Research on Social Work Practice

Issues in Differential Response

Ronald C. Hughes, Judith S. Rycus, Stacey M. Saunders-Adams, Laura K. Hughes and Kelli N. Hughes

The online version of this article can be found at:

Additional services and information for ***Research on Social Work Practice*** can be found at:

Email Alerts: http://rsw.sagepub.com/cgi/alerts

Subscriptions: http://rsw.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> OnlineFirst Version of Record - Jan 9, 2013

What is This?

Issues in Differential Response

Ronald C. Hughes[1], Judith S. Rycus[1], Stacey M. Saunders-Adams[1], Laura K. Hughes[1], and Kelli N. Hughes[1]

Abstract

Differential response (DR), also referred to as alternative response (AR), family assessment response (FAR), or multiple track response, was developed to incorporate family-centered, strengths-based practices into child protective services (CPS), primarily by diverting lower risk families into an assessment track rather than requiring the traditional CPS investigation. Since the program's inception, researchers have conducted several large evaluation studies of DR programs, and a large body of research and program literature has been published touting the success and benefits of DR. In response to significant concerns about the clarity and consistency of DR program models and the validity and generalizability of its associated research, the authors undertook a comprehensive, three-pronged evaluation to provide guidance in shaping the ongoing development of DR programs. This report summarizes the methodology, findings, conclusions, and recommendations from the review.

Part I: Introduction

Differential response (DR), also referred to as alternative response (AR), family assessment response (FAR), or multiple track response, was developed concurrently with other systemic reforms to incorporate family-centered, strengths-based practices into child protective services (CPS). The shared goal of these reforms was to enhance the ability of families to safely care for their own children while limiting, as much as possible, an agency's use of mandated protective authority.

One impetus for these reforms was the concern that CPS agencies too often relied on interventions described as inflexible, adversarial, and unnecessarily intrusive in their work to protect abused and neglected children (Kaplan & Merkel-Holguin, 2008; Schene, 2001). In 1974, the federal government enacted the Child Abuse Prevention and Treatment Act (CAPTA), which set the tone for more frequent use of intrusive intervention in CPS. CAPTA required states to implement specific procedures to receive and respond to allegations of child abuse or neglect to ensure children's safety. States were required to adopt definitions of child maltreatment that were congruent with the CAPTA definition, ". . . at a minimum, any recent act or failure to act on the part of a parent or caretaker which results in death, serious physical or emotional harm, sexual abuse or exploitation, or an act or failure to act which presents an imminent risk of serious harm" (Stoltzfus, 2009, p. 3).

Historically, child protection agencies removed many at-risk children from their families and placed them into foster care as a primary means of ensuring their safety (McGowan, 2005). Another impetus to use out-of-home placement as a primary safety intervention was that federal funding supported costs for foster care placement but not for in-home family services. Subsequently, in the late 1970s, research determined that foster care placements were often unstable and were inherently traumatic for many children (Fanshel & Shinn, 1978; Jones, 1978). Further, children in foster care often experienced multiple placement changes, and many would emancipate with few independent living skills, nowhere to go for assistance and support, and no permanent family connections (Courtney et al., 2007; Kushel, Yen, & Courtney, 2007). This situation—referred to in the child welfare literature as foster care drift—spawned a national permanency planning movement (Child Welfare Information Gateway-a; Hartley, 1984; Pike, 1976; Ward, Mallucio, Hamilton, & Fein, 1982) that sought to elevate permanence for maltreated children to a level of importance equal to that of child safety.

Passage of the Adoption Assistance and Child Welfare Act of 1980 (PL 96-272) was a strong impetus for large-scale system reform to achieve permanence for children in the child welfare system. The legislation required states to make *reasonable efforts* to prevent the removal of children from their families and to quickly reunify children in placement with their own immediate or extended families or permanent alternative

[1] North American Resource Center for Child Welfare, Institute for Human Services, Columbus, OH, USA

Corresponding Author:
Judith S. Rycus, North American Resource Center for Child Welfare, Institute for Human Services, 1706 East Broad Street, Columbus, OH 43203, USA
Email: jsrycus@aol.com

families (PL 96-272). New service models were developed to preserve, strengthen, and empower families to enable them to safely care for their own children (O'Neill Murray & Gesriech, 2004). These service programs required high levels of collaboration between child welfare agencies and community-based service providers, and they promoted intervention methods based on engagement, empowerment, and partnership with families. In the early 1990s, Missouri enacted legislation to make its CPS system more family-centered and responsive to family needs, thereby becoming a front-runner in the evolution of family-centered reform efforts (Key Informants; Waldfogel, 1998). Interventions such as family preservation (Chaffin, Bonner, & Hill, 2001; Hutchison, Lloyd, Landsman, Nelson, & Bryce, 1983; Yoo & Meezan, 2001); intensive home-based services Frankel (1988), wraparound services (Behar, 1986; Rosenblatt, 1996; VanDenBerg & Grealish, 1996), family-centered, neighborhood-based (FCNB) services (Armstrong, 1997; Child Welfare Information Gateway-c), models of partnering with families (Child Welfare Information Gateway-b), and solution-focused interventions (Berg, 1994) were developed, piloted, and promoted throughout the country. In 1997, the federal government enacted the Adoption and Safe Families Act (ASFA), which reaffirmed and expanded programs to preserve and support families but which also clarified the reasonable efforts provisions of PL 96-272 to ensure that children's health and safety would be the foremost criteria when making decisions to remove or reunify children (P.L. 105-89). The refocus on child safety was strengthened when in 2000, safety was included as a primary outcome indicator in the Child and Family Service Reviews (CFSR), the federal government's process for monitoring the performance of state child welfare systems. Fiscal sanctions could be enacted for states that failed to meet CFSR performance standards. Although the legal focus on permanence for children did not change, renewed emphasis on child safety bred caution in some states and organizations about family-centered approaches that could, if not judiciously applied, de-emphasize the use of mandated protective authority in the interest of family engagement and empowerment. The confluence in both law and policy of these compelling practice principles reflected one of the most prevalent and difficult inherent tensions in CPS practice: that is, how to promote engagement, empowerment, and support of families without compromising the essential need and capacity to use more authoritative interventions to protect children at high risk of serious harm when less authoritative methods could not.

Most child welfare professionals welcomed attempts at systemic reform to promote both safety and permanence for maltreated children. Intervention models that employed "family-friendly," strengths-based, and collaborative approaches exemplified fundamental social work principles of family empowerment, self-determination, and partnership. Emerging practices relied on a consistent and collaborative family–caseworker relationship in which families were active participants in assessing their own strengths, needs, and

problems and in identifying relevant service resources. Throughout the 1990s, many states incorporated family-centered practice principles into their revised child welfare practice models. As a result of seminal work by Ohio's state and county public child welfare agencies, these principles were also articulated in knowledge and skill competencies and training curricula that formed the foundation of child welfare specializations in professional social work education and competency-based in-service training systems throughout North America (Institute for Human Services [IHS], 1994; IHS, 1990. Over nearly two decades, tens of thousands of child welfare caseworkers, their supervisors, agency managers, and university social work students were trained to utilize family-centered and strengths-based approaches in their work with families. However, even with widespread and intensive education and training, these preferred practices were not always systematically or effectively implemented by the child welfare organizations that espoused them. A 2005 literature review on implementation science by Fixsen and colleagues delineated the inherent challenges and impediments to implementing and sustaining *any* new program or practice in large, complex, bureaucratic organizations, even with organizational support (University of South Florida, 2005). The barriers they identified, including limited financial and staffing resources, legal barriers, a lack of leadership, and organizational inertia, were equally applicable impediments to implementation of family-centered practices in child welfare agencies.

In this challenging practice environment, the promise of DR was its capacity to strengthen, support, and empower families without compromising children's safety. DR proposed to serve families with different needs in different service tracks, which would enable child protection agencies to intervene authoritatively with families when it was most needed and to use less intrusive interventions with everyone else. DR advocates leveraged the experiences of pioneer states such as Missouri and Minnesota to give credence to the DR approach and to drive legal, organizational, policy, and practice changes in state governments and local child welfare agencies to support DR program implementation. The eventual development of a national advocacy team and access to significant federal and foundation resources to support the initiative together made DR one of the more widely replicated child welfare reform efforts in recent history.

Throughout the almost 20 years since the idea of DR was first conceived, the IHS/North American Resource Center for Child Welfare (IHS/NARCCW) has promoted, supported, and provided training on principles of strengths-based, family-centered practice in all facets of child welfare. The IHS/NARCCW developed and delivered extensive in-service training in many of the states, Canadian provinces, and agencies that were implementing DR programs. Many of these systems revised their formal child welfare practice models to underpin and support the reforms necessary to promote these fundamental social work practices. The IHS/NARCCW also worked throughout North America to establish competency-based in-service training systems as an essential

infrastructure to institutionalize and sustain family-centered practices in the child welfare workforce.

Over time, we developed significant concerns about many of the ongoing DR implementation and evaluation efforts. Our concerns were related to a lack of clarity and consistency in the DR program model as well as questionable validity of its evaluative research and subsequent claims about DR's outcomes and effectiveness. We were also concerned about the possibility of potentially serious unintended consequences of DR programming related to children's safety, particularly in jurisdictions that had abandoned fact-finding, risk assessment, and ongoing safety planning with families in alternative tracks in an effort to remain "family friendly."

In early 2010, we initiated a comprehensive research and literature review to address some of these concerns. Our goal was to identify the strengths and limitations of the DR model as it was being promoted and implemented and to provide sound, objective information to further DR's positive reform efforts. This report summarizes our methodology, findings, conclusions, and recommendations.

Issues and Questions

The research described in this document was undertaken to answer several questions about DR, specifically:

- Is DR a clearly articulated, standardized, replicable model of child welfare practice with formal written policy and program guidelines to guide users in implementation?
- What are the primary characteristics and practice requirements of this model as it is being promoted and implemented?
- Has DR programming been implemented with consistency and fidelity across pilot and implementation sites?
- What evaluation research has been conducted on DR programs, of what type, and what is its level of rigor and quality? What can be validly concluded from this research about DR program effectiveness and outcomes?
- Is there reliable and valid data about the safety of children served in alternative tracks?

We gathered data from multiple sources, using a variety of data collection strategies, and we triangulated the data to derive our conclusions. We also developed internal and external strategies to ensure objectivity in our data collection, analysis, and documentation. To ensure the accuracy of our work, we asked five highly regarded child welfare researchers from outside our organization to critically assess the final draft document. We also produced individual summaries of each research study, which are included Appendix A of this document.

Methodology

To implement this study, we collected and assessed data from three primary sources: (1) published articles and program descriptions outlining the philosophy, concepts, and practice principles underlying DR programming; (2) formal research and program evaluations that had been conducted in DR program implementation sites; and (3) interviews with key informants from states that had current DR programs or that had previously implemented DR programs and discontinued them.

Literature and Research Assessment

The research studies and published literature included in this review were identified by searching several online databases, including EBSCOHOST (accessed through OhioLink, Ohio's university library search interface); Google Scholar; the Cochrane Collaboration Registry; and resources provided by the American Humane Association (AHA); and the National Quality Improvement Center on Differential Response in Child Protective Services (NQIC-DR). The search terms we used to identify relevant sources of conceptual and practice literature were "Alternative Response" OR "Differential Response" OR "Multiple Response" AND "Child Abuse" OR "Child Neglect" OR "Child Maltreatment."

The specific database, search terms, and search strategies used to locate research and evaluation reports differed for each database and were based on the requirements of each search utility. Search terms used to access research and evaluation reports are outlined in Table 1.

We included both published and unpublished research reports in our review, and we obtained original evaluation reports produced by the researchers who conducted the studies rather than relying on other authors' interpretations of study results.

We used a snowball method to identify additional resources from bibliographic citations in procured documents. We reviewed DR-related newsletters, conference brochures and presentations, training curricula, workshop handouts, resources listed on the NQIC-DR website, and special issues of the AHA's publication, *Protecting Children*, that were devoted to DR-related articles. We identified additional research reports by asking key informants during telephone interviews to provide or direct us to reports from research that had been conducted in their state or jurisdiction. A total of 39 program, practice, and conceptual publications and 18 research reports were ultimately included in our review.

Key Informant Interviews
Sample

Our key informant sample included all the states we could identify that either operated a DR program at the time of sampling or had previously operated a DR program but had discontinued it. These states were identified through federal and state documents and websites, reports generated by AHA, and from the attendance roster from the National Differential Response Conference held in 2008 in Columbus, OH. Once the states were identified, we attempted to contact the most appropriate respondents, many of whom were identified either in

Table 1. Databases and Corresponding Search Terms

Database	SearchTerms
EBSCOHOST (includes the following): Academic Search Complete Medline Psychology and Behavioral Sciences Collection Social Work Abstracts SocINDEX Sociological Collection Social Sciences Abstracts	((Differential OR Multiple OR Alternative OR (Family Assessment)) AND Response) AND (child AND ((protective service) OR welfare))
Cochrane Register	(child protection) AND [(Alternative Response) OR (Differential Response) OR (Family Assessment Response)] AND (abuse OR neglect OR maltreatment) in Title, Abstract or Keywords
Google Scholar- Social Sciences	child AND (Differential Response) AND ((protective OR welfare)AND services) child AND (Alternative Response) AND ((protective OR welfare)AND services) child AND (Family Assessment Response) AND ((protective OR welfare)AND services)

AHA reports or on the 2008 conference roster. Ultimately, 35 states were identified as having some involvement with DR programs. We included respondents from state agencies and/or from local implementing agencies to ensure a variety of perspectives on both program design and implementation. We attempted to contact potential respondents by telephone and e-mail. If a person of interest was no longer available, we requested contact information for the most knowledgeable person/persons to talk with about their DR program, and we were usually directed to an appropriate respondent. We also identified some of the respondents by e-mailing the state office responsible for child welfare services at their general inquiry address. Of the original 35 selected jurisdictions, 7 did not respond to our queries, and 8 were considering DR but had not yet implemented it. Our final key informant sample included respondents from 20 states, all of whom completed interviews.

Interview Protocol

Key informant interviews were conducted using a semistructured interview format guided by a standardized questionnaire with 47 primarily open-ended questions. The broad categories for inquiry included program description, track assignment, staffing, staff training, substantiation and documentation, case process, program evaluation, implementation issues and challenges, and general feedback. A copy of the interview protocol is included in Appendix C.

