Recent changes in the health care environment have placed an increasing emphasis on empirically supported therapies (ESTs; Chambless & Hollon, 1998; Kendall, 1998; Spirito, 1999) for psychological interventions. In response, the American Psychological Association (APA) established a task force to begin the process of determining which psychotherapies consistently lead to the positive outcome required to meet criteria established for ESTs (APA, 1995). This emphasis on ESTs has affected the current and future direction of both the science and practice of pediatric psychology as evidenced by the recent special series on ESTs in the *Journal of Pediatric Psychology* (JPP). Consequently, all of psychology treatment’s eggs are placed in one basket, the “outcome basket,” as evidenced by the recent EST literature (Chambless et al., 1996; Crits-Christoph, Frank, Chambless, Brody, & Karp, 1995; Wampold et al., 1997). Such heavy emphasis on outcome data can be problematic, however, as threats to the validity of inferences from such outcome data have largely been overlooked (Wampold, 1997). One such serious threat to validity of outcome data is response shift.

Concern over response shift’s potential impact on quality of life (QOL) research led to a recent special edition of the journal *Social Science and Medicine* addressing this problem. Given that outcome data are now being used for many reasons by legislators, third party payers, and psychologists, the stakes for psychology are very high. We must take steps to ensure that our outcome data are interpreted correctly.
This article will discuss potentially serious threats to the validity of outcome data, with an emphasis on response shift, and present one statistical method that may be used to determine the effect of response shift on pediatric outcome data.

Response shift has been defined as the change in one's internal standard, that is, “subjects’ basis for determining their level of functioning on a given dimension” (Sprangers & Hoogstraten, 1989, p. 265). More recently, Sprangers and Schwartz refined the definition of response shift as referring to a “change in the meaning of one’s self evaluation of a target construct as a result of: (a) a change in the respondent’s internal standard” (scale recalibration in psychometric terms); “(b) a change in the respondent’s values (i.e., the importance of component domains constituting the target construct); or (c) a redefinition of the target construct (i.e., reconceptualization)” (1999, p. 1508). Campbell and Stanley (1963) defined instrumentation as “autonomous changes in the measuring instrument which might account for an O1-O2 difference” (p. 9). That is, response shift might be considered a form of instrumentation due to internal change in the observer (self-report).

To illustrate one example of response shift, some (Breetvelt & Van Dam, 1991) suggested that response shift is a particular problem with QOL data collected from cancer patients. They noted that studies examining QOL in cancer patients have often reported an incongruity between the degree of negative feelings self-reported by the patients and those they appear to be experiencing (Breetvelt & Van Dam, 1991). The underreporting does not appear to occur as frequently in terms of reporting somatic symptoms such as fatigue or nausea but, rather, seems to be most common with attitudes or emotional variables. One explanation proposed by Breetvelt and Van Dam is that cancer patients make a shift in their choice of reference group in terms of the status of their own illness. Thus, although their level of complaints or well-being may change, their evaluation scores may not change because they would use another reference group for comparison. For instance, assume that a child is diagnosed with cancer. As the disease progresses, the child increasingly encounters other ill children, some of whom are at more advanced stages of illness. When asked to rate himself or herself, the child uses the children with more advanced stages of illness as a reference group for making self-ratings.

Other findings that can be interpreted in terms of response shift would include aggressive children’s exaggerated self-ratings of competence and relatedness with others (Diener & Milich, 1997; Hughes, Cavell, & Grossman, 1997). Numerous additional examples are found in studies that suggest response shift may play an important role in adapting to illness (see Sprangers & Schwartz, 1999). Furthermore, Wilson (1999) has illustrated how response shift is a part of typical clinical care and may even be viewed as a specific type of placebo effect. All of this suggests that response shift may be an important phenomenon that may occur in any field where self-reports are used. Given the lack of response shift research with pediatric populations, many issues regarding the response shift phenomenon remain unknown. For example, a child’s proneness to response shift may be a function of age. Type of illness may also affect the development of response shift with some experiencing sudden shifts and others gradual shifts.

The purpose of this article is to discuss the response shift phenomenon and to address related methodological issues. We introduce growth modeling as a potentially useful statistical tool in determining what, if any, effect response shift may have on outcome data. This discussion stems from the increasing use of QOL outcome data in supporting legislation and third party reimbursement. An extended discussion of conceptual issues concerning response shift and related concepts precedes the conceptual foundations of growth modeling. Then, for illustrative purposes, we demonstrate growth modeling, using longitudinal data (Clay et al., 1995) of adjustment to pediatric chronic illness.

