

Assessing Change in Families Following the Home-Start Parenting Program: Clinical Significance and Predictors of Change

Jessica J. Asscher

Maja Deković

Peter Prinzie

Jo M.A. Hermanns*

Abstract: The aim of the present study was to examine whether improvements obtained after a home-based parenting intervention reflected meaningful and significant changes for a sample of Dutch mothers. The results showed that Home-Start mothers demonstrated reliable changes in well-being and enhanced parenting behaviors compared to both a comparison and a norm group of mothers. At posttest, a substantial number (39 – 84%) of the Home-Start mothers functioned at a level equivalent to that of a community group. The most reliable improvement was found with mothers experiencing the most severe problems at pretest, whereas the most recovery was reported for mothers with the fewest problems at pretest.

Key Words: clinically significant change, Home-Start, mothers, parenting behavior, parenting support program, reliable change.

Background and Significance

The first 5 years of life are of crucial importance for the development of both a sense of emotional security and the acquisition of self-regulation skills later in life (Repetti, Taylor, & Seeman, 2002). Parenting behavior has been treated as a theoretical and empirical determinant of emotional well-being in early childhood and beyond (Papp, Cummings, & Schermerhorn, 2004). In general, parenting styles that encompass unpredictable parental behavior, lack of responsiveness and warmth, harsh discipline, and a lack of supervision of the child's activities have been linked to a host of negative outcomes for the child (Olson, Ceballo, & Park, 2002). These negative child outcomes include the development of antisocial behavior, social rejection, academic failure, and membership in deviant peer groups later on in life (Ehrensaft et al., 2003).

Given the well-documented association between parenting competence and child outcomes (see, e.g.,

Belsky, 1984), numerous early intervention and parenting support programs have been developed and implemented to counter potential negative outcomes and to support healthy developmental progress in families with young children. The long-term goal of many clinical interventions conducted by (mental health) professionals has been to ameliorate family dysfunction and behavioral problems of the child (Osofsky, 1998). In addition to clinical programs, a large number of home-visiting parenting support programs have been developed. These home-based parenting support programs show particular promise in that the method of delivering services to families in their own homes offers opportunities for more personalized service, which not only aids families but also increases program-retention rates (see reviews by Bilukha et al., 2005; Diamond & Josephson, 2005). Moreover, it is hypothesized that parents feel more at ease in their own homes (McGuigan,

*Jessica J. Asscher is a postdoctoral fellow in the Faculty of Social Sciences in the Research Centre Psychosocial Development in Context at the Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands (j.j.asscher@uu.nl). Maja Deković is a full professor in the Faculty of Social Sciences in the Research Centre Psychosocial Development in Context at the Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands (m.dekovic@uu.nl). Peter Prinzie is an associate professor in the Faculty of Social Sciences in the Research Centre Psychosocial Development in Context at the Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands (p.prinzie@uu.nl). Jo M.A. Hermanns is a full professor in the Faculty of Social and Behavioural Sciences at the University of Amsterdam, P.O. Box 94208, 1090 GE Amsterdam, The Netherlands (j.m.a.hermanns@uva.nl).

Katzev, & Pratt, 2003) and that these programs offer the opportunity to reach socially or geographically isolated people.

According to Thompson, Kropenske, Heinicke, Gomby, and Halfon (2001), another important feature of many home-based programs is their reliance on volunteers for staffing. There are several reasons why volunteers are used in this form of parenting intervention. First, volunteers are typically perceived by clients as more accessible and less threatening than professionals (Kelleher & Johnson, 2004), which may result in people feeling more at ease, leading to increased responsiveness to the intervention and less dropout. A second reason is the relatively low cost of programs staffed by volunteers, allowing for the provision of this service for many families. As a result, volunteer-based family support programs are widespread; for example, the Home-Start program is active in 17 countries on five continents and in the Netherlands has 52 locations (Home-Start international website: <http://www.homestartinternational.org/>). Given the fact that many families are served by volunteer-based family support programs, it is of great importance to evaluate the effectiveness of these programs. At this time, however, there is limited empirical evaluation of volunteer-based family support programs (Barnes, MacPherson, & Senior, 2006). The limited number of studies that have examined the effectiveness of these programs often suffer from a variety of methodological shortcomings, making it difficult to establish the efficacy of volunteer staffed interventions aimed at improving parenting. These shortcomings, which are common in therapeutic evaluation research (Bamberger, Rugh, & Mabry, 2006), include the use of only postintervention self-reports (Hermanns, Van de Venne, & Leseman, 1997) and the use of small samples (e.g., Kelleher & Johnson). Moreover, the results have been mixed; that is, positive effects were shown only for some of the assessed outcomes (Barnet, Duggan, Devoe, & Burrell, 2002; Johnson et al., 2000; Kelleher & Johnson).

The present study focused on the Home-Start Program, a worldwide volunteer-based home-visiting program designed to support parents with young children. Using self-report and observational data, we examined the program's short- and long-term effectiveness with regard to a broad range of family outcomes (maternal well-being, parenting behaviors, and child problem behavior). According to its developers, Home-Start aims to improve maternal

well-being, which supposedly results in a decrease of negative parenting behaviors, resulting eventually in a reduction of child behavior problems. Lower perceived parental well-being (i.e., lower competence and higher depressive moods) has repeatedly been associated with less adequate parenting (Dawson et al., 2003) and maternal perception of child difficulties (Coleman, 1999).

In the present study, the following outcome variables were studied; first, maternal responsiveness was examined because it promotes healthy development (Domitrovich & Bierman, 2001; Pettit, Bates, & Dodge, 1997) and is inversely related to child behavior problems (Johnston, Murray, Hinshaw, Pelham, & Hoza, 2002). Second, consistency was examined, as it acts as a buffer against the negative effects of stressors (Wolchik, Wilcox, Tein, & Sandler, 2000) and is inversely related to the development of child behavior problems (Owens & Shaw, 2003; Solomonica-Levi, Yirmiya, Erel, & Oppenheim, 2001). Third, the levels of negative control were assessed because negative discipline techniques and high levels of rejection are associated with higher levels of disruptive child behavior (Thompson, Hollis, & Richards, 2003). Finally, we evaluated child behavior problems and negative child behavior because the ultimate aim of early intervention is the improvement of the behaviors for prolonged periods of time, and child behavior problems during the early years are believed to have lasting effects and can be influenced by maternal well-being as well as by parenting behavior.

Home-Start

The Home-Start Program aims to improve family functioning by means of assisting and emotionally supporting parents (e.g., by listening and talking to the mothers, helping out with housekeeping, or occasionally taking care of children). The program works with volunteers who visit mothers once a week for half a day during a period of 6 months, on average. Provision of social support by the volunteers aims to increase maternal well-being. Increased maternal well-being is assumed to result in more positive parenting behavior, which in turn ought to lead to a reduction of behavioral problems in children. Previous evaluations of the Home-Start intervention in both the United Kingdom and the Netherlands showed mixed results. Increased maternal well-being and competence, improved social networks, and

improved parenting behavior have been reported (Frost, Johnson, Stein, & Wallis, 2000). However, Barnes et al. (2006) found no evidence for enhanced parenting, organization of the home environment, or more appropriate use of health services in families that participated in Home-Start as compared to families in a region where Home-Start was not offered. They did report, however, a reduction in parent-child relationship difficulties. McAuley, Knapp, Beecham, and McCurry (2004) reported some improvements in both the Home-Start and the comparison group; therefore, enhanced functioning could not be attributed to Home-Start.