Interview Process

We collected interview data between August 2010 and January 2011. The interviewers explained and clarified the purpose of the study and the intended use of the data. Key informants were told that the data would be used for a research project to ultimately guide child welfare practice and to enhance training as it relates to DR. Participants were informed that the interview would take between 30 and 60 min, and they were

given the option to schedule their interview at a convenient time. Some of the respondents asked whether they could include colleagues in their interview to allow for a variety of perspectives, and we scheduled teleconferences to accommodate these small group interviews. In three instances, key informants asked whether they could review the topics and questions prior to the interview, and we provided this information. We began each interview with the same questions, but as the interviews progressed, the responses of the informants guided the order in which we asked the remaining questions. We encouraged participants to elaborate on their answers to elicit more complete information. Interviews were from 45 min to an hour in length and in all but one case were conducted by telephone. One interview was conducted in person. Some respondents asked to review the questionnaire following the interview to ensure that we had accurately recorded their data. We e-mailed completed questionnaires to these respondents with instructions to provide us with corrections as needed. These corrections were incorporated into the data.

We elected to maintain the confidentiality of all key informants, even though when asked, most of them did not feel that confidentiality was a concern. The assurance of confidentiality was intended to encourage key informants to provide honest, detailed, and accurate information. We did not evaluate any state individually, nor did we discuss any information about client families, individual case circumstances, or the work of individual staff.

Validity of Data

The project staff used several strategies to ensure the accuracy of key informant data and to limit subjectivity. During telephone interviews, the interviewers used low inference (concrete and unambiguous) descriptors and captured direct quotes from respondents whenever possible. Data gleaned from the interviews was qualitative in nature and was compiled using only thematic analysis. We did not attempt any quantification of interview responses.

Data Analysis and Conclusions

The research method was designed to triangulate data from multiple sources, thereby increasing the accuracy of the information and findings. The research reports were reviewed independently by three experienced child welfare researchers, two of whom were external to IHS/NARCCW. The final written report was reviewed by all members of the research team and by selected external reviewers to further ensure its accuracy. Names and affiliations of collaborating research specialists/reviewers are listed in the authors' notes at the end of this article.

Part II: Findings

Finding #1: DR Programs do not Adhere to a Uniform, Standardized Practice Model, nor are Programs Implemented Consistently Across Sites

Our research determined that while underlying principles of DR have been widely articulated in the DR literature, there is considerable variation in how DR programs have been defined, designed, and implemented across states and agencies and over time. Many of these differences were also documented in the *National Study on Differential Response in Child Welfare*, a survey of 27 states and 2 counties completed in 2006 by the Child Welfare League of America and the AHA (Merkel-Holguin, Kaplan, & Kwak, 2006). In 2008, Kaplan and Merkel-Holguin reported that states were still using "different terminology, definitions, models, services, and service providers, all while shaping the evolution of the approach" (p. 6). They further contended that a clear understanding of the DR approach has been challenged by the lack of common terminology and definitions.

We identified the following differences in how DR programs had been designed and implemented across jurisdictions.

Implementation sites used different numbers of tracks designed to serve different types of families. The DR program literature specifies that DR was originally designed to provide an alternative intervention for child abuse and neglect reports that had been *screened in* or *accepted* by CPS (Merkel-Holguin, 2005; Rohm, 2008; Schene, 2005). Most DR program sites had implemented such a two-track system that included an investigation track (often referred to as TR, or the traditional CPS track) and an assessment track (often referred to as AR, or the alternative track). However, some DR programs operated more than two tracks. Some had created tracks for families who did not meet the criteria for CPS and had been screened out of the CPS system (Kaplan & Merkel-Holguin, 2008), ostensibly because these families appeared to need preventive and supportive family services. Several sites had added a primary prevention track or a track in which families were referred to community providers to provide both ongoing services and case monitoring (Carpenter, 2007; Child Welfare Information Gateway, 2008; Gilmore, 2010; NQIC-DR, 2009).

There were differences across DR implementation sites in how families were assigned to tracks. There were no uniform criteria used across implementation sites to make track assignments, and DR sites varied in their determination of which families belonged to which track. One often-cited criterion for making track assignments was the level of potential risk identified by the screener, although it was often unclear how risk had been either defined or measured. Other criteria were state legal requirements (NQIC-DR, 2009), the age of the alleged child victim (Conley, 2007), whether the referred family appeared willing to participate in services (Gilmore, 2010), whether the referral alleged serious or imminent harm to a child (Schene & Kaplan, 2007), the presence and number of previous referrals on the family or the prior victimization of a child (Child Welfare Information Gateway, 2008; Gilmore, 2010; NQIC-DR, 2009), the source of the referral (reports from mandated reporters were more likely to be assigned to the investigation track; Gilmore, 2010), or the type of alleged maltreatment (Gilmore, 2010). For example, one key informant noted that alleged neglect cases were tracked to the alternative track (AR), while all alleged abuse cases were tracked into investigation (TR). One state required children under age 5 to be referred to the investigation track (Conley, 2007), and most states required all cases of alleged sexual abuse, serious physical abuse, serious neglect, or child fatalities to be assigned to the investigation track (Kaplan & Merkel-Holguin, 2008; NQIC-DR, 2009). States appeared to be fairly evenly divided in their track assignments of cases of child abandonment, medical neglect, drug exposed infants, and serious mental injury (Kaplan & Merkel-Holguin, 2008).

In most cases, the track assignment decision appeared to be made based on the judgment of an individual screener or caseworker, often with supervisory approval, although in some states it was recommended that a team of professionals make the track decisions (Sawyer & Lohrbach, 2005). The literature suggests that states based their track assignments on the assessed level of risk, but only 6 of our 16 key informants in states operating DR programs reported consistent use of a structured decision tree or other standardized risk or safety assessment tool to guide these decisions (Key Informants).

There was considerable variation in the percentage of cases assigned to the alternative track in different jurisdictions. One research study identified that the percentage of families assigned to the alternative track in different counties ranged from 19% to 70%, suggesting different criteria and patterns of track assignment even within a single state (Institute of Applied Research, 2010). Other states exhibited differences in the percentages of families assigned to the alternative track, with 47% in Minnesota (Institute of Applied Research, 2004b), 61% in Virginia (Commonwealth of Virginia, 2008), and 71% in Missouri (Institute of Applied Research, 2000).

There were differences among jurisdictions as to when the track assignment decisions were made. Most agencies made track assignments at the time of screening the initial referral (Institute of Applied Research, 1997, 2004a, 2004b, 2010; Key Informants). Some jurisdictions waited until an initial contact had been made with a referred family or until after a home visit and initial safety assessment had been completed (Key Informants). Still others made track assignments only after a more thorough safety and risk assessment had been completed (Key Informants), and one state reportedly made the track assignments only after an "investigation" of all cases (Key Informants).

There was inconsistent use of assessment protocols or tools in the alternative track either to evaluate child safety or to inform case planning and service delivery. Some jurisdictions reported that they mandated the use of standardized safety, risk, and family assessment protocols such as Structured Decision Making™ (Children's Research Center, 2005) or similar decision-making frameworks to guide safety, risk, and family assessment activities (Institute of Applied Research, 2010, 2004b). However, many jurisdictions lacked standardized tools or assessment criteria, and both data collection and data interpretation were often based on the individual judgment of screeners/caseworkers, some with input from their supervisors (Key Informants). Kirk (2008) concluded that while the DR literature stressed the "central importance of comprehensive family assessment as an integral component of differential response . . . information is scant with respect to how to conduct those assessments or case practice tools to support it" (p. 73).

There were inconsistencies in how services were delivered to families. In some jurisdictions, families in the assessment track were served by CPS agency caseworkers (Institute of Applied Research, 1997, 2004b, 2010; Key Informants). In other jurisdictions, the CPS agency contracted with community agencies to both assess family needs and provide or coordinate follow-up services (Key Informants). In a few states with multiple tracks, some families were referred for follow-up to community agencies without any further CPS agency involvement (Conley & Duerr Berrick, 2010; Key Informants).

Discussion

Based on our assessment, DR is a compilation of practice philosophy and inconsistently defined practice principles rather than a clearly delineated program intervention model. This appears to have been at least partially intentional. Kaplan and Merkel-Holguin (2008) contend that because of the complexity inherent in any assignment schema, and given the varying circumstances of individual families and children, flexibility in implementation is one of the DR program's strengths. They further contend that clinical judgment and discretion of individual caseworkers are of great importance, as there are "*few hard and fast rules* that cannot be altered given the

practice wisdom of a specific worker and the approval of a supervisor" (p. 11).

Arguments in support of flexibility in program implementation and service delivery should not be misconstrued as a valid rationale to compromise consistency in the design, implementation, and evaluation of a program model. The success of any program model depends on both theoretical and applied consistency in its definitions, structural design, implementation strategies, and performance standards. In our study, we found very little consistency in how DR reform efforts were undertaken across jurisdictions. There was no consistency in definitions or program components; no clearly defined standards; no established procedures; and wide variances in implementation strategies—in other words, no consistent DR model. Each implementing jurisdiction essentially began its program development without benefit of a standardized program manual to guide implementation and to promote fidelity in implementation. Therefore, implementing jurisdictions often differed in significant ways from previously implemented reform efforts as well as among and within programs or agencies in a single jurisdiction. These differences were not recognized as a fatal threat to effective model development. We did not find support in the DR literature that the model was shaped by empirical science and built incrementally based on the experiences of prior pilot sites with the capacity to test and standardize processes that best achieved desired outcomes. Without a consistent program model which is implemented with fidelity across jurisdictions and comparably evaluated, it is impossible to draw valid conclusions about the effectiveness, benefits and limitations of DR.

Finding #2: Methodological Problems in the DR Research Limit Confidence in Research Findings and Conclusions

Eighteen research reports of DR programs were identified for inclusion in our analysis. Two of these studies used exclusively qualitative methods. Of the 16 research studies that included at least some quantitative analysis, 2 may be classified as observational, 3 as experimental (1 was a follow-up study), and 11 as quasi-experimental. In our review, we identified significant problems in research methodology and implementation that presented threats to internal, external, and construct validity, thereby calling into question the reliability and accuracy of many of the claims and conclusions made in these studies.

Internal validity refers to a study's ability to demonstrate that a variable or phenomenon being evaluated is really responsible for the study findings and outcomes, rather than some other unidentified confounding variable. We identified threats to internal validity in all of the studies in this body of research. The scope and scale of these threats to internal validity varied among the individual studies. In several of the studies, DR implementation had been paired in time with some other large-scale policy or practice change. One example was the entire system reform that occurred in North Carolina. In this case, DR implementation occurred simultaneously with a new

practice model, the introduction of new interventions such as family team meetings, shared parenting meetings in placement cases, a redesign of in-home services, and a more strengths-based structured intake decision-making process (Comer & Vassar, 2008). With changes occurring concurrently and none being effectively isolated in the research, it is impossible to know which change strategies or interventions may have influenced the study outcomes. Further, the definition of the "DR intervention" actually included a variety of interventions that differed among jurisdictions, including screening decisions, track assignment, initial caseworker contact, safety assessment, safety planning, risk assessment, comprehensive family assessment, identification of service needs, service referral, and service delivery, none of which were the same among all the studies we reviewed. The research did not isolate these processes and evaluate them independently, so a claim that DR was responsible for observed outcomes fails to clarify exactly which elements, or groups of elements, had which specific influences, if any, on which observed outcomes.

Threats to external validity were present in all of the studies in our review. External validity refers to the degree to which study findings can be generalized to populations other than the study sample. The absence of uniformity and consistency in programming across DR implementation sites would, itself, be sufficient to caution that no study's outcomes can be generalized to other sites. Further, none of the studies in our analysis used random sampling procedures to select the study sample. The two experimental studies and one follow-up study did not use random sampling to select the study sample, but they did randomly assign members of the initial study sample to the two comparison groups (Institute of Applied Research, 2004b, 2006, 2010). Because of pervasive threats to external validity in all of these studies, we cannot with confidence assert that the study samples were representative of the entire population of families served in the child welfare system, thereby preventing generalization of these studies' findings to other program implementation sites.

Construct validity refers to clarity and consistency in the definitions and measures of both the interventions and the outcomes being evaluated. The primary factor affecting construct validity in much of the DR research was that DR and its components were not clearly or consistently defined or implemented, making it impossible to isolate exactly what factors were included in the variables being measured. As indicated earlier, the implementation sites studied by this body of research had considerable latitude in deciding how to define and implement the components of their DR programs. As a result, key programmatic and measurement constructs such as "child safety," the "assessment track," "engagement," "traditional track," and "family satisfaction," differed both within and between studies. Therefore, we cannot be certain that the measures used in these studies accurately represented the variables being researched, nor can we assume that a study's findings would be relevant to other jurisdictions or useful in shaping ongoing and future reform efforts.

While some of the studies controlled for differences between the experimental and control groups, most did not. For instance, some of the studied jurisdictions allowed caseworkers in their agencies to choose whether to work in the traditional or the alternative track. As a result of self-selection, the two groups of caseworkers were likely not equivalent, potentially resulting in unequal treatment of the families in the AR and TR tracks for reasons other than DR programming. For example, individuals who chose to work in DR may have had qualities or characteristics, education, or work histories that affected their job performance and its subsequent effects on families. DR workers may have been more or less experienced, more or less committed to the child protection field, perhaps had different skills or had received special training, or may have had differing personal values about families in the child welfare system and their capacity to change. Any of these variables could have influenced outcomes and confounded the research. However, these individual and group differences were neither clearly identified nor isolated and measured, making it unclear whether DR program components or qualities of the caseworkers—or even some other factor—was associated with the observed outcomes, whether positive or negative.

An additional concern is related to how initial track decisions were made. In 13 of the 18 studies we reviewed, screeners had attempted to separate high-risk families from families at low- or moderate-risk in order to make track assignments. The initial assignment of families to tracks was generally based on screener recommendations, but as noted earlier, these decisions were often made without the use of standardized tools or criteria, raising question about the equivalence of the families within each of the tracks and, therefore, of the two research comparison groups. It is entirely possible that the comparison groups exhibited potentially significant differences from the onset of the research, including different risk potentials and different propensities to recidivate, both of which were primary outcome variables measured in many of the research designs. In four of the studies (Institute of Applied Research, 2004b, 2006, 2010; Ruppel, Huang, & Haulenbeek, 2011) families screened in as eligible for AR were then randomly assigned to AR or TR tracks, thereby potentially improving equivalence between the two study groups, but only if the screening criteria were well defined, consistently applied, and valid measures of risk.

Other methodological issues threatened the equivalence of the study groups. Many jurisdictions reportedly had provided intensive training to caseworkers in the alternative track in competencies related to engagement, assessment, and service planning but did not provide this training to workers in the traditional track (Institute of Applied Research, 1997, 2004b, 2010). These competencies are fundamental to social work practice and, therefore, are essential to effective work in either track. Social workers in the TR track may have been at a disadvantage because they had not been as well trained, and there is no way to know whether they might have demonstrated increased positive outcomes if they had also received this essential training in core skills.

Many of the tools and instruments in this body of research presented potential threats to internal validity. In most of the reports we reviewed, there was no discussion of how the instruments, questionnaires, and scales used in the study had been tested to ensure either validity or reliability. In one study, some of the instruments were scaled with only positive or negative choices, forcing respondents to choose one or the other, even if they were uncertain, ambivalent, or indifferent, which potentially skewed the data toward the poles (Institute of Applied Research, 2010). Failure to validate study measurement instruments increases the potential for ambiguity and measurement error.

