementing the initiative and/or 2) other events or circumstances in the community that will also impact the outcomes of interest.

An example of the first problem can be observed using the sample shown in Exhibit 1. It may be that the population of children with a case goal of adoption differed with regard to the children’s age and ethnicity from the time that baseline data were established to the time that post-implementation data were collected. Because children’s age and ethnicity have been found to be associated with length of time to adoption, differences in these factors from before and after implementation would reduce the ability to answer key evaluation questions with any level of certainty.

The second common reason for noise would be the result of other events taking place in the community that only impact the before or after implementation population. For instance, most privatization initiatives will require at least one year before meaningful data can be generated regarding outcome measures. During that time period, there could be several external factors that might impact evaluation findings. For example, using the program

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

described in Exhibit 1 as a model, there may have been an economic downturn during the interval between baseline and follow up data collection that might affect the interest of families in adoption. Or, during that timeframe, the Federal Government may have implemented a program designed to increase adoptions by providing monetary incentives to families. Since both of these factors could potentially affect both the number and the timeliness of adoptions, it would not be possible to say with any certainty that the outcomes observed were the result of the privatization initiative rather than these other simultaneous influences.

When a pre-post design is the only possible design to use, evaluators attempt to reduce the noise by establishing a baseline using several years of historical data and averaging across those years. While this may reduce some of the noise caused by variations in population, it does not address the noise caused by external conditions and is not sufficient to increase the level of certainty with which causal inferences can be made.

2. Comparisons of two or more populations served under different conditions

This design often is referred to as a “comparison group” design. A fairly common type of evaluation design for privatization initiatives compares performance of public and private agencies who are serving two groups of children during the same time period. Some states fund both public and private agencies in the same jurisdictions to provide case management services to families and children. In these cases, the performance of the public agency could be compared to the performance of the private agencies, as long as the caseloads served were comparable. On the other hand, if all child welfare services are privatized in a given county in the state, a comparison would be made between comparable counties that have similar key features, such as characteristics of the children served, size of the population served, and urbanicity.

This design has less potential for noise than the pre-post design because both agencies are serving clients at the same time. Therefore, if there is an economic downturn that occurs during the time of the implementation of the initiative, it should affect both types of agencies equally. It also has less potential for noise than the pre-post design because concerted efforts are made to ensure the similarity of the two groups being compared with regard to key factors such as characteristics of the clients and the community. However, in the second example (where comparisons are made between counties), there is still a substantial potential for noise due to differences between the geographic locations of the experimental and comparison groups. It is unlikely that there would be situations in which both the populations served and the characteristics of the locations are identical. For example, during the initiative, a new judge may be appointed in the comparison county who, unlike the prior judge, does not support termination of parental rights unless an adoptive home has already been identified for a child. This change in court practices may lengthen the time required to complete an adoption. Therefore, evaluation findings with regard to timeliness of adoptions may be a result of the change in court personnel rather than a result of the privatization initiative. If this design is used, the site selection process should focus on ensuring that the sites are as similar as possible on key contextual variables.

3. Comparisons of two or more populations in the same location that are randomly assigned to different conditions

This type of design often is called an “experimental” design because it is similar to the research designs used in medical field tests. For a privatization initiative, an experimental

design has many of the same components as the comparison group design. There are clearly distinguishable conditions (for example, a private agency and a public agency), and the implementation takes place over the same time period in both conditions. However, the experimental design differs from the comparison group design in one important way. In the experimental design, the population that receives the privatization initiative (the experimental group) and the population that does not receive the privatization initiative (the control group) are *randomly* assigned to one of these groups. It is the randomness of the assignment decision that results in the groups being highly similar with regard to characteristics such as age, gender, race/ethnicity, length of time in foster care, and reason for entry into foster care. This reduces the potential for noise due to differences in client characteristics.⁶

*In an **experimental design**, eligible members of the target population are randomly assigned to either the treatment (or experimental) group or to the control group.*

Because the potential for noise in the experimental design is lower than in the other designs, the experimental design permits statements of causality to be made with a higher level of certainty. Although the experimental design often is more complex to implement because of the random assignment component and may be more costly to implement, the increased level of certainty associated with answering key evaluation questions often is well worth the extra complexity and costs.

