

4 Adequacy of intervention design and implementation



It is important to interpret evaluation findings in the context of the strengths and weaknesses of the intervention design and implementation, as well as the strengths and weaknesses of the evaluation methodology. If an intervention has not been designed well or implemented according to plan, any evaluation of the program will be misrepresentative of the program it purports to be evaluating (Mrazek and Brown 2002).

Dryfoos (1990) suggests that effective programs are “high-dose” and involve a structured curriculum. These issues, as well as participation rates, program integrity (the extent to which the program was delivered as intended) and drop-out rates/attrition are discussed in relation to the five intervention clusters below.

Dosage of programs

Dosage refers to the amount of intervention provided. Encompassed within the term dosage are the concepts of intensity and duration. For example, participants may receive the same “dosage” from an intensive intervention implemented over a short duration, and a less intensive intervention over a longer duration.

In early childhood interventions, the evidence supports the notion that “more is better” (Berlin, O’Neal and Brooks-Gunn 1998: 8). Interventions that are high on intensity and duration are thought to be more effective than those that are less intense, and run for a shorter duration. For example, Reynolds (1994) suggests that intervention effects are stronger and more lasting for programs that are of three to four years duration, compared to those of only one year duration. Other researchers have also suggested that programs of short duration have limited effects and that effects are more likely to be sustained for programs that are intensive and continue into the school-age years (Brooks-Gunn 2003; Fonagy 2001). The dosage of programs included in the current report was highly variable.

Cluster 1: targeted, child focused, centre based, preschool age. Children in cluster 1 interventions typically received a high dose of intervention, receiving at least part-day child care or education five days a week for the most part of a year. The Perry Preschool Project and the High/Scope preschool curriculum study also included weekly or fortnightly home visits. Interventions were offered for at least one year, with some children receiving the intervention for up to six years.

Cluster 2: targeted, parent focused, home visits, all ages. Parents in cluster 2 interventions also received a reasonably high dose of intervention. Contact with parents was at least weekly, although this varied from phase to phase. For example, the Elmira PEIP began with weekly home visits, reduced to fortnightly visits, went back to weekly visits during the six weeks after birth and then gradually became less frequent over time. In terms of duration, programs ranged from weekly sessions over 16 weeks to regular contact over three years.

Cluster 3: targeted, family economic/welfare focused, all ages. Interventions in cluster 3 were typically medium-dose. Although the intensity was low (actual person-to-person contact was minimal, and took the form of a meeting with a case manager who provided support and assistance in finding employment), programs ran for approximately two years.

Cluster 4: targeted, holistic, various locations, all ages. Participants in cluster 4 typically received a high-dose intervention, although the nature of the intervention (for example, centre based and

home visiting), varied from family to family according to need (for example, Early Head Start, Sure Start and NEWPIN). Services targeted at children were usually the most intensive (often five days a week), while parent services were less intensive (weekly to fortnightly). The programs operated from 22 weeks to five years.

Cluster 5: universal, various foci, various locations, all ages. The intensity of programs in intervention cluster 5 varied widely, even within the same program. For example, the Triple P program ranged from very low dose (parenting information communicated via the media) to weekly parent training sessions over ten weeks, while the PAT program lasted three years.

Participants in clusters 1 and 4 interventions received the highest dose of intervention, suggesting that these interventions may be the most effective (Reynolds 1994; Berlin, O'Neal and Brooks-Gunn 1998).

Participation rates

Involvement in early childhood interventions is usually voluntary, thus participation rates (either full or part-participation) can vary dramatically from program to program. Low participation among those expected to benefit the most from an intervention (children from low-income families, for example) will most likely result in negatively skewed effects (that is, effects will not be as positive as expected), whereas higher participation rates are often associated with better outcomes (see Berlin, O'Neal and Brooks-Gunn 1998). Although what constitutes low participation is not made explicit in the literature, low participation of a target group leads to participation threat to the evaluation design (Mrazek and Brown 2002; Berlin, O'Neal and Brooks-Gunn 1998). Participation rates for programs included in the current report are discussed below.

Cluster 1: targeted, child focused, centre based, preschool age. Participation in cluster 1 interventions was not always reported. Among programs reporting this information, participation rates were reasonably high. The Perry Preschool Project reported a 69 per cent full attendance rate and the High/Scope Preschool Curriculum study reported an 80 per cent participation rate in home visits.

Cluster 2: targeted, parent focused, home visits, all ages. Of the interventions in cluster 2 that reported participation rates, participation was quite low. For example, the Elmira PEIP reported that an average of 23 home visits were conducted between birth and two years, with a range of 0 to 59 (59 visits were specified by the program), while Hawaii's Healthy Start Program reported that very few families were visited weekly, as intended. Other programs, for example HIPPI, had difficulty determining participation rates.

Cluster 3, targeted, family economic/welfare focused, all ages. Given that financial incentives were used to encourage participation, the rates of participation in the programs in cluster 3 were typically near 100 per cent.

Cluster 4, targeted, holistic, various locations, all ages. Participation rates were not often reported for cluster 4. The available information indicated extreme variation in participation rates.

Cluster 5: universal, various foci, various locations, all ages. Participation rates in cluster 5 interventions were not easy to measure (for the media communication strategy of Triple P, for example). However, problems with non-attendance at groups and lack of success in phone contacts were documented. Participation in the Cuyahoga program was very high during the first three months; however no additional information was available. Participation information was not available for PAT.

In all clusters except cluster 3, participation was variable. The higher rates of participation in cluster 3 are most likely due to the low intensity of the interventions and the financial incentives for participation. The evaluations of those interventions with very low participation rates need to be interpreted with caution, as they are based on participants who did not receive the full intervention or did not participate in the intervention at all. In addition, there may be systematic differences between families who participated in the intervention and those that did not that could bias evaluation findings.

