ADVANCES IN MOTIVATION AND ACHIEVEMENT
VOLUME 12

NEW DIRECTIONS IN MEASURES AND METHODS

PAUL R. PINTRICH
MARTIN L. MAEHR
Editors

8. THE USE OF MULTI-LEVEL MODELING TO STUDY INDIVIDUAL CHANGE AND CONTEXT EFFECTS IN ACHIEVEMENT MOTIVATION

Janis E. Jacobs and D. Wayne Osgood

Researchers studying either the developmental or the educational aspects of achievement motivation typically are interested in the changes and fluctuations in motivation that occur over time and in different contexts. How are changes in a child's performance in math related to changes in motivation to continue in math? Does an intervention related to reading affect talented students' engagement in reading in the same way as it affects average students' engagement? If intrinsic interest in soccer increases over time, how does it affect interest in school activities over the same time period? Are changes in children's engagement in school dependent on the school they attend or their classroom teacher? Although these questions differ in their emphases (e.g. group differences, context differences, relationships between constructs), they all share the goals of tracking within-individual changes in achievement motivation or of considering context effects that may impact motivation. Indeed, most of the theories of achievement motivation that have been applied to children or to educational contexts assume that changes in beliefs over time will be related to changes in achievement choices, performance, interest, or some other motivational outcome (e.g. Bandura, 1989; Eccles et al., 1983;

New Directions in Measures and Methods, Volume 12, pages 277–318.

ISBN: 0-7623-0819-2

Skinner, Wellborn & Connell, 1990; Weiner, 1985). In addition, researchers universally acknowledge that achievement is embedded within the contexts of schools, classrooms, or neighborhoods (even when they do not measure or control for the particular context).

Despite theoretical and practical interest in questions about individual change and context effects, few studies of the development of achievement motivation have actually measured either one. This has been due, in large part, to limitations imposed by the availability of statistical techniques that would allow researchers to focus their attention on these issues. The goal of this chapter is to describe potential uses of multi-level modeling to address the following three common issues in achievement motivation research: (1) differences between groups (e.g. gender) in patterns of change over time; (2) relationships between changes in one construct and changes in another (e.g. between changes in perceptions of ability and changes in how much a student likes a task); and (3) influences of context on motivation (e.g. schools, classrooms, or neighborhoods). To provide a backdrop for these analytical questions, we begin with a brief review of relevant assumptions implicit in theoretical positions about achievement motivation. We then describe the ways in which multilevel modeling can be used to address these issues when examining within-individual change and the effects of contextual variables, giving examples in each case.

ASSUMPTIONS UNDERLYING ACHIEVEMENT MOTIVATION MODELS

Assumptions Related to Within-Individual Change

The study of achievement motivation has focused on why individuals choose one achievement task over another, the reasons that they persist at that task, and the qualitative nature of achievement choices, as well as the individual's actual achievement or performance. Those interested in the *development* of motivation extend the questions to include changes in task choice, rationales, qualitative experiences, or persistence that occur with age and over time. Major theoretical models of motivation processes typically suggest that achievement beliefs of various types are important determinants of achievement outcomes. Underlying most of these models is the assumption that an individual's current experiences with a task determine expectancies and beliefs that will affect future motivation for similar tasks (e.g. Atkinson, 1957, 1966; Eccles, Futterman, Goff, Kaczala, Meece & Midgley, 1983; Feather, 1982; Bandura, 1977, 1989; Schunk, 1984). Although different theories emphasize different mechanisms (e.g. attributions, self-efficacy, expectancies, goals), all rest on the notion that, as an individual develops, earlier behaviors and beliefs related to a task will be linked to motivation to engage in the same task in the future. Thus, researchers testing these theories typically want to examine within-individual change over time.

Cognitive Models

Over the past 30 years, cognitive models have been the most prominent in the area of achievement motivation (e.g. Bandura, 1989; Covington, 1984; Kukla, 1972; Nicholls, 1984; Schunk, 1984; Weiner, 1979). In each of these models, perceptions of competence and expectations for future success have been used to explain achievement behaviors. Weiner (1979, 1985) suggested that attributions for success and failure on a particular task would make an individual more likely to either pursue or ignore the task in the future. Although Weiner's work did not focus on development, several other researchers have used parts of his theoretical framework to explain children's developing achievement choices (e.g. Nicholls, 1984; Schunk, 1984; Skinner, Wellborn & Connell, 1990). Others have focused primarily on children's perceptions of their abilities and their future expectancies as a means of maintaining motivation. Covington (1984) suggested that children attempt to maintain a sense of self-worth by maximizing attributions to ability following success and minimizing them following failure. Similarly, Nicholls and Miller (1984) gave ability perceptions a primary role as an explanation for children's achievement motivation. Other cognitive approaches have focused on the role of perceptions of efficacy as a motivational force. Children's feelings of efficacy have been related to activity choice, willingness to expend effort on a task, and persistence (e.g. Bandura, 1989; Schunk, 1983, 1984, 1990).