Sometimes, child welfare agency staff view random assignment as denying needed services and therefore, believe that the existence of a control group is unfair to children, or even detrimental to children who do not receive the new intervention. However, in a privatization initiative, no client is denied services. Children may receive different services under different conditions, but there would not be any client who would not receive services. Also, at the outset of an initiative, it is not known whether the new intervention will produce better results than the existing service set — that is what the evaluation is trying to determine. In addition, from a broader perspective, the kinds of information available from an evaluation using an experimental design will benefit all children in the child welfare system because it identifies what works and what does not work with regard to achieving outcomes related to children’s safety, permanency, and well-being.

III. Implementing the Evaluation

Once the conceptual pieces of an evaluation are completed, a program manager can attend to the more practical aspects of implementing an evaluation. The key roles of the program manager in implementing the evaluation are ensuring that all relevant parties are well-informed about the evaluation and about their evaluation-related roles and responsibilities, if relevant, and that they are carrying out their responsibilities appropriately.

A. Making Sure Key Stakeholders Are Adequately Informed about the Evaluation

The more informed key administrators, program staff, and other stakeholders are about an evaluation, the more likely they are to actively support its implementation. As noted previously, program managers should involve key stakeholders in developing the conceptual

⁶ It should also be noted that in most situations when a comparison group design is used, the treatment and comparison sites are in different geographic locations. On the other hand, a random assignment model will always take place in the same community. Therefore, using a random assignment model will also reduce the amount of noise that results from other events happening in the community studied.

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

pieces of the evaluation—that is, the logic model, evaluation measures, evaluation questions, and evaluation design. Once these are developed, they should be shared with other stakeholders, who should be given the opportunity to ask questions and express any concerns. The program manager should engage key stakeholders both from within and external to the agency in discussions about the conceptual components of the evaluation and how it will be implemented. This would include at a minimum the following stakeholders:

- *Private and public agency staff:* It is very important that the staff involved in providing services to children as part of the initiative understand why the privatization initiative is being implemented, the goals of the initiative, and the goals of the evaluation. The more positively staff view the evaluation, the more likely they are to meet their responsibilities with regard to the evaluation.
- *Judges and other court personnel:* As program managers in the child welfare field are well aware, the courts have a considerable role in determining the outcomes for children who come into contact with the child welfare system. Yet, many initiatives in child welfare are implemented without incorporating judges and other court personnel in the process of developing and evaluating the initiative. Involving court personnel in developing the logic model will ensure that court personnel fully understand why the initiative is being implemented and the goals and objectives of the initiative and the evaluation. In addition, if they are not brought into the evaluation planning they can inadvertently or accidentally undermine the evaluation; for instance, through violations of the random assignment process.
- *Key external agency service providers:* The success of any particular child welfare program also depends to a large extent on the availability of, and access to, services for both children and parents. Bringing the service community into discussions of the logic model and the evaluation implementation process will help them be aware of the evaluation and their role in it, if any.

B. Making Sure that Program Staff and Evaluation Team Members are Well Informed about Their Roles and Responsibilities in the Evaluation

Everyone involved in an evaluation (including an external evaluation professional) must be aware of his or her responsibilities and how those responsibilities fit into the overall scope of the evaluation. This can be facilitated by bringing all relevant parties together to discuss the evaluation and review evaluation roles and responsibilities. The responsibilities of various evaluation participants would include the following:

- *Data collection and entry:* Some individuals, usually the case managers, will be responsible for collecting and entering data regarding the characteristics of each child, the types of services provided, the duration of services, etc. Program managers should make sure that staff members responsible for data entry understand how the data they are collecting will be used in the evaluation and why it is critical to the evaluation that the data be collected and reported in a systematic and consistent manner.

Additional information about the context in which the initiative is operating and the barriers to and facilitators of implementation and evaluation are most likely to be collected through interviews with key stakeholders. These interviews should be done by members of the evaluation team, preferably individuals who are not employees of either the public or private agency. A professional evaluator who is experienced in qualitative data collection and

interview instruments is the best resource for developing the interview instruments and analyzing the data. This interview process should be monitored to ensure consistency of the interview method and accuracy of reporting.

Prior to implementing the evaluation, it will be important to make sure that data are being collected for all of the data elements necessary to calculate the implementation measures, outcome measures, and context measures.