**Conceptual Models of Response Shift**

Sprangers and Schwartz (1999) presented a theoretical model of how response shift would affect health-related QOL evaluations as a result of changes in one’s health. The model consists of a catalyst, antecedent characteristics, and mechanisms or processes that accommodate the catalyst, which lead to a response shift of one’s self-evaluation of his or her perceived QOL. The model begins with a catalyst, which refers to a change in health status (not necessarily due to a treatment). Antecedents are stable or dispositional characteristics of the individual (e.g., personality or gender) that affect the response to the catalyst and play a role in terms of the mechanisms used to accommodate the catalyst. Mechanisms are behavioral, cognitive, and affective
Constructs Related to Response Shift

Gibbons (1999) presented a mediation model where significant life events lead to a change in the social comparisons one makes, which then may lead to a response shift. One may make numerous types of social comparison changes, for example, downward and upward comparisons, selective self-focus, and changes in social comparison over time. Downward comparisons consist of making comparisons with others who are worse off and are potentially adaptive when improvement is possible or control attainable. Downward comparisons may be detrimental, even when control is possible (Gibbons, 1999).

Gibbons (1999) also notes that cognitive changes, such as those often experienced by depressed clients or by those who are ill and have a poor prognosis, may lead to a negative response shift. Additionally, other variables may also play a moderating role in how social comparison affects response shift. For instance, optimism is a variable that may reduce one’s tendency to make lower social comparisons; that is, it may serve as a deterrent to making downward comparisons.

Social comparison appears to be a potentially useful concept in accounting for response shift. Other concepts may prove to be more or less central in the development of a response shift. For instance, Dweck’s (1986) model of self-concept appears to hold promise as another mediational construct relevant to response shift. In addition, other investigators (Rapkin, 2000) have noted the complex relationship between personal goals and QOL, suggesting that changes in personal goals are fundamentally involved in how individuals appraise their QOL. In a QOL study with AIDS patients, the relationship between personal goals and response shifts was examined. Rapkin (2000), using a series of hierarchical regressions, found response shift effects in numerous aspects of QOL in addition to other variables, such as initial QOL, changes in goals, number of goal attainment strategies, effects of stressors, and indicators of active coping.

Research Design Issues Related to Response Shift

The main dilemma for the applied researcher is to be able to show whether or not a response shift occurred, and if one did occur, to be able to assess its impact on the results. This is addressed here as an...
identifiability problem. Addressing the issue of when the response shift occurred could be beneficial in the interpretation of outcome data. Showing that a response shift occurred is not as informative as being able to detect when the shift occurred.

Most of the techniques discussed in the literature focus on addressing the question of whether or not a response shift occurred and how one would detect it (e.g., Schwartz & Sprangers, 1999), rather than the more informative approach of attempting to determine when such a shift occurred and its form or shape. Given the lack of response shift research, however, it may be reasonable in some instances to limit one’s investigation to simply detecting a response shift.

There are two basic alternatives for detecting a response shift; the first is to address the response shift phenomenon from a research design perspective. The second alternative is to address response shift from a statistical framework. In terms of research design, for instance, one avenue for detecting when a response shift occurred would be a research design in which multiple randomly assigned groups were all given preintervention assessments. This would be followed with postintervention assessments in which the multiple times of assessment would be systematically varied by group. Such a design could detect when a response shift occurred. To detect a response shift, one would expect that all the groups would exhibit a response shift, such as a sudden increase in quality of life ratings, within the same time frame.

Ideally, any research design used to detect a response shift would involve one or more control groups. Unfortunately, control groups are difficult or impossible to obtain in many instances. Furthermore, control groups do not always provide complete control over other threats to internal validity, particularly for response shift. In addition to such research design considerations, it may prove useful to not only collect information on the construct of interest, for instance QOL, but also on self-concept or self-elements and comparisons of self vis-à-vis others (Gibbons, 1999). Adding both design components and appropriate variables could increase the chance that response shift would be detected.

At this point, comment on a common method to detect response shift seems relevant. Frequently, response shift has been discussed in terms of comparisons between pretest and posttest. Using “then” ratings (within a retrospective pretest-posttest design), for example, has been proposed as a way to measure response shift (Howard, Schmeck, & Bray, 1979). A “then test” requires that at posttest, one completes a self-report measure on one’s view of oneself before the intervention. In essence, the then test is merely a second, retrospective pretest. Thus, both the posttest and the then test are completed at the same time and assumed to be completed with the same internal measurement standard. By comparing the posttest and then test scores, one can then eliminate treatment-induced response shift effects, allowing one to detect unconfounded treatment effects. A comparison of the pretest and then test scores reflects an estimate of the response shift effects in amount and direction. One may even compare subjective outcomes with objective outcomes for further evidence of a response shift.