The goal of each of these approaches was to identify the cognitive processes (e.g. perceptions of ability, self-efficacy) underlying achievement motivation. Typically, the processes under study were isolated by using experimental methods that manipulated various conditions, allowing the researcher to compare groups of people who had received different instructions, feedback, or outcomes. Such experimental designs were ideal for testing the relationships between variables, thus providing empirical evidence for the role of cognition in achievement motivation. They did not, however, allow researchers to examine assumptions that were implicit in all of the theories regarding the within-individual relationships between variables over time. The need to examine change over time became more acute when the motivation models from the adult literature were applied to children and adolescents, who are expected to be maturing cognitively, as well as experiencing change in social and school environments as they develop. Laboratory experimentation is valuable for addressing many aspects of motivation and achievement; however, that method is not useful for studying longer term stability or enduring change, which are so central to developmental models of achievement motivation.

Multi-construct Models

As research on the development of achievement motivation has matured, the focus of many process models turned to constructs that bring together motivation and cognition as a joint influence on learning, such as self-regulated learning or metacognition (e.g. Borkowski et al., 1990; Borkowski & Muthukrisna, 1995). Others have focused on the role played by individuals' achievement goals in determining achievement behaviors (e.g. Ames & Archer, 1988; Dweck & Leggett, 1988; Maehr, 1984; Maehr & Braskamp, 1986; Wentzel, 1989). Generally, goals have been defined as the purposes children might have for engaging in a task and researchers have delineated a variety of ways to categorize such goals. For example, Nicholls (1984) defined ego involvement versus task involvement; Dweck and Elliott (1983) described performance versus learning goals; and Maehr (1984) added four types of personal incentives, ranging from intrinsic to extrinsic, including mastery, ego, social, and external reward goals (such as money). Pintrich and Schrauben (1992) developed a model that brought together components from several traditions, including concepts derived from goal theory (e.g. expectancies, values, and affect) as well as metacognitive and self-regulatory concepts. Multi-construct models go well beyond single-construct, cognitive explanations of achievement by encompassing a broader range of explanatory variables. Naturally, the emphasis in the empirical work has typically been on the relations between constructs, but researchers have not succeeded in focusing analyses on the within-person relationships that occur over time, which are implied by the theories.

Newer research also includes a central role for individuals' interests, values, or goals that first appeared in Atkinson's work (1957, 1966). He was the first to highlight the concept of values in his expectancy-value theory by developing an algorithm to describe the way in which expectancies for success and incentive values interact to produce achievement motivation for a specific task. Feather (e.g. 1988, 1992) broadened Atkinson's conceptualization by defining values as a set of stable, general beliefs about what is desirable. He integrated Rokeach's (1979) approach to values by arguing that they are a class of motives that affect behavior by influencing the attractiveness of different goals, and consequently, motivation to attain these goals. He confirmed these ideas by showing that values and expectancies are positively related for academic decisions and decisions to join political groups, suggesting that such decisions are influenced by more than the perceived difficulty of the task (Feather, 1982, 1988). His work was with college students, however, and shed little light on the origins of task values.

Eccles and her colleagues have elaborated and tested an expectancy-value model of activity choice that focuses on the social psychological influences affecting choice and persistence (e.g. Eccles, 1987; Eccles et al., 1983; Eccles, Adler & Meece, 1984; Eccles & Wigfield, 1995; Meece, Parsons, Kaczala, Goff & Futterman, 1982; Meece, Wigfield & Eccles, 1990). According to this model, the key determinants of choice will be the relative value and perceived probability of success of each available option. Expectancies and values are assumed to influence performance and task choice directly; and to be influenced by task-specific beliefs such as self-perceptions of competence; perceptions of the task demands; and the child's goals (both short- and long-term) and self-schemas. These social cognitive variables, in turn, are influenced by the child's perceptions of other people's attitudes and expectations for them, by gender roles and activity stereotypes, and by their own interpretations of their previous experiences with achievement outcomes. Finally, the child's perceptions are influenced by the greater cultural milieu, socializers' beliefs, their own aptitudes or talents, and their previous achievement-related performances.