Clinical Significance of Change Following Intervention

In most of the studies in the field of evaluation of family programs, researchers compared the intervention and the control groups on mean levels of outcome measures, a common method to examine program effectiveness (Gomby, Culross, & Behrman, 1999). Usually, a program is only considered effective on the basis of Significant Group \times Time interactions, which would indicate significant changes in the intervention group and no changes in the control group. However, when conducting this kind of analysis, researchers tend to overlook the most important question concerning the effectiveness of a program, which is whether or not changes that have been achieved are meaningful; that is, whether or not program changes correspond to “real” change in everyday life (Kendall, 1999). A second shortcoming of comparing means between groups is that these tests do not provide information on variability in response to the treatment within the sample (Jacobson & Truax, 1991). In addition, an often-overlooked reality in this context is that despite positive expectations, interventions might also have harmful effects on families. For example, MacMillan et al. (2005) reported higher levels of child abuse in an intervention group than in the control group. If some families improve and some deteriorate and evaluators only study group means, the program might seem noneffective, while in fact it is effective for a selection of the participants but harmful for others. This risk is especially prevalent when nonprofessionals are the service providers because the level of professionalism of the service providers is often mentioned as a core characteristic of program success (Schoenwald, Sheidow, & Letourneau, 2004).

Aims of the Present Study

In our previously published article (Asscher, Hermanns, & Deković, 2008), we examined the effectiveness of Home-Start by using a classical approach, that is, we compared Home-Start mothers, as a group, to a comparison group of mothers who reported similar levels of stress and need for parental support. Results showed significant effects for maternal competence but no effect for depressive moods. Mixed results were found for parenting behavior. These classical analyses, focusing on groups means, did provide valuable information at the group level, but they did not provide information about possible variability among mothers in response to Home-Start; neither did they allow for examination of the clinical significance or the relevance of change.

In the present study, we were interested in determining the clinical relevance of changes that occurred as a result of the Home-Start program. Specifically, our first research question was whether or not the amount of change that has occurred—presumably because of the treatment—was large enough to be meaningful (i.e., show improvement; Ogles, Lunnen, & Bonesteel, 2001). Rather than focusing on the group-level analysis, as in our previous article, we examined individual-level change patterns to determine the percentage of mothers who remained stable, reliably improved, or reliably deteriorated over time.

In ascertaining the clinical relevance of any changes, our second research question was whether or not those who participated in the Home-Start program returned to a community level of functioning after treatment (i.e., recovery; Ogles et al., 2001). In other words, we investigated whether or not the level of functioning of Home-Start participants could be distinguished from a community sample at posttest. Only if families have reached a community's level of functioning can the changes be considered clinically significant. Given that recovery is only possible for those participants who were below the community's mean level of functioning at pretest, these analyses concern only a subgroup of the Home-Start group. In order to examine this issue and to supplement tests of clinical significance (Nixon, Sweeney, Erickson, & Touyz, 2003; Sonuga-Barke, Daley, Thompson, Laver Bradbury, & Weeks, 2001), we included in the present study an additional group of mothers, who were not included in our previous article, as a social validation comparison group. This additional group, a community sample group, allowed us to compare

the outcomes of both “support-needing” groups (Home-Start and comparison group) to the outcomes of a randomly selected sample.

We expected that the Home-Start group would show more reliable change than both comparison groups on all assessed domains, but given the aim of Home-Start, we hypothesized that the change would especially be salient in the domain of maternal well-being. Moreover, we expected that these changes would remain observable until the 6-month follow-up. Furthermore, we expected that part of the Home-Start mothers would recover to a community level of functioning.

The third research question of the present study, also not examined in our previous article, was whether the initial problem level and degree of change were related. Specifically, we were interested in whether the severity and number of the experienced problems at pretest was related to the degree of demonstrated change. Previously, contradicting findings have been reported on the relationship between initial level of problems and program effects. Whereas some claimed that families in the most severe condition before start of the intervention were most likely to benefit (e.g., Olds & Kitzman, 1993; Reid, Eddy, Fetrow, & Stoolmiller, 1999), others (e.g., MacLeod & Nelson, 2000; Webster-Stratton, 1996) reported the opposite; that is, effects were smaller for programs aimed at families with highest level of problems. Given these conflicting earlier findings, no specific hypothesis was formulated; instead, this question was investigated from an exploratory stance.

Method

Participants and Procedure

Information on participants and procedures will be summarized briefly. For a more elaborate description, please refer to our previous study (Asscher et al., 2008).

We compared a group of mothers who received support from the Home-Start parenting support program ($n = 66$) with a group of mothers who reported a similar level of parental stress, but who did not receive any form of volunteer or professional support ($n = 58$). In addition, a randomly selected group of mothers was recruited as a social validation comparison (i.e., norm) group ($n = 41$) to supplement tests

of clinical significance (Nixon et al., 2003; Sonuga-Barke et al., 2001).

In order to obtain comparison groups, 1,000 parents with a child between the age of 1.5 and 3 years were sent a short questionnaire assessing parental stress (Dutch version of Parenting Stress Index—Short form; De Brock, Vermulst, Gerris, & Abidin, 1992). In addition, the following questions were asked: “Do you need support regarding parenting every now and then?” (Yes/No), “If this support were to come from a volunteer who’d come to support you three hours each week, would you make use of this service?” (Yes/No), “How often do you find your child to be more difficult than other children?” For the last question, there were four answer categories, ranging from *hardly ever* (1) to *almost always* (4). Three hundred and seventy-five parents returned the questionnaire. From this large pool of families, the comparison group ($N = 58$) was selected. The two criteria used to include families in this group were as following: (a) maternal stress levels above the “normed” mean for nonclinical groups as assessed by the Parenting Stress Index ($M \geq 2.48$) or (b) at least two of the three additional questions answered in ways that indicate stress or need for support or both (“need support every now and then,” “would make use of support,” and “child often/almost always perceived to be more difficult than other children”).

For the parents who agreed to participate, an appointment was made for the first home visit (T1) and the first questionnaire was sent. At the end of the visit, an appointment was made for the second visit (T2), on average 6.8 ($SD = .93$) months later. At the end of T2, an appointment was planned for the follow-up visit (T3) on average 5.77 ($SD = .84$) months after T2.

