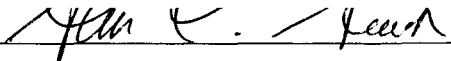


AN ABSTRACT OF THE THESIS OF

Blythe K. Kneedler for the degree of Master of Science in Human Development and Family Studies presented on May 4, 2001. Title: The Efficiency of a Retrospective Pretest: How Does the Parenting Ladder Perform?

Redacted for Privacy

Abstract Approved: 

Alan Acock

As program participants integrate new knowledge and understanding into previous standards regarding attitudes or perceptions, these altered standards violate the assumption of consistency inherent to the pretest-posttest method for measuring change over time. If this violation, or response-shift bias, is not controlled results can be misleading, which carries serious implications for program evaluations. This study extends the literature on retrospective pretests by addressing the efficiency of using a specific retrospective measure as an alternative to a traditionally measured pretest in controlling for response-shift bias. The Parenting Ladder is a measure is used as a traditional pretest, posttest, and retrospective measure of parental self-perceptions of knowledge, confidence, resources, social support, stress, and coping skills in Oregon's Healthy Start program evaluation. Using secondary data from 167 firsttime mothers who were enrolled in Oregon's Healthy Start program for a minimum of six months, this exploratory analysis found evidence of response-shift bias. This means that the

mothers reported significantly lower levels of parental functioning in retrospect, when compared to initial ratings. The validity of the Parenting Ladder self-ratings were considered in relation to family functioning behaviors as observed by program staff. The retro/post measure of change assessed by the Parenting Ladder correlated more closely to the observed pre/post change over time, assessed by the Family Progress Scale. More research is needed with both the Parenting Ladder and the Family Progress Scale. Using a different measure to assess Parenting Ladder validity, rather than an observed measure such as the Family Progress Scale, is recommended.

The Efficiency of a Retrospective Pretest: How Does the Parenting Ladder
Perform?

by

Blythe K. Kneedler

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Masters of Science

Presented May 4, 2001
Commencement June 2001

Master of Science thesis of Blythe K. Kneedler presented on May 4, 2001

APPROVED

Redacted for Privacy

[Signature]
Major Professor, representing Human Development and Family Studies

Redacted for Privacy

[Signature]
Chair of Department of Human Development and Family Sciences

Redacted for Privacy

[Signature]
Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

[Signature]
Blythe K. Kneedler, Author

ACKNOWLEDGEMENT

The author would like to sincerely express the deepest gratitude to everyone that has made this thesis possible. The author would like to elaborate by thanking Dr. Clara Pratt for her astute help and patience throughout the thesis process. The author would also like to express a deep gratitude to Dr. Alan Acock for his patient guidance, kind support, and helpful assistance. The author would like to convey her thankfulness to Dr. Pat Moran for her insightful guidance and editing, and Dr. Sam Vuchinich for his support.

The author would like to further offer heartfelt appreciation for Dr. William McGuigan who first brought the retrospective pretest to the author's attention, and later helped with advice, compiling a dataset, and editing. The author appreciates extensive knowledge of Dr. Aphra Katzev, who was kind enough to share her expertise with the Oregon Healthy Start program, evaluation, and data.

The author is grateful that data for this study were provided by the Oregon Healthy Start Evaluation (1994 – 2000), conducted by the Oregon State University Family Policy Program, under contract for the Oregon Commission on Children and Families. The author would also like to express her gratitude for all the Oregon Healthy Start families, volunteers, and staff members for their participation, dedication, and hard work.

The author would like to sincerely express heartfelt gratitude to family, friends, faculty and staff members, as well as fellow students who helped her make it through. The author would like to articulate the deep sense of indebtedness felt

towards her parents for their unconditional love and support. The author will always be grateful for the inspiration Dr. Tammy Henderson is, and for the warm support, listening ear, and helpful advice of Ms. Rachel Ozretich. Thank you everyone. It has been an illuminating experience.

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION	1
Response-Shift Bias	3
Retrospective Pretest	5
Current Study	8
Research Questions	9
METHOD	10
Oregon Healthy Start	10
Participants	11
Procedure	13
Measures	14
Analysis Strategy	19
RESULTS	24
Stage One: Principle Component Analyses	24
Stage Two: Response-Shift Bias	26
Stage Three: Validity	28
DISCUSSION	31
BIBLIOGRAPHY	33

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1 Mothers' Characteristics (N = 167).....	11
2 Measures Completed at Pre and Post Periods.....	14
3 Principle Component Loadings for the Parenting Ladders (N = 167)	24
4 Principle Component Loadings for the Family Progress Scales (N = 167)...	26
5 Correlational Validity Results (N = 167).....	29

The Efficiency of a Retrospective Pretest: How Does the Parenting Ladder Perform?

INTRODUCTION

Within social programs, resources are inevitably scarce (Rossi & Freeman, 1993), so program evaluations must be efficient as well as valid in assessing proposed outcomes. To maximize available resources, program evaluators must carefully consider what the best method is for gathering the most valid information (The Joint Committee on Standards for Educational Evaluations, 1994). To assess change in participant knowledge over time, the prettest-posttest method might be considered because collecting pretest information beforehand can be useful in tailoring program delivery and services, as well as providing a pretest measure for later evaluations.

Unfortunately, traditional pretest measures are not always obtainable. For example, when programs work with high-risk populations, or when an evaluation was not designed before a program began, or when delaying services may be considered unethical, pretest periods may not be possible. For programs relying on participant self-report information for accountability purposes, a retrospective pretest may be a preferred alternative, because retrospective pretests can be added to posttest forms and administered within the posttest time period. By using retrospective measures, then, programs can gather pretest information from participants without the additional pretest time period.

One example of a retrospective measure being used is the Parenting Ladder, which is a brief measure of parental self-perceptions used in the Oregon Healthy Start program evaluation as a conventional pretest, posttest, and retrospective pretest. This study will focus on whether the traditional pretest and the retrospective version produce equivalent results. The key question is whether the administration of both the retrospective pretest and the traditional pretest is redundant (meaning they produce equivalent results), necessary (they produce different results, both deemed necessary), or whether only one is needed (e.g., the traditional pretest or retrospective version allows evaluators to more closely correlate perceived change with observed change).

In deciding the type of measure to use, both observational and self-report measures have their benefits and limitations. Observational measures can provide objective and precise behavioral information, but they are limited to readily observable behaviors. They may be complicated, and they may require extensive training (Friedman & Haywood, 1994). Alternatively, participants are uniquely qualified to report on their own private thoughts, feelings, and behaviors, so introspective tools such as self-report measures can allow evaluators more freedom in assessing domains affected by program participation. Self-report measures are usually simple in construction, and can be equal to, or superior to, more costly behavioral measures in gathering participant feedback, and in assessing the participant's perception of program effectiveness (Howard, Maxwell, Wiener, Boynton, & Rooney, 1984).

When self-report measures are used within the pretest-posttest method, however, a problem can occur because of the assumption of consistency. Specifically, the method assumes that both the instrument used, and the way it is used, will not change over time. To meet the assumption, then, not only does the measuring tool, itself, need to remain the same, but so too does the respondent's internal standards. This means that how respondents understand and interpret questions asked, cannot change over time. Using self-reported information to measure participant change can be problematic, because as participants attain further knowledge or experience, and integrate new understanding into previous attitudes and standards, those altered standards violate the consistency assumption of the pretest-posttest method. The physical instrument may have remained constant, but the way the instrument was used did not.