Closely related to the work on values, is the recent research on the concept of "interest" (Alexander, Kulikovich & Jetton, 1994; Hidi, 1990; Renninger, Hidi & Krapp, 1992; Renninger & Wozniak, 1985; Schiefele, 1991; Tobias, 1994), in which researchers differentiate individual and situational interest. Individual interest is a relatively stable evaluative orientation toward certain domains; situational interest is an emotional state aroused by specific features of an activity or a task. Two aspects or components of individual interest are distinguishable (Schiefele, 1991, 1996): feeling-related and value-related valences. Feeling-related valences refer to the feelings that are associated with an object or an activity itself – feelings like involvement, stimulation, or flow (Csikszentmihalyi, 1988; 1990). Value-related valences refer to the attribution of personal significance or importance to an object.

The evidence just reviewed suggests that critical theoretical and empirical links have already been established between self beliefs and achievement motivation. Much less is known about the long-term changes in self-competence beliefs and values during childhood and adolescence, and what is already known is based primarily on cross-sectional or short-term longitudinal data. Most studies have focused on changes in the factorial structure or mean-level changes in children's competence and value beliefs. Factor analysis has been used to demonstrate that even very young children have well-differentiated beliefs in different domains (e.g. Eccles, Wigfield, Harold & Blumenfeld, 1993; Harter & Pike, 1984; Marsh & Hocevar, 1985) and across achievement constructs (e.g. Nicholls, 1978; Nicholls, Patashnick & Mettetal, 1986).

Analytic strategies that compare mean levels of competence beliefs across ages or across time generally find declines in children's self-assessments as they get older across a variety of domains (e.g. Blumenfeld, Pintrich, Meece & Wessels, 1982; Dweck & Elliott, 1983; Eccles & Midgley, 1989; Wigfield, Eccles, Mac Iver, Reuman & Midgley, 1991; Marsh, 1989). None of these studies has been able to control for within-person stability, however, potentially masking or inflating the rates of change in children's beliefs and attitudes. Small perturbations may be magnified and subtle trends may go unnoticed in cross-sectional and short-term longitudinal studies. Thus, it is critical to use methods such as multi-level modeling to chart changes across longer time periods if we want a comprehensive picture of the development of children's achievement beliefs.

Assumptions Related to Context Effects

Earlier research in both the adult and developmental literature attempted to find universal processes that would be able to explain human behavior; however, when such models failed to receive empirical support in diverse contexts, the focus shifted. In recent years, models have emphasized differences in both processes and outcomes due to context or to individual differences. Context differences have been included in the achievement motivation work in different ways. Some studies have simply demonstrated the existence of similar processes in different subject domains (e.g. math, science, sports) by analyzing data separately in each domain. Others have emphasized the power of context by including it as a moderator variable.

These studies have shown that level and trajectories of achievement beliefs vary by domain (e.g. Jacobs et al., in press; Eccles et al., 1993; Wigfield et al., 1996) and that home and school contexts make a difference. For example, a variety of studies have shown that classroom climate is important for achievement motivation (Dunkin & Biddle, 1974; Fraser & Fisher, 1982; Trickett & Moos, 1974) and that school organization and school policies affect teacher and classroom practices, which in turn affect student motivation (e.g. Bryk, Lee & Holland, 1993; Anderman & Maehr, 1992; Maehr & Midgley, 1996). Researchers have found that teachers can affect motivation by the type of material presented, the amount of work assigned, the way the work is presented, and instructional style (Ames, 1992; Pintrich et al., 1993; Stipek, 1996). School level variables that have an impact on motivation include ability grouping (e.g. Fuligni, Eccles & Barber, 1995; Pallas et al., 1994), comparative performance evaluations (e.g. Ames, 1992; Mac Iver, 1987; Rosenholtz & Rosenholtz, 1981); school culture or climate (e.g. Bryk, Lee & Holland, 1993), and grade transition points (e.g. Eccles & Midgley, 1989; Simmons & Blyth, 1987).

Theoretically, school and classroom contexts affect individual motivation because student learning is embedded or nested within a particular context (one student attends a particular school or is in a given classroom). Nevertheless, very few studies have analyzed context effects by using nested designs and the statistical techniques designed for them. This is especially true of longitudinal studies. At best, previous studies have used context as a control variable or as a grouping factor (e.g. ability grouped vs. mixed ability classes). This is due, in large part, to the fact that most studies are not large enough to use classrooms as the cases and that analysis strategies that involve nested, longitudinal designs have not been available.