The two “need for support” groups (Home-Start and comparison group) were equal on all demographic variables (ethnicity, age and gender of the target child, number of children in the family, and health problems, number of life events), except age of the mother, $F(1, 102) = 10.19$, $p < .01$, educational level ($\chi^2 = 4.24$, $p < .05$), and marital status ($\chi^2 = 10.51$, $p < .01$). Home-Start mothers were younger, less educated, and more often single than comparison mothers. When both need for support groups were compared to the norm group, the groups were once again equal on most demographic variables. It did turn out that the mothers of the need for support group were younger than the mothers of the norm group, $F(1, 137) = 5.91$,

$p < .05$, were not as highly educated ($\chi^2 = 8.17$, $p < .01$), had more often experienced three life events or more ($\chi^2 = 7.13$, $p < .01$), and were more likely to be single moms ($\chi^2 = 13.95$, $p < .001$). For detailed information with regard to the within group statistics, see also Asscher et al. (2008).

Home-Start: Description of the Intervention

Home-Start support is a volunteer-based parenting support program for families with at least one child under the age of 6 and that experience difficulties in childrearing. Volunteers visit the families once a week for half a day. Families can approach Home-Start through health clinics, social workers, child protection, and self-referral. Home-Start volunteers attend a 3-day training program in which they are taught to be supportive in a nondirective way. In addition, volunteers receive supervision once a month and attend a training day once a year. There are no additional educational requirements. In the present study, 20% of the families used Home-Start for less than 3 months and 20% of the families used Home-Start ranging from 3 to 5 months. Fifty-six percent of the families ended Home-Start after 5–9 months and only 4% of the families used Home-Start for longer than 9 months. The mean number of visits per month was 3.49 ($SD = .82$) with an average duration of 2.4 hr ($SD = .46$). In our sample, both the intensity and the content of intervention were comparable to the way Home-Start is conducted commonly in The Netherlands (Galama & van Rij, 2004). For a more detailed description of the Home-Start program, see our previous article (Asscher et al., 2008).

Instruments

All instruments that were used in the present study have been used in previous studies (e.g., Deković, Janssens, & van As, 2003) and have adequate psychometric qualities (Asscher et al., 2008).

Maternal well-being. Maternal *depressive mood* was measured with the Parenting Stress Index—Revised (Gerris et al., 1993). Parental self-esteem with regard to parenting (perceived parenting *competence*) was assessed with the competence subscale of the Dutch version of the Parenting Stress Index (Abidin, 1983; De Brock et al., 1992).

Parenting behaviors. Assessment of parenting behavior included self-reports as well as observations. The first self-reported measure, *responsiveness*, was

assessed with a subscale of the Nijmegen Parenting Questionnaire (Gerris et al., 1993). The second self-reported dimension of parenting behavior, *the negative control*, was assessed with the Parenting Dimensions Inventory (Slater & Power, 1987). A composite negative control score was computed by standardizing and summing up the scores on permissiveness, ignoring love withdrawal, physical punishment, and exercise of power.

In addition to the questionnaire data, observational data were collected. Research staff coded the Coder Impressions Inventory (CII; Webster-Stratton, 1998) immediately after a home visit. The following parenting constructs of the CII were used in this study: *harsh parenting* represented negative and hostile parenting and *parental warmth* measured affectionate and warm parenting behavior. The observers each had approximately 30 hr of training with videotapes and live observations using the CII until an interobserver agreement rate of at least 80% was achieved.

Child problem behavior. Assessment of child problem behavior also included both maternal reports and observation. Mothers were asked to report on their child's behavioral problems on the Dutch version of the Child Behavior Check List/2-3 (Achenbach, 1992). For the present study, only the *externalizing behavior problems* scale was used.

The CII (Webster-Stratton, 1998) was used as an observational measure of child behavior in interaction with the mother. Observers rated the child's behavior during the entire visit. The *negativity* scale was used in the present study to measure the amount of negative behavior such as aggressive or detached behavior.

Analytic Strategy

In order to examine our first research question (whether the amount of change is meaningful), the Reliable Change Index (RCI) was calculated for each case and for each dependent variables using Jacobson and Truax's (1991) method. In this method, a pretest score is subtracted from a posttest score and this number is then divided by the standard error of difference (*Sdiff*) between the two test scores. The *Sdiff* is derived from the standard error (SE) of measurement using the following formula: $Sdiff = \sqrt{(2(SE)^2)}$. If the calculated RCI is greater than 1.96 or smaller than -1.96 ($p < .05$), then the change is large enough to be reliable. Then, we calculated the percentages of cases that improved ($RCI > 1.96$),

deteriorated ($RCI < -1.96$), or stayed the same ($-1.96 > RCI < 1.96$) at pretest as well as at follow-up. We refer to *improvement* when there was an increase in desirable outcomes (i.e., maternal competence, responsiveness, and warmth) or a decrease in negative outcomes (i.e., depressive mood, negative control, harshness, child externalizing behavior problems, and negativity). When there was a decrease in desirable outcomes or an increase in undesirable outcomes, we use the term *deterioration*.

Next, we examined whether there were significant differences between the Home-Start, comparison, and norm groups in the percentage of cases that either improved or deteriorated by conducting the chi-square tests on raw frequency scores. Because of the small percentage of cases that showed deterioration, we combined the cases that showed no change with cases that showed deterioration and compared this group to the group who showed improvement. Two analyses were conducted: one for short-term effects (posttest comparison = T1 – T2) and one for long-term effects (follow-up comparison = T1 – T3).

To examine the second research question (whether, after treatment, Home-Start mothers were able to function at a level equivalent to that of the community, i.e., recovery), we used the approach suggested by Jacobson and Truax (1991). Recovery is adjudged to have occurred when scores at posttest cross the clinical cutoff score, which is at the midpoint between the means of the clinical and normal population (Sonuga-Barke et al., 2001). Therefore, we first calculated the clinical cutoff score ($[\text{pretest mean of the norm population} + \text{pretest mean of intervention group}]/2$). After that we calculated the percentage of Home-Start mothers who scored above the cutoff point at posttest and follow-up. For Home-Start mothers who, at pretest, were already functioning at a community level, no recovery was expected, and therefore, they were excluded from these analyses.

Finally, to examine our third research question (whether or not the initial level of experienced problems is related to the degree of reliable change), we first divided the mothers into three groups: mothers who showed no reliable change ($n = 33$), mothers who showed reliable change on only one of the outcome variables ($n = 18$), and mothers who showed reliable change on two outcome variables or more ($n = 11$) at posttest. Then, we carried out multivariate analyses of covariance (MANCOVA's) on the

pretest scores of the outcome measures by the group classification of reliable change (no change, change on one variable, and change on two variables or more), with maternal age, education, and marital status as covariates. We carried out the same analyses to examine which mothers were likely to show recovery at posttest. Again, we divided the mothers into three groups: mothers who did not show recovery on any variable ($n = 11$), mothers who showed recovery on only one of the outcome variables ($n = 22$), and mothers who showed recovery on two or more outcome variables ($n = 33$).

Results

Preliminary Analyses

In the current data set, 9.6% of the data were missing. By estimating the maximum likelihood function at the individual level (Schafer & Graham, 2002), data from all participants ($n = 165$) could be included in these analyses, regardless of their pattern of amount of “missingness.” Expected Maximization is a maximum likelihood procedure that uses iterations to impute missing values that are based on all available data. Schafer and Graham recommended this procedure, under the assumption that data are missing at random, as being a highly efficient way to optimally use the available data.