Discussion

Field-based research cannot be as well controlled or as rigorously implemented as research conducted in laboratory settings. A primary deterrent to achieving rigor in field-based research is the presence of variables in the study environment that cannot be identified, isolated, or measured, yet which may strongly influence both program implementation and study outcomes. To claim or imply that an intervention being evaluated is responsible for observed outcomes without fully considering the potential role of these other variables can provide a distorted and exaggerated picture of the state of our knowledge about an intervention's effectiveness.

Both researchers and consumers must use caution when interpreting research findings or using research data to inform critical policy, program, and resource decisions. It is concerning that in most of the studies we reviewed, the claims presented in the research reports frequently overreached what might have legitimately been concluded, considering their many methodological limitations.

Finding #3: There is Insufficient Data to Confirm the Safety of Children Served in Alternative Tracks

The potential to compromise children's safety remains a primary concern when a CPS program chooses to forgo investigations or other forms of targeted case fact-finding in child maltreatment cases. The DR program literature clearly contends that child safety is of prime importance in both AR and TR tracks, and that the DR research has sought to confirm the safety of children served in alternative tracks. Some of the research included in our analysis did, in fact, attempt to evaluate child safety and concluded that children served in alternative tracks were as safe as children served in traditional, investigation tracks (Center for Child and Family Policy, 2009; Institute of Applied Research, 2004b). Based on our analysis, there is insufficient data to warrant such conclusions. This does not necessarily mean that children in AR tracks are unsafe. It does mean that child safety is not being uniformly assessed, accurately measured, or fully addressed in either DR programming or research.

Our review identified several references in the DR literature that raised the issue of safety of children served in alternative

DR tracks. In one study, 37% of those surveyed believed that AR "probably" or "definitely" increased children's safety. About a third (30.5%) of surveyed community service providers expressed "some" concern about the safety of children served in alternative tracks, while 8.4% expressed "major concern" (Institute of Applied Research, 2004b). Schene (2005) suggested that "many of the families diverted from CPS . . . are not adequately assessed or served by the community, and children continue to be at risk" (p. 6), especially when the CPS agency closes its case. She also recommended that community agencies become more sensitive to protection issues so that emerging risks in the families they serve can be rapidly recognized and addressed.

Discussion

In order to ensure that children served in alternative tracks are safe, safety and risk assessments must be an ongoing casework activity that is incorporated at all stages of case assessment and service planning. Further, the tools used to assess risk of both imminent and future harm must include pretested, validated, and standardized criteria, or safety assessments may not be valid. This is particularly important in DR programs, since the level of risk drives the decision of whether to bypass an investigation and serve a family in an alternative track. Inaccurate safety assessment can potentially lead to inaccurate track assignments.

There are many factors operating in DR programming that potentially contribute to inaccurate assessments of children's safety. The first concern is whether child safety can be accurately assessed at the time of screening. Information about the family being referred is generally gathered by screeners without the benefit of direct contact with family members or alleged child victims. At times, referents may have limited or inaccurate information about the family and may not know what information is most relevant to report. A screener's assessment of children's safety must therefore be considered preliminary, even in the best of circumstances.

The screening function in child maltreatment agencies was never intended to make a clear determination about child maltreatment or risk. A screener's recommendation to accept a case is an affirmation that there is sufficient potential that a child is at risk to warrant agency follow-up, but that additional case fact-finding is necessary to confirm or rule out the presence of threats to child safety. Yet, screeners in DR programs typically make recommendations for case acceptance, establish the priority for agency response, and recommend a track assignment based on information collected in a referral telephone interview—information that is typically limited in scope and depth and potentially in accuracy.

This is exacerbated by a lack of consistency in the criteria used to make track assignments in many DR programs. Our key informants reported that rather than being directly measured, the level of risk in a family was often inferred from the type of maltreatment alleged in the referral. For example, in many jurisdictions, reports of sexual abuse, severe physical abuse,

or child fatalities were automatically tracked to investigation. How serious or severe the maltreatment needed to be to warrant an investigation was not clearly defined. Some states referred all very young child victims to the investigation track, presumably because young children are more vulnerable and, therefore, at higher potential risk of harm. However, this does not consider the possibility of protective capacities in a family that might quickly eliminate any safety threats, therefore obviating the need for an investigation. One state reportedly tracked all neglect referrals to AR and all abuse referrals to TR (Key Informants), even though at least as many children die from neglect as from abuse (AHA: Child Welfare Information Gateway, 2012). Another criteria reportedly used by some agencies for track assignment was the position or role of the person who made the referral. Shusterman, Hollinshead, Fluke, and Yuan (2005) and Yuan (2005) reported that referrals from parents, caregivers, friends, schools, or children were more likely to be assigned to the alternative track, whereas reports from community professionals were more likely to be assigned for investigation. Yuan (2005) also noted that some states may have policies against tracking families to AR if they were referred by certain mandated reporters. We know of no research that correlates the type of referent with child safety.

Adding to the challenge of accurately assessing safety at the time of screening is the apparent latitude given to staff making the track assignment decisions. Some of our key informants reported that their jurisdictions used formal safety protocols or standardized screening tools to assess children's safety at the time of intake. Others referred only to legislation, administrative rule, or general practice guidelines, none of which was designed to assess children's safety in individual cases. Research reported over a decade ago by English, Wingard, Marshall, Orme, and Orme (2000) stressed the necessity of establishing consensus in definitions of serious or imminent risk of harm, since track assignments in DR programs were generally based on these constructs, and a lack of consistency in definitions could undermine the accuracy of track decisions. The researchers' concerns originated from some serious initial allegations that had been classified as low risk and subsequently referred to the alternative track—and, in the words of the researchers, reflected "a serious disregard for the health and well-being of these children or the potential for serious harm" (English, Wingard, Marshall, Orme, & Orme, 2000, p. 387). Waldfogel (1998) similarly asserted that the key to DR program effectiveness is an agency's ability to accurately sort cases by level of risk.

Further contributing to inaccuracy and inconsistency of track decisions is the lack of training and skill of some of the personnel in screening positions. Because of the inherent complexity of making child safety determinations in the screening environment, all screeners should be skilled clinicians with well-developed interviewing and assessment skills. In some agencies, particularly those with limited staff resources, administrative assistant staff sometimes cover the screening function and gather initial referral information before transferring it to casework or supervisory staff to make decisions (Key Informants).

A more critical issue is that child safety can fluctuate, perhaps significantly and often within relatively short time frames, in response to changes in family composition or circumstances, environmental conditions, or the presence, capacities, and stability of primary caregivers. If track assignment is based solely on an assessment of a child's condition at the time of referral, children identified as currently safe may actually be at high risk of future harm, and these children may be referred to the alternative track without recognition of the longer term safety concerns. Since formal risk assessment to estimate the potential for serious future harm is not a standard feature of screening, screeners cannot always accurately determine how much protection a child may need in the near or distant future. This is because formal risk assessments—particularly the more reliable and valid actuarial risk assessments—are difficult to complete fully or accurately without face-to-face family contact and a deeper exploration of a family's circumstances, including the presence and dynamics of previous maltreatment (Baird & Wagner, 2000; Baird, Wagner, Healy, & Johnson, 1999; Rycus & Hughes, 2003).

Inaccurate assessments of children's safety and risk of future harm not only affect the accuracy of track assignments—they also affect all subsequent case decisions. This is of considerable concern in a dichotomous tracking system such as DR, where everyone expects that families determined to be at "high risk" will be tracked to investigation, and "lower risk" families will be tracked to AR. Therefore, if the information gathered by the screener does not reach the threshold necessary to trigger an investigation, the family may automatically be perceived as "lower risk," and by inference, the children thought to be "safer" than those children referred for investigation. Caseworkers may not be as vigilant in watching for and identifying emergent safety threats or a reduction in a family's protective capacities.

Other factors inherent in DR philosophy and policy may also affect children's safety. A defining characteristic of the alternative DR track is that caseworkers neither substantiate allegations of child maltreatment, nor do they confirm the perpetrator. If *substantiate* were defined solely as legally documenting an alleged maltreatment incident, and if *not confirming a perpetrator* were narrowly defined as not entering perpetrators' names into a Central Registry database, AR caseworkers could comply with these requirements and still complete an accurate fact-finding assessment of both immediate child safety and future risk. However, there are practical differences in how these core DR principles are operationalized by the agencies and caseworkers who implement them, with many claiming it unnecessary (and inherently disrespectful) to push families to talk about an alleged maltreatment incident or to determine who was responsible for its occurrence.

DR research reports and program literature provide examples of minimizing, and even discouraging exploration of specific incidents of child maltreatment in the interest of remaining family-friendly and supportive. For example:

With AR we identify strengths, we don't focus on an incident that brought the family to child protection . . . . (A caseworker, quoted in a study completed by Institute of Applied Research, 2004b, p. 49)

The allegation is not the reason for services. It is the reason we are coming to the home, but it's more about what we can do to help . . . . Sometimes I will tell them, ''We don't really have to talk about that [the allegation]—What can I do to help you?'' (A caseworker, quoted in a study completed by Institute of Applied Research, 2009, p. 34)

Researchers Loman and Siegel (Institute of Applied Research, 2004b) voice the importance of child safety in their report, but they downplay the need to fully explore and assess prior incidents of child maltreatment for both case planning and longer term risk reduction. They state

It [AR] allowed families ''to take a good look at their family dynamics and how it operates as a system'' and focuses on ''looking forward with the family rather than looking back at an incident.'' (p. 82)

The key informants in our study reported that case documentation was normally based on the needs of the family and, in many cases documenting the details and circumstances of maltreatment was not required. It was a fundamental understanding of these caseworkers that inquiring about the dynamics of a particular maltreatment incident, determining who was responsible, and identifying the child victim/victims was not considered a routine part of information gathering or fact-finding, and in many instances gathering this information was overtly discouraged.

We also found many claims related to children's safety in the DR research, including the assertion that children served in the AR track were as safe as or safer than children served in the TR track (Institute of Applied Research, 2004b; Schene, 2008) The DR research has attempted to measure child safety in a variety of ways, but none of the studies used standardized and validated measures or safety assessment protocols implemented by well-trained raters.

In this body of research, one often-used measure of child safety was recurrence rates—that is, the percentage of cases that receive a new report and/or a substantiation of maltreatment within a designated follow-up period. In these studies, recurrence rates observed in AR families were compared to recurrence rates observed in the TR comparison group. One study that evaluated data from the National Child Abuse and Neglect Data System (NCANDS) found that for all but one state in their six-state evaluation, the rates of maltreatment recurrence were comparable for children receiving an alternative response and children receiving an investigation (Shusterman, Hollinshead, Fluke, & Yuan, 2005). Other studies have documented similar recidivism rates between lower risk AR cases and higher risk TR cases (Center for Child and Family Policy, Duke University, 2009; Institute of Applied Research, 2000, 2004b). Both DR research

reports and the associated DR program literature use this data to support their contention that being served in an AR track—foregoing identification of a perpetrator and fact-finding about the dynamics of child maltreatment—does not compromise children's safety (Institute of Applied Research, 2000, 2004b; Schene, 2008).

Presuming that families are accurately assigned to tracks based on a valid assessment of risk level, and given the restrictions on the families that are even eligible for assignment to AR (families at low to moderate risk), one would expect maltreatment recurrence rates in this group to be low to begin with, regardless of the level of service provided. We would also expect that their recurrence rates would be lower than the recurrence rates in families tracked to TR, simply because high risk is defined as a higher estimated likelihood of maltreatment recurrence (Baird & Wagner, 2000; Rycus & Hughes, 2003). Even in those studies in which the subjects were randomly assigned to AR (experimental) and TR (control) groups, the study sample was not randomly selected: both study subgroups were drawn from an initial sample of lower risk families who had been identified as eligible for AR (Institute of Applied Research, 2004b, 2006, 2010). Thus, one would expect to see comparable recurrence rates in the two comparison groups, regardless of whether they were served in the AR or TR track. Finally, because of typically low rates of maltreatment recurrence in lower risk families, any attempt to obtain accurate statistical comparisons about recurrence rates would be confounded by the low base rate of maltreatment incidents in these groups. For all the above reasons, lower observed recurrence rates in AR families may not be a meaningful measure of the effectiveness of the AR program in ensuring children's safety.

Some of the studies we reviewed attempted to augment their measures of recurrence rates by using caseworkers' ratings of children's safety. The caseworkers' opinions about changes in safety status were sampled at the beginning and at the end of the service period. Improvements in these ''before and after'' ratings were then presented as evidence that children's safety had improved under AR (Institute of Applied Research, 2004b). This type of measure is weak at best and is very prone to bias, particularly since standardized and objective safety assessment protocols were not always used to inform the ratings. Caseworkers, who had invested considerable time and resources into serving these families, cannot be considered objective or disinterested evaluators, and the potential for bias in their ratings can significantly distort research findings.

The inherent challenges in gathering accurate child safety information at screening would be of less concern if caseworkers in the alternative track quickly and routinely continued the assessment of both imminent and future risk for all their cases. In traditional CPS investigations, caseworkers are expected to conduct in-person interviews with all family members, meet and talk with alleged child victims and siblings, gather information from collateral sources, screen for other types of maltreatment, gather in-depth data to assess risk, and initiate safety plans when needed. As noted above, our research

suggests that these activities are not always implemented in the alternative track and in fact, may often be discouraged (Institute of Applied Research, 2004b, 2009). Only half the states represented by our key informants used standardized safety and risk assessment protocols with families served in the alternative track, and only four states (Minnesota, Ohio, North Carolina, and Vermont) and some California AR counties reported use of risk and safety assessment protocols that had been empirically validated. The fact that families assigned to the alternative track have already been designated "lower risk" may, itself, be a disincentive to adequate assessment of specific family dynamics related to risk, particularly if this is perceived as "off putting" to families. Further, it is unlikely that such in-depth scrutiny occurs in community-based service agencies, since their staff may not receive then necessary specialized training nor be given tools with which to monitor risk or to deal with more serious safety concerns as they emerge.

More important, the principles that underlie DR programming may prevent a thorough assessment of risk and safety from occurring in alternative tracks. DR's stated preference that workers focus on family needs rather than incidents of maltreatment could clearly discourage practitioners from having the sometimes difficult conversations with families that are necessary to fully assess risk and to address safety concerns.

The term incident focused emerged in the late 1990s on the heels of the development of the first child welfare risk assessment protocols. At the time, eliciting information about a family's maltreatment history and exploring the specific details of a particular maltreatment incident were purported to alienate families and shut down progress in making constructive changes. Rather than exploring the past, caseworkers were encouraged to focus on protecting children from future harm, and a conceptual dichotomy resulted in which practitioners were encouraged to become risk focused rather than incident focused.