Unanswered Questions

We would like to move now into a detailed discussion of the use of multi-level modeling to examine within-individual change and context effects – issues that are important to our understanding of the development of achievement motivation. Before turning to the general discussion of analysis of change; however, we would like to set the stage by previewing the research questions that we will examine in our examples at the end of the chapter.

First, how do changes over time in one group relate to changes in another group? The specific example that we will consider is how changes in girls' ability beliefs and task values are related to changes in boys' ability beliefs and task values. Do girls' and boys' beliefs change at a consistent rate or do the beliefs of one gender drop at a faster rate than those of the other? Much of the research to date has emphasized changes between groups with development (e.g. boys versus girls, high ability versus low ability). The goal of such studies is to examine whether or not group membership (e.g. gender) is related to increasing differences over time. For example, if boys are more motivated than girls to work on math and science tasks in elementary school, does the gap in achievement increase as they get older? If so, is that due to motivation? Many authors have found gender differences in self-perceptions of ability (e.g. Eccles et al., 1984; Jacobs, 1991), in task values (e.g. Eccles et al., 1993), and in attributions for success and failure (e.g. Dweck & Goetz, 1978; Stipek &

Gralinski, 1991). Also, some research has suggested that these differences are exacerbated in early adolescence due to the intensification of sex role beliefs (Eccles, 1994; Eccles et al., 1993; Harter, 1982). Nevertheless, most of this research has been done with cross-sectional data or with short-term longitudinal studies that fail to account for within-individual stability and change over time. Similar analysis issues arise in attempts to compare motivational changes between ethnic and racial groups or between groups differing by ability level (e.g. Dauber & Benbow, 1990; Graham, 1992; Hare, 1985; see Cooper & Dorr, 1995 and Graham, 1994 for reviews).

The second question we will consider is how changes over time in one motivational construct are related to changes in another. In other words, how does one set of attitudes change over time, and how does it influence the development of another? Our example is an examination of the ways in which changes in task values are affected by changes in perceptions of competence. Do children begin to value reading less as their perceptions of competence in that area declines? Do they value math more if they come to feel more competent in that arena? Most studies have not directly studied the impact of changes in one set of beliefs on changes in another set of beliefs or behaviors over time because the appropriate longitudinal data and statistical techniques have not been available.

The third question we will consider is the effect of context on motivation. As an empirical example, we will look at the effects of attending schools that differ in average parental education level on individuals' perceptions of educational opportunities. Do children who attend schools with higher average parental education levels view their educational opportunities more positively than those who attend other schools? Although we might expect the educational level of one's own parents to impact motivation, there also may be a school-level effect that operates beyond the individual level. As noted earlier, context impacts individual motivation because learning is embedded or nested within a particular setting: each student attends only one school. Nevertheless, most of the research on motivation has focused on individual-level variables while neglecting context effects.

ANALYSIS OF WITHIN-INDIVIDUAL CHANGE

Standard Approaches

The basic research design for studying within-individual change is the longitudinal panel study, in which a single sample provides data on a series of occasions extended over time. For instance, researchers might assess interest in reading each month during the third grade or self concept of math ability annually in the fourth through seventh grades.

Repeated-measures Analysis of Variance (ANOVA)

Within the experimental tradition, the standard approach to analyzing change over time is repeated-measures analysis of variance (Judd & McClelland, 1989). This statistical model provides a test of statistical significance for the differences between means on several occasions. As with all ANOVA approaches, the standard use of the repeated-measures model is the diffuse significance test for any and all differences across the several means. Even so, the general framework is also amenable to more focused comparisons (Judd & McClelland, 1989) such as tests for polynomial trends and contrasts between specified sets of occasions (e.g. before versus after an intervention). Repeated-measures ANOVA can also incorporate between-subjects factors, which allows researchers to address questions such as whether boys differ from girls in their patterns of change in motivation or achievement.

Repeated-measures ANOVA is a true analysis of within-individual change in two different senses. First, because the same individuals provide data at every wave, within-individual change is the only possible source of difference in the means over time. Second, the ANOVA significance test is built on a clear distinction between variation within people over time and stable individual differences. This is apparent in the ANOVA model of Eq. 1 (Bailey, 1971), in which each respondent, i, is observed on several occasions, j:

$$Y_{ij} = \mu + \lambda_j + (\varepsilon_{i\cdot} + \rho_{ij}) \tag{1}$$

The term $\varepsilon_{i\cdot}$ is the individual's mean over time (expressed as a deviation from the grand mean, μ), and it serves as a residual term capturing stable individual differences. The second residual term, ρ_{ij}, is the portion of the outcome measure, Y_{ij}, that remains after subtracting the grand mean, the mean deviation at time j, λ_j, and the individual's mean over time. This second residual term, which reflects only unexplained within-individual change over time, provides the error term for the significance test for mean differences over time.