Response - Shift Bias

If the consistency assumption is violated, meaning posttest questions are interpreted differently than pretest questions, results may be effected by what has been referred to as "response-shift bias", in reference to the systematic shift in perspective experienced by respondents that bias results (e.g., Howard, Ralph, Gulanick, Maxwell, Nance, & Gerber, 1979). This change in interpretation can happen because lack of understanding of the construct being measured at pretest (Sprangers & Hoogstraten, 1989; Aiken & West, 1990). Erroneous perceptions at pretest may result in inflated self-report ratings as participants may overestimate

their knowledge or skills. Through program exposure, participants may develop a greater awareness of their functioning within the domain of interest (Howard et al., 1979). The pretest-posttest design assumes that participants have not reconceptualized the construct being measured. Without accounting for participants' newly acquired knowledge and recalibrated standards, there is no control for response-shift bias in the pretest-posttest design. (Howard et al., 1979).

A broad range of training and treatment programs have found evidence of a response-shift bias, including those targeting abstract concepts such as leadership, assertiveness, dogmatism, and communication skills (Howard & Dailey, 1979; Howard et al., 1979; Bray & Howard, 1980). Howard and his colleagues first found evidence of response-shift bias after participants' verbalized perceptions of positive change, which contradicted their findings of insignificant or negative program effects, which were measured using the pretest-posttest design. Gutek and Winter (1992) also demonstrated how mistaken conclusions can be drawn when response-shift bias is not recognized and controlled in evaluating measures of attitudinal consistency. According to Gutek and Winter (1992), previous studies attributed consistency in job attitudes to individual traits or dispositions. But if individuals are predisposed to be satisfied or unsatisfied, then consistency in job satisfaction ratings would be expected regardless of whether individuals changed jobs, companies, or both. Once Gutek and Winter (1992) designed studies to specifically account for possible response-shift bias among job changers, no consistency in job attitudes was found.

For program evaluations, a mistaken conclusion regarding program effects can compromise program credibility. Evaluations reporting small or negative program effects can be lethal for programs. Thus, accuracy is essential. If a program evaluation design does not control for the presence of response-shift bias, the program could be inaccurately represented as being minimally effective, or completely ineffective (Howard et al., 1979; Bray, & Howard, 1980; Bray, Maxwell, & Howard, 1984; Howard, G. S., & Dailey, 1979).

Retrospective Pretest

A measure is considered to be retrospective when individuals provide information regarding the past. For example, after a specific period of program intervention, participants may complete a self-report measure in reference to current perception of ability (posttest), and then complete the same measure in reference to their previous level of functioning (retrospective pretest). Some researchers (e.g., Gutek & Winter, 1992; Howard, 1980) recommend including retrospective measures in programs targeting participant self-perceptions which chose to use self-report measures to assess perception consistency or change over time. With a retrospective pretest (Campbell & Stanley, 1963), participants can utilize consistent standards in rating potentially altered perceptions.

It should be noted that while programs using retrospective measures can control for biases due to response-shift, other mechanisms such as implicit theories of change (Conway & Ross, 1984), or demand characteristics may

manifest, and not be controlled. Serious concerns such as these deserve consideration. For example, participants may reconstruct past perceptions to support current (potentially invalid) theories they hold for themselves (See Conway & Ross, 1984). Mechanisms such as demand characteristics may lead participants to report artificially lowered retrospective ratings in a mistaken effort to please program providers or staff.

Alternatively, implicit theories of change (Conway & Ross, 1984) may lead participants to believe that change should have occurred. In either scenario, measures of program effects become invalid. Investigating more alternative hypothesis to explain response-shift bias, researchers have also considered self-report methodological confounds such as subject response styles, social desirability, systematic memory distortion, and impression management.

Those alternative hypotheses were found to be unsupported (Howard et. al., 1979; Manthei, 1997). The retrospective pretest was actually found to be a way for participants to report more "honestly" about the pretest period (e.g., Howard et. al., 1979; Manthei, 1997). For example, potentially stigmatizing topics such as drug abuse may not be reported honestly if there is a concern for consequences. Using retrospective pretests helps control for social desirability factors evident in pretesting conditions by removing the threat of consequences (Manthei, 1997). Retrospective pretests have also been found to be time efficient (Brooks & Gersh, 1998), cost effective (Robinson & Doueck, 1994), and robust across procedural differences (Sprangers & Hoogstraten, 1987).

Studies comparing traditional pretest measures to retrospective pretest measures have consistently found equal or superior validity in retrospective pretest-posttest analysis of estimated treatment effects (e.g., Howard, Ralph, Gulanick, Maxwell, Nance, & Gerber, 1979; Rhodes & Jason, 1987). With the presence of response-shift, the retrospective pretest was the most powerful method (comparing analyses of [a] post scores only, [b] post minus pre, [c] minus retrospective pretest, [d] post covaried by pre, and [e] post covaried by retrospective pretests; See Bray, Maxwell, & Howard, 1984). Not all programs using retrospective pretests have found evidence of response-shift bias, however (e.g., Sprangers & Hoogstraten, 1989). When the retrospective pretest was used in place of the traditional pretest in the analysis of treatment effects, the results (of no treatment effects) were the same (Nicholson, Belcastro, & Gold, 1985). In other words, in the presence of participant change (response-shift bias), the use of the retrospective pretest allows researchers to control for response-shift bias. Without the presence of participant change, however, results produced by the conventional and the retrospective pretest were the same.

In summary, it appears from the literature that, at minimum, retrospective pretests are as valid as conventional pretests in assessing longitudinal change in participant self-ratings. Use of the retrospective pretest in place of the conventional pretest offers programs some real advantages, particularly in relation to evaluation efficiency. For some programs, retrospective pretests may be incorporated into the conventional pretest-posttest procedure, resulting in a better

assessment of change (Howard et al, 1979), and providing a more complete picture of program effects (Rhodes & Jason, 1987). The use of all three may enhance the accuracy of data gathered, but programs limited by resources, may find the administration of the traditional pretest to be redundant if the retrospective pretest can provide a relatively equivalent baseline.

Current Study

This study focused on whether the traditional pretest and the retrospective versions of the Parenting Ladder produced equivalent results. This study was conducted in three steps. First, the principle component analyses of the Parenting Ladders were conducted. Second, Parenting Ladder results were assessed for evidence of response-shift bias. Third, the validity of the Parenting Ladder was examined by correlating it with the Family Progress Scale, which is a measure of family functioning completed by program staff at time 1 as a pretest, and at time 2 as a posttest, within the Oregon Healthy Start program evaluations.

This study used repeated measures on one group. Lacking a control group, threats to internal validity such as history and maturation are not controlled for. Using repeated measures, however, allows participants to serve as their own control group and mitigates these limitations by reducing subject variability.