- *Random assignment:* If an experimental design is implemented, one or more individuals will be responsible for randomly assigning the cases to the treatment or control groups. Using the example of establishing performance-based contracts for adoption services, under random assignment, staff will be responsible for assigning an eligible child (for example, one whose case goal changes to adoption) to either the experimental group (the private agency) or the control group (the public agency). To ensure randomization, each assignment is based on either a randomized number system or a lottery system that is established prior to the evaluation. A professional program evaluator can help develop one of these systems and the site staff responsible for ongoing assignment need only follow the system.

It can be difficult to maintain a truly randomized design. A number of evaluations have been initially established as experimental designs in which, over time, the individuals responsible for randomly assigning children to treatment or control conditions did not continue to place children into the correct group according to the randomization system. Randomization breaks down when those tasked begin to assign a child to the experimental condition if they felt that the child “really needed” that service. Consequently, it is very important for program managers to ensure that the individual or individuals responsible for this function understand how important it is that assignment to groups be randomized. In addition, individuals responsible for assigning cases should be informed about the benefits to all children of conducting a rigorous evaluation that produces meaningful information about the effects of various programmatic efforts.

- *Quality assurance:* Often, several members of an evaluation team will be responsible for quality assurance with regard to the random assignment process (if that procedure is being implemented), data collection and entry, and fidelity to the model being evaluated. The implementation of all features of the program and the evaluation must be carefully observed and documented. With regard to data collection, systematic procedures must be developed for the collection of all new data, and both the process and quality of data collection must be monitored on a regular basis.

Quality assurance refers to a process or set of procedures developed to ensure that particular aspects of a program, or, in this case, an evaluation, are being implemented as intended.

For the most part, data on the outcome objectives will be available from the state’s management information system and will not need to be collected through a new process. However, even in this situation, it is important that program managers ensure that case managers in both the public and private agencies are entering data into the state’s management information system (MIS) in a timely and accurate manner. Information about the accuracy and meaningfulness of data in the state’s MIS system should be available from state agency data staff and they should be included in both the development of the outcome measures and the quality assurance process.

Data pertaining to the implementation objectives are most likely to be collected by case managers in the private and public agencies. Some of these data may be available from the state’s current MIS, but it may be necessary to design new data collection instruments. If

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

this is done, the data collection instruments should be designed so that they are useful for both program and evaluation purposes, to avoid increasing the burden for case managers.

With regard to fidelity to the model, it would be important to ensure that implementation objectives are being met on an ongoing basis. For example, if one implementation objective is to provide at least 32 hours of adoption preparation services to all children, then part of the quality assurance process should be to monitor whether that is occurring. If problems in implementing the model as intended are caught early enough, then corrective action can be taken so that the attainment of outcomes can be more effectively linked to the program model.

- *Data analyses:* The program manager also must designate some members of the evaluation team to be responsible for determining the data analyses that will be needed to answer evaluation questions. The type of data analyses will depend on both the nature of the question, the type of evaluation design implemented, and the number of relevant variables.

There are multiple statistical tools that can help with data analyses. There are quantitative analyses, such as analyses of variance (ANOVA), t-tests, multiple regression analyses, and hierarchical linear modeling. These analyses can answer questions about the size of the observed differences in performance on the outcome measures between groups or over time and can examine the relative contributions of various factors to observed differences. For example, one evaluation question might be whether children served by the privatization initiative (the experimental group) spend less time in foster care prior to reunification than children served by a public agency? A t-test or ANOVA could be done to identify the size of the differences between the groups with regard to the time in foster care prior to reunification. Subsequent analyses can be done to identify the strength of the relationship between these variables. For example, what percentage of the observed variation in time in foster care can be attributed to the differences in the independent variable (whether the child was served by a public or private agency).

There also are qualitative data analyses that can be done, such as chi square tests of independence between variables and log linear analyses when multiple variables are considered. These tests usually are used for qualitative data, such as the educational background of the case managers, the race/ethnicity of case managers, and the demographic characteristics of children. For example, one evaluation question may be, is the timeliness of adoption related to the educational background of the case manager? A chi square analysis of independence could be done that would examine the number of adoptions occurring in 24 months in which the case managers had a Master's degree in social work, a Bachelor's degree in social work, or no social work education.

Decisions regarding data analyses should not be undertaken without the assistance of a professional evaluator. The determination of appropriate data analyses will depend upon the evaluation design, the logic model, and the evaluation questions.