The old dichotomy of risk versus incident seems to have reappeared in DR rhetoric but with a subtle shift. Alternative track caseworkers are now instructed to focus on family strengths and needs rather than seeking information about maltreatment incidents. Loman and Siegel reference this DR principle when they contend that the alternative track focuses on "looking forward with the family rather than looking back at an incident" (Institute of Applied Research, 2004b, p. 82). This message appears to have shifted assessment in alternative tracks away from actively exploring the dynamics of maltreatment with families.

The fallacy of this approach is that it is not possible to accurately assess risk of maltreatment recurrence without understanding the nature and dynamics of prior incidents of child maltreatment. Developers of actuarial and other research-based risk assessment models widely confirm that a prior history of child maltreatment is the single factor most highly correlated with future maltreatment (Baird & Wagner, 2000; Children's Research Center, 2003; Children's Research Center, 2009; Johnson, 2004; Sledjeski, Dierker, Brigham &

Breslin, 2008). Risk assessment protocols that do not incorporate this information have reduced predictive validity.

Equally important, exploring maltreatment dynamics with family members is not an end in itself but is part of a broader family assessment designed to inform subsequent service planning and delivery. If the specific factors, conditions, and dynamics that elevate risk in a family are not accurately identified and fully understood, it is significantly more difficult to provide relevant services to remediate these conditions to prevent maltreatment recurrences (Rycus & Hughes, 1998). Caseworkers cannot help parents strengthen their protective capacities and reduce risk if the parents do not recognize or understand the safety threats operating in their family and how these are expressed. If workers do not know who was caring for a child when maltreatment occurred, or the circumstances of its occurrence, it is difficult to help families construct and implement safety plans to avoid similar situations in the future. This essential information is only available through an exploration and assessment of the dynamics of prior maltreatment, and it is the caseworker's job to help family members explore and better understand these difficult topics in an environment characterized by honesty, understanding, compassion, and support.

Effective social work relies on building trusting worker–family relationships, which enable caseworkers to empower families, build on their strengths, address their needs, and help them ensure their children's safety well into the future. Such relationships allow for the honest and realistic exploration of family needs, strengths, and concerns within a context of trust and respect, where family members can participate as partners. Social work is based on the premise that clients have both the capacity and the right to benefit from honest and forthright communication. These fundamental underlying principles make social work—and specifically strengths-based social work—the preferred methodology to bring about constructive personal and interpersonal change. Avoiding unpleasant discussions may make both caseworkers and families feel more comfortable, but depending on case circumstances, it may be dishonest, condescending, and ultimately less helpful toward a goal of helping families acquire specific strategies to reduce risk and ultimately become more effective parents.

Assessing children's safety is an evolving process. Each contact with a family is a targeted opportunity to continue the case fact-finding that is essential for identifying risk of both imminent harm and future harm. For as long as CPS has involvement with any family, caseworkers must remain constantly open and alert for information that could change their assessment of a child's safety and when called for, must begin immediate safety planning. By establishing a precedent from the first contact with families that risk factors, maltreatment dynamics, and family protective capacities are important and ongoing topics of discussion, caseworkers can remove much of the stigma and can empower parents to recognize and consider emerging risks and to participate in developing safety plans to address them.

Based upon our analysis, we conclude that the many claims found in DR research studies and program literature related to children's safety are overreaching. Given the lack of consistent definitions, criteria, and tools to assess safety, the methodological problems in the research, the lack of consistent exploration and fact-finding in many AR tracks about prior maltreatment, and the inconsistent construction and implementation of DR programs from state to state, it is not possible to conclude that a DR model exists that can ensure that children's safety is not compromised in alternative tracks. Based on currently available data, we simply cannot know whether being served in an alternative track is undermining children's safety.

Finding # 4: DR Programs Appear to Prioritize Allocating Services and Resources for Families in Alternative Tracks

One of the potential disadvantages of DR is the allocation of already-limited fiscal and service resources to serve populations of lower risk children and families. DR was originally designed as an alternative to investigation for accepted reports of child maltreatment (Schene, 2005). The intent was to prevent lower risk families from being subjected unnecessarily to a CPS investigation, to allow more flexibility when offering services to these families (Child Welfare Information Gateway, 2008), and to more quickly link families with needed services (Key Informants). However, as DR reform has evolved, the targeted recipients in many jurisdictions have been broadened to include families whose CPS reports were not accepted by CPS agencies, but who were believed to be in need of supportive or protective services nonetheless. In some jurisdictions, alternative tracks have been created primarily to provide prevention services to at-risk families without any CPS involvement (Child Welfare Information Gateway, 2008; Conley & Duerr Berrick, 2010; Johnson, Sutton, & Thompson, 2005). This raises concern about the consequences of allocating resources to families in situations not sufficiently serious to warrant CPS intervention, when most CPS agencies operate in an economic environment that limits, often severely, the availability of resources to serve higher risk families and children.

There is considerable support in the DR literature for including lower risk and non-CPS families in DR's target populations. In 2005, Schene reported, "No state has adequate resources to respond to all the accepted reports of child maltreatment with a careful consideration of risks and needs," so families with less severe allegations may be simply closed after an investigation, and ". . . children are reported multiple times before a careful, face-to-face assessment is conducted" (p. 4). She contends that the rate of service provision in traditional CPS is characteristically low, with fewer than 30% of maltreatment reports being substantiated, and even fewer cases being opened for services. She concludes that the CPS system provides minimal services to families except in more serious situations where out-of-home placement is needed, and she views this as a primary rationale for the

development of DR (Schene, 2005). The federal government concurs that one goal of DR has been to serve low- or moderate-risk families who might not qualify for services under traditional CPS programming (Child Welfare Information Gateway, 2003). Similarly, Connell, Bergeron, Katz, Saunders, and Tebes (2007) contend that unsubstantiated CPS referrals may still represent maltreatment cases that would benefit from services and that DR is an appropriate intervention for these families.

Seven of the states in our key informant interviews operated more than two tracks in their DR programs. The additional tracks were often managed by community service providers, sometimes independent of CPS monitoring or involvement. In some of these tracks, if families declined services, no additional CPS action was taken. Thompson, Siegel, and Loman (2008) relate that Minnesota's DR program was a means to "respond early to a broader set of families and prevent conditions that are harmful to children" (p. 23). This program targeted the 60% of reported families who were screened out of the child welfare system because there was no specific alleged incident of child maltreatment or because there was insufficient information to warrant a formal CPS response. California reportedly developed a DR track to provide early intervention to families who would not typically receive services (Bagdasaryan, Furman, & Franke, 2008; Thompson, Siegel, & Loman, 2008). In these programs, DR was identified as a means of serving families who were historically underserved or closed out of services. This is consistent with the stated DR goal of front-loading services for lower risk and traditionally underserved families, even though many families in traditional CPS may not receive the services they need because the CPS agencies and communities that serve them lack the resources to effectively do so.

The practice of diverting CPS resources targeted to serve higher risk families and using them instead to serve lower risk families is troubling, when it is at the expense of the higher risk, core populations that CPS agencies are mandated to serve, particularly since many high-risk children depend on CPS intervention to ensure their safety and well-being. This in no way suggests that prevention and early intervention services for lower risk families are not needed—they absolutely are. In the best of circumstances, both child abuse prevention and family support services would be available to all families that needed them, in all communities. However, in an environment of chronically limited resources, CPS agencies should carefully consider the consequences of diverting their resources to serve lower risk families in alternative tracks, when families in traditional tracks cannot be adequately served because of insufficient fiscal, staffing, and community service resources.

Finding #5: DR Literature Misrepresents Traditional CPS to Enhance an Alternative Response Model

A primary concern in both the DR research and its associated program literature is the misrepresentation of traditional CPS to enhance the alternative track. In many documents we

reviewed, the traditional, or investigation track was often characterized by pejorative descriptors. The NQIC-DR describes TR as "inflexible," "adversarial," and "not able to provide sufficient services" (2009, p. 1). Kirk (2008) claims that TR "is, by its very nature, judgmental, legalistic, and adversarial" (p. 71). Kaplan and Merkel-Holguin (2008, p. 7) contend that TR is "an approach where parents are 'investigated' in a quasi law-enforcement method" that is "intrusive" and "threatening" to families. Other authors follow suit and assign similar attributes to TR (Institute of Applied Research, 2004b, 2009; Schene, 2001, 2005) In none of these sources did we find any empirical evidence cited to support these claims.

Kaplan and Merkel-Holguin (2008) also characterize AR as categorically different from TR by contrasting the "values" of the two approaches. They claim that AR values "Engagement versus adversarial approach"; "Services versus surveillance"; "Label of 'in need of services/support' versus 'perpetrator'"; "Encouraging versus threatening"; "Identification of needs versus punishment"; and "Continuum of response versus one size fits all" (p. 7). This paragraph alone would lead a reader to conclude that traditional child welfare practice was and is an adversarial, antagonistic, threatening approach to families, with primary goals of surveillance, of labeling, shaming and punishing offenders, and of using a single intervention strategy, devoid of services, with every family served.

Discussion

Most CPS professionals can provide examples of patently bad practice that is consistent with the above descriptions. However, this characterization does not define traditional CPS. The qualities attributed to TR in the above depictions represent a compilation of practice problems and shortcomings that exist to varying degrees in various places but not in the simplistic and harmful characterization presented. The negative and unprofessional qualities so widely ascribed to traditional child protection in this body of literature are stereotypic generalizations.

This negative portrayal of TR does offer a way to exemplify and amplify the presumed benefits and qualities of DR reform. By attributing patently negative characteristics to traditional CPS, DR proponents can build a paradigm that dichotomizes the traditional and alternative tracks and then relegates them to mutually exclusive categories, thereby creating a "good–bad" comparison. This strategy creates a straw man by presenting distorted, exaggerated, or misrepresented information about TR that can then be used to discredit it and highlights the more positive attributes of the AR track.

As with all straw man constructs, the AR versus TR paradigm can have powerful negative consequences. Straw man depictions are easy to tear down, and it is easy to rally support against them. And because this distorted representation is not an accurate representation of TR, accepting the distortion as fact prevents a legitimate, balanced assessment of the real strengths and limitations of both AR and TR. Further, it also obscures the many inherent commonalities in the two approaches. More accurate characterization of the AR and TR tracks, in philosophy and in fact, is essential if we are to construct testable models of practice, precisely because an honest and balanced assessment of current practice is a fundamental prerequisite and provides the foundation for progressive change and improvement.

Fundamentally, AR and TR have much in common. For the past 20 years, the mainstream CPS literature, university social work education, and state and provincial CPS in-service training programs have promoted an approach to working with families that much more closely parallels the values and methods of AR than the distortions and misrepresentations portrayed as traditional practice. As stated earlier, during the past two decades, family-centered, strengths-based practice principles provided the foundation for revised child welfare practice models that drove child welfare reform initiatives in many states and provinces. The fundamental positive tenets of DR are neither new nor revolutionary in child welfare practice. However, as is true with any large-scale system change, success in achieving a service system based on these principles depends largely on agency leadership and commitment, the availability of financial and staffing resources, legislative and policy support, and the skill of administrators in implementing and sustaining complex structural changes to make best practice possible. These elements have not always been present in the environments where such change is attempted. In this context, the DR reform movement has been a positive force in promoting infrastructure change to support family-centered and strengths-based interventions for families served in child welfare systems.

The original and most identifiable difference between the alternative and traditional tracks is the elimination of a formal investigation for families with less serious maltreatment allegations and a presumed lower risk of future harm to children. The stated rationale for creating the alternative track is that investigations undermine a caseworker's capacity to engage and partner with families because investigations are overly intrusive and designed to establish blame. While investigations have the potential to be inherently more confrontational than friendly home visits by a supportive and engaging family advocate, primary purpose of a social work CPS investigation is fact-finding, not faultfinding. Even if caseworkers—perhaps in collaboration with law enforcement—must sometimes use forensic interviewing strategies, evidence-preserving data gathering, legal intervention, and monitoring strategies to ensure that children are not re-abused or neglected, a skilled social work investigator can often do this without harming or alienating family members.

Fact-finding regarding child maltreatment dynamics is only the first step in serving family members where child maltreatment has been identified. The more compelling question is, what happens to families after such fact-finding has been completed? The characteristics defined as the "core

values'' of AR by Kaplan and Merkel-Holguin (2008, p. 7) also form the core values of all family-centered child welfare interventions, and they are especially important for families in the traditional track. With few exceptions, families served in CPS agencies have a right to receive strengths-based, empowering, supportive, and developmental child welfare services regardless of the seriousness of the referral allegations or prior instances of maltreatment. The goal of all child welfare interventions is to help families build their protective capacities and enable them to safely care for their own children. Only when this is clearly impossible, should we consider more intrusive and disruptive interventions such as out-of-home placement.

We must question why this principle is not more widely acknowledged in the DR literature. Are engagement, empowerment, and strengths-based services inappropriate for higher risk families? Is there valid data to confirm that higher risk families cannot be engaged, that they do not have strengths and protective capacities, or that they cannot participate in devising solutions to their own needs and problems? If family-centered, strengths-based practices are preferred because the profession believes them to be more effective and ethical than using unilateral authority to enact family change, don't the majority of members of higher risk families deserve the same opportunity to benefit from these modalities as do members of low- to moderate-risk families? And even when it is not possible or appropriate to engage perpetrators of serious child maltreatment into collaborative working partnerships, shouldn't we provide this opportunity to nonoffending parents, children, and extended family members? In some very severe cases, the ongoing use of mandated authority and legal intervention will clearly be necessary to ensure children's safety, but these families form only a subset of the families served in the traditional track, particularly when families are tracked to TR based solely on the original referral allegation before more in-depth family assessment data have been collected.

This concern raises other pertinent questions, including, exactly how different are the services provided to families in the AR and TR tracks? And, more importantly, how different *should* they be? In some organizations that claim to have fully implemented family-centered and strengths-based practice models, caseworkers report that they do not view AR as being all that different from what they normally do. In one research study, respondents referred to AR as ''TR light'' (Institute of Applied Research, 2004b, p. 4). It is clear that many respondents in states with a history of adopting family-centered practices do not view AR as significantly different from TR practice, which suggests they understand the potential benefits of family-centered models to helping most families, including those whose initial contact with the child protection agency required an investigative approach (Key Informants.)

Investigating alleged maltreatment is not the only thing, nor even the primary thing that happens in a traditional child protection track in CPS organizations that promote family-centered practice. Once past the investigation, the remainder of work in a TR track should follow pretty much lockstep with the AR track—beginning with a comprehensive and collaborative assessment of family strengths, needs, and risk, followed by safety planning when needed, service planning, service referrals, ongoing follow-up, and reassessment. While some child protection agencies may still complete an investigation and follow-up solely with monthly monitoring home visits, this has not been considered good child welfare practice for a very long time. A caseworker's failure to engage families in the collaborative completion of a strengths-need assessment and development of a relevant service plan is a failure of execution or implementation—not a defining characteristic of traditional child protection. Further, promoting stereotypes about TR communicates that the negative characterizations are not only real, but are necessary and appropriate for high-risk cases, which is patently untrue. Family-centered social work must be our first choice of intervention for all families, regardless of risk level.