Although any study could probably implement “then” ratings, this method has several shortcomings. First, if one’s data set contains numerous occasions of measurement, the use of “then” ratings only measures if a response shift occurred by comparing the pretest, posttest, and then test results. The analysis does not address when the shift occurred or its form. It may be possible to adapt this method so that it could detect when a response shift occurred by comparing the posttest rating with “then” ratings conducted at multiple time points between the pretest and posttest occasions, but this has the potential to become problematic with numerous occasions of measurement. More important, using the then test method may be inappropriate for self-reports of children, whose ability to retrospectively rate thoughts and feelings may be limited by cognitive development. Even use with adults is suspect if response shift has occurred, since a cognitive and affective reorganization internally makes the self-evaluation suspect.

The second alternative for detecting a response shift is to use a statistical approach. One potentially useful method is growth curve modeling; when combined with multiple variables, such as a measure of QOL, self-concept, and ratings of self in comparison with others, it could prove informative. Moreover, such a technique goes beyond the first step of simply determining whether a response shift occurred. The timing and form or shape of the response shift may be modeled as well. In short, using growth curve modeling with variables related to self-concept and assessment of self in comparison with others, in addition to the construct of interest, may enable the interested investigator to model both the timing of the response shift and its shape.
Internal Validity and Response Shift

This section begins with an overview of the more common threats to internal validity and how these relate to response shift. These threats are central to the identifiability of a response shift. If a major threat cannot be excluded, there is serious question about whether a response shift has been identified, either for an individual or within a group. Cook and Campbell (1979) detailed a number of internal validity threats that, to varying degrees, may be alternative explanations to response shift. As in any experimental or quasi-experimental design, there is almost never a completely determined cause assignable to a given study, but logic and probability provide the basis for giving weight to the alternatives. Threats such as history, maturation, and regression are viable problematic alternatives, whereas selection, testing, and mortality are either lesser threats or generally nonexistent in the response shift conditions considered (see Glass, Willson, & Gottman, 1975, for a discussion of threats in time series designs).

History will always be a threat in longitudinal research. It is, perhaps, the most difficult to detect because a specific historical event or events must be discovered that fit the time interval during which the response shift might have occurred. For individuals, such events are quite probable; personal history involves a succession of events, some of which may be trait changing. For a group of individuals, however, history as a threat to internal validity must posit an event that simultaneously affected most or at least some of the individuals enough to appreciably change the measured trait. If all individuals are members of some group, the search for this event may be conducted through interview or observation; if the participants are independent, the likelihood of such an event simultaneously changing the trait will be small unless the event occurs in a common setting where all participants are located, such as a hospital.

Maturation is conceived in the validity literature as occurring in two forms: short- and long-term. Short-term maturation is exemplified by fatigue and learning, whereas long-term maturation deals with psychophysical development, cultural changes, and environmental changes that can affect psychological constructs. In long-term response shift situations when measurements are made several months apart, long-term maturation is potentially important. Long-term maturation may be especially relevant due to the emerging emphasis on long-term outcomes in pediatric psychological intervention research. The key to maturation as an alternative to response shift as an explanation for change in a psychological construct is that maturational changes should be expected for the age and developmental level of the individual. For children, markers for such maturation are reasonably well known in the literature; for adults, there are far fewer definable age periods in which definable maturation is expected.

Regression is a threat characterized by extremity in selection combined with subsequent multiple measurements. Extremity is defined in terms of the populations under study. Whereas a patient undergoing cancer treatment is “extreme” with respect to the general population on health indicators, if one considers the population of cancer patients, the person may be average. A subsequent assessment of health will likely result in a regressed score with respect to the general population, yet no change with respect to the patient population. Furthermore, expected regression phenomena are always directed toward the population mean. Therefore, any response shift that is contrary in direction will be distinguishable from regression. For regression changes and response shifts that are expected to operate in the same direction with measures normed on the general population, regression effects can be computed if the retest reliability is known for the population over the period of time under consideration. The expected regression in standard deviation units is equal to the square root of 1.0 minus reliability. Thus, for a test-retest reliability over 6 months of .7 on a depression scale, the expected improvement due to regression for someone 1 standard deviation above average is .55 standard deviations. For someone 2 standard deviations above average, the expected change toward the mean will be 1.1 standard deviations. Therefore, the response shift can be estimated as the difference between predicted and observed scores on the 6-month follow-up. Clearly, unreliable measures can result in major changes through regression.