Reliable Change

Numbers and percentages of cases that improved, deteriorated, or stayed the same at pretest as well as at follow-up have been presented in Table 1.

Maternal well-being. Differences between Home-Start, comparison, and norm groups in percentages of cases that improved (i.e., reported a decrease in depressive mood or an increase in competence, respectively) or deteriorated (i.e., reported an increase in depressed mood or a decrease in competence) were significant both at posttest (depressive mood $\chi^2 = 9.56$, $p < .01$; competence $\chi^2 = 11.15$, $p < .01$) and at follow-up (depressive mood $\chi^2 = 10.49$, $p < .01$; competence $\chi^2 = 13.61$, $p < .01$). Although 12 – 18% of Home-Start mothers showed improvement and none of them deteriorated, the percentage of improvement in the other two groups varied between 0 and 3% and did not differ from the expected distribution if changes were random

Table 1. *Reliable Change in Maternal Well-Being, Parental Behaviors, and Child Behavior for Home-Start, Comparison, and Norm Group*

Outcome/Group	Posttest						Follow-up					
	Improvement		No Change		Deterioration		Improvement		No Change		Deterioration	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Depressive mood												
Home-Start	8	12	58	88	0	0	10	15	56	85	0	0
Comparison	1	2	57	98	0	0	2	3	56	97	0	0
Norm	0	0	41	100	0	0	0	0	41	100	0	0
Competence												
Home-Start	9	14	40	86	0	0	12	18	54	82	0	0
Comparison	1	2	57	98	0	0	2	3	56	97	0	0
Norm	0	0	57	98	1	2	0	0	41	100	0	0
Responsiveness												
Home-Start	4	6	61	92	1	2	5	8	60	91	1	1
Comparison	0	0	58	100	0	0	1	2	57	98	0	0
Norm	1	2	39	96	1	2	0	0	41	100	0	0
Negative control												
Home-Start	4	6	60	91	2	3	4	6	60	91	2	3
Comparison	0	0	57	98	1	2	1	2	56	96	1	2
Norm	0	0	41	100	0	0	0	0	41	100	0	0
Warmth (O)												
Home-Start	6	9	57	86	3	5	8	12	55	83	3	5
Comparison	1	2	54	93	3	5	1	2	53	91	4	7
Norm	0	0	41	100	0	0	0	0	40	98	1	2
Harshness (O)												
Home-Start	3	5	53	80	10	15	2	4	59	85	5	11
Comparison	1	2	57	98	0	0	4	7	53	91	1	2
Norm	0	0	41	100	0	0	0	0	41	100	0	0
Externalizing												
Home-Start	10	15	56	85	0	0	8	14	57	85	1	1
Comparison	3	5	55	95	0	0	2	5	55	83	1	2
Norm	1	2	40	98	0	0	0	0	41	100	0	0
Negativity (O)												
Home-Start	3	4	58	88	5	8	4	6	60	91	2	3
Comparison	1	2	56	96	1	2	2	3	55	95	1	2
Norm	0	0	41	100	0	0	0	0	40	98	1	2

Note. O = observational measure.

(e.g., roughly 2.5% of the sample would improve, 2.5% would deteriorate, and 95% would not change reliably).

Parenting behavior. Although 6% of Home-Start mothers at posttest showed reliable improvement in self-reported responsiveness, the differences between the groups at posttest were not significant ($\chi^2 = 3.90, p = ns$). At follow-up, 8% of the Home-Start families showed improvement and only

1% deteriorated, whereas in both other groups, none deteriorated and 2% of the comparison group improved. The differences between the groups were, however, again not significant ($\chi^2 = 5.05, p < .10$).

Six percent of the Home-Start mothers showed reliable improvements (i.e., decrease) on self-reported negative control at posttest, but 3% of the Home-Start mothers showed an increase in negative control. In the norm group, none of the mothers

changed, and in the comparison group, none of the mothers improved and 2% deteriorated. These differences were significant ($\chi^2 = 6.11, p < .05$). Even though the percentages of reliable changes in the Home-Start group remained stable, at follow-up differences between the groups were no longer significant ($\chi^2 = 3.66, p = ns$). This was most likely because of the fact that 2% of the comparison group also showed reliable improvement at follow-up.

Observed maternal warmth was significantly different between the groups at the posttest ($\chi^2 = 6.50, p < .05$). The results indicated that 9% of the Home-Start mothers showed an improvement in warmth and 5% showed deterioration in warmth. In the comparison group, only 2% showed improvement and 5% deteriorated. In the norm group, no reliable change in maternal warmth was observed. At follow-up, these differences were still visible ($\chi^2 = 9.56, p < .01$) seeing that 12% of the Home-Start families had improved, whereas the deterioration in the comparison group had increased to 7%.

The results for observed maternal harshness showed no significant difference in improvement (i.e., decrease in harshness) or deterioration (i.e., increase in harshness) between the groups either at posttest ($\chi^2 = 2.38, p = ns$) or at follow-up ($\chi^2 = 1.82, p = ns$). Remarkably, 15% of the Home-Start mothers showed an increase in harsh behavior toward the child (i.e., deterioration) and only 5% showed improvement. At follow-up, only 4% of the Home-Start families had improved, whereas 14% had deteriorated. Also, in the comparison groups, deterioration was visible.

Child behavior. According to the mothers' report, children in the Home-Start group improved more than children in the other two groups. Fifteen percent of the children in the Home-Start group showed improvement (i.e., a decrease in externalizing behavior problems), whereas only 5% of the children in the comparison group and 2% of the children in the norm group showed improvement. These differences were significant both at posttest ($\chi^2 = 6.49, p < .05$) and at follow-up ($\chi^2 = 7.51, p < .05$).

However, these results were not confirmed by the observational data. Only 4% of the Home-Start children showed a reliable improvement (a decrease) in negativity at posttest and 8% showed deterioration (an increase) in negativity during the home observations. In the norm group, no changes were visible, and in the comparison group, 2% improved and 2% deteriorated. These differences were, however, not

significant ($\chi^2 = 2.38, p = ns$). At follow-up, 6% of the children in the Home-Start group showed improvement (a reliable decrease) with regard to negativity, but 3% still showed deterioration (a reliable increase) with regard to negativity. Reliable change was only visible in a small percentage of the children in the comparison and norm groups. The difference between the three groups was not significant ($\chi^2 = 2.64, p = ns$).

In sum, the results indicated that Home-Start mothers, compared to comparison and norm groups, showed reliable improvements in well-being both at posttest and at follow-up. There was reliable change for two aspects of parenting behavior (self-reported negative control and observed warmth) at posttest. However, at the follow-up, only change in warmth was still significant. Finally, there was also a reliable improvement in mother's report of child's externalizing problems, both at posttest and at follow-up.