C. Establishing the Target Population and Evaluation Timeframes

Every evaluation must have a clear target population—that is, the children who are eligible for inclusion in the evaluation-- and an evaluation start time—that is, the date that the evaluation will begin assessing for implementation and outcome objectives. In a pre/post evaluation design that compares the performance of public and private agencies, the target

population for the evaluation is the children who receive services from a private agency or agencies on the evaluation start date and thereafter. This target population is compared to the children served by the public agency prior to the evaluation start date. If the evaluation is using an experimental design, the evaluation start date would be the first day that a child is randomly assigned to either the treatment or control group and the target population would be all children who are served after the start date. Children will be randomly assigned to groups until there are sufficiently large enough samples of cases in each group for comparisons to be made.

If an evaluation is conducted on an initiative that has been implemented for some time prior to the evaluation, then the evaluation would focus only on those children who were served by the initiative after the evaluation start date. For example, suppose a state has been implementing a privatization initiative for over a year without evaluating it. The state child welfare agency is then required by the state legislature to evaluate the initiative. After this requirement, the state child welfare agency determines that it will begin the evaluation on March 1. The children included in the evaluation would be those served on or after March 1.

*The **target population for an evaluation** refers to the children who are eligible for inclusion in the evaluation. This will not necessarily include all children served by a program that is being evaluated. Instead, an evaluation may specify that the target population includes, for example, only children who entered foster care after a particular date, or who had a goal change to adoption after a particular date, or who were of a particular age.*

Because privatization initiatives often involve a dramatic shift in the scope of practice for the private agencies that receive the contracts, it may be best to delay the implementation (but not the preparation) of an evaluation until the programmatic features have been stabilized. It may take several months for a private agency to be fully operational with regard to the privatization model. Usually, there will be new staff to hire and train, or there will be a transition of staff from the public to private agencies. In addition, the private agency will often need to establish new relationships with community service providers, courts, and other community organizations that reflect the private agency's new role within the community. The private agency also may need to adjust to new reporting requirements and fiscal responsibilities and make other adaptations in its operations. Because all of these transitional activities can affect performance on outcomes, the evaluation of the effectiveness of the initiative would be more meaningful if it did not begin until after a sufficient amount of time has been allowed for the transition to be completed. Although children may be served during this transition period, they would not be included in the evaluation. If a decision is made to delay the evaluation until the private agency has had an opportunity to adjust to its new responsibilities and reach a somewhat stable level of operation, then a particular start date would be established and the children served after that start date would be included in the evaluation.

In addition to establishing a start date for the evaluation, program managers will need to establish time frames for assessing performance on the outcome measures. It is important that all stakeholders involved in an evaluation of a privatization initiative be aware that a fairly extensive period of time needs to pass before some of the outcomes can be assessed in a meaningful way. For example, if one of the outcome objectives pertains to timeliness of adoptions and the outcome measure is a longitudinal assessment of an entry cohort with regard to adoptions occurring in less than 24 months of the child's entry into foster care, it obviously will be at least 2 years before data on this measure can be collected, and probably 3 years before it can be meaningfully interpreted. However, if the outcome objective pertains to stability of placements of children in foster care, and the outcome measure is the percentage of children in foster care for 12 months who have only one placement setting, then information about this measure will be available 12 months from the date that the last eligible child is enrolled in the evaluation.

IV. Using Evaluation Information Effectively

A program evaluation usually generates a large amount of data. Most program evaluations involve, at a minimum, the following types of data:

- Data for each of the outcome measures. Typically, each of the outcome measures is calculated by using multiple data elements. For instance, if the outcome measure is “time to reunification,” the evaluation will need to collect the date of case opening, the date of case closing, and the reason for case closing for each eligible child to calculate this measure.
- Data on the implementation measures. Again, multiple data elements may be necessary. For example, if the implementation measure pertains to the number of adoption preparation hours provided to a child within one month of a goal change to adoption, then the evaluation will need to collect the following information: the dates that adoption preparation services are provided to a child, the length of the service, the date of goal change to adoption, and the date that is one month after the date of the goal change.
- Data on the ages, gender, race/ethnicity of children.
- Data on the educational background and experience of program staff.
- Data on caseload sizes and rates of staff turnover during the evaluation.
- Data on characteristics of the agencies.
- Interview data pertaining to the barriers to implementation and the factors that facilitated implementation.

The extensive amount of data generated for an evaluation can be overwhelming to a program manager. This is one reason why it is important to establish a set of evaluation questions prior to implementing the evaluation. These questions are useful in determining how the data will be used and prioritizing the kinds of data analyses to be conducted. This is one evaluation area where assistance from a professional evaluator is invaluable.