Testing, instrumentation, selection, and mortality are not likely to be serious alternatives to response shift. Testing effects have been shown to have almost no effect on psychological performance beyond a 1-month period (Willson & Putnam, 1982). Instrumentation due to fatigue or familiarity can conceivably affect responses if the assessment is based on observation by clinicians.
Response shift is a form of instrumentation when assessed by self-report in which the internal standard has changed. Selection is not a threat to within-subject design issues, such as response shift, for the same subjects are being considered at two time points. Similarly, mortality is unlikely to be an issue, for response shift is a within-subject effect assessed on whoever survives to the subsequent measurement periods.

**Modeling Growth and Response Shift**

Within a growth modeling framework, one task facing the researcher is to construct growth models that incorporate a response shift and then to compare these models with ones that do not include a response shift component. Including a response shift element in one’s growth model involves several steps that will typically result in testing several competing models. The first question to be asked is, does the investigator think it prudent to investigate the possibility of a response shift? If so, the next step is to determine when the response shift occurred.

After addressing when the response shift occurred, an additional consideration is to determine the form of the response shift. This component is typically overlooked in discussions regarding the use of pretest, posttest, and then test ratings. Such a response shift may take a variety of forms depending on, among other things, the length of time between measurements, the variables measured, and the type of participants in a particular study. Willson (1982) discussed several types of learning curves that can easily be adapted to the discussion of response shift (additional examples are discussed in Glass et al., 1975). The simplest form of a response shift is a one-time increase (or decrease) in response, represented as a spike in a graph of longitudinal data. Such spikes are transitory responses to internal or external events that have no lasting effect on the subsequent responses.

Another simple response shift is a permanent change in level. For example, a simple response shift could occur at a specified point with the subsequent occasions of measurement all containing the new shift in the respondent's norm. Such a shift could occur if, for instance, a person with a condition that limits mobility undergoes a rapid change in his or her values from favoring activities that required a lot of physical movement to activities that required less physical movement. As a result of such a shift, the person feels more alive and begins to enjoy life rather than focus on those activities that required physical mobility. A self-report QOL instrument would indicate the individual felt his or her QOL had improved, although physically no changes had occurred. This response shift in one's values could be modeled as a simple step function (Figure 1). Other examples of potential response shifts are: ramp functions, declining ramp functions, exponential decay functions, or curvilinear functions. The point is that many different types of response shifts can be modeled.

The approach proposed here addresses the response shift phenomenon by using a statistical method. This approach does not require a control group. When a control group is available, the statistical modeling methods proposed here can be extended to address the potential effect of response shift on outcome data more thoroughly.

Modeling the point of response shift is included in the growth model through specification of the point of the shift. If a shift occurred and is correctly specified in time, the fit of the model can be superior to the fit at any other time point. This incre-
mental difference in fitting a model is the basis for the statistical method presented here.

Method

In the effort to describe growth modeling, a real world data set was used to illustrate the technique. The purpose of this section is to discuss growth modeling, not to provide substantive results. We will discuss the concepts and process of modeling various response shifts without excessive statistical detail. More specific statistical information can be found in Willet and Sayer (1996), and an example of LISREL (Joreskog & Sorbom, 1993) program code, the statistical program used here, is available from the authors upon request. Our intent is to introduce the concepts and not the finer details of growth modeling. For example, the sample size in the illustrative data set is smaller than typically desired, yet the value of using real data seemed to outweigh the benefits of generating a simulated data set.

Description of Illustrative Dataset

The illustrative data set for this article is part of a larger study examining the impact of juvenile diabetes (JD) or juvenile rheumatoid arthritis (JRA) on the psychological responses of children and their families using longitudinal data (Frank, Hagglund, et al., 1998; Frank, Thayer, et al., 1998). Indices of adjustment included measures of both behavior problems and emotional distress on four occasions: at time of diagnosis, and 6, 12, and 18 months later. (Please refer to the articles cited above for more details regarding the samples and procedures.)