Research Questions

The six fundamental research questions, grouped by stage of analysis, are:

1. Stage One: Principle Component Analyses - Is there one shared construct within the Parenting Ladder pretest, posttest, and retrospective measures?
2. Stage Two: Response-Shift Bias (a) - Is there evidence of response-shift bias? Is there a significant mean difference between the traditional Parenting Ladder pretest ratings and the retrospective Parenting Ladder ratings?
3. Stage Two: Response-Shift Bias (b) - Is there a significant mean difference between Parenting Ladder - pre/post change scores and Parenting Ladder – retro/post change scores?
4. Stage Three: Validity (a) - Is there a relationship between any version of the Parenting Ladder and the corresponding version of the Family Progress Scale?
5. Stage Three: Validity (b) - Is there a significant relationship between Parenting Ladder - pre/post change scores and Family Progress Scale –pre/post change scores?
6. Stage Three: Validity (c) - Is there a significant relationship between Parenting Ladder - retro/post change scores and Family Progress Scale –pre/post change scores?

METHOD

Oregon Healthy Start

Participants were drawn from the Oregon Healthy Start program, a voluntary home visitation program currently serving 18 of Oregon's 36 counties (Katzev, Henderson, and Pratt, 1999). New parents are screened into the program within their child's first month using the Hawaii Risk Indicator Scale - a checklist of 15 risk characteristics associated with potentially higher family stress levels such as being single, or low-income (Pratt, Katzev, Moran, & Eddy, 1995). Further assessment of families with two or more risk characteristics is done using the Kempe Family Stress Interview, which measures the level of stress in the family on a range of 0 to 100, with scores over 25 being associated with higher risk of negative outcomes such as child maltreatment. (Murphy, Orkow, & Nicola, 1985; Orkow, 1985). Families with higher levels of stress are offered intensive services, depending on available resources, to prevent negative child and family outcomes (See Katzev, Pratt, with Henderson & McGuigan, 1999).

Long-term intensive services last up to five years. Trained professionals (home visitors) provide family support services during weekly home visits. Services include child development screening, parent education, and referrals. Home visitor training, based on the Healthy Families of America model, includes one week of basic training (approximately 35 hours over four and a half days), as well as wrap-around training which is ongoing over their first months on the job (approximately 60 hours over a six month period). Wrap-around training consists

of individual modules, reading, and discussing topics such as baby care, health, and safety with a supervisor. On average, the program serves families for approximately one year (Katzev, Pratt, Henderson, & McGuigan, 1999).

Participants

Data were collected from mothers enrolled in the Oregon Healthy Start program for a minimum of six months as part of ongoing program evaluation efforts. The mothers' data were used because a large portion of the fathers' data was either missing, or was reported by the mother for the father (See Table 1).

Table 1

Mothers' Characteristics (N = 167)

	<u>n</u>	<u>M</u>	<u>SD</u>
Age in years	164	20.7	5.1
Income ^a	99	882.7	575.8
Education in years completed	154	11.0	2.4
Hawaii Risk Indicator	101	4.6	1.9
Kempe Family Stress Interview	154	42.1	13.3

Note. Full data were not available on all demographic and risk variables. These variables were used descriptively in this study, not as key dependent or independent variables.

^aGross monthly household income in dollars.

Participants reported an average of four or five risk characteristics on the Hawaii Risk Indicator Scale ($\underline{M} = 4.56$, $\underline{SD} = 1.93$, $\underline{n} = 101$). Kempe Family Stress Interview scores ranged from 5 to 85 ($\underline{M} = 31$, Median = 40; $\underline{n} = 154$). The mothers' ages ranged from 14 to 40 ($\underline{M} = 20$, $\underline{SD} = 5.1$) at time of the child's birth. Education ranged from 0 to 14 years ($\underline{M} = 11.0$, $\underline{SD} = 2.4$). Gross monthly household income in dollars ranged from \$100.00 to \$2,600.00 ($\underline{M} = \$883$, $\underline{SD} = 575.8$). Thirty-six (21.6% of the sample) mothers were married, five (3%) were divorced, and 124 (75.2%) had never been married. Ethnic identities of the babies were obtained from birth certificates, and used because the information was complete. One hundred and thirty-six (81%) of the babies were described as White, two (1.2%) as Hispanic, two (1.2%) as American Indian, 23 (13.8%) as Asian or Pacific Islander, and four (2.4%) as African-American.

The large number of White children in this sample is reflective of the Oregon population in general (86.6% White; U.S. Census Bureau, 2000). The relatively large number of Asian or Pacific Islander children, and comparatively small number of Hispanic children in this sample is a serious limitation however. According to the January, 2000 Healthy Start Status Report, 73% of 2,486 children receiving intensive home visitation were reported as being White, while 23% were Hispanic or Latino, 1% were African-American, 1% were Asian or Pacific Islander, and 1% were American Indian or Alaskan Native (Katzev, Pratt, Grobe, & McGuigan, 2000). These sample data cannot be considered as representing any larger population. Participants were selected according to

completeness of data. Only 167 of the mothers had completed every item on all five measures of interest (three Parenting Ladders and two Family Progress Scales). Without established psychometrics for the measures of interest, it was considered critical that all participants have complete data.

Procedure

This study is based on secondary data on families who participated in the Oregon Healthy Start program for a minimum of six months. Upon enrollment into the program, parents complete an initial Parenting Ladder (Parenting Ladder – pre). After a family has been enrolled in the program for a month, their home visitor assesses family functioning behaviors using the Family Progress Scale (Family Progress Scale – pre). While there may be a lapse of a month between the two measures, they are both considered to be pretest measures for evaluation purposes.

When the child is six months old, parents complete a second Parenting Ladder (Parenting Ladder – post), as well as the retrospective version (Parenting Ladder – retro). Also after six months, home visitors complete a second Family Progress Scale (Family Progress Scale – post). A parent's sense of understanding, confidence, ability, resources, stress, coping, and support, assessed in the Parenting Ladder, may indicate overall parental self-efficacy. In turn, parental self-perceptions may be related to overall family functioning as assessed by the Family Progress Scale.

Table 2

Measures Completed at Pre and Post Periods

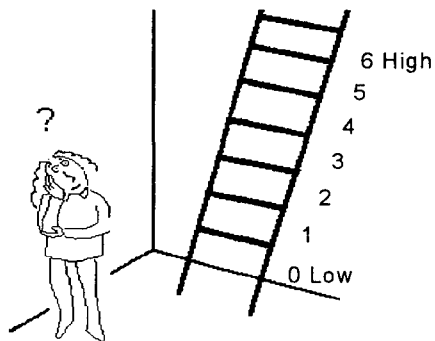
Measure	Rater	Time of Collection	
		Pretest	Posttest
Parenting Ladder			
Parenting Ladder - pre	Parent Self-Report	X	
Parenting Ladder - post	Parent Self-Report		X
Parenting Ladder - retro	Parent Self-Report		X
Family Progress Scale			
Family Progress Scale –pre	Home Visitor	X	
Family Progress Scale –post	Home Visitor		X

Note. The Parenting Ladder - pre is completed before child is one month old. Family Progress Scale - pre is completed within a family's first month of program participation. Parenting Ladder posttest and retrospective measures are completed when child is six months old. Family Progress Scale posttest is completed after six months of program participation.

Measures

Parenting Ladder. The Parenting Ladder was designed to quickly assess a wide variety of information relating to parenthood with one set of global questions. By adding the image of the ladder, the intention was clarify the instrument for parents with low reading comprehension levels, or language barriers (see Figure 1).

Figure 1

Oregon Healthy Start Parenting Ladder

Where are you on the ladder?