In general, a first analysis of the data collected would examine whether there were significant differences between the experimental and comparison/control groups in their performance on the outcome measures. Once this is completed, additional analyses could be conducted to understand why one group outperformed the other or to see if there were differences in outcomes among subgroups of the sample population. In cases where significant differences were obtained between the treatment and control groups, it is important to determine whether there were differences in the characteristics of the children served in both conditions. Even when random assignment is used to assign children to the different conditions, it is still necessary to validate the similarity of the groups with the data.

It also is important to determine if the implementation objectives of the privatization initiative were achieved—that is, whether the initiative was implemented as intended. If the program was implemented as planned that is, all implementation objectives were attained, then some connection may be made between the programmatic features and the outcomes. If the initiative was not implemented as intended, but desired outcomes were still attained, then it would be important to identify what aspects of the initiative, if any, may have produced the desired outcomes. If the evaluation does not find significant differences in performance between the experimental and control or comparison groups, then similar analyses would be useful to

learn more about why no differences were found. Remember, an evaluation is done to determine whether a new program outperforms another and understand why this did or did not happen.

Although statistical analyses of quantitative data can provide a broad array of information to assist in interpreting key findings, content analysis of contextual information can provide a broad understanding of why a particular initiative may or may not have worked as intended. An analysis of potential barriers and facilitating factors is particularly relevant to addressing this issue. For example, for the privatization initiative focusing on children with a case goal of adoption, it may be found that that caseload sizes and, consequently, service intensity differed between the public and private agencies. Continuing with this example, stakeholder interviews may indicate that when the public agency began to consider privatizing case management for adoption-related child welfare services, public agency adoption caseworkers and supervisors became concerned about their job security and began to migrate to the private agency. The public agency was unable to fill these positions because they could not guarantee the caseworkers and supervisors long-term employment. This resulted in public agency staff carrying larger caseloads and consequently seeing children less often than they would have under the normal operating conditions of the public agency. Consequently, the evaluation compared private agency performance with the performance of a public agency operating under unusually adverse circumstances.

Finally, one of the key decisions for program managers and administrators is whether evaluation findings support continued or expanded implementation of the initiative. This is not always an easy decision. For instance, in the example given above regarding loss of staff in the public agency, the program manager must decide whether, given the situation that occurred in the public agency, the evaluation findings represent a good test of the effectiveness of the privatization initiative, or whether the unusual conditions in the public agency resulted in comparisons that were not meaningful. These decisions often can be made in conjunction with advice from a professional evaluator regarding the strength of the overall evaluation findings in the context of various contextual variables.

V. Cost Evaluations — A Special Consideration

As noted in the introductory section, the primary focus of this paper is on how program managers can prepare for and implement program evaluations that assess attainment of implementation and outcome objectives. Often, however, program managers and administrators, legislatures and others want to know how the costs of privatizing a service or services compares to the costs of delivering services by the public agency. Answering this question is very complex and information on conducting a cost-related evaluation is beyond the scope of this paper.

Despite its complexity, a cost-related evaluation is an important component of any evaluation and is particularly important for privatization initiatives. This is because for many years, states and jurisdictions sought to privatize services with the expectation that in the long run, providing services through a private agency would cost less than providing the same services through a public agency. Today, many in the field seem to have a more realistic expectation that improving outcomes will likely cost at least as much, and potentially more, than current funding levels. For this reason, it is important to collect cost-related information and conduct a cost-related evaluation to understand *how costs are related to outcome achievement*.

The process of establishing a final cost, or attaching a dollar amount to providing services, is again highly complex, particularly in the child welfare field. The costs of providing

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

services in the private agency may be easier to establish because it may be possible to use the dollar amounts provided through the contract as the cost. However, in some privatization initiatives, there have been unanticipated costs for private agencies (or anticipated costs that were supported by other than public contracts), which in some situations were quite extensive. That said, it is generally more difficult to establish the full cost of care within public agencies than within private agencies because, for instance, many administrative functions that support the child welfare division are carried out by departments within the larger social services agency. Parsing out these costs and expenditures is very complex. In short, program managers will need assistance from an evaluator with extensive expertise in cost analysis to determine both individual and aggregate costs of delivering particular types of services, such as foster care services, case management services, counseling services, services to parents, and services to children. **Appendix B** lists a number of resources that provide additional information on this topic, which can be consulted if cost considerations are an important aspect of a privatization initiative in your community.