Analytic Procedure

The statistical method proposed here to examine potential response shift is an extension of Willet and Sayer’s (1996) work that combined growth modeling and structural equation modeling (SEM). They were interested in presenting a statistical method that could assess individual change not only in one domain but also across several domains, with the possibility of determining if the changes in each domain were related to each other. This was accomplished using LISREL (v8.0) to conduct the analysis using SEM (Joreskog & Sorbom, 1993). Typically, one uses the covariance matrix of repetitions of the variables as the major input. The unique aspect of the model used here is that, in this application, the investigator specifies a priori the growth variables and response shift parameters, such as change in level discussed earlier.

The variances for mean level and growth (response shift for our application) can be estimated to inform the researcher about heterogeneity in model fit of individuals around the group-fitted parameters. A small variance for the shift parameter, for example, will suggest that the estimated shift is quite similar for all subjects. The square root of these variances can be meaningfully interpreted in relation to the parameter estimate as a standard deviation of individual’s level or change. For example, a shift of 5 points on a scale might have an associated 3-point standard error. This would suggest, based on asymptotic normality, that most individual shifts can be expected to occur in a .95 confidence interval approximately from \( \pm 1.96 \) to 11 points, with most shifts expected to be positive. Whether this is meaningful and important will depend on the nature of the measure. A similar analysis might be done for the mean level if it is a useful exercise.

In the study reported here, the two dependent variables of interest, behavior problems and negative affect, are discussed. Behavior problems were not analyzed because this variable consisted of parent ratings of their ill child. Such ratings could exhibit a response shift, but because these ratings were by observers, they were not used. However, with the self-report ratings of negative affect one would assume that a response shift within the respondent could have occurred. Contrast coding (Kirk, 1995) was used to represent the response shift. In Figure 1, for instance, if the middle curve titled Step Function consisted of four time points and the response shift had occurred after the first time point, one would model the shift with a +3 for the first time point and a −1 for each time point after the response shift. In many instances, the investigator may not know when the response shift occurred. However, by running several different models with the response shift modeled at different time points, the investigator may determine if and when the shift occurred. This was our approach in the illustration here. As with all statistical tests, there must be a correction for the Type I error rate to avoid overinterpretation of chance differences.

A final note before turning to the results must be made regarding two key characteristics of the illustrative sample. Two types of diseases were pres-
ent in the ill children. There are no data to suggest that children with these two diseases would have similar response shifts. This issue could be investigated empirically but was not done here to maintain focus on the heuristic value of the analysis. The second issue is that a response shift may vary by age of the child. This too could be empirically investigated and remains an avenue for future investigations. The SEM approach can easily represent these additional variables of interest.

Results

Illustration

As noted, no a priori rationale for when the response shift should have occurred was made, and only one variable, negative affect, was tested for a potential response shift. Three different models were tested that assumed the response shift occurred as a step function. These models portrayed the response shift at three different time points, after the first, second, and third time points. The best fitting model suggests that a response shift would have occurred after the third time point, \( \chi^2 (4) = 5.63, p = .23 \). The response shift parameter had an associated \( t = -1.97, p < .05 \), indicating that this parameter was statistically different from zero.

Given that there appeared to be a linear trend in these data, a second series of models was tested with a linear trend included in the models. Including the trend component in the model produced different results. Overall, these models fit much better with the best fitting model, \( \chi^2 (3) = 1.76, p = .62 \). With the trend component included in the models, no response shift component was statistically significant. These results suggest that including the trend component removed any potentially significant response shift effect (modeled as a step function). This highlights an important and possibly confounding effect in detecting response shifts. The presence of a trend in one’s data is an additional feature that may make the detection of response shifts difficult. As a shift will change the values of scores after its occurrence, this can be misinterpreted as a trend. The chi-square for model fit, however, should be smaller for the response shift than the trend line, if in fact a shift occurred.

A final point to note is that there may be times when an investigator is interested in the difference between each individual’s possible response shift and the mean response shift. Growth modeling in SEM specifies variances for individuals (disturbances and residuals) and is not generally used to estimate means. However, point estimates contain information about variability between each individual and the average response shift. SEM can provide point estimates for the shift parameters. This can be helpful in determining if the majority of individuals experienced similar response shifts based on the deviation from the mean response shift at any given point.

