Attachment A

Comments in Response to the Federal Register Notice

**ADD RESPONSES TO PUBLIC COMMENTS ON DDPIE (5-8-2009)
PUBLISHED AT 43 FED. REG. 10,051**

INTRODUCTION

The Administration on Developmental Disabilities (ADD) requested comments on the proposed Information Collection Activity for the Developmental Disabilities Program Independent Evaluation Project ("DDPIE"). This request for comments published at 43 Fed. Reg. 10,051 on March 9, 2009 specifically asked for comments on evaluation instruments (e.g., interview protocols and self-administered questionnaire), that would be used in an independent evaluation of the three national programs funded by ADD: The State Developmental Disabilities Councils (DDCs), State Protection and Advocacy Systems (P&As); and University Centers for Excellence in Developmental Disabilities (UCEDDs). Comments were requested on the following:

A.   Whether the proposed collection of information is necessary for the proper performance of the functions of the agency, including whether the information shall have practical utility;

B.   The accuracy of the agency's estimate of the burden of the proposed collection of information;

C.   The quality, utility, and clarity of the information to be collected; and

D.   Ways to minimize the burden of the collection of information on respondents, including through the use of automated collection techniques or other forms of information technology.

ADD received 23 separate comments from this first federal register notice. Comments came from the following sets of organizations and individuals:

▪   National membership organizations for the DDCs, P&As, and UCEDDs respectively [3]

- Grantees (e.g., DDCs, P&As, UCEDDs) [13]
- Members of the public [7]

There was little variability in the comments received by ADD and, in many cases, the comments were a duplication of those submitted by other entities.

Overall, the commenters valued the importance of accountability and using data to document program outcomes. However, it is fairly evident from the comments received that ADD's grantees continue to have opposition to the approach for the independent evaluation as well as a significant misunderstanding of the design for the study. Of greatest concern is the apparent misperception by ADD grantees of the scope of the study and related cost. The specific purpose of the study is to conduct a national program evaluation that examines grantee impact using rigorous scientific methods. It is not the purpose to examine or re-tool the current reporting system used by grantees. The two activities are clearly distinct from one another. The independent evaluation will use the scientific process to study program impact at the national level, whereas the reporting system focuses on monitoring, thereby allowing for individualization by the programs. Moreover, ADD has made a relatively small investment in this study. The total cost of Phase 1 of the project, which was designed to develop the evaluation tools, was .004% of the total funding for the three programs being evaluated. ADD estimates that it would significantly increase the cost of the study to simultaneously overhaul the current reporting system and test it for reliability and validity for use in an independent program evaluation. What follows is a more detailed summary of the comments received from the first federal register notice and ADD's response to the comments.

COMMENTS AND RESPONSES TO COMMENTS

A.   Whether the proposed collection of information is necessary for the proper performance of the functions of the agency, including whether the information shall have practical utility

There was near unanimous consensus among the commenters that an evaluation is an important and valuable endeavor for measuring the outcomes of the programs. Therefore, it can be concluded that the commenters in general indicated that the information collection is necessary. However, there was unanimous disagreement on the approach taken by ADD in conducting this particular independent evaluation. With regard to practical utility, the commenters in general expressed concern about the design of the evaluation and the validity of the proposed information collection instruments. The most frequently cited concerns regarding the design of the independent evaluation were the following:

- The design should be based on participatory research methods.
- The scope of the study should incorporate ADD and the training and technical assistance contractors.
- For the P&A program, the design should be based on the evaluation for the Protection and Advocacy Program for Individuals with Mental Illness.

More of the comments focused on the proposed instruments and related measurement matrices (key functions, benchmarks, and indicators). The commenters in general felt that the proposed instruments and related measurement matrices:

- Do not accurately capture the programs' operations and functions and impose a "one size fits all" to programs that are required to contextualize their work within the State or territory.
- Impose requirements for the programs not specified in the authorizing legislation.
- Are redundant with other data collection required by ADD, and ADD should use data currently collected by grantees for the independent evaluation.
- Impose an additional burden for information on programs already burdened by current reporting requirements.
- There are concerns that findings from the independent evaluation will be used to report on or penalize individual programs.

Below are responses to the comments summarized above.

A-1. ***The design should be based on participatory research methods.***

Response:

ADD values participatory research methods and promotes the appropriate use of this investigative approach with its grantees. To the extent possible, ADD has made significant effort to infuse participatory methods into the evaluation design. However, there are several factors that preclude this study from being carried out in a fully participatory manner. These factors are discussed in more detail below.

In addition to its grantees, ADD has several other stakeholders that are invested in the research design for this study. Full participation by all stakeholders would have been very complicated and beyond the scope of the study. Therefore, ADD and its contractor, Westat, had to carefully balance participation by various stakeholders in the research process and the extent of participation by the various stakeholders.

ADD recognizes that the ***initial overall research design*** for the independent evaluation – a performance-based research design – was not developed in a participatory manner. In order to understand how the study design was developed, it is important to review the history of the project and part of the impetus for conducting the evaluation.

The research design for the independent evaluation was developed within the parameters of the Program Assessment Rating Tool (PART) guidance for conducting independent evaluations. This guidance did not support participatory research design methods. Instead it called for experimental or quasi-experimental research designs for independent evaluations. As outlined in the RFP for the project, ADD discussed that an experimental research design was not feasible for the DD Network programs and identified performance-based research design as a more feasible yet still rigorous scientific method that would meet the PART criteria. The information that

follows is excerpted from the RFP outlining the rationale for the research design.

Background on Independent Evaluations

The Office of Management and Budget (OMB) provides guidance to Federal agencies for conducting independent evaluations. The guidance from OMB can be found in the Program Assessment Rating Tool (PART) (see http://www.whitehouse.gov/omb/part/ for additional information). In 2003, ADD conducted a PART self-assessment under OMB guidance.

PART is a systematic method of assessing the performance of program activities across the Federal Government. The PART is a diagnostic tool; the main objective of the PART review is to improve program performance. The PART assessments help link performance to budget decisions and provide a basis for making recommendations to improve results.

The PART is composed of a series of questions designed to provide a consistent approach to rating programs across the Federal Government, relying on objective data to assess programs across a range of issues related to performance. The PART holds programs to high standards. Simple adequacy or compliance with the letter of the law is not enough. Rather, a program must show it is achieving its purpose and that it is well managed.

One aspect of the PART focuses on questions related to independent evaluations. The PART review seeks documentation that independent evaluations of sufficient scope and quality are conducted on a regular basis or as needed to support program improvements and evaluate effectiveness and relevance to the problem, interest, or need. The purpose of this question is to ensure that the program conducts non-biased evaluations on a regular or as-needed basis to fill gaps in performance information. These evaluations should be of sufficient scope to improve planning with respect to the effectiveness of the program.

The PART guidelines state that the most significant aspect of program effectiveness is impact—the outcome of the program, which otherwise would not have occurred without the program intervention. PART guidelines recommend that, where feasible, programs should measure program impact using experimental designs, which use random samples representing control and treatment groups. However, the PART guidelines recognize that these studies are not suitable or feasible for every Federal program. A variety of evaluation methods,

such as quasi-experimental studies, may need to be considered because Federal programs vary so dramatically. Quasi-experimental studies use techniques such as pre- and post-tests to assess the impact of the intervention.

The ADD Independent Evaluation

The structure and intent of the three ADD programs create significant challenges for carrying out independent evaluations based on experimental or quasi-experimental research designs. The structure of the ADD programs does not lend itself to conducting randomized trials or pre- and post-tests. ADD programs exist in every State. Many interventions of the DD Network are meant to impact all individuals with developmental disabilities residing in the State rather than one sector of the developmental disabilities population. Moreover, many DD Network interventions are not experimental. For example, P&As bring lawsuits in response to conditions within the State. Such interventions may be reactive rather than pro-active and their effect is felt over time, rather than immediately.

Other conditions that affect the design of the independent evaluation include the variability across the programs funded under the DD Act. Given that the DD Network responds to needs of individuals with developmental disabilities and their families in a State, the programs are characterized by great variability. For example, DD Network programs can choose to address one or more areas of emphasis (e.g., quality assurance, education and early intervention, child care, health, employment, housing, transportation, recreation, and other services available or offered to individuals in a community, including formal and informal community supports that affect their quality of life). Some may choose to address one area of emphasis (e.g., child care) while other may address 5 areas (e.g., quality assurance, education and early intervention, health, employment, and housing). This flexibility in the law creates great variability in terms of program implementation and impact on individuals with developmental disabilities and their families. This level of variability across the DD Network programs poses measurement challenges and increases the likelihood of sampling errors.

Finally, the developmental disabilities population is unique and small making it difficult to gather representative, random samples. The nature of the population poses challenges to gathering a national, representative sample that contributes to the validity and reliability of the research design for measuring program impact. Potential threats to internal validity include small sample size, mortality, regression, diffusion, and selection bias. Threats to external validity include selection and program interaction and setting and program interaction.

Design of the ADD Independent Evaluation
The limitations affecting the design of an independent evaluation project require that ADD consider non-experimental methods to determine program impact. The non-experimental design for any independent evaluation conducted to assess DD Network program impact must be rigorous and use techniques that ensure the validity and reliability of the research. ADD believes that for a non-experimental independent evaluation to be rigorous, it must follow performance-based approaches to measuring program impact. Performance-based approaches offers ADD a vehicle for reporting program accomplishments with particular attention on progress towards pre-established goals – the goals being those outlined in the purpose and principles of the DD Act.