Drop-out rates

“Drop-out rates” refer to participants who began, but did not complete the full intervention. High drop-out rates pose an attrition threat to the evaluation design (Mrazek and Brown 2002), and can result in positively skewed findings, as participants who drop out may do so because they have not found the intervention acceptable or useful. However, it is often possible to statistically account for drop-out rates in an evaluation by comparing characteristics of participants who dropped out to the characteristics of participants who continued with the intervention, and controlling for any differences between the two groups. Drop-out rates for the programs reviewed in the current report are outlined below.

Cluster 1: targeted, child focused, centre based, preschool age. Drop-out rates were not reported for interventions in cluster 1.

Cluster 2: targeted, parent focused, home visits, all ages. Of the cluster 2 interventions that reported drop-out rates, the rates were quite high, ranging from 40 per cent to 69 per cent.

Cluster 3: targeted, family economic/welfare focused, all ages. Given the financial incentives, as well as the mandatory nature of some of the programs, drop-out rates in cluster 3 were close to zero.

Cluster 4: targeted, holistic, various locations, all ages. Drop-out rates in the interventions in cluster 4 were high, ranging from 24 per cent to 67 per cent.

Cluster 5: universal, various foci, various locations, all ages. Drop-out rates were not reported for any of the interventions in cluster 5.

In summary, drop-out rates when reported were generally high across all clusters, except for those interventions in cluster 3. This reflects the fact that these interventions were often mandatory, or involved some type of financial incentive for participation.

Program integrity

Program integrity refers to the implementation of the program according to its design (that is, the same content, delivered in the same way). The quality of program implementation is perceived to influence program effectiveness (Shonkoff and Phillips 2000), and poor implementation leads to an implementation threat to the evaluation design (Mrazek and Brown 2002). Staff qualifications, staff to child ratios and staff turnover are aspects of program implementation that can affect program integrity. Schorr (1997), for example, suggests that higher qualified and more experienced staff result in greater program effectiveness (see also Berlin, O’Neal and Brooks-Gunn 1998). Tomison and Wise (1999) suggest that professional staff are particularly necessary when dealing with very vulnerable families, or where there is a risk of child maltreatment.

Although it is far more difficult to evaluate programs that do not follow a strict curriculum (such as programs in cluster 5 which are tailored to family and community needs), flexibility in program delivery may be necessary to meet the specific needs of individuals and families. The integrity of programs under review in this report are discussed below.

Cluster 1: targeted, child focused, centre based, preschool age. Cluster 1 interventions were generally highly standardised and of high quality. Low staff to child ratios (ranging from 1:5 to 1:8 depending upon children’s ages) were employed, and staff were highly qualified. Staff turnover rates, however, could have been improved upon. The Perry Preschool Project reported that ten teachers occupied four positions over five years and the High/Scope Preschool Curriculum study reported that new teachers were appointed in the second year of the study. Three of the six interventions involved set programs. The remaining three, although guided by strict protocols, were not implemented in the same manner from site to site, meaning that there was not one consistent program to evaluate.

Cluster 2: targeted, parent focused, home visits, all ages. In most cases, the staffing of programs in cluster 2 involved paraprofessionals (lay people trained specifically to implement the program), often

from the same community as participants. Three programs were staffed by professionals; the Elmira PEIP, Baby HUGS and Project 12-ways. Although most interventions had some set guidelines, most were not standardised but administered according to individual or community need.

Cluster 3: targeted, family economic/welfare focused, all ages. Interventions in cluster 3 were staffed by employees of the relevant social services department, and were typically trained social workers. Program content varied from participant to participant, with some participants attending very few sessions and others attending quite a number. However, implementation was successful in terms of applying the specified financial incentives or disincentives.

Cluster 4: targeted, holistic, various locations, all ages. Cluster 4 interventions ranged from highly structured and standardised (for example, Incredible Years) to highly unstructured and non-standardised (for example, Even Start and SESS). Staff to child ratios in child care centres were good to very good; usually about 1:3 for infants and 1:6 for preschool aged children. Professionals or paraprofessionals most often implemented the programs and typically received ongoing training and/or supervision. For example, all staff (including drivers and cooks) of the Syracuse FDRP received two weeks of intensive training each year.

Cluster 5: universal, various foci, various locations, all ages. The interventions in cluster 5 were typically administered by professionals, although, with the exception of Triple P, the program content varied from participant to participant (for example, Cuyahoga).

In summary, most interventions were administered by professional or paraprofessional staff, which is likely to enhance intervention effectiveness (Schorr 1997). Most interventions had some form of flexibility in-built, although most followed some form of protocol. The most standardised interventions were found in cluster 1. Evaluations of the interventions that were not standardised need to be interpreted with care, as an implementation threat to the design may be present (Mrazek and Brown 2002).

Summary

Overall, very few of the programs reviewed in this report were adequate in all areas of design and implementation. As suggested above, although the adequacy of interventions within clusters was variable, the design and implementation of interventions in cluster 1 appear to be the most adequate across all aspects of design and implementation. Although dosage levels were high, other aspects of design and implementation were quite poor for cluster 2 interventions. Although most aspects of design and implementation were adequate for interventions in cluster 3, intensity levels were low. Given the great variability in cluster 4, it is difficult to draw any conclusions about the adequacy of targeted, holistic interventions. Similarly, it is difficult to draw conclusions about the adequacy of universal interventions in cluster 5 because of the limited information available about their design and implementation.