



AGENDA ITEM: 7

DATE OF MEETING: January 24, 2013

ACTION: _____

INFORMATION: X

OPTIONS FOR LONGITUDINAL RESEARCH

SUMMARY

During the July 2012 quarterly meeting of the California Children and Families Commission (First 5 California), Commissioners requested an assessment of the feasibility of longitudinal study designs for evaluation of State Commission-sponsored programs. Accordingly, this brief addresses the following topics:

- Longitudinal designs commonly used in research;
- Select longitudinal studies related to early childhood education;
- Applicability and feasibility of longitudinal designs relevant to evaluation of First 5 California programs; and
- Additional considerations for commissioners.

First 5 California staff welcomes further commentary and discussion on these topics.

BACKGROUND

Longitudinal Designs Commonly Used in Research

Longitudinal study designs involve the collection of data from the same individuals at different points in time in order to measure change across time. This contrasts with cross-sectional study designs that collect data at a single point in time from many individuals who differ in some pre-determined characteristic of interest (e.g. children of different ages or grades). Under a broad definition, many research activities such as surveys or ethnographies employ repeated cross-sectional designs and may be considered to be forms of longitudinal research (Bernard 2000).

Longitudinal studies can be separated into three basic types: trend, cohort, and panel study designs--and can vary in time dimension (prospective or retrospective). Table 1 beginning on page 9 provides an overview of designs.

- Trend designs are repeated cross-sectional points of data collection administered at regular intervals to detect changes in the population. Some examples are the Gallup Poll, the American Community Survey by the US Census Bureau, or the California Health Interview Survey (CHIS). Surveys, as one example of trend designs, examine *different samples of the population* at each interval. A limitation of sampling is that the researcher cannot be sure if observed change is due to

differences between samples or to actual changes in the population (see Babbie 2004 and Bernard 2000).

- Cohort studies examine change within a specific subset of the population over time by focusing on individuals who share some common temporal characteristic or experience (e.g. year of birth) (Goldstein 1968 and Babbie 2004). Data are collected across all members of the cohort and aggregated (Bauer 2004 and Ludlow et al. 2011). This approach reduces variability by grouping together individuals who may have similar experiences.
- Panel studies are the most powerful of longitudinal designs because they examine and track changes in the same group of individuals over time (Babbie 2004). Good examples of panel design in early childhood education research include landmark studies such as the Abecedarian Project and the High/Scope Perry Preschool study (these studies are experimental in design as well). The longitudinal panel design allows the researcher to better show cause-and-effect relationships between treatment and outcome variables and is the best design to examine individual development and change (Ludlow et al. 2011). Panel designs may be naturalistic, experimental, or a combination of both as in regression discontinuity designs (Schochet et al. 2010). By contrast, cross-sectional research is limited to examination of differences between participants, whereas panel research allows for examinations of change within and between individuals (Taplin 2005).

Implementing longitudinal research is challenging. Difficulties include costs associated with follow-up and tracking of participants through the course of the study; extensive data management associated with large-scale and long-term data collection efforts; attrition, or the risk that participants drop out of the study thereby reducing the representativeness of the original sample; and the necessary time commitment by researchers to implement a study from sample recruitment to final analyses (Ludlow et al. 2011, Anstey and Hofer 2004, and Taplin 2005).

When tracking changes through time, longitudinal studies may look forward or backward. Cohort and panel study designs are usually referred to as *prospective* study designs because they follow participants into the future and track changes as they actually occur (Taplin 2005). But, another type of longitudinal study design is the *retrospective* design in which researchers look into the past of a particular sample of people. Such designs are subject to limitations of availability of records and reliability of respondents' memories (Goldstein 1968). *Retrospective* studies may be somewhat less rigorous than *prospective* studies, but may be easier to conduct because they do not involve many of the drawbacks of the longitudinal panel design: attrition, cost, and duration of the study.

Select Longitudinal Studies Related to Early Childhood Education

Longitudinal designs have been used successfully in early childhood education (ECE) research. Research designs and findings for two longitudinal studies on children, and one study on teacher retention, are summarized below.