As a performance based independent evaluation, a primary objective of Phase I of this project will be the development of measurement matrices. The measurement matrices will be dynamic tools with criterion-referenced performance standards to assess the extent to which a DD Network entity exhibits the purpose and principles of the DD Act and affects outcomes for people with developmental disabilities. Using numerical indices, the measurement matrices will be able to gauge the impact of the individual DD Program and the collaborative work of the Network. To this end, the measurement matrices will function as scoring rubrics for assessing program impact.

Despite the fact that the initial design for the study as outlined in the original RFP was not developed in a participatory manner, ADD included participatory aspects into the project design. In addition, ADD and its contractor for the study, Westat, made adjustments to the original research plan within the first months after the start date to make the study more participatory. The participatory aspects of the project are outlined below:

Working groups
One significant adjustment to the original research plan was the decision to use working groups to develop the measurement matrices (e.g., key functions, benchmarks, and indicators). In the original research plan, Westat had intended to visit a small sample of programs to collect information for developing the measurement matrices. However, after receiving feedback from the DDCs, P&As, and UCEDDs, the plan was modified to convene working groups representing the grant programs that would provide for a

more participatory process in developing the measurement matrices. Four working groups were formed: (1) DDC, (2): P&A; (3) UCEDD, and (4) DD Network collaboration. Through the working groups, input was sought through a series of in-person and teleconference meetings. Because of this adjustment to the original research plan, the project to develop the measurement matrices and related evaluation tools took an additional year to complete.

Advisory Committee
ADD infused participation into the research design through the Advisory Committee. The Advisory Committee included representatives from the grant programs as well as other key stakeholders (e.g., individuals with developmental disabilities, family members, national organizations). The Advisory Committee met throughout the three years of the project both in-person and through conference calls. The Advisory Committee was critical to providing feedback to Westat on the study and the measurement matrices.

Feedback from the Grantees
Another participatory aspect of the project has been solicitation of feedback from the grantees on the measurement matrices. As described later in this section, ADD and Westat conducted presentations at the national technical assistance meetings for the grantees on several different occasions. In doing so, grantees were provided multiple opportunities to learn about the evaluation and to provide feedback to ADD and Westat on the study.

In 2007, Westat established a process for soliciting feedback from the grantees on the draft measurement matrices. This included establishing, at the request of grantees and at a fairly significant cost, an e-room for individual stakeholders to access the draft measurement matrices and provide comment. In addition, Westat held several conference calls for grantees to provide feedback and engaged program participants in presentations at all three annual meetings of the national associations. It is important to note that despite these efforts, there was limited participation by the grantees in both the e-room and the conference calls.

Pilot Study
The original research design included a plan to conduct a pilot study for verifying and determining the reliability of the measurement matrices and related evaluation tools. This step in the plan allowed for another opportunity for program participation. To this end, grantees were able to contribute to the study by serving as a test-site and by giving feedback to Westat.

Validation Panels

As part of the original research design, validation panels comprised of various stakeholders (e.g., grantees, individuals with developmental disabilities, family members, national organizations) were convened to review the draft indicators in the measurement matrices to confirm the strength of the draft indicators to measure the DDC, P&A, UCEDD programs and DD Network collaboration. The validation panels offered the first opportunity to receive feedback from a **_large_** number of non-program staff (i.e., self-advocates and family members) in addition to program representatives. The validation panels indicated a high level of agreement that the draft indicators are valid.

Information Sharing

Key to an individual's ability to participate is knowledge. Information about the project has been shared throughout the three-year time period to keep the programs informed and knowledgeable about the study, thereby enhancing opportunities for participation. The chart below is an overview of the ways in which ADD has kept programs informed about the project.

October 2005	ADD announces to grantees the award of a contract to conduct Phase 1 of the independent evaluation.
November 2005	ADD and Westat conduct listening sessions on three different occasions to introduce the study to grantees. ADD meets with the Executive Directors of the three membership organizations (NACDD, NDRN, AUCD) to provide an overview of the project and its purpose. NACDD provides this information at their meeting in the Fall of 2006.
2006 - 2008	ADD Updates include information about the independent evaluation. The ADD website is updated to provide information about the study.
2006	ADD co-presents with Westat at the national technical assistance meetings for the P&As, DDCs, and UCEDDs. The presentations provide an overview of the study and outline key aspects of Phase 1.
2007	Westat presents at the national technical assistance meetings for the P&As, DDCs, and UCEDDs draft measurement matrices and outlines the process for soliciting feedback.
2008	ADD and Westat provides updates on the independent evaluation to grantees at the national technical assistance meetings for the P&As, DDCs, and UCEDDs.

It is important to note that ADD has to carefully balance its interest in being participatory with its interest in ensuring that it can defend the design and process for conducting the independent evaluation to OMB. OMB

expectations for independent evaluations are clear: such studies are to be objective and neutral and follow the scientific method; quality indicators that are in many ways in stark contrast to participatory research methods. The further the independent evaluation strays from these quality indicators for independent evaluations, the more difficult the position ADD is in when it needs to defend its programs.

While ADD appreciates the DD Network grantees' eagerness and deep interest in actively guiding the research design for the independent evaluation, the importance of independence in this situation cannot be underscored. ADD feels that it has achieved a healthy balance between being participatory while at the same time allowing for the study to be carried out in a relatively independent manner.

A-2. The scope of the study should be expanded to include ADD and the training and technical assistance contractors.

Response:

The Program Assessment Rating Tool (PART) provided the context for the design of the independent evaluation and was a major contributing factor to the conceptual framework for this study. Under the PART, there was less concern about the role of the agency and technical assistance providers in the performance of the programs. Therefore, ADD did not include these elements in the original design.

However, after implementation of Phase 1, ADD received feedback that it would be beneficial to include the agency and the TA contractors in the scope of the study. Various efforts were made to change the study to incorporate these elements; however, such attempts were invariably unsuccessful, mainly because the intensive nature of the participatory process to develop the measurement matrices left little time to consider how to broaden the scope of the study.

A-3. For the P&A program, the design should be based on the evaluation for the Protection and Advocacy Program for Individuals with Mental Illness program evaluation.

Response:

The P&As have recently undergone an evaluation of the Protection and Advocacy for Individuals with Mental Illness (PAIMI) Program funded by the Substance Abuse and Mental Health Services Administration, Center for Mental Health Services (CMHS). Since the start of DDPIE, ADD and Westat have engaged staff from the PAIMI program to coordinate study activities. This has included meeting with PAIMI staff to learn more about the evaluation, including PAIMI staff on the Advisory Committee, and reviewing materials from the PAIMI evaluation. This has helped to share information and to the extent possible create linkages to the PAIMI evaluation.

While effort has been made to share information about the two evaluations and although the National Disability Rights Network and P&A grantees would like to use the same model and basic questionnaire of the PAIMI evaluation for the DDPIE, there are a number of reasons why this would not be possible. Most particularly and importantly is that the two designs are significantly different and DDPIE has entirely different research objectives. The basic intent of the PAIMI evaluation was different than that of the DDPIE. The PAIMI evaluation did not examine outcomes; instead, it was developed to examine two things: (1) the degree to which grantees take on significant issues facing people with serious psychiatric disability; and (2) the effort invested in securing change given available resources, <u>as opposed to achievement of particular outcomes</u>. DDPIE objectives include the development of benchmarks and indicators; the use of a framework that incorporates structural, process, output, and outcome indicators; the rigorous measurement of the indicators that were identified; the development of performance standards as a way to gauge and evaluate the findings once data are available from measuring the indicators; and the ability to roll up evaluation findings to the national level. None of these objectives would be achieved with the PAIMI evaluation.

Once the PAIMI evaluation research objectives were developed, PAIMI evaluators developed an evaluation methodology and tool that were consistent with the CMHS research objectives. DDPIE evaluators developed an evaluation methodology and tools that were consistent with the DDPIE research objectives. Because the DDPIE research objectives are different from the PAIMI research objectives, it is not possible to use either the PAIMI model or the PAIMI data collection tool.

Moreover, the methodology of the PAIMI evaluation was less rigorous than that of DDPIE. Its sampling methodology was purposive, which will not be the case in the full-scale DDPIE. In addition, PAIMI evaluation data collection was not consistent (e.g., only four programs provided data on clients). It will be the intention of the DDPIE to collect data in the same way from the same types of individuals, in approximately the same numbers from each participating grantee in each grantee program.

A-4. **_The proposed data collection instruments and related measurement matrices (key functions, benchmarks, and indicators) do not accurately capture the programs operations and functions and impose a "one size fits all" to programs that are required to contextualize their work within the State or territory._**

Response:

The comments received to date provide very general feedback on the validity (e.g., the instruments measure what they purport to measure) of the proposed measurement matrices and related evaluation tools. Without more specificity in the comments, it is difficult to know what is not valid, most particularly since the evaluation tools were developed in a participatory manner. This included soliciting input from the grantees, conducting a pilot study to test the interview protocols, and review of the indicators by validation panels where there was general agreement that the measurement matrices accurately reflected what the programs do and what should be measured. While it is true that the self-administered questionnaire developed as result of the pilot study was not tested for reliability and

validity, there are plans to conduct cognitive testing on this instrument prior to field implementation.

ADD and Westat certainly understand and appreciate the uniqueness of each program. This characteristic of the programs was clearly conveyed throughout Phase 1 and is substantiated in the DD Act, which promotes the value of tailoring the work of the programs to the conditions in the State and, more importantly, the expressed needs of individuals with developmental disabilities and their family members. For example, each program is required to solicit feedback from key stakeholders and base its plan on data driven strategic planning.