A related concern is the following assertion from the DR literature. Kaplan and Merkel-Holguin claim that caseworkers working in AR reported being more satisfied with AR because they were performing ''real'' social work (2008, p. 12). It may be easier to engage and partner with families who have less complex histories and less challenging issues, but ''real'' social work is the professional capacity to work successfully with those families whose needs are not so easily met or whose problems are not so easily remedied. The implication that it is not possible to do ''real'' social work with families served in TR tracks is a misleading overgeneralization. It is precisely this aspect of professional child welfare social work that sets it apart from the other disciplines serving maltreated children and their families, and which makes it the appropriate profession to serve as the lead provider in the child welfare field of practice.

The negative portrayal of TR to promote AR serves no useful purpose. It undermines the contributions of child welfare professionals who have worked to integrate principles of family-centered and strengths-based practice into their day-to-day work in the only track available to them at the time and often in systems with infrastructures that were more restricting than enabling.

Finally, we found many additional examples of inaccurate, unsupported, and unfounded promotional claims about DR programming in both DR research reports and the associated program literature. A more detailed delineation of some of these can be found in Appendix B.

PART III: Preserving the Benefits of DR Reform
Conclusions and Recommendations

The DR advocacy movement has produced many benefits for the child welfare field, and these should not be lost or discounted in light of the previous discussion regarding the

problems in DR research and reform. Of greatest importance, DR has been a powerful stimulus for child welfare organizations to transform their organizational structures, policies, and procedures, and in some cases, their laws, to be more in line with principles of family-centered, strengths-based practice—principles which have formed the foundation of "best practice" in the child welfare field for perhaps 20 years, even though they have yet to be universally and fully implemented.

The last two decades have produced a variety of replicable service models to strengthen and support families. All have been attempts to engage, empower, and partner with families and involve them in collaborative planning to improve their lives and keep their children safe. For all the reasons mentioned earlier, these programs have not always been fully implemented nor sustained over time. Some agencies that started down this road may have devolved into more reliance on investigations and court involvement to deal with the increased federal emphasis on child safety, or as a response to a local child death or community pressures, or to deal with always-prevalent liability concerns. The erratic implementation of family-centered practice also reflects the always-present dilemma facing all child protection agencies: How can we be family-friendly, engage and empower families, support them, and build their strengths, while retaining the authority and legal capacity to act quickly and definitively to ensure children's safety when such intervention is needed? And how do we decide which approach is most appropriate, while ensuring children's safety at all times?

Although many of the underlying concepts of DR are neither unique to the DR approach nor new to the child welfare field, their fundamental message is entirely sound: Let us not put families through forensic investigative protocols and court involvement if we do not need that level of intervention to help them keep their children safe and prevent future maltreatment. What is different about DR is the growing capacity of its advocates to engage and move whole cadres of child welfare professionals, policy makers, and governmental administrators to implement DR programs. In this respect, their work clearly represents a significant strategic success. DR marketing and implementation strategies have successfully convinced organizations to become more open and receptive to larger system change when previous attempts at reform did not. DR can also provide a springboard for child welfare agencies to critically evaluate their own performance and practices and to explore different approaches to their work. However, the demonstrated power of this sophisticated marketing strategy should further emphasize the critical importance of rigorous and valid outcome research and open, transparent, rational public discourse to ensure that our marketing and implementation efforts are truly a force for constructive positive change.

For DR to survive and thrive, and to capitalize on the gains already made, the DR initiative must address several challenges.

The first is to determine what, exactly, is DR. If DR is intended to be a standardized program or intervention, different in significant respects from other family-centered CPS interventions, these differences must be defined and clearly articulated so they can be modeled, evaluated, and replicated. Definitional clarity of program components and interventions will make large-scale and ongoing research possible to determine what outcomes can be expected from these interventions and will also enable researchers to compare outcomes both within and across research studies and implementation sites. Until then, DR can only be considered a strong restatement of a family-centered practice philosophy and a series of disconnected and as yet unproven attempts to implement this philosophy in child welfare organizations by using various tracks to move families through the child welfare system.

Presently, the only distinctly novel program component of DR is that some families may progress directly from screening to a comprehensive family assessment without undergoing a formal investigation. If this is the primary substantive difference of the DR model, utilization of several service tracks may be only one of several potentially valid approaches to achieving this outcome. The essential program change would be the capacity to divert families believed to be at high risk to complete an investigation prior to assigning them to a family services caseworker for a comprehensive family assessment and service planning. This might be an effective model of implementation of the DR philosophy without resorting to different tracks, as it would ensure the same family-friendly, culturally competent and strengths-based services to members of high-risk families who could benefit as much as the lower risk or moderate-risk families who are currently the targets of service in alternative tracks. It would also allow local agencies to better use limited staff and fiscal resources by serving a pool of families who most need our help, while continuing to advocate with community partners to take more responsibility for serving and strengthening other families.

If progressive reform and quality improvement are developmental objectives, whatever DR evolves to be, "it" should be clearly defined, both philosophically and programmatically and manualized into a program guide that will promote more fidelity and consistency in program implementation. The closer we can get to this type of program standardization, the sooner research will be able to provide valid data regarding the program components and activities that warrant standardization. If program implementers do not provide an environment designed to support effective evaluation research, our "best available" data will be of limited utility in helping to shape and emend strategies to achieve our desired outcomes.

The second challenge is to be absolutely certain that we are not compromising child safety in our legitimate enthusiasm to engage and partner with families. Child welfare organizations must be able to negotiate the slippery slope that balances family engagement with child protection. Slipping too far to one side relegates us to being a governmental intervention that can tear families apart in the name of child safety. Too far to the other side, we become a supportive community cushion for needy families that fails to identify children at heightened risk

and denigrates our fundamental responsibility to act quickly and definitively to protect them. In this context, DR means knowing with reasonable certainty which level of intervention is best for which families, by taking sufficient time at the front end to make the most accurate assessment of both imminent and future risk and by planning interventions best suited for the family's situation, and by continuing fact-finding related to risk factors and protective capacities as long as a family is being served.

In the past, many child protection workers based their child safety and casework decisions on accumulated experience, judgment, and intuition (Rycus & Hughes, 2003). Based on the DR research we reviewed, apparently many still do, thereby opening the door to bias and misjudgment (Gambrill, 2011). The child welfare profession has developed standardized assessment and decision-making protocols that can help agencies assess risk and ensure child safely reliably and with good validity. Decision-making frameworks, such as Structured Decision Making™ (Children's Research Center, 2005) provide validated protocols, including decision trees for use at screening, actuarial risk assessment instruments, and safety assessment technologies to promote safety planning. These tools give caseworkers a much greater capacity to assess risk at all stages of case involvement and to make treatment intervention decisions accordingly. The other necessary component to making good case decisions is training for all staff in the fundamental and essential competencies for family-centered, strengths-based practice. For caseworkers, this must include a thorough understanding of how to balance the dual responsibilities of family engagement and child protection, and the casework skills necessary to help families identify their unique risk factors and to build their strengths and protective capacities to mitigate these conditions. Any child welfare reform effort that discounts the need for such balance could have a significant negative effect on meeting our responsibilities of due diligence in securing child safety.

The third challenge is to be empirically sound in our efforts to obtain and utilize good data as we go forward, always relying on the best available evidence to guide our programming decisions and increasing the amount and quality of data available to us as the program evolves. To do this requires a thorough understanding of the research that has been completed to date on DR programs and recognizing that in spite of our investment of time and resources and the availability of many published research studies and reports, we still lack a strong evidence base that confirms the capacity of DR, as it has been implemented, to reliably produce our desired outcomes in multiple jurisdictions over time, or to provide a confident path to better practice. As indicated earlier, DR research is in its early stages, and research ethics require the open and transparent communication of this fact. This does not suggest that we wait for conclusive research data before implementing family-centered practices. But we should avoid the wholesale and costly retooling of child welfare infrastructures when we lack data to confirm which approaches are necessary, effective, or most efficient. DR researchers and

advocates must accurately communicate the current state of our knowledge—and our ignorance—and must help jurisdictions design methodologically sound program evaluations that can build a better knowledge base to support future development. DR advocates should also temper their enthusiasm for marketing and promoting the large-scale implementation of specific DR methodologies before we know better what we should be asking people to implement, especially when we lack the data to ensure that these changes will lead to desired outcomes. In this article, we have tried to make the point that the current body of research supporting claims of safety and improved outcomes for children in DR programs is, at best, inconclusive, and at worst, misleading. Our hope is that evidence of our collective uncertainty about the effects of these reform efforts would prompt the humility that is a precondition to more rigorous research evaluations and more cautious approaches to change.

Conclusion

DR is a well-intended attempt to provide the best possible services to all families served by the child welfare system and to modify our responses so they are appropriate for each family's needs and circumstances. Any program that promotes constructive change in the child welfare system to achieve this end deserves consideration and support. It also deserves our commitment to a scientific and rational process that will promote objective, transparent, valid, and reliable development of practice and policy models for child welfare reform.

Appendix A

Summaries of Research Studies Reviewed for This Report

The following summarizes each of the research studies reviewed for this report. Data sheets with more detailed information about the methodology of each study will be provided upon request.

California

Conley, A. & Duerr Berrick, J. (2010). Community-based child abuse prevention: Outcomes associated with a differential response program in California. *Child Maltreatment, 15*(4), 282-292. Retrieved from http://cmx.sagepub.com/content/early/2010/07/16/1077559510376236

This is a very well-written study, but it has serious limitations, all of which were clearly acknowledged and articulated by the authors. The study has many problems with selection bias, and when coupled with a lack of data on comparison group families, it is impossible to arrive at definitive conclusions regarding program effectiveness. The differential response (DR) program evaluated by this study differed from most other DR initiatives in that it targets families that have been screened out of rather than

having been accepted for CPS services. Therefore, even had the authors been able to draw conclusions based on group comparisons, the results would not have been applicable to most DR designs.

Franke, T., Bagdasaryan, S., & Furman, W. (2011). Differential Response in rural counties: Path differentiation, service receipt and case disposition. *Journal of Public Child Welfare, 5*(1), 1-22.

This study reviewed DR practices in 11 rural California counties. In total, there were only 90 cases included in the study, an average of 8.2 families per county. The authors reported disappointing results in terms of services provided to families served in DR. However, given the differences in operations across the 11 counties and the very small size of the cohort, there is little substantive information to be gained from this study.

Virginia

Commonwealth of Virginia, Office of Social Services (2008). *Evaluation of the Differential Response System.* Richmond, VA. Retrieved from http://www.dss.virginia.gov/files/about/reports/children/cps/all_other/2008/differentialresponsesystem_evaluation_annualreport_2008_12-08.pdf

This study is descriptive in nature, and while it is well written and provides a comprehensive overview of the Virginia DR program, it did not attempt to evaluate the relative effectiveness of DR.

Kentucky

Huebner, R.A. (2005). Program Evaluation of the Multiple Response System: Kentucky [Unpublished report].

Huebner, R. A., Durbin, L., & Brock, A. (2009). Program evaluation of the multiple response system: Kentucky department for community based services. Retrieved from http://chfs.ky.gov/nr/rdonlyres/ba81ef9b-572d-4c59-ab4e-597eee3b0935/0/evaluationofthemultipleresponsesysteminkentucky_09.pdf

These studies are described as formative, or process evaluation studies, rather than summative, or outcome evaluation studies. They describe the Kentucky Multiple Response System program in considerable detail, but they do not include any outcome information and cannot provide valid information regarding the effectiveness of the state's DR approach.

Oklahoma and Kentucky

The Urban Institute. (2006). *Families' connections to services in an Alternative Response System.* Washington, DC: Zielewski, E.H., Macomber, J., Bess, R., & Murray, J.

As the title of the report implies, this study was a review of services provided in alternative response programs. Reviews were conducted in Oklahoma and Kentucky. The authors identified strengths, problems, and barriers encountered in both states, and they discuss implications for policy and program development. The study did not evaluate the impact of alternative response.

Washington State

State of Washington Department of Social & Health Services (2005). Alternative Response Systems Program Progress Report. July 1, 2003-June 30, 2004. Olympia, Washington: Brandford, C., Brummel, S.C., Clark, T.K., Filatova, N. & Skinner, J.

This study is a descriptive review of practices observed in the Washington state alternative response program. Satisfaction surveys were conducted and 6-month outcomes were tracked, but there was no attempt to determine whether AR was more effective than traditional child protection services. The researchers found a considerable lack of fidelity in implementation of program requirements and offered several recommendations for improving AR practice. This study provides little information on the effectiveness of AR programs.

West Virginia

ACTION for Child Protection. (1998). *Family options initiative: Pilot evaluation report.* Albuquerque, N.M.: Costello, T.

This evaluation of West Virginia's Family Options Initiative was observational in nature and, therefore, adds little to the knowledge base regarding the effectiveness of DR programs.

North Carolina

Center for Child and Family Policy, Duke University (2009). Multiple Response System (MRS) evaluation report to the North Carolina Division of Social Services (NCDSS).

The evaluation of North Carolina's Multiple Response System (MRS) used a quasi-experimental evaluation design with matched counties as experimental and comparison groups. There were 10 initial pilot counties and 9 comparison counties. Information gathered from a second wave of counties that adopted the MRS system (referenced as Wave 2 counties) was then compared to findings from the original sample. The study is well presented, and it summarizes data from very complex analyses (interrupted time series analyses) and extensive survey data, all of which were reported in the study

appendices. Only summaries of the data were presented in the body of the report.

A variety of process and outcome measures were presented. Process measures included the timeliness of agency responses to reports and time devoted to cases from acceptance of the referral to making a finding. The latter measure was referred to as front-loading services. Child safety was measured by combining rates of subsequent substantiation and subsequent assessments that determined there was a need for services. In addition, annual rates of substantiation/need for services were reported for the total population of the counties that comprised each study cohort.

The research team found very little difference in any measures between the pilot and control counties, and, in general, outcomes for the Wave 2 counties were not as positive as those observed in either the pilot or control counties. On process measures such as timeliness of response to reports of maltreatment and front-loading of services to new cases, the control counties actually did somewhat better than both the experimental and Wave 2 counties. The authors' contention that "data suggest MRS is increasing safety" appears to be overly optimistic. When general statistics from the three groups of counties are compared, there is very little change over time in the number of assessments conducted per 1,000 children for either the experimental or control groups. However, shortly after the percentage of cases placed in MRS in the Wave 2 counties increased substantially, the overall rate of assessments increased about 16% over the next 2 years. While this should raise some concern, it was largely ignored in the report.