VI. Conclusion

The child welfare field has struggled to improve outcomes for many years, in part because it lacks a sound evidence base for its interventions. New ideas are tried and discarded and tried again without evidence about what works to achieve desired outcomes. Methodologically sound program evaluations are the only way to build the evidence base to achieve improved outcomes for children and families in the child welfare field. Without methodologically sound program evaluation, any statements about the relationships between programmatic features and outcomes are nothing more than hypotheses. Every agency and organization wants to deliver quality services. Program evaluation will help program managers and administrators understand whether or not programs and services achieve this goal. Information collected from the evaluation will help determine which activities to continue and build upon and which to change to improve the effectiveness of the program.

Understanding and incorporating the results of a program evaluation is one the most important steps to take in designing programs and improving outcomes. Public and private agencies must work to consistently use data to make decisions, improve programs and systems, and provide the highest quality services to the families and children in their care. While implementing an evaluation will add to the cost of an initiative, this cost is minimal when compared to the cost of continuing over a period of years to provide services or service approaches that do not achieve desired objectives with regard to children's safety, permanency, and well-being.

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

Appendix A Evaluation Resource Guides and Web links (if available)

Outcome Accountability: An Evaluation Toolkit. This site provides a range of resources to develop an evaluation plan and offers links to other resources. Website:

www.friendsnrc.org/outcome/toolkit/index.htm

Outcome-Based Evaluation: A Training Toolkit for Programs of Faith. This site targets its evaluation training to faith-based organizations and programs. PDF version:

www.fastennetwork.org/Uploads/2F3325EC-7630-425B-8EDF-847AAA69BE76.pdf

Planning and Evaluation Resource Center (PERC). This site provides tutorials on conducting self evaluations with a focus on youth programs. Website: www.evaluationtools.org/

The Program Evaluation Kit. Herman, Joan L., Ed., Sage Publications. Newbury Park. Includes nine books written to guide and assist practitioners in planning and managing evaluations.

1. Evaluator's Handbook
2. How to Focus an Evaluation
3. How to Design a Program Evaluation
4. How to Use Qualitative Methods in Evaluation
5. How to Assess Program Implementation
6. How to Measure Attitudes
7. How to Measure Performance and Use Tests
8. How to Analyze Data
9. How to Communicate Evaluation Findings

The Program Manager's Guide to Evaluation. This is a detailed evaluation guide for program managers in social services. PDF version:

http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval/reports/pmguide/pmguide_toc.html

United Way of America – Outcome Measurement Resource Network. This site provides outcome measurement tools and links to other sites with similar tools. Website:

<http://national.unitedway.org/outcomes/>

User-Friendly Handbook for Mixed Method Evaluations. This is a detailed handbook on the use of a range of research designs and data collection strategies and includes sample data collection guides. Web version: www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/start.htm

W.K. Kellogg Foundation – Evaluation Handbook. This handbook provides information for people with a range of evaluation experiences who seek to conduct evaluations without external support. PDF version www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf

Appendix B

Examples of Program Evaluations with Cost Components as well as Guides to Developing Cost Estimates with Web Links (if available)

Emspak, Frank, Roland Zullo, and Susan J. Rose. 1996. *Privatizing Foster Care Services in Milwaukee County: An Analysis and Comparison of Public and Private Delivery Systems*. Milwaukee, WI: The Institute for Wisconsin's Future.

www.wisconsinfuture.org/publications/otherpubs/FosterCare.pdf

Kee, James. 1999. *At What Price? Benefit-Cost Analysis and Cost-Effectiveness Analysis in Program Evaluation*. The Evaluation Exchange. Vol. V, No. 2/3.

www.gse.harvard.edu/hfrp/eval/issue14/pp1.html

Kornfeld, Bob, and Laura R. Peck. 2003. *The Arizona Works Pilot Program: A Three-Year Assessment – Executive Summary and Full Report*. Prepared for the Arizona Department of Economic Security. Cambridge, MA: Abt Associates.

McConnell, Sheena, Andrew Burwick, Irma Perez-Johnson, and Pamela Winston. 2003. *Privatization in Practice: Case Studies of Contracting for TANF Case Management*. Submitted to the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. Washington, DC: Mathematica Policy Research.