Nevertheless, common features exist across the programs. In order to conduct a national program evaluation, common features of the programs must be identified for measurement purposes. Despite the unique qualities of the programs, the DD Act outlines key activities for each program. This perhaps is most clearly identified for the UCEDDs in the form of core functions. For national measurement purposes, Westat – with assistance from working groups – was able to classify these essential activities for each of the programs into key functions. These key functions were further described in benchmarks that serve as general standards or key expectations for the programs. The indicators are what would get measured. An example of a key function, benchmark, and indicator for the DD Council follows:

KEY FUNCTION: **State Plan Development**

BENCHMARK: **1.1 DD Council State Plans represent key issues, priorities, and needs of people with developmental disabilities and their families.**

INDICATORS: 1.1.1 The process DD councils use to develop the State Plan:

- Ensures wide and varied input from those who are knowledgeable about the needs of people with developmental disabilities and their families;
- Enables those in both urban and rural communities to participate;
- Provides opportunities for those who have never participated before to participate, particularly those who are typically unserved or underserved;
- Enables people with developmental disabilities to participate;
- Includes the use of feedback about current programs and activities (e.g., from participants in DD council-funded programs; staff or grantee feedback); and
- Includes the use of reliably collected, timely, and valid data

The example above outlines a measurement structure that describes common expectations for the DD Councils in developing the State plan that is general enough to be applied broadly. Moreover, the qualitative nature of the design will allow grantees to express their uniqueness in relation to the key function, in this case the development of the State Plan.

Using this exemplar, it is not clear how this measurement structure imposes a one size fits all approach, most particularly when the measurement matrices were developed with input from grantees, tested in a pilot study, and confirmed by validation panels. Without more specific examples of key functions, benchmarks, and indicators that do not apply broadly, it is not clear where the revisions are needed.

A-5. ***The proposed instruments impose requirements for the programs not specified in the authorizing legislation.***

Response:

One of the challenges with regard to setting the scope of the independent evaluation has been determining what gets measured. To this end, there have been on-going discussions regarding the extent to which the study should examine activities that go beyond the scope of what the DD Act requires for the programs. Given that the study is not a monitoring tool, there is more latitude to measuring the accomplishments and impact of the programs. However, it is important that any measurements developed do not appear to impose new requirements for the programs.

Certainly, throughout the development process, ADD and Westat have tried to carefully strike a balance between the two. However, the comments suggest that this was not fully achieved. Nevertheless, without more specific comments on the items that impose new requirements on the programs, it is difficult to know where changes are needed, most particularly since representatives from the programs contributed to the development of the measurement matrices, which were confirmed by validation panels.

For example, indicators accepted by the validation panels related to *increased* choice, control, participation, access, and satisfaction among consumers with developmental disabilities and their families. Yet, one commenter noted the DD Act does not require the DD Network to demonstrate its work has resulted in increases in choice, control, etc. Instead, the Act requires describing and measuring *improvements* in choice, control, participation, and access. However, the term 'improvements' is much too broad to be measured and is inappropriate for the independent evaluation. The question then becomes: It there a better verb than *increased* to describe *improvements*? ADD would certainly welcome suggestions for this particular example.

A-6. ***The proposed information collection instruments are redundant with other data collection required by ADD;***

therefore, ADD should use data currently collected by grantees for the independent evaluation.

Response:

The majority of commenters felt that the proposed data collection instruments were redundant with other data collection required by ADD. The commenters felt that the study can use current data collected by the grantees and reported to ADD using the program performance reports (PPRs) and the monitoring and technical assistance review system (MTARS). The commenters felt that the current information collected from grantees is: (a) Sufficient, reliable, and valid; and (b) enough of a burden that additional reporting requirements should not be placed on grantees. A related comment was that an evaluability assessment should have been conducted prior to initiating the study. In doing so, critical infrastructure improvements could be identified prior to new data being collected.

ADD and Westat agree that it would be ideal to use the current data and theoretically, it is logical to do so. As part of Phase 1, Westat explored whether current data from the PPRs and MTARS could be used for the independent evaluation and found that the: (1) Data are not always available and, (b) the data that are available are not reliable (i.e., collected in the same way by each grantee). More detailed information about these conclusions appears under C.

Westat's findings obviously raise questions about the utility of the current data to provide information for the independent evaluation. It leads to many good reasons not to use current data, most importantly, that in doing so, ADD would be supporting an evaluation that is potentially biased and based on unreliable data. As such, it would be very difficult to defend the methods for and the quality of the findings from the independent evaluation to OMB and others.

To better understand the need for collecting information using the proposed instruments, it is important to clearly distinguish the current data collection activities (e.g., PPRs, MTARS) and the DDPIE. Distinctions can be made in

three areas: (1) purpose; (2) design features, and (3) information collection procedures. These are discussed in more detail below.

Purpose

The PPRs serve two purposes: (1) Monitoring ***individual*** grantee activities by collecting narrative information about progress on goals; and (2) Annually collecting quantitative data on specific output and outcome (e.g., GPRA) measures that provide a broad portrait of what grantees are accomplishing. This quantitative data is reported out in a number of ways. The data submitted for the GPRA measures are compiled into aggregate form for ***national*** reporting purposes. Some of the output measures are also reported by ADD as a ***national*** aggregate figure. Other output measures are reported on a State-by-State basis.

The purpose of the MTARS is for conducting a more in-depth review of ***individual*** grantees to determine compliance with the DD Act and to improve individual grantee programs. The MTARS includes a peer review process in which a team reviews information submitted by the grantee to determine compliance, technical assistance needs, and areas of program innovation. This information is not reported out at the national level; however, ADD does make some of the program innovation information available.

The purpose of DDPIE is to conduct a ***national*** study of the impact of three grant programs funded by ADD: (1) DDCs, (2) P&As, (3) UCEDDs, and (4) collaboration among the three programs. It will use qualitative procedures to produce findings that provide a national portrait of DDC, P&A, and UCEDD program impact. It is based on the scientific process to ensure the reliability and validity of the data collection procedures. It will use rigorous procedures to assess the impact of the ADD grantee activities.

At the outset of this project, ADD stated in the RFP that it was not the purpose of this study to examine the current reporting mechanisms for ADD's grantees. Instead, the purpose was to use independently developed instruments for determining impact (see the text box below).

It is <u>not</u> the purpose of the independent evaluation to analyze ADD's current measurement system that is used by grantees to report on their activities. Instead, the purpose is to have an objective, outside contractor develop a new and specialized measurement system designed specifically to determine the impact of the programs on individuals with developmental disabilities, state service systems, and on the capacity of service providers and a wide-range of professionals to reach, treat, or assist individuals with developmental disabilities to become more

Design Features

It is also important to distinguish between the designs for the different information collection instruments as research design is critical to ensuring the reliability and validity of the evaluation findings. The PPRs and MTARS were designed for compliance purposes. The PPRs and MTARS were not designed for program evaluation purposes. DDPIE is designed for program evaluation purposes.

The PPRs were designed to ensure grantees are compliant with the reporting requirements outlined in the DD Act and to collect information for the GPRA measures. Because the DD Act requires programs to report on progress in meeting goals and on indicators of progress, the PPRs are designed to capture that kind of information. ADD also uses the PPRs to collect data for the GPRA measures; however, these measures in no way provide a full portrait of the ways in which programs have an impact. Instead, the GPRA measures are for a targeted purpose. Thus, the PPRs were designed to serve multiple purposes thereby providing a window into individual and national level program performance. The PPRs were never designed to document program outcomes in a more in-depth, quantitative and qualitative manner for national program evaluation purposes.

The MTARS tools were designed for monitoring and technical assistance purposes. The information collection tools used for MTARS allow for an in-depth compliance review of the grantees by examining implementation of the requirements for the programs as outlined in the DD Act. It is not a scientific process. Instead, it is a review process for examining compliance tailored to the needs of the program. Furthermore, compliance has not been proven to equate with impact.

DDPIE is not designed for compliance purposes. Instead, the design for the program evaluation is based on the scientific process. To this end, data collection instruments were developed using the scientific process to verify through a pilot study and validation panels that the proposed measurement matrices are valid and that the instruments developed prior to the pilot study were reliable and valid. Moreover, the independent evaluation is free of the confines of measuring program compliance, thereby allowing for a more focused and in-depth review of program impact.

Information Collection Procedures
Finally, and perhaps most critically for the purposes of the discussion about redundancy, it is important to distinguish between the procedures for collecting the information. For research purposes, it is critical that data are collected using rigorous procedures to ensure reliability of the findings.

There is tremendous variability in the way in which each grantee collects and reports data in the PPRs. ADD regularly confirms from grantees, as did Westat, the variability in PPR data reporting methods. For example, Directors from the various programs have told ADD that they input data under erroneous measures to ensure that the data gets 'counted'. In other cases, it is clear from Directors that one program enters the data under one category and another reports the same kind of data under another category.

This does not provide evidence that the data are non-biased. ADD cannot use unreliable data for a program evaluation intended to demonstrate at the national level the impacts of its programs. If data were collected in a consistent manner, then ADD could consider using current data for this study. However, this is not the case.

ADD estimates that it would bear a tremendous cost to add a process to this study that would ensure the data provided by the grantees are reliable. This would require for all 179 grantees further and more in-depth investigation of the current data collection procedures for setting a baseline, development of data collection protocols, training grantee staff on the data collection protocols, and examining data collection procedures post-training to establish reliability.