Recovery

Table 2 shows the cutoff scores for the community level of functioning for each outcome measure, the number of Home-Start mothers who scored below this cutoff scores at pretest (i.e., mothers who were eligible for recovery), and number and percentages

Table 2. *Clinical Significant Change (Recovery) in Parental Well-Being, Parenting Behavior, and Child Behavior at Posttest and Follow-up for Home-Start Group*

Outcome	T1	Posttest		Follow-up		
	Cutoff Score	N ^a	N ^b	%	N ^b	%
Maternal well-being						
Depressive mood	2.40	19	15	79	13	68
Competence	4.76	19	16	84	16	84
Parenting behavior						
Responsiveness	5.06	29	21	72	21	72
Negative control	0.56	29	22	76	18	62
Warmth (O)	1.36	20	12	60	11	55
Harshness (O)	1.36	33	13	39	16	49
Child behavior						
Externalizing	0.59	21	17	81	17	81
Negativity (O)	1.61	23	15	65	12	52

Note. O = observational measure.

^aNumber of mothers who were eligible for recovery (i.e., who scored below the cutoff point at T1).

^bNumber of mothers who recovered (i.e., who scored above the cutoff point at T2 and follow-up, respectively).

of mothers who recovered at posttest and at follow-up. It is important to note that on all outcome measures, a substantial number (ranging from 39 to 84%) of mothers returned to community level of functioning immediately after the Home-Start intervention. Moreover, the recovery appeared to be stable: at follow-up, the percentage of the mothers who recovered ranged from 49 to 84%. However, it must be noted that these percentages only concerned a relatively small subgroup of Home-Start mothers; the

remaining mothers already functioned at the level of the community sample.

Which Families Are Likely to Show the Most Improvement?

Figure 1a shows the pretest scores on each outcome measure mothers who showed no reliable change, mothers who showed reliable change on only one of the outcome variables, and mothers who showed

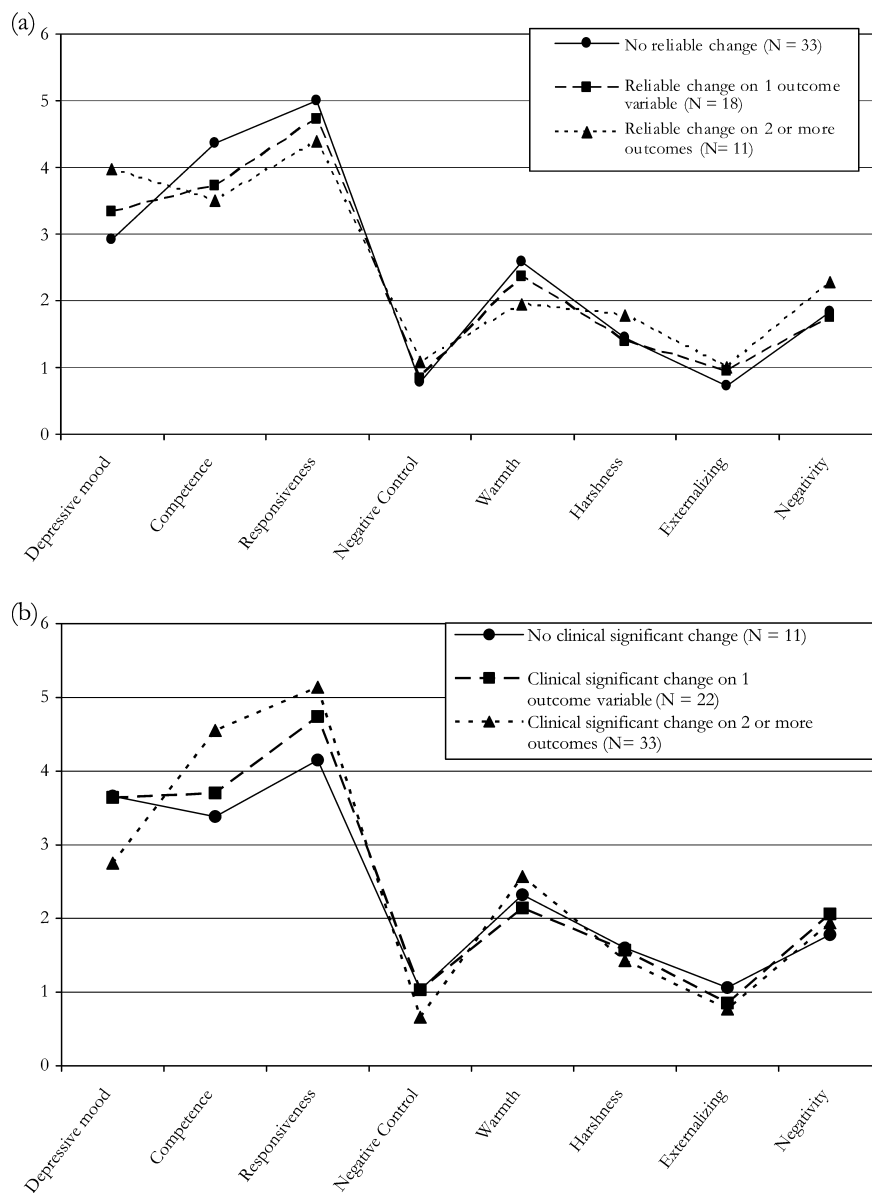


Figure 1. (a) Mean Scores at Pretest for Groups Demonstrating No Reliable Change, Reliable Change on One Outcome Variable, and Reliable Change on Two or More Outcomes in Home-Start Families. (b) Mean Scores at Pretest for Groups Demonstrating No Clinical Significant Change, Clinical Significant Change on One Outcome, and Clinical Significant Change on Two or More Outcomes in Home-Start Families.

reliable change on two outcome variables or more. The MANCOVA revealed a significant main effect of group membership, Pillai's trace = .47, $F(16, 100) = 1.89$, $p < .05$, $\eta^2 = .23$. At univariate level, significant differences emerged for the following dependent variables: depressive mood, $F(2, 61) = 4.02$, $p < .05$, $\eta^2 = .13$, competence, $F(2, 61) = 5.14$, $p < .01$, $\eta^2 = .16$, harsh parenting, $F(2, 61) = 5.06$, $p < .05$, $\eta^2 = .15$, and parental warmth, $F(2, 61) = 5.26$, $p < .01$, $\eta^2 = .16$.

Bonferroni post hoc analyses showed that mothers who did not show reliable change reported less depressive mood at pretest ($MD = -1.09$, $p < .05$) than mothers who changed on two aspects or more. Mothers who did not show reliable change also felt more competent at pretest than mothers who changed on two or more outcome measures ($MD = .88$, $p < .01$). Mothers who did not show reliable change showed more warmth ($MD = .64$, $p < .01$) and less harsh parenting ($MD = -.33$, $p < .05$) at the pretest home visit than mothers who showed change on two or more outcomes measures.

The same analyses were carried out to examine which mothers were likely to show recovery (Figure 1b). The MANCOVA produced a significant main effect for group membership, Pillai's trace = .47, $F(16, 108) = 2.05$, $p < .05$, $\eta^2 = .23$. Univariate tests also showed significant differences in pretest scores for depressive mood, $F(2, 60) = 5.75$, $p < .01$, $\eta^2 = .16$, competence, $F(2, 60) = 10.00$, $p < .001$, $\eta^2 = .26$, responsiveness, $F(2, 60) = 4.27$, $p < .01$, $\eta^2 = .19$, negative control, $F(2, 60) = 3.54$, $p < .05$, $\eta^2 = .11$, and parental warmth, $F(2, 60) = 3.90$, $p < .05$, $\eta^2 = .12$.

