

Notes on the Measurement, Assessment and Outcomes Theme
Final Plenary Panel and Discussion
Child Care Policy Research Consortium
Friday, March 11, 2005
Martha Zaslow

An important feature of the Child Care Policy Research Consortium is that the research that is presented is grounded in child care delivery systems and child care policy. Research questions and hypotheses often have a theoretical basis and orientation, but they are not purely theoretical; they also derive from the reality of policy and practice as these are occurring at the local, state and federal levels and are tested within these contexts.

This grounding in the reality of policy and practice was a consistent feature of the discussions at the Child Care Policy Research Consortium meeting that focused on issues of measurement. Measurement was viewed in the context of service delivery systems, and as actually implemented within them. Indeed, rather than a focus on individual measures, the focus at the meeting was on systems of data collection, and on the ways in which data collection needs to keep up with evolving service delivery systems. Rather than looking in a “decontextualized” way at the features of measures of quality or measures of child outcomes, the discussions at the meeting focused on how measures were affected by the contexts in which they were implemented, and in turn could affect these contexts.

Within this framework, I saw three underlying questions as running through many of the discussions and presentations at the meeting concerning measurement:

- (1) We are changing our systems of data collection. Do these changes have the potential to affect our service delivery?
- (2) Our systems of service delivery are changing. What do we need to do within our data collection to reflect these changes appropriately; to keep up with these changes?
- (3) How does the context in which a research tool is implemented affect the data derived from it?

I briefly summarize some of the discussions from the meeting that focused on each of these questions.

We are changing our systems of data collection. Do these changes have the potential to affect our service delivery?

- **Defining and measuring the content of training.** Important new steps are being taken in attempting to measure the *content* of community-based training as a facet of early childhood professional development. This is in part a response to discussions at previous Child Care Bureau sponsored meetings. As examples, the National Association of Child Care Resource and Referral Agencies (NACCRRA) is collecting data in a national survey of CCR&R directors

regarding the content of training delivered through CCR&Rs. There is a longer-term goal of introducing into NACCRRAware (NACCRRRA's software for data collection) a system for recording the content of training completed by individual child care providers. In addition, the National Registry Alliance, a consortium of state child care registries, has launched a new effort to collect data on the content of training using common terminology across registries in different states. These are changes to data collection that will affect whole data collection systems. Is it possible that collecting data on the content of training in a systematic manner, something that has not been done before, has the potential to alter the training that child care providers actually complete? For example, will it become clear for the first time that training tends to be concentrated in particular areas (such as children's cognitive development), with fewer providers completing training in other areas (such as perhaps child health issues in child care)? Might ongoing documentation of training result in the introduction of a requirement that individual child care providers complete training across differing content areas (a kind of distribution requirement)?

- **Documenting the qualification of trainers.** A change in our systems of ongoing data collection that was also discussed at the meeting was the collection of data regarding the qualifications of those who provide training in community-based settings. We are beginning to see a requirement that training will only be recognized if it is conducted by trainers with certain qualifications. Does the collection of data about trainer qualifications have the potential to affect the quality of training? Will trainers themselves complete further professional development, or will there be more careful selection of trainers?
- **Use of a common identifier for child care providers.** There was much discussion at the meeting regarding use of a common identifier for individual child care providers, using the common identifier to link data across multiple administrative data systems and potentially making it possible for data on professional development to follow individual providers across state lines if they move. Is it possible that individual providers will be better able to progress up a career ladder in early childhood if their professional development records can follow them when they move across states? Again, will changed and improved systems of data collection potentially affect quality?
- **Articulation.** At the meeting we heard about states (for example, Maryland), that are working on articulation between community-based training experiences and college credit in early childhood. Here again we need to ask if the improvement in a system of measurement (here of the way in which community-based training and experience translate into or accumulate into college credit) has the potential to alter our service delivery systems. In particular, is it possible that clarifying the rungs on the ladder of professional development in this way can foster progress up the ladder?

Our systems of service delivery are changing. What do we need to do within our data collection to reflect these changes appropriately?