There is recognized need to revise the current PPR templates and work is underway to change the PPR templates and improve the reliability of the data collection systems. Obviously ADD would not want to invest in developing reliability for a measurement system that needs to be improved. It will take at least two years before the quality of the revised reporting systems is known. ADD cannot further delay collecting information in a reliable manner that reveals critical accountability information. It is much more cost effective to use the instruments developed by Westat given the very modest investment ADD has made in this project to have information collected by trained research staff thereby ensuring reliability of the data. In doing so, ADD will reduce burden for the grantees while at the same time ensure the quality of the data and ultimately the findings.

While the MTARS uses the same tools for grantees to provide information about program compliance (e.g., self-assessment checklists), ADD does not provide specific guidance on how to collect this information, nor does it expect that specific procedures are followed. In addition, various reviewers participate in the MTARS. While ADD trains reviewers on the process, the variation in reviewers leads to variations in the review process. Therefore, there is great variability in the way in which the MTARS is conducted. However, this variation is purposeful and important because it allows for programs to individualize the review process and address areas of need.

Because DDPIE is based on the scientific process and the results will be used to report nationally about the impact of the ADD grant programs, the proposed information collection instruments for DDPIE allows for the data to be gathered in a consistent manner thereby ensuring reliability of the information. Moreover, the proposed instruments were developed over a three year time period and various steps were taken to determine the validity of the instruments. The evaluation will use rigorous research procedures, such as random sampling and strict adherence to study protocols, to collect the data.

ADD recognizes that it is confusing to have many information collection activities occurring simultaneously that are designed to examine program

performance. The table below summarizes the three information collection systems for understanding program performance and the differences and similarities between them.

Characteristics	PPRs	MTARS	Independent Evaluation
Concerned with compliance	✓	✓	
Uses broad based quantitative output and outcome (e.g., GPRA) measures to report results	✓		
Based on the scientific process			✓
Uses rigorous procedures to ensure reliability of the data			✓
Concerned with program improvement	✓	✓	✓
Characterized by individualization and variability	✓	✓	

ADD commends its grantees for setting such high expectations for this project in terms of overhauling the current reporting system. While we appreciate the call to reform the PPR systems now, ADD realistically never envisioned this project to accomplish that objective. It would be beyond the scope and more importantly, the resources of this project to make changes to the current reporting system. ADD certainly recognizes the need to improve its current data collection systems and is engaged in various processes to make such improvements.

The fact that the current PPR system needs to be improved should not prohibit ADD from conducting a program evaluation in the meantime that documents grantee impact. The impetus to develop and implement an independent evaluation of the DD Network programs remains equally, if not more overwhelmingly, urgent and is consistent with the new Administration's intent to restore "responsibility and accountability to government." In the document entitled "Building a High-Performing Government" (available at http://www.whitehouse.gov/omb/budget/fy2010/assets/building.pdf and last accessed on June 9, 2009), the Obama Administration outlines its framework for accountability. Below are key statements from the Administration on federal accountability:

- The Obama Administration will work with the [newly established] PIC [Performance Improvement Council] to fundamentally reconfigure how the Federal Government assesses program performance. A reformed performance and

analysis framework will switch the focus from grading programs as successful or unsuccessful to requiring agency leaders to set priority goals, demonstrate progress in achieving goals, and explain performance trends. In order to break down silos, cross-program and cross-agency goals would receive as much or more focus as program-specific ones.

- As a first step in this process, OMB, during the next few months, will ask each major agency to identify a limited set of high priority goals, supported by meaningful measures and quantitative targets, that will serve as the basis for the President's meetings with cabinet officer to review their progress toward meeting performance improvement targets.

- A reformed performance improvement and analysis framework also would emphasize program evaluation.

As the federal agency, it is ADD's responsibility to provide the public with information about the programs' impact. If ADD postpones the evaluation, it keeps the public waiting for reliable information to show the impact of tax payer's investment. This project fills an important gap in the story of program impact.

A-7. ***The proposed data collection instruments and related measurement matrices (key functions, benchmarks, and indicators) impose an additional burden on programs already burdened by current reporting requirements.***

Response:

ADD acknowledges the amount of time and effort that grantees put into meeting the reporting requirements for the programs through the PPRs and MTARS and that this proposed information collection activity will add to that burden. Given the significance of demonstrating accountability for its programs, ADD feels that the proposed information collection activity is important to its operations. Nevertheless, ADD wants to ensure that the programs do not experience significantly more burden as a result of the

independent evaluation and has made every effort to keep the burden to a minimum. Some factors that contribute to minimizing burden include:

- This information collection process will be a one-time event during the course of the study. Programs will not have to respond annually to the information collection.
- The data are collected by an outside evaluator from the grantees and other stakeholders, mainly through interviews. As well, the data will be analyzed and reported out by the outside evaluator. The interview participants will not be responsible for ensuring the quality of the data, which is a time-consuming process. Rather, they will be responsible for taking time to be interviewed, providing responses to the interview questions, completing a self-administered questionnaire, and helping with organizing interviews. If the programs were to collect and analyze the data directly, there would be a tremendous increase in the burden of this information collection activity.
- Rather than having all programs provide data for this proposed information collection activity, there will be a sample of programs that participate in the independent evaluation. This sample will be up to 20 States. To reduce burden for the programs that participate, ADD will not conduct the MTARS in the same year in the States that are participating in the independent evaluation, thereby reducing the burden for these programs.

ADD agrees that the original burden estimates did not include the time necessary for logistics in setting up interviews and providing the evaluator with information on the program. Therefore, the burden estimates have been revised to take into consideration this level of effort (see below).

A-8. ***There are concerns that findings from the independent evaluation will be used to report on or penalize individual programs.***

Response:

The primary purpose of the independent evaluation is to reveal more in-depth information than is currently available to describe the impact of the DD Councils, P&As, and UCEDDs on individuals with developmental disabilities and their families. This information will be shared with the public and other stakeholders to demonstrate accountability.

With this purpose in mind, it has always been the intent to collect national data and to report findings at the aggregate level. To this end, it has never been the intent of the independent evaluation to report findings for one program or to penalize programs. Instead, data will be kept confidential and analyzed across programs to develop national findings for the programs. Since this is not a monitoring activity but an evaluation, ADD does not intend to receive data and results for individual programs.

ADD agrees that to improve program quality, it must have a plan in place. However, it is difficult to know at this point what that plan should be given that findings are pending from the study. In general, ADD has discussed how to best utilize the information and tools from the independent evaluation. This has included using the measurement matrices to improve the PPR and MTARS. In addition, part of the study will include recommendations to ADD regarding program improvements based on findings from the evaluation.

B. **The accuracy of the agency's estimate of the burden of the proposed collection of information**

All three national organizations and individual programs commented that the actual burden was underestimated and that the actual burden to participate in the DDPIE would be greater than the estimates given. They indicated that they are already burdened with information requirements from ADD. They also thought that programs would be responsible for contacting interviewees who were not program staff, obtaining consent and possibly IRB approval. They feared that such contact with potential interviewees would compromise their relationship with the community. Below are responses to the comments summarized above.

B-1. **The actual burden was underestimated.**

Response:

The original estimate of burden only included an estimate of the hours it would take to respond to questions in data collection instruments (as required by OMB). It did not include hours for preparation for the evaluation (e.g., time to collect, organize, and submit advance materials and materials collected on site; identify key informants, obtain consent; prepare agenda; schedule interviews; make logistical arrangements; and participate in an exit interview). A revised estimate in Tables 1 and 2 include hours originally indicated plus hours for preparation, per program type and per individual grantee program. Table 3 presents a breakdown of additional hours per program type by task.

Table 1. Summary: Estimate of total burden for DDPIE evaluation, per program type for 20 grantees per program type			
Program type	Estimate of burden to administer data collection instruments	Additional estimate of burden*	Estimate of total burden
P&A	800 hours	670 hours	1,470 hours
DD Council	755 hours	670 hours	1,425 hours
UCEDD	510 hours	670 hours	1,180 hours

All programs	2,065 hours	2,010 hours	4,075 hours

*includes time to collect, organize, and submit advance materials and materials collected on site; identify key informants, obtain consent; prepare agenda; schedule interviews; make logistical arrangements; and participate in an exit interview

Table 2. Estimate of total burden for DDPIE Evaluation, per grantee program by program type			
Program type	Estimate of burden to administer data collection instruments	Additional estimate of burden*	Estimate of total burden
P&A	40 hours	33.5 hours	73.5 hours
DD Council	37.75 hours	33.5 hours	71.25 hours
UCEDD	25.5 hours	33.5 hours	59 hours
All programs	103.25 hours	100.5 hours	203.75 hours

*includes time to collect, organize, and submit advance materials and materials collected on site; identify key informants, obtain consent; prepare agenda; schedule interviews; make logistical arrangements; and participate in an exit interview

Table 3. Estimate of additional burden for each task in the DDPIE Phase 2—full-scale evaluation, by program type				
Task—DD Council	Number of respondents	Number of responses per respondent	Average burden hours per respondent	Total burden hours
Prepare agenda (including emails and phone calls with contractor staff to schedule a two-day visit, understand selection criteria for interviewees, and identify a topic for the group interview).	20	1	5	100
Track down documents on checklist of materials (including compiling, photocopying, and sending requested documents to contractor).	20	1	10	200
Select interviewees and make arrangements for their participation	20	1	10	200
Review questionnaires prior to visit.	20	1	5	100
Set up video-conference or phone conferences	20	1	3.5	70
Subtotal			33.5	670