Where are you on this ladder?	Low						High
a. Your knowledge of how children grow and develop?	0	1	2	3	4	5	6
b. Your confidence that you know what is right for your child?	0	1	2	3	4	5	6
c. Your ability to help your child learn?	0	1	2	3	4	5	6
d. The amount of stress in your life right now?	0	1	2	3	4	5	6
e. Your ability to cope with the stress in your life?	0	1	2	3	4	5	6
f. Your resources, like money, food, and transportation?	0	1	2	3	4	5	6
g. The amount of helpful advice or moral support you get from other people?	0	1	2	3	4	5	6

The Parenting Ladder is a 7-item self-report measure with a range of 0 (low) to 6 (high). Each of the seven Parenting Ladder items assess a different variable targeted by the program. For the sake of brevity these items (listed in the order found on the Parenting Ladders; See Figure 1) will be referred to as knowledge, confidence, ability, stress, coping, resources, and social support.

The seven items range from the concrete (amount of resources) to the more abstract (amount of stress), yet it is reasonable to believe they could be joined by an underlying construct of parental self-efficacy beliefs. Parental self-efficacy beliefs are the perceptions a parent has regarding her¹ own competence in terms of the complex, and very demanding role of parent (Coleman & Karraker, 1997).

Because the Parenting Ladder is used as a pretest, a posttest and a retrospective pretest measure, there are slightly different instructions on each. Both the Parenting Ladder - pre and the Parenting Ladder - post ask a mother to look at the ladder and record current self-perceptions. The retrospective Parenting Ladder asks a mother to think back, and consider where she was on the ladder at the time that the child was born. The Parenting Ladder - retro is identical to both the Parenting Ladder - pre and the Parenting Ladder - post other than the question alterations that were necessary to place it in the past tense.

In terms of face validity, the inclusion of each of the seven Parenting Ladder items (knowledge, confidence, ability, resources, stress, coping, and support) appears logical. According to Wolfe (1993), knowledge, skills, and support boost

¹ Because 100% of the sample is female, feminine pronouns will be used in reference to the participants.

parents' "coping abilities, [and] increases their ability to resist forces that oppose their healthy adjustment (i.e., stress)". Having knowledge of infant growth and development enables a mother to have reasonable expectations of her child.

Parents may have expectations of a child that are too high and unrealistic if they have little knowledge of what children are developmentally capable of.

Unrealistic expectations are associated with developmentally inappropriate demands or rules for behavior. Alternatively, due to lack of appropriate expectations, a parent may not provide enough structure and challenge for the developing child. It is possible that a parent's lack of understanding of child development is an underlying issue in child maltreatment (Thompson, 1995).

In contrast, understanding a child's needs enables a parent to better adjust to the parenting role, and feel proud and confident of the child's growth and accomplishments (Thompson, 1995). Maternal confidence and knowledge have been found to have an interaction effect on parent-child interactions (Conrad, Gross, Fogg, & Ruchala, 1992). Specifically, what mothers know and understand relates to the level of confidence felt within the parental role, and subsequently affects how they interact with their children.

Research has demonstrated that parental resources, stress levels, and coping skills effect parental ability. For instance, Wolfe (1993) reported that low levels of resources, low levels of coping skills, and high levels of stress threatened effective parenting, because parents were preoccupied and were less likely to attend to their children. Time and energy spent accessing resources and coping

with stress translated into time and attention diverted from the child. Leventhal (1996) noted that families such as those served by the Oregon Healthy Start program spend much of their time and attention on issues such as housing, transportation, and the organization of their lives. Further, self-perceptions of low economic well-being contribute to psychological distress (Fox & Chancey, 1998). Social support has been found to be a buffer for stress (Kotch et al., 1997; Rodgers, 1993), but the parent-child relationship is more affected by a parent's psychological resources, than external sources of support (Belsky, 1984). Therefore, it is how a parent deals with financial and personal issues that effects the parent-child interaction.

At-risk families are particularly vulnerable to stress. According to Spicer and Franklin (1994), the frequency with which parents experience hassles directly affects the verbal aggression and violence directed toward their children. The use of undesirable parenting strategies (including punishment, inconsistency, parental-coldness, sensitization, and rejection-oriented behavior) is both directly and indirectly affected by parental stress (Rogers, 1993).

Family Progress Scale. A second measure used in this study is the Oregon Healthy Start Family Progress Scale, which is a 10-item observational measure of family functioning behaviors completed by home visitors after roughly one month of program participation, and again after six months of program participation, based on family behaviors observed during home visits. The 10-item scale asks

home visitors to rate how frequently the family (a) maintains a stable home life, (b) keeps appointments or call to reschedule, (c) provides child(ren) with adequate and appropriate food, clothing, and shelter, (d) handles routine child-related, household, and family responsibilities appropriately, (e) provides nurturing care for the child(ren), (f) engages in positive parent-child interactions, (g) uses positive guidance and discipline strategies, (h) creates a developmentally appropriate learning environment for child, (i) makes use of positive social support system or person(s) other than home visitor, and (j) makes use of a needed community resources or public support services. Scores reflect the frequency of observed parental behaviors, ranging from 0 (not at this time) to 4 (almost always).

Analysis strategy

Principle component analyses. The reliability of a measurement instrument refers to the degree of consistency in results produced by the instrument, upon repeated usage. One of the assumptions of reliability is that the items on the measurement scale are measuring one construct, and are doing so equally (Carmines & Zeller, 1979). If a violation of this assumption is suspected, principle component analysis can be used to identify the likelihood that the scale is measuring more than one construct. A principle component analysis is one multivariate statistical approach used to define common underlying constructs by examining how variables are interrelated. It is designed to cope with data

that measure a single construct unequally, or measure more than one construct (Carmines & Zeller, 1979). Before reducing the Parenting Ladder to one global measure, it is necessary first to test whether there is justification for doing so.

A principal component analysis was completed for each of the three Parenting Ladders to test whether there was a common construct among the three.

Finding a common construct would simplify later analysis. Those principle component analyses addressed the first research question.

1. Stage One: Principle Component Analyses - Is there one shared construct within the Parenting Ladder pretest, posttest, and retrospective measures?

Principle component analyses were also conducted on the two Family Progress Scales. Identifying one shared construct in the Family Progress Scale measures would simplify the analysis. The results of all principle component analyses will be presented in the results section.

Response-shift bias. Once a principle component was identified, a significant difference between traditional Parenting Ladder pretest means and retrospective Parenting Ladder pretest means was tested for using a paired t-test. That comparison addressed the second research question.

2. Stage Two: Response-Shift Bias (a) - Is there evidence of response-shift bias? Is there a significant mean difference between traditional Parenting Ladder pretest ratings and retrospective Parenting Ladder ratings?

A significant difference between the two pretests would mean that there was a response-shift bias between the pretest period (intake), and the posttest period (6 months). The difference in scores would indicate that the mothers had shifted in their perception of efficacy with respect to the time that their child was born. If there is no significant difference, that would mean that there was no response-shift bias. Further, it would mean that using either of the pretest versions of the Parenting Ladder would produce similar results.

Change scores were calculated by subtracting the prettest or retrospective pretest score from the posttest score, then compared using paired t-tests to address the third research question.

3. Stage Two: Response-Shift Bias (b) - Is there a significant mean difference between the Parenting Ladder - pre/post change scores and the Parenting Ladder - retro/post change scores?