The Abecedarian Project

The Abecedarian Project focused on measuring the long-term developmental outcomes of children who participated in a high quality preschool program. The Abecedarian Project was a longitudinal panel study (prospective randomized trial) involving the comparison of two randomized groups (experimental treatment and control). The theoretical underpinnings of the project were drawn from general systems theory and ecological models that view human development as situated within a complex interplay of physical, social, and psychological systems and institutions such as the child, the family, the school system, and the state (Bertalanffy 1975, Bronfenbrenner 1986 in Campbell and Ramey 1994, Campbell et al. 2002). The main hypothesis addressed whether or not these systems could be enhanced to produce lasting positive developmental effects: “coming from this ‘improved’ environment, the child should enter school with a greater degree of school readiness and an enhanced likelihood of success.” (Campbell and Ramey 1994:684)

One hundred eleven infants from low-income families were originally recruited for the study and randomly assigned to either an experimental (treatment) or control group. Later in the study, approaching kindergarten entry, these two groups were again randomly assigned to either school-age intervention programs or control groups producing four total groups. A high quality curriculum was created focusing on language and pre-literacy skills to boost cognitive, perceptual-motor, and social development (Sparling and Lewis 1979 in Campbell and Ramey 1994). Children received on-site healthcare and parents were encouraged to take part in the preschool program as well. These children were (and continue to be) followed well into adulthood.

Intellectual outcomes persisted through 7, 12, and 15 years of age. Children in the treatment group continued to achieve significantly higher scores on various cognitive and intelligence tests at age 18. By age 21, treatment group participants had significantly more years of education, differed significantly in levels of employment, were more likely to be engaged in skilled jobs, were significantly less likely to have used marijuana, were almost three time more likely to attend a 4-year college, and had obtained significantly higher scores in reading, mathematics, and cognitive tests (Campbell et al. 2002, Campbell and Ramey 2007). Total cost of the program was estimated to be around \$67,000 per child, but estimated benefits amounted to \$158,278 per child (Campbell and Ramey 2007).

High/Scope Perry Preschool

The High/Scope Perry Preschool Project was a longitudinal panel study (prospective randomized trial) that examined the long-term effects of high-quality preschool on underprivileged children. One hundred twenty-three African American children, considered to be at high risk for school failure, were randomly assigned to treatment and control groups. Data were collected from both groups annually until age 11, and at ages 14, 15, 19, 27, and 40 (Schweinhart 2005). Perry Preschool was a full-day, full-year intensive quality preschool program. In terms of the research: “The scientific strength of

[the] study, its ability to assess preschool program effects even many years later, is due to an experimental design..."(Schweinhart 2007).

Program participants achieved greater school readiness and higher levels of school success (Schweinhart 2001). At age 27, the treatment group enjoyed more stable living arrangements; reported less reliance on social services; had significantly fewer lifetime arrests for violent crimes, property crimes, and drug crimes (Schweinhart 2005). At age 40, they had outperformed the control group on measures of level of education, median annual earnings, and home ownership; had spent significantly less time in prison; and experienced significantly lower levels of unemployment (Schweinhart 2005). In 2000 dollars, Perry Preschool cost an average of \$15,166 per participant, but the estimated return on that investment was calculated at \$258,888 per participant. Most of the saving came from the fact that male program participants generated 41% less crime costs (over \$700,000 less) (Schweinhart 2005, Schweinhart 2007).