Task—P&A	Number of respondents	Number of responses per respondent	Average burden hours per response	Total burden hours
Prepare agenda (including emails and phone calls with contractor staff to schedule a two-day visit, understand selection criteria for interviewees, and identify a topic for the group interview).	20	1	5	100
Track down documents on checklist of materials (including compiling, photocopying, and sending requested documents to contractor).	20	1	10	200
Select interviewees and make arrangements for their participation	20	1	10	200
Review questionnaires prior to visit.	20	1	5	100
Set up video-conference or phone conferences	20	1	3.5	70
Subtotal			33.5	670

Task—UCEDD	Number of respondents	Number of responses per respondent	Average burden hours per respondent	Total burden hours
Prepare agenda (including emails and phone calls with contractor staff to schedule a two-day visit, understand selection criteria for interviewees, and identify a topic for the group interview).	20	1	5	100
Track down documents on checklist of materials (including compiling, photocopying, and sending requested documents to contractor).	20	1	10	200
Select interviewees and make arrangements for their participation	20	1	10	200
Review questionnaires prior to visit.	20	1	5	100
Set up video-conference or phone conferences	20	1	3.5	70
Subtotal		5	33.5	670
Total of additional estimated total			100.5	2,010

Task—UCEDD	Number of respondents	Number of responses per respondent	Average burden hours per respondent	Total burden hours
program burden for all three programs				

With the addition of burden hours for tasks other than administration of data collection instruments (e.g., to collect, organize, and submit advance materials and materials collected on site; identify key informants, obtain consent; prepare agenda; schedule interviews; make logistical arrangements; and participate in an exit interview), the total estimate of burden to conduct the DDPIE on all 20 grantee programs is 4,075 hours – 1,470 hours for the P&A, 1,425 for the DD Councils, and 1,180 for the UCEDDs. Per grantee,[1] the total burden would be 73.5 hours for a P&A grantee, 71.25 hours for a DD Council grantee, and 59 hours for a UCEDD grantee. These total estimates are distributed among program staff, program staff interviewees, and interviewees external to the program.

These estimates of burden are based on pilot testing the materials in nine DD Network programs. Time to complete the self-administered questionnaire is based on time to complete similar questions during the face-to-face interviews in the pilot study that are now incorporated into the self-administered questionnaire.

B-2. ***Programs are already burdened with information requirements from ADD.***

Response:

Although commenters note that grantees are already burdened with information requirements from ADD, as noted above, data collections to meet those requirements have different purposes. Programs are responsible for collecting, analyzing, and reporting data through current reporting tools, which provide data at the programmatic level. The DDPIE will not be used to

---

[1]There will be up to 20 P&A grantees, up to 20 DD Councils, and up to 20 UCEDDs participating in the full scale evaluation.

evaluate individual grantee programs. Instead, it will generate data ADD needs to obtain a picture of performance and the impact of DD Network programs at the national level. Standardizing existing evaluation tools to meet this national evaluation need would take away the flexibility of programs to report what they want to report to meet other requirements. Existing reporting tools capture the unique characteristics and accomplishments of individual programs. Although every attempt has been made to capture the unique characteristics and accomplishments of individual grantees, the DDPIE has the added responsibility of collecting, analyzing, and reporting data in a consistent and standardized format so that data can be rolled up to a national level for accountability purposes and to inform national decision-making.

B-3. *Programs would be responsible for contacting interviewees, obtaining consent, and possibly IRB approval.*

Response:

Although programs will be asked to identify and arrange interviewees for the DDPIE, DD Network programs will not be asked to bear any associated costs. Evaluators will provide supports, reimburse expenses for supports and accommodations; and reimburse expenses for travel for interviewees and personal aides, as needed. The need to reimburse for respite or childcare will be unlikely, since evaluators will complete interviews by phone for those for whom travel to the site is difficult. However, such reimbursement will be provided, if necessary. Evaluators will ask programs to try to select interviewees who live in near proximity to the program site, and the evaluators will travel to a location more easily accessible to the interviewee, if necessary. In addition, evaluators will send copies of all materials in advance to the sites. Grantees are free to make these materials available to external stakeholders (e.g., self-advocates who participated in DD Network programs). However, there is no need to prepare interviewees in advance, except to clarify the purpose of the evaluation (which is NOT to evaluate the individual grantee program). In fact, preparing external stakeholders any

further than that would be seen as compromising the validity of data collection.

Consistent with Westat's IRB requirements, the independent evaluators will be responsible for obtaining formal consent from individual interviewees and participants in in-person and group interviews. Westat will prepare all consent forms, explain participant rights, clarify instructions, and answer all questions before obtaining written consent for in-person interviews and discussions and oral consent for phone/video conference participation.

Regarding IRB approval, Westat obtained approval for the pilot study of the evaluation from its own IRB and will do the same for the full scale evaluation. Should grantee programs also feel the need to obtain IRB approval, Westat will facilitate that process as much as possible (e.g., by providing the package of materials sent to its own IRB). The DDPIE met the Westat's IRB criterion for an expedited IRB review ("research on individual or group behavior, including perception and social behavior or research employing survey, interview, focus group, program, human factors evaluation, or quality assurance). We expect that will be the case for the full scale evaluation at Westat and other IRBs.

B-4. ***Obtaining interviewees for the purpose of the independent evaluation would compromise relationships with the community.***

Response:

We do not believe that participation in DDPIE will be an intrusion on participants or compromise in any way a grantee program's relationship with the community. Evaluators do not need to interview the same people who have provided evaluation feedback previously. Also, in the pilot study, consumers and collaborators were pleased to be able to provide input on their experiences with the programs. DDPIE pilot sites reported that pilot site participation was a learning experience for the program.

C.   The quality, utility, and clarity of the information to be collected.

As was discussed under A, many of the comments received related to the utility as well as quality and clarity of the information to be collected. Below is a summary of the various comments received about the quality, utility, and clarity of the information to be collected:

All three programs indicated that the information collection is redundant with current data collection (e.g., MTARS, APR, and planning documents), and all thought that using or revising current reporting and evaluation tools would be more efficient.

The commenters stated that this data collection effort will force conformity on program data collection and reporting efforts when flexibility is preferred. Some wanted an evaluability assessment conducted first to identify infrastructure factors within the ADD that need to be resolved prior to conducting the evaluation.

Others stated that the proposed evaluation materials do not enjoy endorsement of the programs they are to measure and that throughout the development process, there was little agreement on their validity. They felt that additional work is needed to resolve those problems prior to acceptance of the tools and conduct of the evaluation.

Other comments received stated that the evaluation materials have been substantially revised, but the revised tools were not subject to assessment of their reliability and validity.

Others expressed concern that the DDPIE data collection will produce information of limited to no practical utility and that portions of the proposed collection of information are of poor conceptual quality, and their relationship to the DD Act is tenuous at best. The commenters stated that ADD should use working groups established by ADD and establish additional groups as needed to address improvements in existing tools and ensure a more participatory process to make sure the evaluation tools meet ADD evaluation needs. There is concern about what DDPIE measures and what DDPIE does not measure.

Others stated that there are flaws in the DDPIE evaluation approach; that DDPIE uses older evaluation methodology that should be re-considered in favor of a more participatory model that could be tailored to the complexities of the programs' systems change work and unique state responses. Some commenters stated that the program has performance measures but needs efficiency measures to assess cost-effectiveness urging that the focus be on positive outcomes, not process and methodology, and should show how the programs are making a difference. There was concern expressed about what DDPIE measures and what it does not measure and that a prescriptive approach will alienate DD programs.

Some stated that there would be more utility in delaying the independent evaluation until the reauthorization of the DD Act is concluded so the evaluation can be tailored to the new law.

Some stated that measurement of collaborative endeavors with other programs should reflect outcomes of collaboration rather than mandating a specific collaboration process that might or might not be appropriate for all states and territories.

Most felt that the evaluation should recognize diversity among the programs, such as funding for minimum allotment states.

Some comments stated that he evaluation should recognize that the programs are independent entities, and they may be affiliated with other entities (e.g., as a state agency, as part of an institute of higher education), whose policies and structures may vary or be driven by grants and contracts the programs secure to support their operation and activities.

The section below provides responses to the comments received, including responses to specific

C-1. *DDPIE data collection tools are redundant.*

Response:

The pilot study directly addressed the issue of using existing data for the evaluation. It was concluded that existing information is not available in a consistent format usable for an independent evaluation rolled up to the national level.

In the pilot study, the evaluator reviewed the following reports submitted to ADD by pilot study programs to determine whether data from these reports would be able to answer the questions in the pilot study questionnaires:.

- State Plans (from New Mexico, Ohio, and Wyoming DD Councils),
- Statements of Goals and Priorities (from Ohio, New Hampshire, and Alaska P&As),
- 5-year Plans (from Iowa, New Hampshire, and Los Angeles, California UCEDDs),

- Program Performance Reports (PPR) (from New Mexico, New Hampshire, Ohio, and Wyoming DD Councils and from Ohio, New Hampshire and Alaska P&As),
- Annual reports (from Iowa, New Hampshire, and Los Angeles, California UCEDDs),
- Background information in preparation for Monitoring and Technical Assistance Review System (MTARS) visits conducted by ADD from Los Angeles, California UCEDD, and
- Data from the National Information and Reporting System (NIRS) (for Iowa, New Hampshire, and Los Angeles, California UCEDDs).

Through a crosswalk of the indicators that would be measured in the DDPIE with existing data provided to ADD in various reports, the evaluator identified the approximate location of data in several reports to ADD that might be able to answer the evaluation questions. The evaluator incorporated the existing data from those reports into questionnaire binders.