Further, the overall rate of substantiation/needs services remained almost constant in the experimental group from 2003 to 2007, while it declined steadily over the same period in control counties. Again the rate of substantiation/needs services in Wave 2 counties was higher than the rates found in the other two groups. Finally, the percentage of children assessed and reassessed within 6 months was almost exactly the same in both pilot and control counties in 2007–2008 as it had been in 1996–1997. MRS may not have compromised safety, but the study data do not support the finding that child safety was improved under MRS. In fact, the data indicates that MRS had little to no impact on child safety.

Minnesota

Institute of Applied Research. (2004). Minnesota Alternative Response Evaluation: Final Report. St. Louis, MO: Loman, L.A., & Siegel, G.

Institute of Applied Research. (2006). Extended follow-up study of Minnesota's Family Assessment Response: Final report. St. Louis, MO: Siegel, G.L., & Loman, L.A.

The two Minnesota studies have received a great deal of attention for the following reasons: (1) they incorporated random assignment methodology; (2) they addressed many different issues; and, (3) they have served as a model for subsequent evaluations completed in other states. Our review, however, found fundamental flaws that severely limit the value of the study findings as put forth by the authors.

The primary problem with these studies is that it is impossible to determine precisely what is being evaluated. The premise behind alternative response (AR) is that the process of investigating allegations of child abuse and neglect is adversarial and does not provide a positive atmosphere for helping families address the factors that led to CPS referral. It has been hypothesized that a process focusing on assessing family circumstances and offering them assistance would enhance the potential for effective family engagement and successful resolution of issues without compromising child safety. However, to test whether such an approach to serving families is as effective as traditional child protection practice, all other conditions must remain equal. Instead, in the Minnesota programs being evaluated, families served in the alternative track were provided with additional resources and services that were not as readily available to families in the comparison group. Funds were also made available to help families in the AR track pay their rent and to purchase appliances and other necessities. If these extra services resulted from cost savings generated by the AR approach itself (e.g., savings from a reduction in costs incurred in traditional practice such as court costs), they could be justified. However, much of the funding for these extra services came from a foundation grant, and these funds were allocated entirely to help families in the AR track rather than being provided equally to both the experimental and control groups. Hence, any improvements noted in the experimental group could potentially be attributed to the extra resources provided to the AR group rather than differences in how an agency responds to an allegation. In effect, these studies are more an evaluation of the effects of enhanced services than of an alternative response to allegations of maltreatment

Furthermore, when evaluating the impact of services, opportunity costs must be examined. Research in other fields clearly demonstrates that focusing resources on high-risk cases results in better overall outcomes.

Providing enhanced services to high-risk families, rather than the relatively low-risk households assigned to AR, may have been a much better use of these resources.

Study results were further compromised by additional training in interviewing and family engagement skills provided to AR workers but not to TR workers. If the principal goal of the study was to compare the results of two approaches to responding to allegations of child maltreatment, then workers in both study groups should have been comparably equipped to work with families. There may also have been selection bias at play; how workers were selected for the experimental AR program was not explained, but if they volunteered, they may well have been workers who were more flexible and more committed to family engagement and problem solving than their counterparts in the control group.

Many other issues were identified. Chief among these were (1) a lack of needed controls, such as a uniform follow-up period for examining recurrence rates; (2) the cost analysis seemed particularly questionable because the lack of a uniform follow-up period . . . up to 1,200 days . . . meant that a few outliers could have had a significant impact on cost comparisons; (3) questionable research strategies—e.g., beginning the follow-up period at the end of the service period; (4) findings that seem incongruent, such as families without prior referrals having done worse than families with prior referrals; and (5) assumptions about risk levels in each cohort—assuming that AR reduced risk immediately because families appeared to be more cooperative.

Finally, a comment seems warranted on the tone of the written reports. Whenever possible, evaluators should be completely disengaged from the subject of the research and objective about the data. This is not always possible. In some instances, comprehensive knowledge of the subject matter is beneficial. However, when this is true, it is important that evaluators acknowledge potential bias and take extra measures to ensure objectivity. This is not the case in these evaluation studies. Throughout these reports, the research team seems to wholly embrace AR, and they frequently present ambiguous results in ways that support the AR approach. After a comprehensive review of these studies, we found nothing in the data that demonstrates that AR is more effective than TR. The presence of many exogenous variables in the research simply precludes any definitive findings regarding the AR approach.

Missouri

Institute of Applied Research. (1997). The Missouri Family Assessment and Response demonstration: Final evaluation report. St. Louis, MO: Siegel, G.L., & Loman, A.

Institute of Applied Research. (2000). The Missouri Family Assessment and Response demonstration impact evaluation: Digest of findings and conclusions. St. Louis, MO: Siegel, G.L., & Loman, L.A.

Institute of Applied Research. (2004). Differential Response in Missouri after Five Years: Final Report. St. Loius, MO: Loman, L.A., & Siegel, G.L. Retrieved from http://www.iarstl.org/papers.htm#ancC5

The Missouri evaluations were conducted by the same research team that conducted the Minnesota evaluations. The Missouri design was quasi-experimental, using matched pilot and comparison groups. It appears that the experimental counties volunteered for the project, which introduces the potential for selection bias. Unlike the Minnesota studies, it does not appear that additional resources were provided to families in the pilot counties. There is no information provided on how pilot counties were chosen. If they volunteered, there may be considerable selection bias present.

A tremendous amount of data was reported, but much of it was of little consequence to the central goals of the Missouri AR program. This makes it very difficult to determine exactly what had been accomplished by AR. Safety is, of course, the primary concern of child protection services, and we found the findings related to safety to be more ambiguous than was reported in the research reports. Recurrence rates increased during the study period in both the experimental and comparison counties, but the rate for the pilot counties was lower than that reported for the control counties (37.7% vs. 40.4%). The difference, however, was due to new reports for less serious types of maltreatment (e.g., lack of necessities, lack of supervision, and lack of proper concern about education). This could well have been more an artifact of where staff were located than anything else. In some of the pilot counties, staff were stationed in schools, and in one county, they were placed at the welfare office. These staff reportedly worked closely with teachers and welfare staff. Hence, reports that came into the hotline in the comparison counties may have gone instead directly to workers in the pilot counties who were placed in the schools and the welfare office. The numbers reported appear large enough to account for all of the differences noted in recurrence rates between pilot and comparison counties.

Even more importantly, the five-year follow-up study found that a substantially higher percentage of children were subsequently removed from their homes in the pilot counties, suggesting negative outcomes of child safety.

Finally, the question of opportunity cost must be addressed. A lot of CPS services in the pilot counties

went to families reported for low level child maltreatment. While there is little question that these families need services, targeting high-risk families may have produced better overall results. In sum, many of conclusions of the report are not totally supported by the evidence presented. It would have been just as easy to conclude that AR has had little impact and could have, in the long run, jeopardized child safety.

Ohio

Institute of Applied Research (2010). *Ohio Alternative Response Evaluation: Final Report.* St. Louis, MO: Loman, L.A., Filonow, C.S., & Siegel, G.

This study was completed by the same principal investigators who authored the Missouri and Minnesota reports. This study employed random assignment in 10 Ohio counties, assigning cases eligible for the alternative track to either alternative or traditional CPS services. Minor but significant differences in case outcomes all favoring alternative response were reported. While this report is better organized and somewhat easier to understand than either the Missouri or Minnesota studies, many confounding issues are present that diminish the value of the study.

First, additional resources were available to families assigned to the alternative track, including funding of $1,000 per family; another $50,000 in foundation support was allocated to each AR county. As a result, extra services were provided to families served in the alternative track, including help with rent, appliances, and daily expenses. AR workers were also provided with additional training and technical assistance. These advantages make it impossible to determine precisely what generated any reported improvements in child safety, family satisfaction, or any other outcome measure.

Second, there was at least one significant difference in the study cohorts that could also have accounted for the modest difference noted in recurrence and placement rates. Safety issues at the time of intake were noted for only 25.4% of families in the experimental group but were found in 33.2% of families in the control group. Safety concerns drive most placements, and because safety and risk are correlated, it is likely that the expected rate of subsequent maltreatment reports would have been higher for the control group anyway, based on inherent group differences.

Further, nearly 4% of the cases randomly assigned to the alternative track were subsequently reassigned to the traditional track. While there was no explanation provided, it may be that these cases were not appropriate for the alternative track in the first place. If no similar review was conducted for cases in the control group who were later deemed eligible for the alternative track, group equivalence may have been further compromised.

Finally, there is again the issue of opportunity cost. If comparable extra services and support had been provided to higher risk families, it is possible that better overall results could be attained for families served in the traditional track. The extra resources and services provided to families in these studies conflate the results, making it impossible to attribute any observed outcome to the primary AR concept of conducting family assessments rather than investigations.

Nevada

Institute of Applied Research. (2009). *Differential Response in Nevada: Implementation and impact evaluation after 32 Months.* St. Louis, MO: Siegel, G.L., Shannon, C.F., & Loman, L.A.

This evaluation was conducted by the same team that conducted the Minnesota, Missouri, and Ohio evaluations. The Nevada study used a quasi-experimental design, using cases that were eligible for DR but were assigned to traditional services as a comparison group. Research methods incorporated worker and family surveys to augment data from Nevada's UNITY (SACWIS) system.

Two things differentiate the Nevada system from other DR programs. First, nearly all alternative track cases were served by private community-based organizations. Caseloads for the Alternative Track were set at 15 per caseworker. Second, DR has been used sparingly in Nevada. In 2009, only 8.8% of all referrals went to the alternative track. Originally, only cases rated as Priority 3 (cases which generally had no child safety concerns) were eligible for referral to the alternative track. Priority 3 families comprise the least serious of all referrals.

The Nevada program was a prime example of a system that allocates a higher level of service resources to cases that presented a rather low risk to the children. Even though there was no mention in the report of additional resources available to families served in AR, we were able to estimate the size of AR caseloads by extrapolating data from several charts and tables. The average AR caseload appeared to be fewer than 10 cases per worker, which was less than half the size of caseloads in the TR track. This translates into more time, more services, and more expense allocated to Priority 3 cases.

The Nevada evaluation exhibits the same problems found in all the studies conducted by this research team. These researchers are clearly supporters of DR programs, and their opinion carries over to the interpretation of

outcome data. If an outcome measure favored the comparison (traditional response) group, the research team usually attempted to look for an explanation. However, outcomes that favored the alternative track were often accepted at face value without further analysis, even when they seemed to be an anomaly. As an example, in this study, the rates of new reports observed during the follow-up period were nearly identical for both the experimental and the comparison groups. However, new investigated report rates were considerably higher in the comparison group and substantiation rates were somewhat higher. Our review of the data surfaced a number of potential reasons these rates would be higher for the comparison group. First, the two groups were not entirely comparable at the outset. The comparison group, while exhibiting a slightly lower rate of prior reports, had a higher rate of prior investigations, substantiations, and removals, all of which indicated histories of more serious maltreatment. Such histories certainly make it more likely that many new reports would trigger investigations. Second, some comparison group families were selected from counties where DR was not available for at least part of the follow-up period, making investigation of new reports more likely. Other families were selected from areas where DR caseloads were full. If this continued into the follow-up period, it would also make investigations more likely for comparison group families. Finally, a much higher percentage of children in the alternative response group were adolescents (35.6% vs.18.2%). More of these children would have "aged out" during the follow-up period. Taken together, these trends would have reduced all measures of recurrence in the experimental group. Comparing the new report rates to new investigation rates within each of the two groups should have spurred further analysis. Instead, it was left to stand as an indicator of DR success.

Again, we believe that despite higher levels of services going to Alternative Track families, there is no evidence that the DR program led to better outcomes. Putting these resources into higher risk cases may well have proved a better overall strategy.

New York

Ruppel, J., Huang, Y., & Haulenbeek, G. (2011). Differential Response in Child Protective Services in New York State: Implementation, initial outcomes and impacts of pilot project. Retrieved from http://www.ocfs.state.ny.us/main/reports/CPS%20Differential%20Respon se%20Evaluation%20Final%20Report_%20Jan%202011 .pdf

The evaluation of New York's Family Assessment Response (FAR) program combined results from surveys of staff and families with outcome information obtained from New York's child welfare information (SACWIS) system. The researchers used a combination of a quasi-experimental evaluation design and random assignment. Control groups were established for only two counties; one was based on random assignment, the other used matching techniques and propensity scoring (which was not fully explained in the evaluation report). The measured outcomes included new reports, investigations, substantiations, and petitions.

The research team analyzed differences in outcomes reported for the two groups, and they did denote possible alternative explanations for the observed differences other than the introduction of FAR. Overall, there appeared to be no difference in outcomes in the FAR and control groups. Some measures favored the FAR group, while others favored the control group. The evaluation results were limited (based on data from only two counties) and ambiguous, thus providing little evidence of the effectiveness of the FAR program.

As was true for other DR evaluation studies, it was impossible to determine precisely what was being evaluated in New York. First, the pilot counties volunteered to participate in the program rather than being randomly selected. Volunteer status suggests inherent differences—potentially stronger county leadership, greater flexibility, and a higher interest in testing new concepts—increasing the likelihood of bias that could affect the outcomes. More importantly, FAR offered many advantages that were not available to families or to caseworkers in the traditional response track. These advantages included additional training and coaching for FAR staff, quality assurance checks to help ensure program fidelity, funding from the Office of Children and Family Services and from foundations to provide additional wraparound services to FAR families, and education of and outreach to community-based programs that would work with FAR cases. With all of these additional resources available, it would be very surprising if family and worker satisfaction were not significantly higher in the FAR program than in the TR track. Since supportive, family-centered services should improve these measures regardless of the track chosen, one could easily conclude that the extra services, not the AR program itself, produced the improvements. Providing these services to higher risk families may have, in fact, produced better results.

National Child Abuse and Neglect Data System (NCANDS) Six-State Review

Fluke, J. D., Shusterman, G. R., Hollinshead, D., & Yuan, Y. T. (2005). *Alternative responses to child maltreatment: Findings from NCANDS* (Washington DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, 2005). Retrieved from http://aspe.hhs.gov/hsp/05/child-maltreat-resp/

The most interesting data are this report can be found in addendum materials. Information was provided on DR programs from six states, including several states that were subjects of the evaluation reports described above. Data are provided comparing 6-month recurrence rates for cases assigned to both the AR and TR tracks as well as for cases that were investigated but not substantiated. Recurrence data are further delineated into three categories: those investigated and substantiated; those investigated but not substantiated and those assigned to the AR track. These data allow the reader to weigh the impact of past decisions on CPS response when a new allegation is reported.