If there is no significant difference found between pre/post change scores and retro/post change scores, that will indicate that the retro/post and the pre/post methods function similarly. If retrospective pretest scores are equivalent to the traditional pretest scores, and change scores demonstrate similar change patterns over time, then the retrospective pretest results will be considered to be equivalent to the traditional pretest results.

Validity. The third stage of this study examined the construct validity of the Parenting Ladder. Validity of a measurement scale refers to the extent of its ability

to measure the construct it is designed to measure. The construct validity of a measurement scale refers to its degree of consistency with other scales (Carmines & Zeller, 1979). In other words, scales that measure related concepts should demonstrate the appropriate correlation.

Parental perceptions have been linked to parenting behaviors, including parent-child interaction (Conrad , Gross, Fogg, & Ruchala, 1992). Parenting style is reflective of other attitudes and behaviors, which effects a child (Newcomb & Loeb, 1999). Proponents of home visiting programs have also seen the link between how parents feel about themselves and their situations, and how the parent actually cares for a child (Leventhal, 1996).

The relationship between maternal self-reports of ability were compared with objective ratings of observed ability. Specifically, the viability of the Parenting Ladder was considered in relation to the Family Progress Scale. To compare the Parenting Ladders to the Family Progress Scales on the same metric scale, the Family Progress Scale scores were transformed mathematically from a 5 - point scale to a 7 – point scale. Correlational analysis tested the construct validity of the Parenting Ladders by comparing the relationship between the traditional Parenting Ladder pretest and the retrospective Parenting Ladder pretest with the Family Progress Scale pretest, and the Parenting Ladder posttest with the Family Progress Scale posttest. These analyses addressed the fourth research question.

4. Stage Three: Validity (a) - Is there any relationship between a specific version of the Parenting Ladder and its corresponding version of the Family Progress Scale?

Research question five addressed the relationship between the change scores on the Parenting Ladders and the change scores on the Family Progress Scales. Change scores were calculated by subtracting the pretest or retrospective pretest score from the posttest score. If Parenting Ladder -retro/post change scores correlate more strongly with Family Progress Scale - pre/post change scores, this would mean that change assessed using the retrospective pretest was more in agreement with observed behavioral change than change assessed with the traditional pretest. If the reverse is true, that Parenting Ladder – pre/post change scores correlate more strongly with Family Progress Scale – pre/post change scores, then change assessed using the traditional pretest-posttest method was more in agreement with observed behavioral change. A correlational analysis between Parenting Ladder - pre/post and Parenting Ladder - retro/post with Family Progress Scale - pre/post addressed the fifth and sixth research questions.

5. Stage Three: Validity (b) - Is there a significant relationship between Parenting Ladder - pre/post change scores and Family Progress Scale – pre/post change scores?
6. Stage Three: Validity (c) - Is there a significant relationship between Parenting Ladder – retro/post change scores and Family Progress Scale – pre/post change scores?

RESULTS

Stage One: Principle Component Analyses

A principle component analysis was conducted for each of the Parenting Ladder versions to test whether they individually measured one or more constructs (see Table 3).

Table 3

Principle Component Loadings for the Parenting Ladders (N = 167)

Item	Parenting Ladder Component Matrix		
	Pre	Post	Retrospective
Knowledge	0.63	0.64	0.66
Confidence	0.76	0.77	0.79
Ability	0.76	0.80	0.76
Coping	0.44	0.53	0.66
Basic Resources	0.56	0.56	0.56
Social Support	0.68	0.68	0.50
Alpha	0.71	0.73	0.73

The stress item was first reverse coded because within this context, higher stress levels would be in opposition of parental skills and abilities. Later, the stress item

was dropped from the scale because it was the strongest contributor to a second construct on all three Parenting Ladders. Without the stress item, a second set of analyses was done. The pretest and retrospective versions continued to show a second construct, but the six remaining items were constrained into one because that second construct was inconsistent (or nonexistent in the posttest). This was done to simplify later analyses.

Eigenvalues (and second eigenvalues, where appropriate) for the three versions were as follows: pretest 2.52 (1.08); posttest 2.70; and retrospective 2.64 (1.23). Parenting Ladder principle component loadings ranged from 0.44 to 0.76 on the pretest; 0.53 to 0.80 on the posttest; and 0.50 to 0.79 on the retrospective version. Alpha values for the pretest, posttest, and retrospective were 0.71, 0.73, and 0.73, respectively.

After an initial principle component analysis of the 10-item Family Progress Scale was conducted, the social support and community resources items were dropped from the scale because they contributed strongly to second construct. The items on keeping appointments and positive guidance were dropped because they contributed strongly to a third construct. Principle component analysis identified a single construct for the remaining six items. Family Progress Scale principle component loadings were between 0.62 and 0.87 for the pretest, and between 0.64 to 0.86 for the posttest. Alpha values were 0.83 and 0.86 for the pretest and posttest, respectively (See Table 4).

Table 4

Principle Component Loadings for the Family Progress Scales (N = 167)

Item	Family Progress Scale	
	Component Matrix	
	Pre	Post
Stable home life	0.63	0.64
Food	0.62	0.65
Routine	0.80	0.77
Nurturing	0.87	0.85
Interaction	0.84	0.86
Developmentally appropriate environment	0.74	0.82
Alpha	0.83	0.86

Stage Two: Response-Shift Bias

The means (and standard deviations) of the Parenting Ladder pretest, posttest, and retrospective version were 5.30 (0.83), 5.63 (0.73) and 4.83 (0.94), respectively, on a 7 point scale. Using a two-tailed, paired t-test, there was a significant difference between participant self-ratings on the conventional pretest and their self-ratings on the retrospective pretest, $t(166) = 6.80$, $p < .001$. This is evidence of response-shift bias, because it is estimated that participants rated

themselves almost half a point (0.47) lower in retrospection, compared to initial self-reports (95% confidence interval from 0.33 to 0.61).

Parenting Ladder change scores were calculated by subtracting pretest ($M = 5.30$, $SD = 0.83$) or retrospective pretest ($M = 4.83$, $SD = 0.94$) ratings from posttest ratings ($M = 5.63$, $SD = 0.73$). Parenting Ladder – pre/post change scores ranged from -1.83 to 2.83 ($M = 0.33$, $SD = 0.75$), Parenting Ladder – retro/post change scores ranged from -2.17 to 3.17 ($M = 0.80$, $SD = 0.82$). Using a two-tailed, paired t-test to compare change scores yielded a significant difference between Parenting Ladder - pre/post change scores and Parenting Ladder - retro/post change scores, $t(166) = 6.80$, $p < .001$. The difference between the two sets of change scores is estimated to be 0.47, with the Parenting Ladder – retro/post reporting more positive participant change (95% confidence interval from 0.33 to 0.61). While both methods found evidence of significant participant change over time, the traditional pre/post measure and the retro/post measure produced significantly different estimates of how much the mothers changed. Using two-tailed, paired t-tests, the traditional pre/post method estimated the change in self-ratings to be 0.33, a significant difference over time, $t(166) = 5.67$, $p < .001$. The retro/post method showed stronger results, however, estimating the difference to be 0.80, $t(166) = 12.509$, $p < .001$. In other words, after six months of participation in the Oregon Healthy Start program, mothers retrospectively rated their initial parental knowledge and skills lower than they had originally reported on the pretest.