The evaluators entered data from reports to ADD in Access databases and Excel spreadsheets to fit the indicators in questionnaires. For some of the reports, such as the MTARS Checklists, copies were made of relevant portions. Evaluators integrated hard copies of the data from the reports into the questionnaires, attempted to work with interviewees (usually the Executive Director) to use the data from reports, and also collected data from programs participating in the pilot study using the DDPIE data collection instruments.

In many cases, the existing data was incomplete, out of date, not related to the dates of interest, or, as Westat found, not specifically related to the evaluation question. There was considerable inconsistency in definitions used by each program and differences in formatting and contents of each report. Although much of the data was useful as background, it was not useful for answering the questions in the evaluation data collection instruments, which related specifically to the benchmarks and indicators that had been developed. Moreover, because of the inconsistency in definitions and data collection methodology, data would not be considered reliable

enough to meet OMB requirements, and it would not be possible to combine program data for roll-up to the national level.

The following summarizes findings from the pilot study on each DD Network Program regarding the use of existing data.

DD Councils. The evaluators reviewed a variety of reports from DD Councils, including State Plans, PPRs, and background information for MTARS visits and found that:

1.   Both the State Plan and PPR provided basic background information (e.g., Goals and Objectives in the State Plan, Council members and their affiliations, Comprehensive Review and Analysis, and collaborators). However, such information did not help evaluators fully understand approaches a DD Council used to achieve its goals and objectives and the impacts of the Council's work on people with developmental disabilities, their families, service providers, or State systems. The "process" information (e.g., how the State Plan was developed and projects carried out) was not always included in the State Plan or PPR. In addition, the performance targets of the State Plan and outcomes reported in the PPR tended to focus on outputs rather than impacts on people with developmental disabilities and/or their families.

2.   The MTARS report was not available for all DD Councils because some Councils did not have an MTARS visit in recent years. The report on the MTARS visit was compliance focused and was about 2-3 years old. Therefore, the utility of the report was limited.

3.   DD Council projects, activities, and project outcomes included in the PPR were organized as required by Area of Emphasis (e.g., childcare, education, employment, housing, health, and transportation). Without obtaining additional information from staff, it was difficult to determine how DD Council projects, activities, or outcomes were related to specific key functions.

P&As. The evaluators obtained each pilot study P&A program's SGP and PPR, background information for MTARS visits where it existed, and a variety of materials from the programs (e.g., policies, procedures, forms). Evaluators also examined each program's website. Existing data, reports, and other materials from programs were helpful as background material and documentation of policies and procedures but not as answers to specific data-related questions. There was redundancy when the evaluator already had documentation (e.g., job descriptions, board membership requirements and responsibilities, intake forms) for questions. However, there were no existing data for many other questions. Specific examples follow.

1.   Although all P&A systems are required to develop and submit an SGP to the Secretary of Health and Human Services through ADD each year, these documents varied considerably in format and in what was included in each SGP. The goals in one program's SGP were general (e.g., ensure that individuals with developmental disabilities have access to government benefits), but the objectives were specific (e.g., to assist 10 individuals with developmental disabilities to obtain and maintain SSI/SSDI and related benefits; to conduct eight trainings, etc.). One of the programs in the pilot study presented a narrative that summarized the solicitation effort for this SGP, summarized public comment and P&A system response, listed the objectives and priorities, and targeted the number of cases that would be served. The others did not. The third SGP provided one or more indicators of success (e.g., Will handle a minimum of 30 special education cases; HB 766 and HB 679 will be enacted into law; The number of private attorneys talking special education pro bono cases will increase by at least 10%; Enforcement actions by the state Department of Education against school districts and other education al programs will increase), while the others did not.

2.   The evaluators were able to obtain information from some of the PPRs on:

        ▪   Number of people on the board of directors or oversight group

- Number that were sent a satisfaction survey and the response rates
- Number of clients served by type of disability
- Types of systemic advocacy activities, such as full investigation, monitoring, class-action litigation; information on major areas under consideration; groups likely to be affected; major outcomes during the year; and outputs, such as organized a conference, attended oral arguments, or obtained a court ordered action requested by the P&A, which some P&As referred to as outcomes.
- Number of information and referral services provided
- Number of clients receiving advocacy at the beginning of the fiscal year.
- Number of new or renewed clients.
- Number of publications disseminated

However, for the most part, the numbers could not be rolled up to the national level. For example, the numbers did not all cover the same reporting time period; the P&As defined information and referral differently; the P&As not only defined clients differently, they also defined the services provided differently; sizes, types, and purposes of publications produced and disseminated varied, adding a mix of variables on how to "count" publications. The P&As also differed in data they collected on persons who contacted the P&A for different reasons; some sent satisfaction surveys to all who contacted the P&A regardless of the reason and some sent surveys only to persons who became individual advocacy clients.

3. The PPR did not provide information on targeting unserved or underserved populations or how those populations are reached.

4. Only one P&A was able to submit MTARs documents. Responses to the MTARS checklist sometimes only referenced a specific document without reference to a particular year. In addition, other materials the evaluator received (e.g., an SGP for current year) did not

correspond to the document referred to in the MTARS comment/response column.

5. Program websites provided useful information (e.g., agency priorities, Board composition and nominating committees, and information on the board of directors or other oversight group, including number, names, and meeting schedule). However, each website provided slightly different information.

UCEDDs. The evaluators extracted data from UCEDD reports and NIRS related to the indicators within the pilot study questionnaires. However, recent MTARS data were available only for a selected number of locations at any given time, and much of the information was in narrative form and not comparable from location to location. The following points are findings from the pilot study on UCEDD reports to ADD:

1. All UCEDDs complete annual reports that describe their progress on implementing their grants. However, reports differed widely in types of activities covered and level of detail provided.

2. UCEDDs provide considerable updated information to the National Information and Reporting System (NIRS) data collection system by the end of June each year. However, programs do not interpret definitions consistently and definitions are not precise. The following are examples of problematic definitions:

    ▪ Faculty and staff – UCEDDs can list anyone they choose under faculty and staff on the NIRS database, regardless of their basis for a connection to the UCEDD. Moreover, using existing NIRS data, or even the grantee program website, it is not possible to know which university faculty and staff is considered "UCEDD faculty and staff." Without such information, it would not be possible to reliably measure several indicators:

- o UCEDD faculty[2] is comprised of a variety of disciplines.
- o UCEDD faculty and teaching staff are considered to be effective teachers by their students and peers.
- o UCEDD faculty and staff provide their disability-related expertise to the university.
- o UCEDD faculty and staff publish on their disability research.
- o UCEDD faculty and staff are selected to make presentations on their disability-related research (including public policy analysis and evaluation) at conferences and meetings.
- o UCEDD faculty and staff provide advice on disability related issues to local, state, federal, and international organizations.
- o UCEDD faculty and staff review grants, manuscripts, books, articles, and other types of publications.
- o Disability-related publications authored or co-authored by UCEDD faculty and staff are cited by other researchers.

- ▪ Students – Students are categorized according to the number of hours they spend in the UCEDD program. The category "40 or fewer hours" can include people who observe a program for a few days, students who are earning continuing education units (CEUs), or students who have taken regular classes. The same difficulty applies to students who have spent 150 hours or more in the program. Without a consistent definition of student, the following indicators could not be measured reliably:

  - o UCEDD-developed curricula, courses, and course content prepare students to work with and for typically unserved or underserved populations or communities.
  - o Interdisciplinary pre-service students who completed their course of study work to benefit and affect the quality of life of people with developmental disabilities.

---

[2] UCEDD faculty and teaching staff are individuals with a university or faculty appointment (**tenure, non-tenure or adjunct**) and who have a designated official role with the UCEDD (e.g., at least some proportion of their salary is funded under the UCEDD's budget or a UCEDD grant or contract; works for a university academic department and is released from some of their departmental academic responsibilities in order to work with the UCEDD; is funded by the university fully or partially to be a UCEDD faculty member; works for an academic department but does some work for the UCEDD in addition to their departmental academic responsibilities).

- o   Among those students who participated in a disability studies program, disability is an important component of further education, career or their daily lives.
- o   Continuing education students apply what they learned in UCEDD continuing education courses to their work.

C-2   ***Using or revising current reporting tools would be more efficient than using the data collection tools in DDPIE.***

There is no evidence to support the assumption that it would be more cost efficient and less burdensome to revise current reporting and evaluation tools to collect data on DD Network programs in a standardized format that would allow the data to be rolled up to the national level. Revision of existing tools would need to be very extensive in order to collect data that can be rolled up to the national level. The DDPIE will try to use what it can of data generated through current existing tools. However, experience in the pilot study has shown that very little can be used in its current form.

C-3   ***DD Network Programs and grantees prefer a flexible data collection and reporting.***

Response:

The DDPIE data collection effort is a totally separate data collection effort from existing ones. Therefore, there is no basis for the concern that it will force conformity on current program data collection and reporting when flexibility is preferred. The two data collection activities have different purposes. The DDPIE will gather data from 20 sampled programs in a consistent and standardized format and report data rolled up to the national level to allow ADD to review the performance of the DD Network programs as a whole and use the data to inform decision making. The independent evaluation data collection instruments will not be used to evaluate individual grantees. In addition, a special attempt has been made to solicit information from individual grantees about their unique characteristics and achievements that influence decision making at the local program level.

C-4 ***There should be an evaluability assessment and the tools are not endorsed by the programs.***

Response:

As noted above, the DDPIE is not intended to be an evaluation of ADD and its infrastructure. However, as one commenter acknowledged, "…Phase 1 of the DDPIE does provide a great deal of information that can be used as an evaluability assessment and can guide infrastructure improvements in ADD". The commenter's concern is that ADD has not expressed a commitment to do so. However, obtaining an expressed commitment to address possible ADD infrastructure changes is beyond the scope of the DDPIE.