While the criteria for AR eligibility varied among the six states, it was clear that AR cases were generally lower risk families than those assigned to the TR track. Hence, lower rates of recurrence would be expectable for AR cases. Comparing outcomes of those initially investigated but not substantiated with outcomes for the AR group is interesting for two reasons. First, it is likely that the unsubstantiated group had a somewhat lower risk profile than the substantiated group and is, therefore, a somewhat better comparison to the AR track. Second, it is likely that the vast majority of these cases were closed without services, which sets up a comparison between relatively low-risk families who were provided no services and relatively low-risk families who had been provided with a myriad of services through AR. In two states, Minnesota and New Jersey, AR cases were *more* likely to have had a subsequent substantiation. In Kentucky, Oklahoma, Wyoming, and Missouri, AR cases had a lower rate of new substantiations, but this appears to be largely because referrals for this group more often led to assignment to the AR track, where substantiations were not made. In Wyoming and Missouri, far more cases from the AR track than the TR track were likely to be reassigned to the AR track when a subsequent allegation was received. These two states reported the greatest differences in subsequent substantiation rates between AR and TR cases. Clearly, prior decisions regarding assignment to AR or TR are affecting outcome measures which have subsequently been used to promote the success of AR.

When results are aggregated across states in this study, taking into account the higher percentage of the initial AR cohort that were reassigned to AR when a new allegation was received, they may well demonstrate that providing services to low-risk families is no more effective than closing these cases.

Appendix B

Science or Promotion?

There is considerable discussion in the professional literature about the prevalence of inflated claims and false findings in published medical and social sciences research (Best, 2004; Chalmers, 1990; Coyne, 2009; Gambrill, 2011; Goodman & Greenland, 2007; Ioannidis, 2005, 2008, 2011; Reich, Green, Brock & Tetlock, 2007; Rubin & Parrish, 2007). Many contend that a substantial number of research papers submitted for publication in juried professional journals contain claims that are not supported by the reported data and instead reflect a biased promotion of personal views or agendas (Gambrill, 2005, 2010, 2011; Littell, 2005; Rubin & Parrish, 2007). In an article titled "Ethical Aspects of Outcome Studies in Social, Behavioral, and Educational Interventions," Gambrill (2011) delineates many of these concerns and contends that many research claims are "more advertisements than material designed to forward discovery of what is true, what is false, and what is uncertain," and which "may be more an advertisement for a product promoted rather than a scholarly description" (p. 655). She further delineates the significant harms that occur when flawed studies hype their findings and rely on inflated claims of knowledge to support their conclusions.

Promotion is characterized by the propagation, championing, and support of specific beliefs, ideas, or philosophies. The field of advertising relies on highly developed promotional strategies to market both ideas and products, irrespective of their inherent worth. By contrast, sound empirical research is designed to promote knowledge, accuracy, and truth. It identifies, presents, and explains evidence gleaned from research that does or does not support targeted outcomes. Additional objectives of good outcome research include controlling for bias, eliminating confusion and ambiguity in favor of clarity, divesting emotion and focusing on fact, and ensuring transparency of methods. Commitment and adherence to rigorous research methods and standards for reporting results are intended to achieve these ends.

Table B1. Promotional Claims and Concerns

Promotional claim	Concern
Type: Discouraging critical appraisal of claims	
"The evaluation results are consistent with the belief of Office of Children and Family Services (OCFS) and many local districts in the potential of FAR to provide a more effective and less intrusive approach to child protection" (Ruppel, Huang, & Haulenbeek, 2011, p.vii).	Consumers of outcome research reports would be looking to the research for data-driven research findings, not the "beliefs" of influential supporters. Showcasing the supporting beliefs of influential others discourages critical appraisal.
"We have said before that we do not believe these results are easily replicated. They can be, but only if the necessary effort and intelligent design are applied" (Institute of Applied Research, 2006, p. 4).	This quote inoculates readers against future negative findings, suggesting that future negative findings are "unintelligent," rather than potentially well-reasoned or well-researched alternative disagreements.
Type: Reliance on association and suggestion, such as negative innuendos	
"This finding will be counterintuitive to many, particularly those who hold that a police-like and essentially adversarial investigation is necessary to protect children" (Institute of Applied Research, 2000, p.37).	The authors could have said "those who hold that a more forensic fact-finding approach is often necessary," but "police like" and "adversarial" are more negatively laden terms. This negative framing is often used in advertising to promote emotional support of a claim rather than rational assessment.
"Workers tended to like the fact that AR is not punitive, intrusive, threatening, labeling, and blaming, shaming, antagonistic" (Institute of Applied Research, 2004b, p.81).	The implication is that what AR is ostensibly replacing was categorically all of these things. Again, an example of negative framing, soliciting a negative emotional response rather than the neutral framing necessary for rational and critical appraisal.
Type: Presenting testimonials as scientific evidence	
One outcome research report (Institute of Applied Research, 2000) includes a side bar with numerous quotes from consumers regarding their positive feelings for AR. There were no quotes offered that reported negative opinions.	Cherry-picking responses to support bias and institutionalizing selection bias into research designs are fatal to empirical research. This is advertising and marketing, not science.
One study presented the following quote from a caseworker as evidence in support of DR: "The reaction of families has been nearly all positive. Family assessments always involve the children. We always see the kids. The children we see have always been safe so far. CPS has done a good job screening" (Institute of Applied Research, 2009, p.34).	The opinion of one staff person, with no data to back up the claims, is an example of using testimonials rather than objective data to promote acceptance of a product.
"The results of the evaluation all favored the family assessment . . . over the traditional approach to child protective services" (Institute of Applied Research, 2000, p.39).	In fact, the results in this particular research were nuanced and did not all favor family assessment, suggesting "selective reporting" of research outcomes.
Type: Inflated claims	
"The basic principles and practices of alternative response models are fairly consistent across states." However, in the next sentence the authors state, . . . "the criteria used in determining which reports to include or exclude varies, as do assessment protocols, time frames for service completion, and the degree of latitude local jurisdictions have in defining their DR programs and FAR practices" (Ruppel et al., 2011, p. 15).	This is an important claim as it suggests there is a consistent model of DR practice. (All of the data we provided in Finding #1 of this report demonstrates that this claim is incorrect.) If these inconsistencies are typical, then the only thing consistent would be the inconsistency in practice model structures.
"As the program helps families become more resourceful, and communities become more responsive to the needs of families with children, we hypothesize that more positive effects will be seen in the long term" (Ruppel et al., 2011, p. 15).	This statement follows a finding that the research showed no positive effects in the short term (2 years) and a recommendation to adopt the program in spite of this. Lack of positive effects was construed to be insignificant, claiming the research supported the expectation that the reform would inevitably produce positive results in the near future.[a]

Note. DR = differential response; AR = alternative response; FAR = family assessment response; CPS = child protective services.
[a] The above examples were selected for their explanatory value. The DR outcome research reports reviewed for this study were replete with similar examples. A more detailed delineation is available upon request from the authors.

Motives aside, we found unsupported, inflated, and unfounded promotional claims to be a significant problem in the differential response (DR) research we reviewed. We similarly identified many unsupported claims in the theoretical, descriptive, and program literature about the purported advantages and positive potentials of DR. Additionally, many of the studies in our review failed to fully articulate and explain study limitations and often failed to propose alternative explanations for study findings, thus increasing the likelihood that readers would draw erroneous conclusions not supported by the data. Many claims in this body of literature about the benefits of DR exemplify marketing and promotional strategies rather than objective science.

Gambrill (2011) described several types of inappropriate promotional discourse found in social science outcome research. We used this framework to identify and categorize claims in the DR research that were promotional and propagandistic rather than data driven, and we include a sampling of these claims in Appendix Table B1.

One type of promotion—the use of testimonials—warrants further discussion. Family satisfaction has been one of the most researched DR outcomes, and it is widely presented as evidence of the value of AR over TR. In much of this research, the measurement of family satisfaction has consisted primarily of opinion surveys conducted with families and caseworkers. In the research reports, the reported testimonials are overwhelmingly positive, indicating either the absence of negative opinion from consumers or the cherry-picking of positive responses. In some studies, incentives such as gift cards, cash, or recreational passes were provided to families who did participate in surveys or interviews (Institute of Applied Research, 2004b, 2009, 2010). There was no reported attention given to the reasons other families may not have responded to the surveys, or how their perspectives might have differed.

This type of promotion has no place in outcome research. The scientific and ethical foundations of outcome research, its empirical and ethical legitimacy, and its practical utility are all undermined or discounted by promotional strategies other than the accurate, objective, and transparent demonstration of evidence.

Appendix C

Key Informant Interview Protocol

Program Description
How long have you been operating an "alternative response" program?

What were the factors that created your system's desire to do this?

What title have you given your "alternative response" program, and why are you using the terminology you chose?

What were you hoping to gain by this initiative?

- Problems you were trying to solve?
- Enhancements you were hoping to make?
- Financial considerations?
- Quality of services to families?
- Other?

Who actually provides the alternative response services? Staff in your own agency (the public child welfare agency?) Agencies under contract with your agency? Regular community agencies? Describe all.

Track Assignment
How have you differentiated the "tracks" in your system? How many tracks do you have?

What criteria do you use to decide into which track you assign a case? What are the defining characteristics of each "track"?

At what point in the intake process do you make the track assignment on a case?

- At the time of the initial screening call?
- After a supervisor reviews the screening information?
- After an initial intake visit?
- After completing the intake?
- After completing a safety and/or risk assessment?

Do you make the case track assignment based on:

- The nature of the allegation, as determined by the screener
- Information gathered during an intake interview
- Information from safety assessments
- Information from risk assessment
- Parent's professed interest in receiving services
- Other?

Substantiation and Documentation
Do you use the legal classifications of "abuse" "neglect" "sexual abuse" in your formal case dispositions? For which cases and under what circumstances?

Do you use case dispositions of "substantiated," "unsubstantiated," and "indicated" (or some other terminology)? For which cases?

Do you use a CHINS (Child in need of services) framework rather than categorizing by type of maltreatment?

Do you document exactly what type of maltreatment has occurred; i.e. physical abuse, sexual abuse, or neglect? For which cases in which tracks?

Do you use a different case determination or disposition strategy in your alternative response tracks? What is it?

Do you focus on determining the existence of, nature of, and dynamics of child maltreatment, including identifying

perpetrators? In which tracks? For all cases? For some cases? Not at all?

Does your state have a formal registry? If so, what criteria do you use to assign perpetrators to the state registry?

If you don't have a state registry, or don't record all cases in the registry, how do you monitor repeat referrals on a family?

What strategies do you use to track families that move out of your county or catchment area and communicate with other agencies about prior referrals?

What terminology do you use to refer to persons who have maltreated a child? Perpetrators? Other terms?

Case Processes

Do you have a formal practice model to guide your interventions in different tracks? Can you articulate its basic principles? Do you have practice guidelines or materials that you could share with us?

Do you believe the program operates as it has been defined in the practice model? In what ways do you think implementation in the field is inconsistent with the practice model? Why?

How and when do you focus efforts on engaging families? In which tracks? At what time? Under what circumstances?

Describe any differences in case information gathering activities for the different tracks in your system. Specifically:

- Are there differences in the types of assessment data you collect with families on different tracks? Describe the differences.
- Do you use different types of assessment or investigation methods for families on different tracks? Describe the differences.
- How and when do you use forensic interviewing methods in assessing a family?
- Are there differences in how you do safety assessments and/or risk assessments in the various tracks? What are they?
    - At what point in your case contact with families in different tracks do you begin providing services?
    - How long do you typically keep cases open in your "alternative" tracks?
    - What generally are your criteria for closing these cases?

Staffing and Training

Do you have different expectations for the competence of social workers performing different types of investigations/assessments?

Do you provide different types of training for workers in different tracks?

Program Evaluation

Have you collected data to document the success of program implementation activities? Has the program gone as you expected? Why, or why not?

What outcomes has the program achieved? (Service completion, customer satisfaction, decreased recidivism, decreased placement, increased capacity to problem solve, others)

Have you collected data to evaluate the outcomes of your program? I.e. data about recidivism? Successful completion of services? Reduction of re-referrals?

What does this data tell you?

What have you learned about the cost of the program?

Challenges Encountered

What challenges did you encounter in implementation?

What barriers did you run into that you didn't expect?

How have you tried to resolve or work around these barriers?

What, if any, unanticipated consequences have there been (either positive or problematic)?

General Conclusions

What have been the most obvious benefits of your multi-track system, from your perspective?

Is the program operating in the way it was envisioned?

What would you do differently, in hindsight?

Would you like your information to be anonymous?

States Represented by Key Informant Data

Alaska, California, Florida, Hawaii, Kentucky, Louisiana, Maine, Massachusetts, Minnesota, Missouri, Nevada, New Jersey, New York, North Carolina, Ohio, Oklahoma, Vermont, Virginia, West Virginia, Wisconsin

States Interviewed but Not Included in Key Informant Data

Delaware, Idaho, Iowa, Kansas, Michigan, South Dakota, Utah, Washington

References

ACTION for Child Protection. (1998). *Family initiative: Pilot evaluation report*. Albuquerque, NM: Author.

Adoption and Safe Families Act of 1997, P.L. 105-89

Adoption Assistance and Child Welfare Act of 1980, P.L. 96-272

American Humane Association. (undated). *Fact sheet: Fatalities due to child abuse and neglect*. Retrieved from http://www.americanhumane.org/children/stop-child-abuse/fact-sheets/fatalities-due-to-child-abuse-neglect.html

Armstrong, K. L. (1997). Launching a family-centered, neighborhood-based human services system. *Administration in Social Work*, *21*, 109–126.

Bagdasaryan, S., Furman, W., & Franke, T. M. (2008). Implementation of California's differential response model in small counties. *Protecting Children*, *28*, 40–56.

Baird, C., & Wagner, D. (2000). The relative validity of actuarial and consensus-based risk assessment systems. *Children and Youth Services Review*, *22*, 839–871.

Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999). Risk assessment in child protective services: Consensus and actuarial model reliability. *Child Welfare*, *78*, 723–748.

Behar, L. B. (1986). A state model for child mental health services: The North Carolina experience. *Children Today*, *15*, 16–21.

Berg, I. K. (1994). *Family based services: A solution-focused approach*. New York, NY: Norton.

Best, J. A. (2004). *More damned lies and statistics: How numbers confuse public issues*. Berkeley: University of California Press.

Carpenter, C. (2007). Alternative response. *Children, Families, and the Courts Ohio Bulletin*, *3*. Retrieved from http://www.americanhumane.org/assets/docs/protecting-children/PC-AR-Ohio-child-law-bulletin.pdf

Center for Child and Family Policy, Duke University (2009). *Multiple Response System (MRS)*. (Evaluation Report). North Carolina Division of Social Services (NCDSS). Durham, NC: Lawrence, C. N. and Rosanbalm, K.

Chaffin, M., Bonner, B. L., & Hill, R. F. (2001). Family preservation and family support programs: Child maltreatment outcomes across client risk levels and program levels. *Child Abuse & Neglect*, *25*, 1269–1289.

Chalmers, I. R. (1990). Underreporting research limitations is scientific misconduct. *Journal of the American Medical Association*, *263*, 1405–1408.