Early Childhood Education Workforce Retention in California

A longitudinal study of the ECE workforce in California was conducted by Bridges (et al. 2011) as part of the Child-Care Retention Incentive (CRI) program. First 5 California contributed \$164 million to this incentive program from 2001 to 2004. This study was designed as a retrospective survey that allowed retention to be assessed by various categories (demographics, job characteristics, etc.). The First 5 California program Comprehensive Approaches to Raising Educational Standards, or CARES, and other similar programs such as AB 212, provide stipends to ECE teachers, directors, and classroom aids to encourage them to take advantage of professional development activities in hopes of reducing notoriously high rates of turnover in the ECE field. In the CRI study, 2,783 participant histories were reconstructed for three years of the program (2001 to 2004) to examine professional development and job turnover. A key research question was whether or not CRI was effective in reducing job turnover, which was defined as either leaving the field or moving to a new job within the field. A purposive sample of ten counties that had CRI programs in place was selected in the initial sampling phase. This selection procedure took into account factors such as location of the county, population, urban or rural, eligibility requirements, variations in core program features, and the stipend scale. All program participants in those counties were contacted and invited to take part in the study. Data collection for the study involved 25-minute telephone interviews covering demographics, current job status, education activities, and training activities. To measure retention, respondents were asked specifically about where they were employed, how long they had worked there, and whether they had left the center or the field altogether and their reasoning behind the decision. The retrospective survey design allowed the researchers to make an important distinction between staff who had moved to a different program or position within the ECE field from those who had left the field altogether. Both activities are defined as "turnover" because they both disrupt the continuity of care that is thought to be so crucial to the development of young children.

The researchers found that more than a quarter of classroom aids left their original center within two years, and 12% left the field altogether. Whites were more likely to leave the field at 13% compared to 9% of African American and Latinos; Asians were also less likely to leave the field than Whites. Higher paid participants with more experience were significantly more likely to move within the ECE field; these same participants were also less likely to leave the field. Support also played a role as participants in programs with higher levels of advising were more likely to keep working at their centers. Participants reported high levels of dedication to the field regardless of whether they received a stipend.

DISCUSSION

Applicability and Feasibility of Longitudinal Designs Relevant to Evaluation of First 5 California Programs

How might various longitudinal designs apply to evaluation of current First 5 California programs? What questions would be addressed? Which designs might be practicably feasible? As mentioned in the overview of designs, key considerations are costs and effort related to data management, tracking of individuals, potential subject attrition, and long-term commitment.

With regard to the Child Signature Program (CSP), evaluation data are presently collected for a trend, or repeated cross-section, design. An in-depth cohort or panel design for longitudinal study is not readily feasible.

- First, data necessary for a long-term study on outcomes for individual children are not currently collected. No individual data on children or families are currently collected because the evaluation focuses on improvements in classroom quality. A database and evaluation infrastructure would have to be developed to collect data on individual children and families. Because of the complexities involved in tracking children and families over time, such an effort would require advance planning and, potentially, use of contracted evaluators.
- Further, evaluation priorities should be considered. Because the classroom is the intervention point of the program, it is reasonable for current evaluation efforts to first establish that classroom quality has improved. Data are currently collected at the classroom level, including data from Environment Rating Scales (ERS), CLASS instruments, and other measures. These data are collected annually for selected evaluation classrooms. For CSP, all Quality Enhanced classrooms and a random sample of Maintenance of Effort classrooms comprise the set of classrooms to be evaluated.

With regard to the Teacher Signature Program (CARES Plus), evaluation data are presently collected for a trend, or repeated cross-section, design. Some form of longitudinal study, such as a periodic retrospective cohort design, is potentially feasible.

- First, since teachers are the intervention point of the program, demographic data are currently collected for all individual teachers as they participate annually. These data provide a framework for potential longitudinal follow-up. Data collected includes demographics, work experience, and pre- and post-training CLASS scores for a sample of participants in program Components A and B, and for all participants in program Component D (My Teaching Partner). First 5 California may also add a teacher participant satisfaction questionnaire as a standard element of program participation.
- Additionally, First 5 California might reasonably consider some kind of retrospective evaluation for program graduates in the manner of the Bridges et al. (2011) study. Using existing data collected about teachers who participate in the program, a retrospective cohort evaluation could be constructed using a variety of methods including: email survey, mail survey, or in-person interview. The human subjects' evaluation protocol would need to be modified to inform and request consent of teachers for potential follow-up after their program participation. A retrospective survey, conducted at regular multi-year intervals, might provide data to measure teacher retention in the field of Early Childhood Education, one of the key goals of the program.