It should also be noted that an evaluability assessment is a specific type of evaluation meant to identify whether the evaluation of a specific program is justified, feasible, and likely to provide useful information. It not only shows whether a program can be meaningfully evaluated, but also whether conducting the evaluation is likely to contribute to improved program performance and management. Whereas individual grantee programs may have the luxury to decide when their program is ready to be evaluated, ADD has an obligation to examine the programs in their current state to meet the purpose of the DDPIE -- to examine through rigorous and comprehensive performance-based research procedures the targeted impact of grantee activities funded under the DD Act.

Nevertheless, although the contractor did not conduct an evaluability assessment per se, considerable work went into understanding the culture of the three DD Network programs, existing data collection required by ADD, and other existing infrastructure. Evaluation materials were developed to be sensitive to the culture of the three programs (e.g., most of the questions are open-ended and allow all grantees to select the specific activities and approaches they want to talk about). Moreover, as discussed above, considerable effort was made to understand and utilize existing data. These data were found inadequate for the purposes of the DDPIE. In addition, ADD does not have the luxury to wait for improvements in the existing infrastructure for collecting the data before conducting the DDPIE.

It should also be mentioned that the evaluation materials do not need the endorsement of the programs. ADD recognizes the importance in achieving stakeholder buy-in to an evaluation. However, judging from the opposition the programs expressed throughout the development process, it seems unlikely that the programs will endorse any new evaluation efforts. Evaluation is an expectation of accountability and is unavoidable. The DDPIE development process has involved intense program participation to ensure that the evaluation materials would be relevant to the programs' work and usable to ADD.

The validation panels reviewed and rated the components of the evaluation materials. Their positive ratings confirmed that the evaluation materials were measuring what they should be measuring and the indicators that comprised the materials were measurable. Moreover, when validation panel members made suggestions for wording revisions, Westat made every effort to incorporate those suggestions.

C-4   *DDPIE data collection will produce information of limited or no practical utility.*

Response:

The process for producing the DDPIE materials involved input from many stakeholders directly involved in the DD Network programs (e.g., ADD staff, DD Network member working groups, validation panels, an advisory committee, pilot site participants, and, importantly, consumers with disabilities and their families). To say the information has limited or no practical utility, is of poor conceptual quality, and has a tenuous relationship to the DD Act is to repudiate the work of those who contributed so much time and expertise to development of the materials.

The DDPIE development process was highly participatory. It consisted of eight steps: 1) Obtaining background on DD Network programs; 2) Establishing and working with the project's Advisory Panel and DD Network Working Groups in person and by telephone and web cast; 3) Developing

draft measurement matrices with input from these same groups; 4) Obtaining feedback from programs; 5) Conducting the pilot study; 6) Synthesizing pilot study findings and obtaining feedback from the Advisory Panel; 7) Validating the materials through a two-day validation panel process; and 8) Final synthesizing and reporting to ADD.

The evaluators sought and received feedback on the materials throughout the process. The pilot study of DDPIE materials resulted in findings related to the evaluation tools themselves, logistics for data collection, and existing data. As a result, the evaluators revised the tools and presented them to validation panels for review and input and incorporated their recommendations into benchmarks, indicators, and examples of performance standards that will pertain to the full-scale evaluation.

The validation panel meetings provided the opportunity for large numbers of self-advocates and family members to respond. Although some of the same issues on the necessity of the evaluation surfaced during large group discussions, feedback on the benchmarks and indicators was generally positive and constructive and enabled the evaluator to make additional meaningful revisions to the evaluation tools. In addition to formalized participatory development, Year 2 DDPIE activities consisted of providing DD network programs in each state with opportunities to provide feedback and comments on the evaluation and draft documents, including inviting programs to participate through an e-room, established to allow ease of participation and response by individual programs and other interested parties.

C-5: ***The DDPIE evaluation models use outdated methodology that focuses on process and methodology.***

Response:

The DDPIE is based on two evaluation models that are old, but considered by evaluation experts to be "tried and true." These are the Discrepancy Evaluation Model and the Open Systems Model. The Discrepancy Evaluation Model asks and compares answers to two basic questions: What is? and

What should be? For the DDPIE, one might ask, for example, "Does a particular 'score' on how DD Council State Plans represent key issues, priorities, and needs of people with developmental disabilities and their families meet the standard set by the DDPIE process? Based on a comparison between the expected level of performance (the standard) and the measured level of performance, discrepancies are identified. These discrepancies can be either positive or negative, thus identifying areas of strength or weakness. Identifying strengths and weaknesses facilitates informed decisions about program needs and, ultimately, the quality of the program.

The DDPIE development process was also consistent with the concepts contained in an Open System Model. In such a model, effectiveness is defined as the relationship between the outcomes achieved and the processes used to affect those outcomes. Program efficiency is evaluated through a comparison of inputs and outputs. When outcomes are satisfactory, or remain constant, the relative efficiency can be assessed among different approaches to building capacity to serve people with disabilities. In this model, efficiency is only relevant if positive outcomes are realized. DDPIE materials look at important inputs, processes, outputs, and outcomes relevant to the DD Network programs. By identifying performance standards related to those inputs, processes, outputs, and outcomes, the DDPIE will be able to perform a useful review, over time, of the national DD Network's efficiency and effectiveness.

Throughout DDPIE Phase 1, the evaluators used an Open System Model to conceptualize all aspects of the project and guide the development of the evaluation materials (which include benchmarks, indicators, and examples of performance standards), data collection, and analysis. An Open System Model of evaluation contains four elements. Structures (or inputs) are those resources that are needed to set processes in motion and keep them running. In the context of the DDPIE, examples of inputs include the DD Act, sections in the Act that establish components of the DD Network and areas of emphasis on behalf of the target population (people with developmental disabilities), and funds that flow from the Act.

Processes are those event sequences and arrangements of staff, services, and resources needed to achieve the intended result(s) (e.g., processes that are set up to implement the activities of each DD Network component and their collaboration). When inputs are in place and processes are functioning as intended, outputs and outcomes are produced. Outputs, often referred to as "products," are the "units" produced by processes supported by given inputs. Examples of products in the context of the DDPIE are policy changes, increased capacity, and legislative and policy compliance.

Outcomes are the intended results of creating certain outputs or products. Outputs reflect the success of DD Network component efforts to improve the capacity of service providers to more effectively serve people with developmental disabilities. However, they say little about the ultimate goal of improving the employment, housing, transportation, health, and other outcomes of those with developmental disabilities and their families. Therefore, outcomes are an important aspect of the Open System Model. Outcomes represent the overarching goals of the DD Act. If the proposed pathway is correct, then the outputs become the inputs needed to produce the outcomes expected by the ADD-funded grant programs. There are two types of outcomes – short-term (or intermediate) and long-term. Two examples of long-term outcomes are reflected in employment rates of people with developmental disabilities and their earning levels. An example of short-term outcomes consists of changes in legislation, policy, or community practice as a result of systemic advocacy efforts. The DDPIE primarily examines short-term outcomes because they can be linked more directly to the efforts of the DD Network programs.

The DDPIE contains all four types of indicators – inputs, processes, outputs, and outcomes. Outcome indicators are the most difficult to measure. However, they will be taken seriously by ADD. Inputs, processes, and outputs are also important to measure in order to examine some of the barriers to achieving outcomes.

In addition to the two models described above, the DDPIE was guided by the professional evaluation standards promulgated by the American Evaluation Association. We do not consider these standards to be outmoded. These

standards are categorized as utility, feasibility, propriety, and accuracy. Utility standards ensure that an evaluation is useful and meets the information needs of the intended users. Feasibility standards foster a realistic, prudent, diplomatic, and frugal evaluation. Proprietary standards ensure that an evaluation is conducted legally and ethically and that there is proper regard for the welfare of those involved in the evaluation and those affected by evaluation results. There are 12 accuracy standards that relate to the evaluation reporting and conveying technically adequate information about the positive aspects of the program being evaluated.

C-6 The DDPIE will not collect reliable and valid data.

Response:

The most important aspect of the pilot study was to determine whether reliable and valid data could be collected to measure the indicators that had been developed. The evaluator designed data collection instruments to be consistent with the draft versions of benchmarks, indicators, and examples of performance standards and examined every benchmark, indicator, and performance standard to determine whether, and to what extent, the items needed to be revised or eliminated, based on the pilot study.

Analysis of program visits and interviews consisted of extracting the findings for each indicator from transcripts and incorporating findings into a spreadsheet. The evaluator examined findings for each indicator in each key function in each DD Network Program and for collaboration. Using pre-determined criteria, the evaluator revised the benchmarks, indicators, and performance standards and made them ready for distribution to validation panels. Based on input before the pilot study, the evaluator greatly reduced the number of questions on collaboration and asked the collaboration questions of all 10 pilot sturdy programs. Descriptions of collaboration and activities practiced varied.

C-7 There is concern about what the DDPIE data collection measures.

Response:

A driving force behind DDPIE Phase 1 was a need to develop indicators of quality, instead of those that measure compliance. DDPIE started with hundreds of input, process, output, and outcome indicators in an effort to identify those indictors that contribute to high program quality and achievement of impact. The number of indicators was reduced dramatically based on feedback from programs on each iteration of indicators. Then, to ensure that the DDPIE was measuring what it should measure, a pilot study was implemented, and validation panels reviewed and provided feedback on the materials. Throughout the validation process, validation panelists were reminded to ask this key question:

Is this indicator important to look at to determine the impact of DD Councils/P&As/UCEDDs on people with developmental disabilities, family members, state systems, and/or service providers?