Child Welfare Information Gateway. (2003). *Reducing re-referral in unsubstantiated child protective services cases: Research to practice*. Washington, DC: U. S. Department of Health and Human Services, Children's bureau. Retrieved from http://www.childwelfare.gov/pubs/focus/researchtopractice

Child Welfare Information Gateway. (2008). *Differential Response to Reports of Child Abuse and Neglect*. Washington, DC: U.S. Department of Health and Human Services, Children's bureau. Retrieved from http://www.childwelfare.gov/pubs/issue_briefs/differential_response/differential_response.pdf

Child Welfare Information Gateway. (2010). *Family engagement*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau. Retrieved from http://www.childwelfare.gov/pubs/f_fam_engagement

Child Welfare Information Gateway. (2012). *Child abuse and neglect fatalities 2010: Statistics and interventions*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau. Retrieved from http://www.childwelfare/gov/pubs/factsheets/fatality.cfm

Child Welfare Information Gateway. (a) (undated). *Concept and history of permanency in U.S, child welfare*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau. Retrieved from http://www.childwelfare.gov/permanency/overview/history.cfm

Child Welfare Information Gateway. (b) (undated). *Partnering with parents*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau. Retrieved from http://www.childwelfare.gov/preventing/promoting/partnering.cfm

Child Welfare Information Gateway. (c) (undated). *Family-centered, neighborhood-based foster care*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau. Retrieved from http://www.childwelfare.gov/famcent ered/overview/approaches/fostercare.cfm

Children's Research Center. (2003). *Evaluation of Michigan's foster care structured decision making case management system*. Madison, WI: Author.

Children's Research Center. (2005). *Structured decision making™ in child welfare*. Madison, WI: Author.

Children's Research Center. (2009). *North Carolina department of health and human services risk assessment validation: A prospective study*. Madison, WI: Johnson,, K. & Bogie, A.

Comer, D. P., & Vassar, D. (2008). Six principles of partnership: Building and sustaining system-wide change. *Protecting Children*, *23*, 96–104.

Commonwealth of Virginia, Office of Social Services. (2008). *Evaluation of the differential response system*. Richmond, VA. Retrieved from http://www.dss.virginia.gov/files/about/reports/children/cps/all_other/2008/differentialresponsesystem_evaluation_annualreport_2008_12-08.pdf

Conley, A. (2007). Differential response: A critical examination of a secondary prevention model. *Children and Youth Services Review*, *29*, 1454–1468.

Conley, A., & Duerr Berrick, J. (2010). Community-based child abuse prevention: Outcomes associated with a differential response program in California. *Child Maltreatment*, *15*, 282–292.

Connell, C. M., Bergeron, N., Katz, K. H., Saunders, L., & Kraemer Tebes, J. (2007). Re-referral to child protective services: The influence of child, family, and case characteristics on risk status. *Child Abuse & Neglect*, *31*, 573–588.

Costello, T. (2008). *Proceedings from the 3rd annual conference on Differential Response in Child Welfare: Differential response* [PowerPoint slides]. Columbus, OH: National Resource Center for Child Protective Services.

Courtney, M., Dworsky, A., Cusick, G., Havlicek, I., Perez, A., & Keller, T. (2007). *Midwest evaluation of the adult functioning of former foster youth: Outcomes at age 21*. Chicago, IL: Chapin Hall Center for Children, University of Chicago.

Coyne, J. C. (2009). Keynote article: Are most positive findings in health psychology false. . . or at least somewhat exaggerated? *The European Health Psychologist*, *11*, 20–23.

English, D. J., Wingard, T., Marshall, D., Orme, M., & Orme, A. (2000). Alternative responses to child protective services: Emerging issues and concerns. *Child Abuse & Neglect*, *24*, 375–388.

Fanshel, D., & Shinn, E. (1978). *Children in foster care: A longitudinal investigation*. New York, NY: Columbia University Press.

Fluke, J. D., Shusterman, G. R., Hollinshead, D., & Yuan, Y. T. (2005). *Rereporting and recurrence of child maltreatment: Findings from NCANDS*. Washington, DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, 2005. Retrieved from http://aspe.hhs.gov/hsp/05/child-maltreat-rereporting/

Franke, T., Bagdasaryan, S., & Furman, W. (2011). *Differential response in rural counties: Path differentiation, service receipt and case disposition*. Journal of Public Child Welfare, *5*, 1–22.

Frankel, H. (1998). Famil6-centered home-based services in child protection: A review of the research. *Social Service Review*, *62*, 137–157.

Gambrill, E.(2005). Decision making in child welfare: Errors and their context. *Children and Youth Services Review*, *27*, 347–352.

Gambrill, E. (2008). Providing more effective, ethical services: The philosophy and process of evidence-based (informed) practice. In D. Lindsey & A. Shlonsky (Eds.), *Child welfare research: Advances for practice & policy* (pp. 51–65). New York, NY: Oxford University Press.

Gambrill, E. (2010). Evidence-informed practice: Antidote to propaganda in the helping professions? *Research on Social Work Practice*, *20*, 302–320.

Gambrill, E. (2011). Ethical aspects of outcome studies in social, behavioral and educational interventions. *Research on Social Work Practice*, *21*, 654–663.

Gilmore, D. A. (2010). *Issue brief 1: Beyond investigations: Current innovations in responding to reports of child maltreatment*. Retrieved from www.differentialresponseqic.org

Goodman, S., & Greenland, S. (2007). Why most published research findings are false: Problems in the analysis. *PloSMed*, *4*, 168.

Hartley, E. K. (1984). Government leadership to protect children from foster care "drift." *Child Abuse & Neglect*, *8*, 337–342.

Huebner, R. A. (2005). *Program evaluation of the multiple response system: Kentucky*. Unpublished report. Retrieved from http://site.americanhumane.org/site/DocServer/ProgramEvaluationKY.pdf?docID=4530

Huebner, R. A., Durbin, L., & Brock, A. (2009). Program evaluation of the multiple response system: Kentucky department for community based services. Retrieved from http://chfs.ky.gov/nr/rdonlyres/ba81ef9b-572d-4c59-ab4e-597eee3b0935/0/evaluationofthemultipleresponsesysteminkentucky_09.pdf

Hutchison, J. R., Lloyd, J. C., Landsman, M. J., Nelson, K., & Bryce, M. (1983). *Family centered social services: A model for child welfare agencies*. Oakdale, IA: National Resource Center on Family Based Services, University of Iowa School of Social Work.

Institute for Human Services (1990). *Universe of Child Welfare Competencies*. Columbus, OH: Author.

Institute for Human Services (1994). *Core Curriculum for Child Welfare Caseworkers*. Columbus, OH: Author.

Institute of Applied Research. (1997). *The Missouri Family Assessment and Response demonstration: Final evaluation report*. St. Louis, MO: Author.

Institute of Applied Research. (2000). *The Missouri Family Assessment and Response demonstration impact evaluation: Digest of findings and conclusions*. St. Louis, MO: Author.

Institute of Applied Research. (2004a). *Differential Response in Missouri after Five Years: Final Report*. St. Loius, MO: Author. Retrieved from http://www.iarstl.org/papers.htm#ancC5

Institute of Applied Research. (2004b). *Minnesota Alternative Response Evaluation: Final Report*. St. Louis, MO: Author. Retrieved from http://www.iarstl.org/papers.htm#ancC5Loman,

Institute of Applied Research. (2006). *Extended follow-up study of Minnesota's Family Assessment Response: Final report*. St. Louis, MO: Author.

Institute of Applied Research. (2009). *Differential Response in Nevada: Implementation and impact evaluation after 32 Months*. St. Louis, MO: Author.

Institute of Applied Research. (2010). Ohio Alternative Response Evaluation: Final Report. St. Louis, MO: Loman, L. A., Filonow, C. S., & Siegel, G. Retrieved from http://www.iarstl.org/papers.htm#ancC5

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PlosMed*, *2*, e–124. doi:10.1371/jojurnal.pmed.0020124

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*, 640–648.

Ioannidis, J. P. A. (2011). An epidemic of false claims: Competition and conflicts of interest distort too many medical findings. *Scientific American Magazine*, June. Retrieved from http://www.scientificamerican.com/article.cfm?id=an-epidemic-of-false-claims

Johnson, C., Sutton, E. S., & Thompson, D. M. (2005). Child welfare reform in Minnesota. *Protecting Children*, *20*, 55–60.

Johnson, W. (2004). Effectiveness of California's child welfare structured decision making model: A prospective study of the validity of the California family risk assessment. Sacramento: California Department of Social Services.

Jones, M. L. (1978) Stopping foster care drift: A review of legislation and special programs. *Child Welfare*, *57*, 571–580.

Kaplan, C., & Merkel-Holguin, L. (2008). Another look at the national study on differential response in child welfare. *Protecting Children*, *23*, 5–21.

Kirk, R. S. (2008). Development and field testing of a family assessment scale for use in child welfare practice settings utilizing differential response. *Protecting Children*, *23*, 71–87.

Kushel, M. B., Yen, L. G., & Courtney, M. E. (2007). Homelessness and health care access after emancipation: Results from the Midwest evaluation of adult functioning of former foster youth. *Archives of Pediatric Medicine*, *161*, 986–993.

Littell, J.H. (2005). Lessons from a systematic review of Multisystemic Therapy. *Children and Youth Services Review*, *27*, 445–463.

McGowan, B. G. (2005). Historical evolution of child welfare services. In G. Mallon & P. Hess (Eds.), *Child welfare for the twenty-first century: A handbook of practices, policies, and programs*, (pp. 10–40). New York, NY: Columbia University Press.

Merkel-Holguin, L. (2005). Differential Response: A common sense reform in child welfare. *Protecting Children*, *20*, 2–3.

Merkel-Holguin, L., Kaplan, C., & Kwak, A. (2006). *National study on differential response in child welfare*. Washington,

DC: American Humane Association and Child Welfare League of America. Retrieved from http://www.americanhumane.org/assets/docs/protecting-children/PC-DR-national-study2006.pdf

O'Neill Murray, K., & Gesriech, S. (2004). A brief legislative history of the child welfare system. Retrieved from http://www.pewtrusts.org/resource_library_browse.aspx? category+4

Pike, V. (1976). Permanent planning for foster children: The Oregon project. *Children Today*, *5*, 22–25.

Reich, D. A., Green, M. C., Brock, T. C., & Tetlock, P. E. (2007). Biases in research evaluation: Inflated assessment, oversight, or error-typ0e-weighting? *Journal of Experimental Psychology*, *43*, 633–640.

Rohm, A. (2008). Differential response: Progressive child welfare. *Protecting Children*, *23*, 3–4.

Rosenblatt, A. (1996). Bows and ribbons, tape and twine: Wrapping the wraparound process for children with multi-system needs. *Journal of Child and Family Studies*, *5*, 101–117.

Rubin, A. E., & Parrish, D. E. (2007). Problematic phrases in the conclusions of published outcome studies: Implications for evidence-based research. *Research on Social Work Practice*, *17*, 334–347.

Ruppel, J., Huang, Y., & Haulenbeek, G. (2011). *Differential Response in child protective services in New York State: Implementation, initial outcomes and impacts of pilot project*. Retrieved from http://www.ocfs.state.ny.us/main/reports/CPS%20Differential%20Response%20Evaluation%20Final%20Report%20Jan%202011.pdf

Rycus, J. S., & Hughes, R. C. (1998). *Field guide to child welfare*. Washington, DC: Child Welfare League of America.

Rycus, J. S., & Hughes, R. C. (2003). *Issues in risk assessment in child protective services: A policy white paper* (White paper). Columbus, OH: North American Resource Center for Child Welfare. Retrieved from www.narccw.com

Sawyer, R., & Lohrbach, S. (2005). Differential response in child protection: Selecting a pathway. *Protecting Children*, *20*, 44–54.

Schene, P. (2001). Meeting each family's needs: Using differential response in reports of child abuse and neglect. In *Best Practice, Next Practice, Spring*, (pp. 1–14). National Child Welfare Resource Center for Family-Centered Practice. Retrieved September 18, 2006, from www.hunter.cuny.edu/socwork/nrcfcpp/downloads/newsletter/ BPNPSpring01.pdf

Schene, P. (2005). The emergence of differential response. *Protecting Children*, *20*, 4–7.

Schene, P. (2007). *Differential response and data* [PowerPoint slides]. Retrieved from http://powershow.com/view/3e4cf-ZjMyN/Differential_Response_and_Data_powerpoint_ppt_presentation

Schene, P., & Kaplan, C. (2007). *Proceedings from the 2nd national conference on differential response: Getting started with differential response: fundamentals and first steps* (PowerPoint slides). Retrieved from http://www.americanhumane.org/assets/pdfs/children/differential-response/c-dr-conf-2007.pdf

Shusterman, G. R., Hollinshead, D., Fluke, J. D., & Yuan, Y. T. (2005). *Alternative responses to child maltreatment: Findings from NCANDS*. Washington DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. Retrieved from http://aspe.hhs.gov/hsp/05/child-maltreat-resp/

Sledjeski, E. M., Dierker, L. C., Brigham, R., & Breslin, E. (2008). The use of risk assessment to predict recurrent maltreatment: A classification and regression tree analysis (CART). *Prevention Science*, *9*, 28–37.

State of Washington Department of Social & Health Services (2005). *Alternative Response Systems Program Progress Report. July 1, 2003-June 30, 2004*. Olympia, WA: Author.

Stoltzfus, E. (2009). The Child Abuse Prevention and Treatment Act (CAPTA): Background, programs, and funding. *Congressional Research Service, November*. Retrieved from http://www.napcwa.org/Legislative/docs/CAPTACongressionalResearchReport.pdf

Thompson, D., Siegel, G. L., & Loman, L. A. (2008). The parent support outreach program: Minnesota's early intervention track. *Protecting Children*, *23*, 23–29. Retrieved from http://www.differentialresponsequic.org/resources/studies/the-parent-support-outreach.html

The Urban Institute. (2006). *Families' connections to services in an alternative response system*. Washington, DC: Author. Retrieved from http://www.urban.org/uploadedpdf/311397_families_connections.pdf

University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication No. 231). Tampa, FL: Author.

VanDenBerg, J. E., & Grealish, M. E. (1996). Individualized services and supports through the wraparound process: Philosophy and procedures. *Journal of Child and Family Studies*, *5*, 7–21.

Waldfogel, J. (1998). Rethinking the paradigm for child protection. *The Future of Children*, *8*, 104–120.

Ward, D. E., Maluccio, A., Hamilton, J., & Fein, E. (1982). Planning for permanency planning in foster care. *Children and Youth Services Review*, *4*, 223–237.

Yoo, J., & Meezan, W. (2001). Family preservation research: Where we've been, where we should be going. *Family Preservation Journal*, *5*, 25–42.

Yuan, Y. T. (2005). Potential policy implications of alternative response. *Protecting children*, *20*, 22–31.