Discussion

Limitations and Additional Considerations

In addition to a discussion about response shift, an extension of Willet and Sayer’s (1996) growth modeling approach was presented to examine the effects of response shift where one is able to specify both the form of a growth curve and a response shift, either simultaneously across variables or separately for each variable, depending on the nature of one’s investigation. Ideally, one uses existing theory and the empirical findings from previous research to test various competing models. These models may examine the specified response shift occurring at different times and may include different response shift forms. In addition, if the investigation involves multiple groups, one may either investigate growth and the response shift phenomenon separately or simultaneously, or for each group separately. Before statistical modeling, however, the researcher should eliminate logically or statistically the internal validity threats. Only then is it reasonable to attempt to identify a potential response shift. The internal validity threat analysis strengthens the case for response shift, but cannot prove it.

There is an unfortunate lack of understanding about distributions of response shifts in chronically ill treated populations or healthy populations. In terms of the distribution of response shifts in a treated population, if the shift is a common response to some psychological stressor(s), then it should occur in predictable form within samples. That is, a detectable effect should consistently be found within a specific time frame (with effect size, say, of at least .2, corresponding to a small effect in Cohen’s [1988] discussion). If a likely time frame can be specified for these effects, the possibility that one has overinterpreted maturation, instrumenta-
Psychological outcomes such as QOL are increasingly used to guide legislators and third party payers. However, there is little knowledge to date about response shift variation for various psychological outcomes in the general population. Assessing such variation might provide researchers and clinicians guidance for the range of shift likely to be encountered. Changes well beyond the expected range might ultimately be helpful as indicators of more severe change in the variable of interest. As a step toward addressing this issue, the investigator may examine the likely range of responses in a random sample of pediatric patients, given that the average response is an estimated .2 in effect size units. This information will help the researcher get a better feel for what kind of change is occurring and how to interpret it.

Three types of response shift have been noted: changes in one’s internal standards, values, and conceptualization of QOL. It remains for future investigators to determine if and when these three response shifts occur simultaneously. Some have argued that it is questionable whether these three shifts would occur simultaneously in a parallel fashion (Schwartz & Sprangers, 2000). There may be a typical orderly progression of response shifts with a change in values occurring first, followed by a reordering of those values, and, last, a change in one’s internal standards.

Several additional concerns need to be noted. First, correlated errors are potentially troublesome in longitudinal models. Tests for nonzero covariation among errors for the manifest variables can help provide evidence for serial correlation. These can be modeled as autoregressive or moving average processes. Positively correlated errors increase the probability of finding nonexistent response shifts (Type I error). Second, overinterpreting shift occurrence is a distinct possibility due to the potential interference of many competing explanations over time, such as group development (maturation), regression (if initial scores were extreme with respect to the norm population), and real treatment-induced changes. The best course of action would likely be to only model response shift if prior research and theory provide information regarding the form and timing of such response shift. The circularity of the argument is, unfortunately, that observation is needed to inform the theory. Where relevant, researchers might report possible response shifts in the hope of detecting similar effects as a corpus of research is built in a field.

The illustrative results highlight another point. Pediatric patients are often undergoing various changes in their lives. This presents a different challenge than those researching response shift with adults, where many constructs are assumed stable unless changed by events that disrupt one’s homeostasis or system. Detecting a response shift in a sample of children may be more difficult if the construct one is assessing appears to be changing over time. In such a situation, it may be possible to investigate the response shift using interrupted time series methods (e.g., Franklin, Allison, & Gorman, 1996). Clearly, more work should be done to expand this aspect.

This article intended to discuss response shift as a confounding effect that a growing number of investigators are beginning to address. This discussion stems largely out of concerns about how response shift phenomena may potentially compromise the validity of longitudinal QOL data. Growth curve modeling was illustrated as one potentially beneficial technique for investigating the response shift phenomenon when the investigator has a longitudinal data set. It appears that extensions of growth modeling may serve as a potentially useful tool in identifying the impact of response shift, particularly in research on quality of life in pediatric populations. Additional research is needed to investigate the clinical meaning and implications of response shift. For example, one approach would be to collect detailed case histories to rule out other candidate models based on the results from aggregated data. Additionally, one could conduct Monte Carlo studies that would give researchers insights into how response shift would affect data sets and how response shifts can be detected.

Acknowledgments

The original study was supported by grants H133B30339 and H133B80075 from the National Institute of Disability and Rehabilitation Research, U.S. Department of Education, and by NIH grant T32 HD07460–02 awarded to Robert G. Frank.

Received December 7, 1999; revisions received February 7, 2000; accepted June 1, 2001
References


Spirito, A. (1999). Introduction: Special series on empiri-
cally supported treatments in pediatric psychology. *Journal of Pediatric Psychology, 24*, 87–90.