Stage Three: Validity

There appeared to be a ceiling effect with Family Progress Scale pretest and posttest means (and standard deviations) being 6.05 (0.86) and 6.26 (0.74), respectively, on a 7 point scale. There are a couple of possible explanations for the high ratings. For example, the families have allowed the home visitor in to observe. It may be that the home visitor observed families primarily “on their best behavior”. Family Progress Scale – pre/post change scores ranged from –2.10 to 2.57 ($M = 0.20$, $SD = 0.73$).

The correlation matrix of individual Parenting Ladders and their corresponding Family Progress Scales was examined (See Table 5). The only significant relationship between Parenting Ladder versions and corresponding Family Progress Scales was the relationship between Parenting Ladder pretests and Family Progress Scale pretests ($r = 0.21$, two-tailed $p = .08$). The relationship between Parenting Ladder retrospective pretests and Family Progress Scale pretests was not significant ($r = 0.12$, two-tailed $p = .13$), nor was the relationship between Parenting Ladder posttests and Family Progress Scale posttests significant ($r = 0.10$, two-tailed $p = .19$). The low correlations between the Parenting Ladder versions and Family Progress Scale versions may reflect the difficulties in correlating self-report measures to behavioral ones (Howard, Maxwell, Wiener, Boynton, & Rooney, 1980).

Table 5

Correlational Validity Results (N = 167)

	Family Progress Scale		
	(M, SD)		
	Pre	Post	Pre/post change
Parenting Ladder (M, SD)	(6.05, 0.86)	(6.26, 0.74)	(0.20, 0.73)
Pre (5.30, 0.83)	0.21**		
Post (5.63, 0.73)		0.10	
Retrospective (4.83, 0.94)	0.12		
Pre/post change (0.33, 0.75)			0.13
Retro/post change (0.80, 0.82)			0.16*

Note. Family Progress Scales and Parenting Ladders scores can range from 1 to 7.

* $p < .05$. ** $p < .001$.

The correlation matrix of change scores, however, showed a different picture (See Table 5). The Parenting Ladder - retro/post self-rated measure of change was estimated to be more in agreement with the observed behavioral change.

The Parenting Ladder – post/retro change scores were more closely associated with the Family Progress Scale – pre/post change scores ($r = 0.16$, two-tailed $p = 0.05$), as opposed to the Parenting Ladder – pre/post ($r = 0.12$, two-tailed $p = 0.11$). These correlations, although relatively low, show support for using this

retrospective pretest in measuring changed perceptions over time, because without the retrospective Parenting Ladder, self-perceived change over time might have been underreported as 0.33 (Parenting Ladder pre/post change scores mean), rather than 0.80 (Parenting Ladder retro/post change scores mean) on a 7 point scale.

DISCUSSION

Evidence of response-shift bias has been found in measures of job satisfaction (Gutek & Winter, 1992), and a wide array of intervention program evaluations (e.g., Howard et al. 1979; Rhodes & Jason, 1987). A similar bias was evident in the maternal self-ratings of efficacy. This study found a retrospective pretest led to different results than a conventional pretest. Specifically, the pre/post method produced a more conservative estimate of maternal self-perceived improvement, compared to the retro/post method. Considering that the retro/post method found more than twice the improvement that the pre/post method found, the decision of which pretest to use also may decide the amount of improvement found when there is a response-shift. This echoes Howard and his colleagues (1979) who pointed out that program effects can be missed, or underrated without the use of the retrospective pretest.

The difficulty in correlating self-report and behavioral measures has been written about (e.g., Howard, Maxwell, Wiener, Boynton, & Rooney, 1980), and was demonstrated here. The individual measures primarily showed no significant relationship between maternal perceptions of self-functioning and observed behavioral family functioning. The initial family functioning measure (Family Progress Scale pretest) was related more closely to the mothers' initial self-ratings (Parenting Ladder pretest) than to her retrospective self-ratings (Parenting Ladder retrospective pretest), which would seem to suggest a superior validity on the part of the conventional pretest. The post/retro method of measuring maternal

self-efficacy change, however, was more in agreement with the observed behavioral change. This suggests the retrospective method may have superior validity in considering change over time.

To summarize, this study concludes that this retrospective measure allows program staff a perspective of participant change that is different than the perspective gained with the traditional pretest. While a traditional pretest enables a program to be tailored to meet participant needs, a retrospective version allows participants to integrate new knowledge into self-reported change measures. More research is needed on both the Parenting Ladder and the Family Progress Scale measures, as neither have established psychometrics. The validity of the Parenting Ladder certainly needs to be further tested, due to the inconsistent relationship between it and the Family Progress Scale.

It is questionable whether the Family Progress Scale was the right measure to use for validation purposes. In considering the difficulties in comparing self-report measures to behavioral measures (Howard, Maxwell, Wiener, Boynton, & Rooney, 1980), future work validating the Parenting Ladder might be better served by finding a different measure than the Family Progress Scale.

BIBLIOGRAPHY

- Aiken, L.S., & West, S. G. (1990). Invalidity of true experiments: Self-report pretest biases. Evaluation Review, 14(4), 374-390.
- Aronson, E. (Ed.). (1995). Readings About The Social Animal (7th ed.). New York: W. H. Freeman.
- Azar, S. T. (1997). A cognitive behavioral approach to understanding and treating parents who physically abuse their children. In D. A. Wolfe, R. J. McMahon, & R. D. Peters. (Eds.), Child Abuse: New Directions in Prevention and Treatment Across the Lifespan (pp. 79-101). Thousand Oaks, CA: Sage.
- Barratt, M.S., Roach, M. A., Morgan, K. A., & Colbert, K. K. (1996). Adjustment to motherhood by single adolescents. Family Relations, 45(2), 209-215.
- Benjamin, E. R. (November, 1982). Using the "post-then" method of evaluation. Training / HRD, 72.
- Berger, P. L. (1963). Invitation to Sociology: A Humanistic Perspective. New York: Anchor Books.
- Bogensneider, K. (1996). An ecological risk/ protective theory for building prevention programs, policies, and community capacity to support youth. Family Relations, 45(2), 127-138.
- Booth, C. L., Barnard, K. E., Mitchell, S. K., & Spieker, S. J. Successful intervention with multi-problem mothers: Effects on the mother-infant relationship. Infant Mental Health Journal, 8(3), 288-306.
- Bray, J. H., & Howard, G. S. (1980). Methodological considerations in the evaluation of a teacher-training program. Journal of Educational Psychology, 72(1), 62-70.

- Bray, J. H., Maxwell, S. E., & Howard, G. S. (1984). Method of analysis with response-shift bias. Educational and Psychological Measurement, 44(4), 781-804.
- Breaky, G., & Pratt, B. (1991). Healthy growth for Hawaii's "Healthy Start": Toward a systematic statewide approach to the prevention of child abuse and neglect. Zero to Three (Bulletin of National Center for Clinical Infant Programs), 11(4), 16-22.
- Brody, G. H., & Sigel, I. E. (1990). Methods of Family Research: Biographies of Research Projects. Volume II Clinical Populations. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. American Psychologist, 32(7), 513-531.
- Brooks, L., & Gersh, T. L. (1998). Research in brief: Assessing the impact of diversity initiatives using retrospective pretest design. Journal of College Student Development, 39(4), 383-385.
- Cicchetti, D., & Carlson, V. (Eds.). (1989). Child Maltreatment: Theory and research on the causes and consequences of child abuse and neglect. New York: Cambridge University.
- Claussen, A. H. (1991). Physical and psychological maltreatment: Relations among types of maltreatment. Child Abuse and Neglect, 15, 5-18.
- Coleman, P. K., & Karraker, K. H. (1997). Self-efficacy and parenting quality: Findings and future applications. Developmental Review, 18, 47-85.
- Conrad, B., Gross, D. Fogg, L., & Ruchala, P. (1992). Maternal confidence, knowledge, and quality of mother-toddler interactions: A preliminary study. Infant Mental Health Journal, 13(4), 353-362.

- Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. Journal of Personality and Social Psychology, 47(4), 738-748.
- Dumka, L. E., Stoerzinger, H. D., Jackson, K. M., & Roosa, M. W. (1996). Examination of the cross-cultural and cross-language equivalence of the Parenting Self-Agency Measure. Family Relations, 45(2), 212-222.
- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. Child Development, 65, 296-318.
- Farber, N. (1994). Perceptions of pregnancy risk: A comparison by class and race. American Journal of Orthopsychiatry, 64(3), 479-484.
- Feiring, C., Fox, N. A., Jaskir, J., & Lewis, M. (1987). The relation between social support, infant risk status and mother-infant interaction. Developmental Psychology, 23(3), 400-405.
- Fields, M. V., & Boesser, C. (1994). Constructive Guidance and Discipline: Preschool and Primary Education. New York: Macmillan.
- Fisher, P. A., Fagot, B. I., & Leve, C. S. (1998). Assessment of family stress across low-, medium-, and high-risk samples using the Family Events Checklist. Family Relations, 47(3), 215-219.
- Fox, G. L., & Chancey, D. (1998). Sources of economic distress: Individual and family outcomes. Journal of Family Issues, 19(6), 725-749.
- Friedman, S. L., & Haywood, H. C. (Eds.). (1994). Developmental Follow-up: Concepts, Domains, and Methods. San Diego, CA: Academic Press.
- Ganong, L. H., & Coleman, M. (1995). The content of mother stereotype. Sex Roles, 32(7-8), 495-512.

- Goedhart, H., & Hoogstraten, J. (1992). The retrospective pretest and the role of the pretest information in evaluation studies. Psychological Reports, 70, 699-704.
- Gutek, B. A., & Winter, S. J. (1992). Consistency of job satisfaction across situations: Fact or framing artifact? Journal of Vocational Behavior, 41(1), 61-78.
- Hjelle, L. A., Busch, E. A., & Warren, J. E. (1996). Explanatory style, dispositional optimism, and reported parental behavior. Journal of Genetic Psychology, 1557(4), 489-499.
- Hooper, F. H., & Hooper, J. O. (1990). The family as a system of reciprocal relations: Searching for a developmental lifespan perspective. In I. E. Sigel, & G. H. Brody, (Eds.), Methods of Family Research: Biographies of Research Projects. Volume 1: Normal Families (pp. 289-316). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/ post self-reports. Evaluation Review, 4(1), 93-106.
- Howard, G. S. (1981). On validity. Evaluation Review, 5(4), 567-576.
- Howard, G. S. (1982). Improving methodology via research on research methods. Journal of Counseling Psychology, 29(3), 318-326.
- Howard, G. S., & Dailey, P. R. (1979). Response-shift bias: A source of contamination of self-report measures. Journal of Applied Psychology, 64(2), 144-150.
- Howard, G. S., & Maxwell, S. E. (1982). Linked raters' judgements: A more sensitive index of change. Evaluation Review, 6(1), 140-146.

- Howard, G. S., Maxwell, S. E., Wiener, R. L., Boynton, K. S., & Rooney, W. M. (1980). Is a behavioral measure the best estimate of behavioral parameters? Perhaps not. Applied Psychological Measurement, 4(3), 293-311.
- Howard, G. S., Millham, J., Slaten, S., & O'Donnell, L. (1981). Influence of subject response style effects on retrospective measures. Applied Psychological Measurement, 5(1), 89-100.
- Howard, G. S., Obledo, F. H. Cole, D. A., & Maxwell, S. E. (1983). Linked raters' judgements: Combating problems of statistical conclusion validity. Applied Psychological Measurement, 7(1), 57-62.
- Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. Applied Psychological Measurement, 3(1), 1-23.
- Kaplan, L., & Girard, J. L. (1994). Strengthening High-Risk Families: A Handbook For Practitioners. New York: Lexington Books.
- Katzev, A. (2000). Oregon Healthy Start Evaluation Manual. Corvallis, OR: Oregon Healthy Start Evaluation Project.
- Katzev, A. R., Pratt, C. C., with Grobe, D., & McGuigan, W. (January, 2000). Oregon's Healthy Start Effort 1998 - 99 Status Report. Corvallis, OR: Oregon State Family Policy Program.
- Katzev, A. R., Pratt, C. C., with Henderson, T., & McGuigan, W. (May, 1999). Oregon's Healthy Start Effort 1997 - 1998 Status Report. Corvallis, OR: Oregon State University Family Policy Program.
- Katzev, A. R., Pratt, C. C., Henderson, T., & Ozretich, R. (December, 1997). Oregon's Healthy Start 1996 - 1997 Status Report: Part I Family Assessment and Service Delivery. Corvallis, OR: Oregon State University Healthy Start Evaluation Project.

- Katzev, A. R., Pratt, C. C., Henderson, T., & Ozretich, R. (March, 1998). Oregon's Healthy Start 1996 - 1997 Status Report: Part II Performance Outcomes for Children and Families. Corvallis, OR: Oregon State University Healthy Start Evaluation Project.
- Kempe, C. H. (1976). Approaches to preventing child abuse: A health visitors concept. American Journal of Disabled Children, 130, 941-947.
- Kim, S., Crutchfield, C., Williams, C., & Hepler, N. (1994). An innovative and unconventional approach to program evaluation in the field of substance abuse prevention: A threshold-gating approach using single system evaluation designs. Journal of Community Psychology, CSAP Special Issue, 61-78.
- Kitzman, H., Olds, D. L., Henderson, C. R. Jr., Hanks, C., Cole, R., Tatelbaum, R., McConnochie, K. M., Sidora, K., Luckey, D. W., Shaver, D., Englehardt, K., James, D., & Barnard, K. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. Journal of American Medical Association (JAMA), 278(8), 644-652.
- Kotch, J. B., Browne, D. C., Ringwalt, C. L., Dufort, V., Ruina, E., Stewart, P.W., & Jung, J. -W. (1997). Stress, social support, and substantiated maltreatment in the second and third years of life. Child Abuse and Neglect, 21(11), 1025-1037.
- Leventhal, J.M. (1996). Twenty years later: We do know how to prevent child abuse and neglect. Child Abuse and Neglect: The International Journal, 20, 647-653.
- Linn, R. L., & Slinde, J. A. (1977). The determination of the significance of change between pre- and posttesting periods. Review of Educational Research, 47(1), 121-150.
- Mann, B. J., Sylvester, C. E., & Chen, R. (1996). Reliability and validity of the retrospective Childhood and Current Adult Fear Survey (RCCAFS). Journal of Anxiety Disorders, 10(1), 59-72.