<http://aspe.hhs.gov/hsp/privatization-rpt03/>

Nightingale, Demetra Smith, and Nancy Pindus. 1997. *Privatization of Public Social Services: A Background Paper*. Prepared for the U.S. Department of Labor, Office of the Assistant Secretary for Policy. Washington, DC: The Urban Institute.

www.urban.org/publications/407023.html

Scarcella, Cynthia Andrews, Roseanna Bess, Erica Hecht Zielewski, and Rob Geen. 2006. *The Cost of Protecting Vulnerable Children V*. Washington, DC: The Urban Institute.

www.urban.org/publications/311314.html

U.S. General Accounting Office. 1996. *Child Support Enforcement: Early Results on Comparability of Privatized and Public Offices*. Publication no. GAO/HEHS-97-4. Washington, DC. www.gao.gov/cgi-bin/getrpt?HEHS-97-4

_____. 1997. *Privatization: Lessons Learned by State and Local Governments*. Publication no. GAO/GGD-97-48. Washington, DC.

www.gao.gov/cgi-bin/getrpt?GGD-97-48

Vargo, Amy C., Mary Armstrong, Neil Jordan, Mary Ann Kershaw, Jennifer Peraza, and Svetlana Tampolskaya. 2006. *Evaluation of the Department of Children and Families Community-Based Care Initiative Fiscal Year 2004-2005*. University of South Florida.

Westat, Inc., Chapin Hall Center for Children, and James Bell Associates. *Estimating Child Welfare Service Costs: Methods Developed for the Evaluation of Family Preservation and Reunification Programs*. 2002. Washington, DC: USDHHS, Assistant Secretary for Planning and Evaluation. <http://aspe.hhs.gov/hsp/fampres02/index.htm>

Appendix C Federally Established Child Welfare Outcomes and Measures

1. Outcome Measures for the Child Welfare Outcomes Report

- **Child Welfare Outcome 1: Reduce recurrence of child abuse and/or neglect**

Measure 1.1: Of all children who were victims of substantiated or indicated child abuse and/or neglect during the first 6 months of the reporting period, what percentage had another substantiated or indicated report within a 6-month period?

- **Child Welfare Outcome 2: Reduce the incidence of child abuse and/or neglect in foster care**

Measure 2.1: Of all children who were in foster care during the fiscal year, what percentage was the subject of substantiated or indicated maltreatment by a foster parent or facility staff?

- **Child Welfare Outcome 3: Increase permanency for children in foster care**

Measure 3.1: For all children who exited foster care in the fiscal year, what percentage left either to reunification, adoption, or legal guardianship?

Measure 3.2: For children who exited foster care in the fiscal year and were identified as having a diagnosed disability, what percentage left either to reunification, adoption, or legal guardianship?

Measure 3.3: For children who exited foster care in the fiscal year and were older than age 12 at the time of their most recent entry into care, what percentage left either to reunification, adoption, or legal guardianship?

Measure 3.4: Of all children exiting foster care in the fiscal year to emancipation, what percentage was age 12 or younger at the time of entry into care?

Measure 3.5: For all children who exited foster care in the fiscal year, what percentage by racial/ethnic category left either to reunification, adoption, or legal guardianship?

- **Child Welfare Outcome 4: Reduce time in foster care to reunification without increasing reentry**

Measure 4.1: Of all children who were reunified with their parents or caretakers at the time of discharge from foster care in the fiscal year, what percentage was reunified in the following time periods?

- (a) Less than 12 months from the time of latest removal from home
- (b) At least 12 months, but less than 24 months
- (c) At least 24 months, but less than 36 months
- (d) At least 36 months, but less than 48 months
- (e) 48 or more months

Measure 4.2: Of all children who entered foster care during the fiscal year, what percentage re-entered care:

- (a) Within 12 months of a prior foster care episode?
- (b) More than 12 months after a prior foster care episode?

- **Child Welfare Outcome 5: Reduce time in foster care to adoption**

Measure 5.1: Of all children who exited foster care in the fiscal year to a finalized adoption, what percentage exited care in the following time periods?

- (a) Less than 12 months from the time of latest removal from home
- (b) At least 12 months, but less than 24 months
- (c) At least 24 months, but less than 36 months
- (d) At least 36 months, but less than 48 months
- (e) 48 or more months

- **Child Welfare Outcome 6: Increase placement stability**

Measure 6.1: Of all children served in the fiscal year who had been in foster care for the time periods listed below, what percentage had no more than two placement settings during that time period?