Additional Considerations for Commissioners

As context for discussion of longitudinal evaluation for First 5 California programs, Commissioners might consider the following issues:

- Current evaluation designs for both CSP and CARES Plus offer some longitudinal data through repeated cross-sectional data collection. For CSP, the current design can provide information about whether classroom quality is increasing from year to year and how this may differ by classroom resources (Quality Enhanced versus Maintenance of Effort classrooms). For CARES Plus, the current design can provide information about whether teacher effectiveness has increased for individuals (pre/post CLASS assessments in one year) and in aggregate by program component (A, B, and D). Should these current designs be augmented by more costly and complex designs such as cohort or panel designs? For example, answering questions about whether teacher participation in the CARES Plus was the cause of participant children's growth would require experimental designs like the Abecedarian and Perry preschool studies.
- Should State Commission programs seek to replicate the outcomes found by previous academic research? The argument may be made that government agencies might best devote efforts toward developing programs based on demonstrated successes published in peer-reviewed academic research--and then showing that those programs were indeed implemented with fidelity, rather than replicating original research.

- Some evaluation designs require long-term funding. Given the limited terms of Commissioner service, will there be sufficient long term commitment on behalf of the Commission to support longitudinal studies over multiple years? Commissioner turn-over may impact the funding sustainability of long-term evaluation efforts.
- What kind of partnerships might be needed to conduct longitudinal evaluations? Longitudinal evaluations can be complex and will likely require university or other specialized research partners.
- For what length of time should longitudinal studies be conducted? The longer the study, the greater the likelihood of research subject attrition—and costs will increase for data management and locating individuals who are part of the evaluation. More “invasive” location procedures may be needed to follow-up with subjects (parents, children, and teachers).

First 5 California staff welcomes further commentary and discussion by commissioners on these topics.

STAFF RECOMMENDATION

- For the Child Signature Program, an in-depth longitudinal study of individual children is not readily feasible at this point in time.
- For the CARES Plus Program, First 5 California might reasonably consider some kind of retrospective evaluation for program graduates in the manner of the Bridges et al. (2011) study.

ACKNOWLEDGMENTS

This brief was prepared by Robert Dean, MA, and David Dodds, PhD, MPH, of First 5 California. First 5 California gratefully acknowledges comments and suggestions provided by Dr. Donna Bryant, Senior Scientist at the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill, during preparation of this document.

REFERENCES CITED

- Babbie, Earl. 2004. *The Practice of Social Research*. Thompson Wadsworth: Belmonte, CA
- Bernard, Russell H. 2000. *Social Research Methods: Qualitative and Quantitative Approaches*. Sage Publications: Thousand Oaks, CA
- Bertalanffy, L. V. 1975. *Perspectives on General Systems Theory*. New York: Braziller
- Bridges, Margaret, Bruce Fuller, Danny S. Huang, Bridget K. Hamre. 2011. Strengthening the Early Childhood Workforce: How Wage Incentives May Boost Training and Job Stability. *Early Education and Development*, 22(6), 1009–1029.
- Bridges, Margaret, Natasha Dagys, Jennifer Ly, and Bruce Fuller. 2012. Features of Quality in Early Care and Education: Recent, Rigorous Evidence on What Matters Most for Children. University of California, Berkeley, with UCLA Center for Healthier Children, Families, and Communities. Research brief for First 5 California, Sacramento.
- Bronfenbrenner, U. 1986. "Ecology of the Family as a Context for Human Development: Research Perspectives." *Developmental Psychology*. 22:723-742
- Campbell, Frances A. and Craig T. Ramey. 1994. "Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children from Low-Income Families." *Child Development*. 65(2):684-698
- Campbell, Frances A. and Craig T. Ramey. 2007. Carolina Abecedarian Project. Presentation at the National Invitational Conference of the Early Childhood Research Collaborative. Accessed 12/12/12 from:
<http://earlychildhoodrc.org/events/presentations/campbell-ramey.pdf>
- Campbell, Frances A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. 2002. "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project." *Applied Developmental Science*. 6(1):42-57
- Goldstein H. 1968. "Longitudinal Studies and the Measurement of Change." *The Statistician*. 18(2):93-117
- Ludlow, Larry H., Joseph Pedulla, Emile Mitescu Reagan, Sarah Enterline, Mac Cannady, and Stephanie Chappe. 2011. "Design and Implementation Issues in Longitudinal Research." EPAA AAPE Education Policy Analysis Archives. 19(11):1-18
- Schweinhart, Lawrence J. 2005. "The High/Scope Perry Preschool Study Through Age 40: Summary, conclusions, and Frequently Asked Questions". Accessed 12/12/12 from:
http://www.highscope.org/file/Research/PerryProject/specialsummary_rev2011_02_2.pdf
- Schweinhart, Lawrence J. 2007. "Crime Prevention by the High/Scope Perry Preschool Program". *Victims and Offenders*. 2:141-160