- Manthei, R. J. (1997). The response-shift bias in a counsellor education programme. British Journal of Guidance and Counseling, 25(2), 229-237.
- Marcenko, M. O. & Spence, M. (1994). Home visitation services for at-risk pregnant and postpartum women: A randomized trial. American Journal of Orthopsychiatry, 64(3), 479-484.
- Milner, J. S., & Dopke, C. (1997). Child physical abuse: Review of offender characteristics. In D. A. Wolfe, R. J. McMahon, & R. D. Peters. (Eds.). Child Abuse: New Directions in Prevention and Treatment Across the Lifespan (pp. 27-54). Thousand Oaks, CA: Sage.
- Murphy, S., Orkow, B., & Nicola, R. M. (1985). Prenatal prediction of child abuse and neglect: A prospective study. Child Abuse and Neglect, 9, 225-235.
- Newcomb, M. D., & Loeb, T. B. (1999). Poor parenting as an adult problem behavior: General deviance, deviant attitudes, inadequate family support and bonding, or just bad parents? Journal of Family Psychology, 13(2), 175-193.
- Nicholson, T., Belcastro, P. A., & Gold, R. S. (1985). Retrospective pretest-posttest versus traditional pretest-posttest analysis. Psychological Reports, 57(2), 525-526.
- Olds, D. L., Eckenrode, J., Henderson, C. R. Jr., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P., Pettitt, L. M., & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. Journal of American Medical Association (JAMA), 278(8), 637-643.
- Olds, D., & Korfmacher, J. (1997). The evolution of a program of research on prenatal and early childhood home visitation. Journal of Community Psychology, 25(1), 1-7.

- Olds, D., Kitzman, H., Cole, R., & Robinson, J. (1997). Theoretical foundations of a program of home visitation for pregnant women and parents of young children. Journal of Community Psychology, 25(1), 9-25.
- Orkow, B. (1985). Implementation of a family stress checklist. Child Abuse and Neglect, 9, 405-410.
- Peterson, J., & Hawley, D. R. (1998). Effects of stressors on parenting attitudes and family functioning in a primary prevention program. Family Relations, 47(3), 221-227.
- Phelps, J. L., Belsky, J., & Crnic, K. (1998). Earned security, daily stress, and parenting: A comparison of five alternative models. Development & Psychopathology, 10(1), 21-38.
- Pianta, R. Egeland, B., & Erickson, M. F. (1989). The antecedents of maltreatment: Results of the Mother-Child Interaction Research Project. In D. Cicchetti, & V. Carlson, (Eds.), Child Maltreatment: Theory and research on the causes and consequences of child abuse and neglect (pp. 203-253). New York: Cambridge University.
- Pratt, C. C., Katzev, A., Moran, P., & Eddy, L. (1995). Oregon's Healthy Start Effort Status Report: July 1, 1994 – June 30, 1995. Corvallis, OR: Oregon State University, Healthy Start Evaluation Project.
- Pratt, C. C., Katzev, A., Ozretich, R., Henderson, T., & McGuigan, W. (January, 1998). Building Results III: Measuring Outcomes for Oregon's Children, Youth, and Families. Corvallis, OR: Oregon State University, Family Policy Program.
- Preziosi, R. C., & Legg, L. M. (May, 1983). Add "then" testing to prove training's effectiveness. Training, 48-49.
- Reid, W. J. (1993). Fitting the single-system design to family treatment. Journal of Social Service Research, 18(1-2), 83-99.

- Rhodes, J. E., & Jason, L. A. (1987). The retrospective pretest: An alternative approach in evaluating drug prevention programs. Journal of Drug Education, 17(4), 345-355.
- Robinson, E. A., & Doueck, H. J. (1994). Implications of the pre/post/then design for evaluating social group work. Research on Social Work Practice, 4(2), 224-239.
- Rockwell, S. K., & Kohn, H. (Summer, 1989). Post-then-pre evaluation: Measuring behavior change more accurately. Journal of Extension, 19-21.
- Rogers, A. Y. (1993). The assessment of variables related to the parenting behavior of mothers with young children. Children and Youth Services, 15, 385-402.
- Rossi, P. H., & Freeman, H. E. (1993). Evaluation: A Systematic Approach (5th ed.). Newbury Park, CA: Sage.
- Rubin, A., & Babbie, E. (1997). Research Methods for Social Work (3rd ed.). Pacific Grove, CA: Brooks/ Cole.
- Rueter, M. A., Conger, R. D., & Ramisetty-Mikler, S. (1999). Assessing the benefits of a parenting skills training program: A theoretical approach to predicting direct and moderating effects. Family Relations, 48(1), 67-77.
- Salkind, N. J. (1997). Exploring Research (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. American Psychologist, 54(3), 183-203.
- Schilmoeller, G. L., & Baranowski, M. D. (1985). Childrearing of firstborns by adolescent and older mothers. Adolescence, 20(80), 805-822.

- Schwartz, N. (1999). Self-reports: How the questions shape the answers. American Psychologist, 54(2), 93-105.
- Sigel, I. E., & Brody, G. H. (Eds.) (1990). Methods of Family Research: Biographies of Research Projects. Volume 1: Normal Families. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Smith, J. A. (1994). Reconstructing selves: An analysis of discrepancies between women's contemporaneous and retrospective accounts of the transition to motherhood. British Journal of Psychology, 85(3), 371-392.
- Spicer, S. J., & Franklin, C. (1994). Exploratory effects of social support, stress, and locus of control on the conflict tactics of parents at-risk for child maltreatment. Journal of Social Service Research, 19(3/4), 1-22.
- Sprangers, M., & Hoogstraten, J. (1987). Response-shift effects, response-shift bias and a bogus- pipeline. Psychological Reports, 61(2), 579-585.
- Sprangers, M., & Hoogstraten, J. (1988). Response-shift effects, response-shift bias and a bogus- pipeline: A replication. Psychological Reports, 62(1), 11-16.
- Sprangers, M., & Hoogstraten, J. (1989). Pretesting effects in retrospective pretest-posttest designs. Journal of Applied Psychology, 74(2), 265-272.
- Thompson, R. A. (1995). Preventing Child Maltreatment Through Social Support. Thousand Oaks, CA: Sage.
- Toedter, J. J., Lasker, J. N., & Campbell, D. T. (1990). The comparison group problem in bereavement studies and the retrospective pretest. Evaluation Review, 14(1), 75-90.

- Tolin, P. H., Guerra, N. G., & Kendall, P. C. (1995). A developmental-ecological perspective on anti-social behavior in children and adolescents: Toward a unified risk and intervention framework. Journal of Consulting and Clinical Psychology, 63(4), 579-584.
- Tower, C. C. (1996). Child Abuse and Neglect (3rd ed.). Needham Heights, MA: Allyn & Bacon.
- Vaughn, B. E. (1982). Individual differences in qualities of caregiving during the first six months of life: Antecedents in maternal and infant behavior during the newborn period. Infant Behavior and Development, 5, 77-95.
- Wolfe, D. A. (1993). Prevention of child neglect: Emerging issues. Criminal Justice and Behavior, 20(1), 90-111.
- Wolfe, D. A., McMahon, R. J., & Peters, R. D. (Eds.) (1997). Child Abuse: New Directions in Prevention and Treatment Across the Lifespan. Thousand Oaks, CA: Sage.
- US Census Bureau. (2000). Oregon's QuickFacts from the US Census Bureau. Retrieved April 4, 2001 from the World Wide Web:
<http://quickfacts.census.gov/qfd/states/41000.html>