- (a) Less than 12 months from the time of latest removal from home
- (b) At least 12 months, but less than 24 months
- (c) At least 24 months, but less than 36 months
- (d) At least 36 months, but less than 48 months
- (e) 48 or more months

- **Child Welfare Outcome 7: Reduce placements of young children in group homes or institutions**

Measure 7.1: For all children who entered foster care during the fiscal year and were age 12 or younger at the time of their most recent placement, what percentage was placed in a group home or an institution?

II. Outcome measures developed for the second round of the CFSR beginning in FY07

Permanency composite 1: Timeliness and permanency of reunifications

- **Individual Measure C1.1:** Of all children who were discharged from foster care to reunification in the fiscal year, and who had been in foster care for 8 days or longer, what percent were reunified in less than 12 months from the date of the latest removal from home? (This includes the trial home visit adjustment, if relevant.)
- **Individual Measure C1.2:** Of all children who were discharged from foster care to reunification in the fiscal year, and who had been in foster care for 8 days or longer, what was the median length of stay in months from the date of the latest removal from home until the date of discharge to reunification? (This includes the trial home visit adjustment, if relevant.)

Child Welfare Privatization Initiatives —

Assessing Their Implications for the Child Welfare Field and for Federal Child Welfare Programs

- **Individual Measure C1.3:** Of all children who entered foster care for the first time in the 6-month period just prior to the target year, and who remained in foster care for 8 days or longer, what percent were discharged from foster care to reunification in less than 12 months from the date of latest removal from home? (This includes the trial home visit adjustment.)
- **Individual Measure C1.4:** Of all children who were discharged from foster care to reunification in the 12-month period prior to the target year, what percent re-entered foster care in less than 12 months from the date of discharge?

Permanency composite 2: Timeliness of adoptions

- **Individual Measure C2.1:** Of all children who were discharged from foster care to a finalized adoption during the fiscal year, what percent were discharged in less than 24 months from the date of the latest removal from home?
- **Individual Measure C2.2:** Of all children who were discharged from foster care to a finalized adoption during the target year, what was the median length of stay in foster care in months from the date of latest removal from home to the date of discharge to adoption?
- **Individual Measure C2.3:** Of all children in foster care on the first day of the 12-month target period who were in foster care for 17 continuous months or longer, what percent were discharged from foster care to a finalized adoption by the last day of the 12-month target period?
- **Individual Measure C2.4:** Of all children in foster care on the first day of the 12-month target period who were in foster care for 17 continuous months or longer, and who were not legally free for adoption prior to that day, what percent became legally free for adoption during the first 6 months of the 12-month target period? (A child is considered to be legally free for adoption if there is a parental rights termination date reported to AFCARS for both mother and father.)
- **Individual Measure C2.5:** Of all children who became legally free for adoption during the 12 months prior to the target year, what percent were discharged from foster care to a finalized adoption in less than 12 months from the date of becoming legally free?

Permanency composite 3: Achieving permanency for children in foster care for long periods of time

- **Individual Measure C3.1:** Of all children who were in foster care for 24 months or longer on the first day of the target year, what percent were discharged to a permanent home by the last day of the year and prior to their 18th birthday?
- **Individual Measure C3.2:** Of all children who were discharged from foster care during the target year, and who were legally free for adoption (i.e., there is a parental rights termination date for both parents) at the time of discharge, what percent were discharged to a permanent home prior to their 18th birthday?

- **Individual Measure C3.3:** Of all children who either (1) prior to age 18, were discharged from foster care during the 12-month target period with a discharge reason of emancipation, or (2) reached their 18th birthday while in foster care but had not yet been discharged from foster care, what percent were in foster care for 3 years or longer?

Permanency composite 4: Placement stability

- **Individual Measure C4.1:** Of all children who were served in foster care during the fiscal year, and who were in foster care for at least 8 days but less than 12 months, what percent had two or fewer placement settings?
- **Individual Measure C4.2:** Of all children who were served in foster care during the fiscal year, and who were in foster care for at least 12 months but less than 24 months, what percent had two or fewer placement settings?
- **Individual Measure C4.3:** Of all children who were served in foster care during the fiscal year, and who were in foster care for at least 24 months, what percent had two or fewer placement settings?