Post hoc analyses showed that mothers who were best off before the start of Home-Start were most likely to recover. Specifically, mothers who recovered on two or more outcome variables felt more competent at pretest than mothers who did not recover on any outcomes ($MD = 1.19$, $p < .01$) or who recovered on only one of the outcomes ($MD = .85$, $p < .05$). Furthermore, mothers who recovered on two or more outcome variables were suffering from a less depressive mood than mothers who recovered on one ($MD = -.88$, $p < .05$) or none ($MD = -.97$, $p < .05$) of the outcome measures. In addition, mothers who showed most recovery were more responsive at pretest than mothers who did not show any recovery ($MD = 1.00$, $p < .01$). Finally, mothers who showed the most significant clinical change also showed more warmth at pretest

($MD = .42$, $p < .05$) than those mothers who showed significant clinical change on only one outcome measure.

In summary, despite findings for the RCI that those families who were worst off before the start of Home-Start were most likely to obtain meaningful changes, families who were best off before the start of Home-Start were also most likely to return to a normal level of functioning after having participated in Home-Start.

Discussion

Reliable Change

The first research question of the present study was whether the amount of change in maternal well-being, parenting behavior, and child problem behavior, on the basis of comparison of group means and already reported in previous articles evaluating the Home-Start program (Asscher et al., 2008), was large enough to be considered meaningful. Regarding maternal well-being, we found substantial changes in the Home-Start group, in both competence and depressive mood, that were not found in either the comparison or the norm group. These results were even more pronounced at follow-up. These results suggest that Home-Start does indeed increase maternal well-being, one of the focal points of the Home-Start program. It is possible that mothers feel relieved merely by being able to discuss the issues regarding their children with an experienced mother or feel supported, or both, when a volunteer takes care of their children or praises positive maternal behavior.

For parenting behavior, the results were somewhat less clear. When self-report data were examined, most of the reliable improvements were generally found for the Home-Start mothers when compared to the two comparison groups. The improvements in parenting behavior, however, were not reflected by the observational data, apart from observed maternal warmth, which did improve. Observational data further suggest that 15% of the Home-Start mothers showed deterioration (i.e., an increase) in harsh parenting behavior at posttest and in 11% of the cases, this deterioration was still visible at follow-up. It must be noted, however, that this result was not statistically significant in comparing groups. At follow-up, mothers in both comparison

groups also showed an increase in harsh parenting. The increased harsh parenting behavior in all groups may be because of the age of the children involved. Toddlerhood is often marked by the child's increasing noncompliance (Kochanska, 1995), which provokes restrictive behavior on the part of their parents, leading to the child naturally protesting, perhaps then resulting in increased observed harshness in all groups, and so forth.

In general, it seemed that Home-Start mothers did change psychologically, given the reliable positive changes in maternal well-being. However, these psychological changes were only partially reflected by self-reported parenting and observational data. Despite the fact that merely modest changes were found in parenting behavior, mothers did report reliable decrease in their child's externalizing behavior both at posttest and at follow-up. These findings may also be the result of the psychological change in mothers. In their mother's perception, the children have changed in a positive way, although during observations no improvement of behavior may be apparent. It could be that the children have not changed, but that mothers perceive them differently as a consequence of their own improved well-being. However, this change of perception may eventually lead to actual improvements in the children's behavior.

An important question that remains is whether the mothers were suffering from diminished well-being at pretest and therefore perceived their children as more problematic than mothers who felt well, or, alternatively, whether the children were indeed more problematic in the beginning and did mothers therefore experience diminished well-being. These questions reflect a long-standing debate in literature in which the core question is whether or not mothers with a diminished sense of well-being have distorted perceptions of their children, which then lead to an increase of behavioral problems in a child (Field, Morrow, & Adlestein, 1993). Interestingly, the present study suggests the opposite: mothers whose well-being increased perceive their children as less problematic than before, although the observational data do not suggest actual improvements in child behavior. This would suggest that mothers with improved well-being might not see the misbehavior of their children. Equally, it could mean that these changes were real, but that they were not apparent during the short period of the home observation, whereas mothers who spend all day with

their children did note a decrease in behavioral problems.

In conclusion, these results suggest that Home-Start leads to a reliable change for some of the participating families; however, in both comparison groups, there was no more change than can be attributed to chance (roughly 2.5% of the sample). Most changes were as expected. Specifically, in the Home-Start group, the reliable change was mostly in the positive direction (improvement). Furthermore, we found no evidence of deterioration effects, with the exception of harshness. However, we should be cautious not to overinterpret these results because the deterioration in harshness was observed in the "in-need-for-support" groups as well as the Home-Start group and because the Home-Start mothers did not deteriorate significantly more than the other groups.

Recovery

The second research question concerned the clinically significant change (i.e., recovery) and whether or not the Home-Start mothers would be able to function on a community level after treatment. We found substantial percentages of mothers who were able to function at the community level after Home-Start. Unfortunately, none of the mothers who showed meaningful changes also showed recovery and vice versa. This means that mothers who showed the largest improvements (i.e., for whom Home-Start yielded the greatest results) were not the same mothers as those who functioned at a community level after Home-Start. Apparently, there are two types of mothers who participate in Home-Start: first, mothers for whom Home-Start does a lot and who improve quite a bit, but who do not manage to function at a community level at posttest or follow-up, and second, those mothers who did recover to community level functioning, but showed less improvement, most likely because of the fact that they were already functioning closer to the community level before treatment. In order to be able to decide whether or not a program functions and whether or not clinical significant reliable changes are achieved, it is best to combine both outcomes.

Predicting the Change

The third research question was for which families Home-Start would have the largest effects. When we examined this question for the RCI, we found that

those families who were having most problems at preintervention assessment also showed reliable change on most outcome measures. However, when we examined this question for *recovery*, we found exactly the opposite result: those families who were best off before the start of Home-Start were most likely to function at a normal level after the intervention. These results shed some light on the inconsistent results that have been reported regarding the question as to which families are most likely to improve—those who start off as best or those who start off as worst (e.g., Olds & Kitzman, 1993). The answer to the question of what kind of families programs ought to aim for depends on the question of what a program hopes to achieve. If a program hopes to obtain the largest change between pre- and posttest, then it is best to start with families who are worst off at pretest. If, however, a program hopes to get as many families back to functioning at a level equivalent to the average of a community sample, then it is best to focus on families that are closest to average level of functioning.

Limitations

Several limitations of this study should be taken into consideration. We used a relatively small sample size. This is a consequence of our choice to include observational data in the present study. As this is an intensive method of data collection, it was unfeasible to have a larger sample. A relatively small sample size is a shortcoming that we share with many other evaluations of interventions (Weisz, Jensen Doss & Hawley, 2005). Therefore, a replication of this study is needed with a larger sample size to come to clearer understanding of the effectiveness of Home-Start.