Singer, Judith D. and John B. Willett. 1996. "Methodological Issues in the Design of Longitudinal Research: Principles and Recommendations for a Quantitative Study of Teachers' Careers", *Educational Evaluation and Policy Analysis*. 18(4):265-283

Schochet, P., Cook, T., Deke, J., Imbens, G., Lockwood, J.R., Porter, J., Smith, J. (2010). Standards for Regression Discontinuity Designs. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf.

Taplin, Stephanie. 2005. "Methodological Design Issues in Longitudinal Studies of Children and Young People in Out-of-Home Care". NSW Centre for Parenting & Research, NSW Department of Community Services.

Table 1. *Strengths and Limitations of Different Types of Longitudinal Study Designs*

Time Dimension	Study Design	Strengths	Limitations
Prospective	Trend: Repeated Cross- Sectional	Can measure change across time Repeated at regular intervals Flexible Relatively low cost Avoids attrition Common (e.g., survey research)	Measures aggregate change across the population
Prospective	Cohort	Can measure change across time May show time order between treatment and outcome across the cohort Controls for shared effects due to exposures in a similar time period (e.g., birth year)	Measures aggregate change for the cohort, not the population.

Table 1 (continued)

Prospective	Panel Type: Naturalistic Observation	Can measure change at individual level as individuals or local programs categorize themselves “naturalistically” May provide opportunity to show cause and effect relationships as they “naturally” occur	Requires major planning High cost Necessitates a long term institutional commitment Many challenges to tracking subjects throughout life Attrition Higher scrutiny for human subjects ethics protocol (Institutional Review Board) Long-term data storage effort
-------------	--	--	--

Table 1 (continued)

Prospective	Panel Type: Randomized Trial	Can measure change at individual level	Requires a control group. Politically and ethically difficult to implement in a government program.
		Experimental design allows for strong comparisons between treatment and control groups	Requires randomization to control and treatment groups
		Provides opportunity to show cause and effect relationships	Requires major planning
		Generally regarded as the most powerful and scientific type of study design	High cost
			Necessitates a long term institutional commitment
			Many challenges to tracking subjects throughout life
			Attrition
			Higher scrutiny for human subjects ethics protocol (Institutional Review Board)
			Long-term data storage effort
			Might require data collection from sources beyond First 5

Table 1 (continued)

Prospective	Panel Type: Naturalistic and Randomized (Regression Discontinuity)	Combines elements of naturalistic and randomized trial designs Has been used in education evaluation	Requires major planning High cost Necessitates a long term institutional commitment Many challenges to tracking subjects throughout life Attrition Higher scrutiny for human subjects ethics protocol (Institutional Review Board) Long-term data storage effort
Retrospective	Cohort or Panel	Can measure change at the individual level Short term commitment (research is conducted over a short period looking backward in time) Relatively low cost	May not be able to locate all program participants after the period of program participation Depends on availability of internal and external information (e.g., administrative records) Depends on reliability of program participant memories