- **The introduction of quality rating systems in many states.** One noteworthy change in the delivery of child care services is the development of quality rating systems in multiple states to help inform consumers about differing levels of child care quality. This change in service delivery in turn is creating challenges in terms of data collection and measurement. What should be the defining features of differing levels of quality? What are appropriate gradations, supported by the research, that are also important for consumers? There are discussions occurring in multiple states about meaningful distinctions between levels of quality. As one example, in Wisconsin, a particular challenge to measurement growing out of creation of a quality rating system is the need to come up with a reliable and valid way of distinguishing between the first two levels in the system. The second level is for facilities in compliance with licensing requirements, while the first level is for those with some problem with compliance. But what should the dividing line be? Should a minor issue with compliance that is readily addressed and corrected be grouped in the same level with a serious or recurrent problem with compliance? What should qualify as being out of compliance?
- **On-site consultation approaches to improving child care quality.** We are seeing the emergence in multiple states of on-site consultation approaches to improving child care quality. The underlying theory for this approach is the view that practice changes more readily when information is provided in the situation it is relevant to, and when behaviors are modeled and practiced with feedback, rather than only discussed. However the emergence of this approach to supporting child care quality is creating a challenge to measurement. What are the important features of this type of intervention that should be measured? What beyond number of visits is important to document? What are the features of on-site consultation that are important to its effectiveness?
- **Partnerships.** Another change in practice that was discussed at the consortium meeting was the greater reliance on partnerships across child care, Head Start and pre-kindergarten. An interesting challenge associated with this change in service delivery is that each type of early care and education has its own measurement tradition. For example, in terms of the quality of the early care and education environment, child care has tended to rely on the “ERS” scales (the ECERS, FDCRS and ITERS); while pre-kindergarten has tended to rely on the ELLCO to measure early language and literacy practices. An interesting challenge associated with the spread of partnerships is the question of how to develop or select measures of quality that will be respected within all traditions.

How does the context in which a research tool is implemented affect the data that is collected using it?

A theme at the consortium meeting was that it is not just whether an appropriate tool has been developed for a particular purpose, but how it works under ongoing conditions in the field, that will affect data quality.

- **Documenting ongoing training by child care providers: Goals for data collection in light of issues of feasibility and reliability.** As one example, we heard that for the QUINCE evaluation, a grid has been created for collecting detailed data about child care providers' ongoing training experiences. This grid was based closely on recommendations for data collection from the steering committee on defining and measuring professional development in the early childhood workforce. The grid seeks to collect data about multiple aspects of training. But experiences in the field in implementing the grid are raising the question of whether the data we are seeking to collect exceeds what child care providers can recall and report on accurately. We may need to take the "ideal" tool, and revise it in light of experiences with it in the field.
- **Training on the "ERS" scales when used as a tool for quality improvement.** We heard at the meeting that the usefulness of the "ERS" scales for improving child care quality is related to the extent of training in using the ERS scales that child care providers have received. The tool does not exist apart from the circumstances of its implementation. We tend to think about improving reliability of observation for researchers, but the discussions at the meeting raised the issue of precision of observations when the ERS scales are completed by practitioners for purposes of improving quality.
- **High stakes data collection has the potential to undermine the validity of a measurement tool.** There were discussions at the meeting about the potential to change the characteristics of a measure according to how data from the measure would be used. The example given was the pattern of intercorrelations between certain features of child care quality according to whether there were high stakes consequences linked with the measure. Child care ratio might be linked significantly associated with observed quality under naturalistic circumstances. But if a quality rating system uses ratio but not observed quality to differentiate the number of stars or the level of quality presented to the public, the potential exists that child care centers may work to improve their ratios, perhaps even hiring less qualified people than they otherwise might to meet a goal quickly. If this is the case, the normal relationship between ratio and observed quality could be altered or distorted. The need to understand measures within their contexts of implementation was stressed at the meeting.
- **Use of child care registries.** The participants at the meeting also stressed that a strategy for data collection such as a child care registry would only be fully useful if it was implemented widely enough within a state. Similarly, a voluntary quality rating system might not affect consumer choice or eventually the child care market unless a high enough proportion of child care

facilities chose to participate. There was discussion at the meeting of strategies to increase the use and therefore utility of voluntary systems. For example, the use of voluntary child care registries could be increased by using the registry as the basis for public recognition of attainment of new qualifications (for example through announcements in the newspaper). Child care facilities could be encouraged to advertise using categories and levels from the child care registry.

Challenges and promising directions in measurement

Both challenges and promising directions in measurement were identified at the meeting.

Challenges included:

- Protection of privacy when using a unique identifier for child care providers.
- The issue of when it is appropriate to expect a change in a child outcome measure in the context of an intervention to affect child care quality. How long does a child need to be exposed to improved quality before a change in developmental status can reasonably be anticipated?
- The question of whether we are measuring child outcomes in program accountability efforts when we should instead be measuring specific practices in the classroom.
- The issue of building into measures of professional development, not only training participated in or classes completed, but also whether practice in the caregiving environment has changed.
- The need to acknowledge that the stability and precision of child assessments is not as great with younger as with older children.
- The need to continue the work towards better definition and measurement of early childhood professional development. A key next step is to work towards a common core of measures for inclusion in national and state level surveys focusing on the early childhood workforce.

Promising directions included:

- The expansion of data collection on early childhood professional development in market rate surveys, as in the work in Texas.
- A relatively new project focusing on measurement of child outcomes among children with disabilities. The work of this project may also serve as a resource for assessment of children more generally.
- The development of child assessment measures focusing on processes (such as the way a child approaches a problem) rather than outcomes (whether a child gets the “right” answer).
- Work focusing on the potential importance of child care stability, taking a careful look at ways to measure child care stability in more meaningful ways than in the past. As one example, leaving but then returning to care by a grandparent perhaps should not be included in measures of instability.