Furthermore, the sample was not randomly allocated to the treatment or the control group, with a consequence that there were differences between the mothers regarding age, educational level, and marital status. These characteristics are known to influence both parent and child outcomes and could have impacted our results. Although we did control for the effects of these variables in MANCOVA, we were unable to do so when conducting the chi-square tests that we used to answer the first research question (i.e., whether Home-Start leads to meaningful and clinically significant change). However, in our previous article (Asscher et al., 2008), we examined the effectiveness of Home-Start by conducting a more “traditional” type of analysis (i.e.,

the comparison of mean scores of intervention and comparison group). In this analysis, we did control for the demographic variables and found that they did not influence the effectiveness. Still caution is warranted when interpreting these results.

Conclusions and Implications

Notwithstanding its limitations, we believe this study gives insight in changes that were undetected in our previous article (Asscher et al., 2008), in which a classical approach to examine effectiveness was followed. In the previous article, we were able to show that Home-Start mothers, as a group, changed significantly, when compared to comparison group, whereas in the present article, we examined whether this change was large enough to be considered meaningful. Moreover, by adding a norm group in the present article, we were also able to determine whether Home-Start mothers returned to a level of functioning equivalent to that of the community. Finally, the present approach provided the information about variability in response to Home-Start, which allowed us to examine which families were most likely to improve. It was remarkable, then, that there was an important difference between mothers who were likely to show reliable change and mothers who eventually functioned at community level.

The present results can contribute to the optimization of effects of Home-Start in two ways. First, it is important for Home-Start coordinators, volunteers, and policymakers to be aware of the fact that most reliable changes are to be expected for the families that are in the worst condition before the start of Home-Start. Home-Start seems to contribute to the quality of life of a number of these families. However, given the fact that these mothers and children still had a relatively high level of problems after the intervention, it might be necessary to provide extra support for these families. Possibly, the intervention was ended too soon (e.g., because families and volunteer were content too early with their progress) or Home-Start has only limited potential for these families. A solution might be to add one or more other interventions (e.g., parent management training or counseling for the mother) to the work of the volunteer. On the other hand, clinically, the most relevant results (i.e., recovery) were achieved for families with a relatively low level of problems at the start of the intervention. A substantial number

of mothers in these families reach a level of average well-being, average parenting, and/or average child behavior. In general, Home-Start seems suited best for families with a low level of problems at the start. Furthermore and it can stand as a relatively low-cost intervention and attain satisfying results.

A second important consequence of this study is that the findings point to the necessity of regular assessment of progress on relevant aspects of family functioning, so that progress can be related to standardized measures and that choices to continue, stop, or expand intervention efforts can be empirically driven. When evaluations are carried out, it is important to be aware of the fact that there is a possibility that better results may be observed at the follow-up (i.e., not immediately after termination of the intervention). This suggests that even if improvements are not observable immediately after intervention, eventually they may emerge. Finally, although the present study suggests positive and clinically significant changes for the Home-Start families, the benefits should be balanced with costs in order to determine whether or not the program is indeed feasible. Future research should focus on the cost aspect as well.

It is fitting to make a final remark on the use of RCI. Although it has been recommended that researchers report not only the statistical but also the clinical significance of their findings, this unfortunately still is not a common practice (Fidler et al., 2005). The RCI is not intended to replace traditional analyses of group means. It provides additional information by doing several things: it looks within the group, it provides information regarding the frequency with which meaningful change occurs, and it allows examination of both positive and negative change (Hawley, 1995). As Roberts, Caspi, and Moffit (2001) aptly pointed out: "Change can and should be assessed using numerous techniques" (p. 681). Only then are we able to estimate what a program is worth to politicians, clinicians who carry out the program, and—most of all—the families involved.

References

- Abidin, R. R. (Ed.). (1983). *Parenting stress index manual*. Charlottesville, VA: Pediatric Psychology Press.
- Achenbach, T. M. (1992). *Manual for the child behavior checklist/2-3 and 1992 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Asscher, J. J., Hermans, J. M. A., & Deković (2008). Effectiveness of the Home-Start parenting support program: Behavioral outcomes for parents and children. *Infant Mental Health Journal*, 29(2), 95–113.
- Bamberger, M. J., Rugh, J., & Mabry, L. (2006). *Realworld evaluation. Working under budget, time, data, and political constraints*. Thousand Oaks, CA: Sage.
- Barnes, J., MacPherson, K., & Senior, R. (2006). The impact on parenting and the home environment of early support to mothers with new babies. *Journal of Children's Services*, 1, 4–20.
- Barnet, B., Duggan, A., Devoe, M., & Burrell, L. A. (2002). The effect of volunteer home visitation for adolescent mothers on parenting and mental health outcomes: A randomized trial. *Archives of Pediatrics & Adolescent Medicine*, 156, 1216–1222.
- Belsky, J. (1984). The determinants of parenting: A process model. *Child Development*, 55, 83–96.
- Bilukha, O., Hahn, R. A., Crosby, A., Fullilove, M. T., Liberman, A., Moscicki, E., Moscichi, E., Snyder, S., Tume, F., Corso, P., Schofield, A., & Briss, P. (2005). The effectiveness of early childhood home visitation in preventing violence. A systematic review. *American Journal of Preventive Medicine*, 28, 11–39.
- Coleman, P. K. (1999). *Maternal self-efficacy beliefs as predictors of parenting competence and toddlers' emotional, social, and cognitive development*. Morgantown: West Virginia University, Department of Psychology.
- Dawson, G., Ashman, S. B., Panagiotides, H., Hessel, D., Self, J., Yamada, E., & Embry, L. (2003). Preschool outcomes of children of depressed mothers: Role of maternal behavior, contextual risk, and children's brain activity. *Child Development*, 74, 1158–1175.
- De Brock, A. A. J. L., Vermulst, A. A., Gerris, J. R. M., & Abidin, R. R. (1992). *Nijmeegse ouderlijke stress index [Nijmegen parenting stress index]*. Lisse, The Netherlands: Swets & Zeitlinger.
- Deković, M., Janssens, J. M. A. M., & van As, N. M. C. (2003). Family predictors of antisocial behavior in adolescence. *Family Process*, 42, 223–236.
- Diamond, G., & Josephson, A. (2005). Family-based treatment research: A 10-year update. *Journal of American Academy of Child and Adolescent Psychiatry*, 44, 872–887.
- Domitrich, C. E., & Bierman, K. L. (2001). Parenting practices and child social adjustment: Multiple pathways of influence. *Merrill Palmer Quarterly*, 47, 235–263.
- Ehrensaft, M. K., Wasserman, G. A., Verdelli, L., Greenwald, S., Miller, L. S., & Davies, M. (2003). Maternal antisocial behavior, parenting practices, and behavior problems in boys at risk for antisocial behavior. *Journal of Child and Family Studies*, 12, 27–40.
- Fidler, F., Cumming, G., Thomason, N., Pannuzzo, D., Smoith, J., Fyffe, P., Edmonds, H., Harrington, C., & Schnitt, R. (2005). Toward improved statistical reporting. *Journal of Consulting and Clinical Psychology*, 73, 136–143.
- Field, T. M., Morrow, C., & Adlestein, D. (1993). Depressed mothers' perceptions of infant behavior. *Infant Behavior and Development*, 16, 99–108.
- Frost, N., Johnson, L., Stein, M., & Wallis, L. (2000). Home-Start and the delivery of family support. *Children-and-Society*, 14, 328–342.
- Galama, M., & van Rij, C. (2004). *Home-Start knipselkrant [Home-Start collection of newspaper cuttings]*. Amsterdam: Home-Start landelijk steunpunt.
- Gerris, J. R. M., Vermulst, A. A., Van Boxtel, D. A. M., Janssens, J. M. A. M., Zutphen, R. A. H., & Felling, A. J. A. (1993). *Parenting in Dutch families*. Nijmegen, The Netherlands: Institute of Family Studies.
- Gomby, D. S., Culross, P., & Behrman, R. (1999). Home visiting: Recent program evaluations—analysis and recommendations. *Future of Children*, 9, 4–26.
- Hawley, D. R. (1995). Assessing change with preventive interventions: The reliable change index. *Family Relations*, 44, 278–284.
- Hermans, J. M. A., Van De Venne, L., & Leseman, P. (1997). *Home-Start geëvalueerd [Home-Start evaluated]*. Amsterdam: SCO-Kohnstamm Instituut.
- Home-Start International website. Retrieved October 30, 2007, from <http://www.homestartinternational.org/>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Johnson, Z., Molloy, B., Scallan, E., Fitzpatrick, P., Rooney, B., Keegan, T., & Byrne, P. (2000). Community mothers programme—seven year follow-up of a randomized controlled trial of non-professional intervention in parenting. *Journal of Public Health Medicine*, 22, 337–342.