The validation panels were made up of persons with a developmental disability, family members, self-advocates, and others. The others were individuals who were familiar with research and policy, had an understanding of consumer needs and purposes of the programs; had appreciation for outcomes; were directly involved in the DD Network system; or had a proven track record of self-advocacy (e.g., DD Council members; self-advocates outside the programs). There was a mix of urban and rural representation (with some thought to geographic representation) and a mix of senior and junior program staff.

Overall, the ratings of indicators by DD Network program stakeholders on the validation panels demonstrated widespread agreement on the importance of a large number of indicators, with or without changes to wording. All DD Council key function packages contained a total of 44 indicators. There were 10 DD Council raters (one individual had to leave after the first day, so her ratings were not included), for a total of 440 ratings on DD Council indicators. Among the 440 ratings given by DD Council validation panel members, 308 (70.0 percent) were rated as a 1 (important to include); 13

(3.0 percent) were rated as a 2 (not important; do not include); and 108 (24.5 percent) were rated as a 3 (important to include if re-worded). The DD Council validation panel left 11 ratings blank (2.5 percent). There was no clustering of number 2 ratings for each indicator.

For the P&As, there were a total of 45 indicators and seven raters, for a total of 315 ratings. Among those 315 ratings, 230 (73.0 percent) were rated as a 1; 2 (0.6 percent) were rated as a 2; and 71 (22.5 percent) were rated as a 3. Ten (3.2 percent) were left blank. No clustering of number 2 ratings was found for any indicator.

The UCEDD validation panel had 48 indicators and seven raters,[3] for a total of 336 ratings. There were 139 ratings of 1 (41.4 percent), 27 ratings of 2 (8.0 percent), and 150 ratings of 3 (44.6 percent). Twenty (6.0 percent) were left blank.

Finally, the full group of 24 raters rated five collaboration indicators, for a total of 120 ratings on collaboration. Out of 120 ratings, 75 (or 62.5 percent) were rated as a 1; 13 (10.8 percent) were rated as a 2, and 27 (22.5 percent) were rated as a 3. Five collaboration indicators out of 120 were left blank (4.2 percent).

Performance standards will be developed as part of the evaluation tools during DDPIE Phase 2 of the independent evaluation. Performance standards are the benchmarks against which ADD can gauge performance. Without them, interpretation of national data collected on the programs can vary. One person's opinion on what is acceptable for DD Network programs to achieve will be different from another's.

ADD will proceed with the development of the performance standards in Phase 2 using the following steps: 1) Analyze the data collected as part of the full-scale evaluation, 2) Develop draft performance standards based on the range of responses, 3) Establish a validation panel, 4) Implement a rating

---

[3] The group began with eight raters. One person asked that her ratings not be included at the end of the meeting because of her concerns about the evaluation.

process similar to the one utilized with the Phase 1 validation panel, 5) Synthesize data, and 6) Finalize performance standards.

C-8. ***The data collection approach is prescriptive.***

Response:

In the development of benchmarks, indicators, and the data collection tools to measure the indicators, every attempt was made to recognize the uniqueness of program grantees and the particular needs of the jurisdictions they serve. Thus, much of the data collection effort is open-ended and qualitative. Of particular note are the outcome indicators in which respondents are free to select the particular programs and activities they wish to describe.

C-9: ***DDPIE should be delayed until the DD Act is reauthorized.***

Response:

We do not know when the DD Act will be reauthorized or what changes, if any, to anticipate. Even if there are changes to the DD Act, it will be at least two years, if not more, to know the impact of any new or revised requirements for the programs. As noted above, to delay the independent evaluation would be doing a disservice to taxpayers who deserve a periodic accountability of the programs they fund.

C-10. ***Measurement of collaborative endeavors with other programs should reflect outcomes of collaboration rather than mandating a specific collaboration process that might or might not be appropriate for all states and territories.***

Response:

The focus of the DDPIE is on the impact of the three DD Network programs (DD Councils, P&As, and UCEDDs), as well as collaboration among the three programs. Thus, the evaluator developed a set of benchmarks, indicators,

and examples of performance standards, with accompanying data collection instruments to measure indicators of collaboration. The collaboration document contains input, process, output, and outcome indicators. They do not prescribe a specific process for collaboration. The following are the benchmarks and indicators identified for collaboration, as validated by the validation panel.

<div style="border: 1px solid black; padding: 10px;">

COLLABORATION

BENCHMARKS AND INDICATORS

1.   DD Network programs have a process for identifying, planning, and implementing collaborative efforts.

1.1   DD Network programs identify and document common goals on which to collaborate.

1.2   DD Network programs support and encourage collaborative efforts.

2.   DDN programs collaborate to achieve common goals.

2.1   DD Network programs communicate regularly.

2.2 DD Network programs show a united front to the community.

2.3 DD Network programs monitor progress and barriers of collaborative efforts.

2.4 DD Network programs revise goals and objectives, as appropriate, to overcome barriers identified.

3.   Collaboration achieves outcomes that have an impact on people with developmental disabilities and their families.

3.1   DD Network programs collaboratively achieve common goals set by the DD Network programs (e.g., changes in community practice, improved access to services, increase in disability leaders in the community).

</div>

C-11.   *The evaluation should recognize diversity among the programs, such as funding for minimum allotment states.*

Response:

Diversity among the programs was recognized in the pilot study and will be similarly recognized in the full-scale evaluation. Programs were selected for the pilot study on the basis of specific criteria to ensure diversity among program representation. To begin, programs were excluded if they had already participated in another aspect of the DDPIE, such as membership on

the DDPIE Advisory Panel or Working Groups, or if the met ADD exclusion criteria, such as having an upcoming MTARS visit in 2008, being a new UCEDD, or having a new or no executive director. Programs not excluded were stratified according to ADD's inclusion criteria (see table below) and randomly selected within each stratum.

Pilot study inclusion criteria

UCEDDs	DD Councils	P&As
- Has a high percentage of residents that live in rural areas - Is from each geographic region in the United States (west, Midwest, east) - Participated in MTARS in 2005 – 2007 (UCEDD, P&A, and DD Council) - Is and is not a medical school (UCEDD) - Is and is not a LEND program (UCEDD)	- Has a high percentage of residents that live in rural areas - Is from each geographic region in the United States (west, Midwest, east) - Participated in MTARS in 2005 – 2007 (UCEDD, P&A, and DD Council) - At least one program has a minimum allotment - At least one program is its own Designated State Agency (DSA)	- Has a high percentage of residents that live in rural areas - Is from each geographic region in the United States (west, Midwest, east) - Participated in MTARS in 2005 – 2007 (UCEDD, P&A, and DD Council) - At least one program has a minimum allotment - At least one program is in a state agency - At least one program uses a legal model and one uses an advocacy

In selecting states and programs for the full-scale evaluation, we will also stratify by key variables that represent the diversity of the programs – for example, allotment size, rural/urban, medical school vs. non-medical school for UCEDDs, LEND vs. non-LEND program for UCEDDs, is its own designated state agency vs. not for DD Councils, is part of a state agency vs. not for P&As. We will also pay attention to geographic distribution.

We also recognized diversity in development of the benchmarks and indicators and made every attempt to ensure that any indicators included could be measured by any program. When we received feedback on drafts that some benchmarks and indicators were discriminating against minimum allotment grantees, we revised or deleted them.

There are now combinations of both quantitative and qualitative indicators that will accommodate the fact that DD Network programs are complex and

operate differently within the requirements of the DD Act. The evaluation will collect qualitative data that will enable ADD to have an in-depth picture of these programs, their achievements, and the improvements they make to real people's lives. On the other hand, ADD will need to collect quantitative data that will enable ADD to be certain that the qualitative data obtained are illustrative of a large number of programs and not simply best case anomalies.

C-12   *The evaluation should recognize that the programs are independent entities, and they may be affiliated with other entities (e.g., as a state agency, as part of an institute of higher education), whose policies and structures may vary or be driven by grants and contracts the programs secure to support their operation and activities.*

Response:

This comment is a variation on the theme that grantees are unique and this uniqueness must be taken into account. This issue has been addressed elsewhere.

This is also an important issue to remember when developing performance standards, which will take place after data are collected on a sample of programs in 20 states.

D.   **Ways to minimize the burden of the collection of information on respondents, including use of automated collection techniques and other information technology.**

Strategies for reducing burden were described in other the responses above and include:

- This information collection process will be a one-time event during the course of the study. Programs will not have to respond annually to the information collection.
- The data are collected by an outside evaluator from the grantees and other stakeholders, mainly through interviews. As well, the data will be analyzed and reported out by the outside evaluator. The interview participants will not be responsible for ensuring the quality of the data, which is a time-consuming process. Rather, they will be responsible for taking time to be interviewed, providing responses to the interview questions, completing a self-administered questionnaire, and helping with organizing interviews. If the programs were to collect and analyze the data directly, there

would be a tremendous increase in the burden of this information collection activity.

- Rather than having all programs provide data for this proposed information collection activity, there will be a sample of programs that participate in the independent evaluation. This sample will be up to 20 States. To reduce burden for the programs that participate, ADD will not conduct the MTARS in the same year in the States that are participating in the independent evaluation, thereby reducing the burden for these programs.

- The DDPIE will incorporate web-based data entry for the self-administered questionnaire as was done in the PAIMI evaluation.