- Johnston, C., Murray, C., Hinshaw, S. P., Pelham, J. W. E., & Hoza, B. (2002). Responsiveness in interactions of mothers and sons with ADHD: Relations to maternal and child characteristics. *Journal of Abnormal Child Psychology, 30*, 77–88.
- Kelleher, L., & Johnson, M. (2004). An evaluation of a volunteer-support program for families at risk. *Public Health Nursing, 21*, 297–305.
- Kendall, P. C. (1999). Clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 283–284.
- Kochanska, G. (1995). Children's temperament, mother's discipline, and security of attachment: Multiple pathways to emerging internalization. *Child Development, 66*, 597–615.
- MacLeod, J., & Nelson, G. (2000). Programs for the promotion of family wellness and the prevention of child maltreatment: A meta-analytic review. *Child Abuse and Neglect, 24*, 1127–1149.
- MacMillan, H. L., Thomas, B. H., Jamieson, E., Walsh, C. A., Boyle, M. H., Shannon, H. S., & Gafni, A. (2005). Effectiveness of home visitation by public health nurses in prevention of the recurrence of child physical abuse and neglect: A randomised controlled trial. *Lancet, 365*, 1786–1793.
- McAuley, C., Knapp, M., Beecham, J., & McCurry, N. (2004). *Young families under stress. Outcomes and costs of Home-Start support*. York, UK: Joseph Rowntree Foundation.
- McGuigan, W. M., Katzev, A. R., & Pratt, C. C. (2003). Multi-level determinants of mothers' engagement in home visitation services. *Family Relations, 52*, 271–278.
- Nixon, R. D., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent-child interaction therapy: A comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *Journal of Consulting and Clinical Psychology, 71*, 251–260.
- Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review, 21*, 421–446.
- Olds, D., & Kitzman, H. (1993). Review of research on home visiting for pregnant women and parents of young children. *Future of Children, 3*, 53–92.
- Olson, S. L., Ceballo, R., & Park, C. (2002). Early problem behavior among children from low income, mother headed families: A multiple risk perspective. *Journal of Clinical Child and Adolescent Psychology, 31*, 419–430.
- Osofsky, J. D. (1998). On the outside: Interventions with infants and families at risk. *Infant Mental Health Journal, 19*, 101–110.
- Owens, E., & Shaw, D. S. (2003). Predicting growth curves of externalizing behavior across the preschool years. *Journal of Abnormal Child Psychology, 31*, 575–590.
- Papp, L. M., Cummings, M. E., & Schermerhorn, A. C. (2004). Pathways among marital distress, parental symptomatology, and child adjustment. *Journal of Marriage and the Family, 66*, 368–384.
- Pettit, G. S., Bates, J. E., & Dodge, K. A. (1997). Supportive parenting, ecological context, and children's adjustment: A seven year longitudinal study. *Child Development, 68*, 908–923.
- Reid, J. B., Eddy, J. M., Fetrow, R. A., & Stoolmiller, M. (1999). Description and immediate impacts of a preventive intervention for conduct problems. *American Journal of Community Psychology, 27*, 483–518.
- Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin, 128*, 330–366.
- Roberts, B. W., Caspi, A., & Moffitt, T. E. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81*, 670–683.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147–177.
- Schoenwald, S. K., Sheidow, A. J., & Letourneau, E. J. (2004). Toward effective quality assurance in evidence-based practice: Links between expert consultation, therapist fidelity, and child outcomes. *Journal of Clinical Child and Adolescent Psychology, 33*, 94–104.
- Slater, M. A., & Power, T. G. (1987). Multidimensional assessment of parenting in single-parent families. In J. P. Vincent (Ed.), *Advances in family intervention, assessment, & theory* (pp. 197–228). Greenwich, CT: Jai Press.
- Solomonica-Levi, D., Yirmiya, N., Erel, O., & Oppenheim, D. (2001). The associations between observed maternal behavior, children's narrative representations of mothers, and children's behavior problems. *Journal of Social and Personal Relationships, 18*, 673–690.
- Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit-hyperactivity disorder: A randomized, controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 402–408.
- Thompson, A., Hollis, C., & Richards, D. (2003). Authoritarian parenting attitudes as a risk for conduct problems. Results from a British national cohort study. *European Child & Adolescent Psychiatry, 12*, 84–91.
- Thompson, L., Kropenske, V., Heinicke, C. M., Gomby, D., & Halfon, N. (2001). *Home visiting: A service strategy to deliver proposition 10 results*. Los Angeles: Deanna Gomby Consulting.
- Webster-Stratton, C. (1996). Early-onset conduct problems: Does gender make a difference? *Journal of Consulting and Clinical Psychology, 64*, 540–551.
- Webster-Stratton, C. (1998). Preventing conduct problems in head start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology, 66*, 715–730.
- Weisz, J. R., Jensen Doss, A., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review of critique of the evidence base. *Annual Reviews of Psychology, 56*, 337–363.
- Wolchik, S. A., Wilcox, K. L., Tein, J. Y., & Sandler, I. N. (2000). Maternal acceptance and consistency of discipline as buffers of divorce stressors on children's psychological adjustment problems. *Journal of Abnormal Child Psychology, 28*, 87–102.