Limiting the analysis to within-individual change is especially valuable because it eliminates a broad class of alternative explanations for the relationship of an explanatory variable to the outcome. This strategy capitalizes on the repeated assessments to use "subjects as their own controls." Because the effect takes the form of change within a person over time, it cannot be due to any stable characteristics of the person, such as most demographic characteristics, many personality factors, and any constant aspects of the individual's environment, which would include a large share of parent, school, and neighborhood effects. Only other explanatory variables that change over time remain plausible alternative explanations. Indeed, we would argue that the potential to eliminate the contribution of stable factors is the primary advantage of the longitudinal panel research design.

The fundamental limitation of repeated-measures ANOVA is its lack of flexibility. In this statistical model, every respondent provides data on the outcome measure *Y* for each value of each repeated measures factor. A special strength of tightly controlled experimental studies is the possibility of creating a data structure that meets this criterion. For instance, in a laboratory situation it is feasible to insure that each subject responds to all types of stimulus materials under each of several different instructional conditions. In that case stimulus type and mode of instruction would be orthogonal repeated-measures factors, and the data would be amenable to repeated-measures ANOVA.

Without this type of experimental control, repeated-measures ANOVA is of far less value. For the longitudinal panel design, this statistical approach is effectively limited to making comparisons between the times of assessment. Thus, its primary value is to assess developmental trends. Yet the inflexibility of the approach is problematic even for this purpose. For instance, repeated-measures ANOVA does not permit a straightforward analysis of age differences in a panel study unless respondents are uniform in age at the beginning of the study. Otherwise the age trend is buried in the interaction of the within-subjects factor of time with the between-subjects factor of initial age (e.g. Wigfield et al., 1997).

This limitation of repeated measures ANOVA precludes using the longitudinal panel research design to address the much wider range of research questions for which that design is suited. Consider our example of whether children come to place less value on an achievement domain when their perceived competence in that arena declines. A longitudinal panel study should be useful for addressing this question. Though it does not offer the experimental control necessary for certainty about causal order, it does yield sufficient information for determining whether within-individual change on one variable is associated with within-individual change on the other. Repeated-measures ANOVA is unable to capitalize on that information, however, because an explanatory variable such as perceived competence varies freely over time, rather than being constrained to a particular fixed set of values.

Longitudinal Structural Equation Models

In non-experimental research, the typical approach to analyzing data from a longitudinal panel study is structural equation models in which the outcome

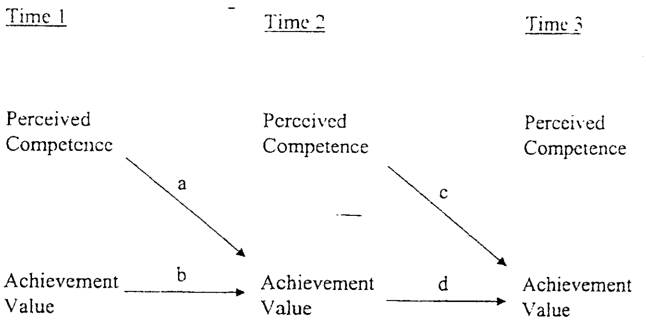


Fig. 1.

variable at each occasion is a function of explanatory variables at the preceding wave as well as the prior level of the outcome variable. Figure 1 illustrates this approach with a three wave panel study, using our example of the impact of perceived competence on value for achievement. This path model corresponds to a pair of regression equations that would provide estimates of the causal paths:

$$\begin{aligned} Value_{2i} &= a\ Competence_{1i} + b\ Value_{1i} + e_{2i} \\ Value_{3i} &= c\ Competence_{2i} + d\ Value_{2i} + e_{3i} \end{aligned} \tag{2}$$

Kessler and Greenberg (1981) provide a thorough introduction to structural equation models of this sort.

This approach overcomes the primary limitation of repeated-measures ANOVA (in this case, competence beliefs and values) because it yields an estimate of the longitudinal relationship between two uncontrolled variables. Thus, this structural equation approach focuses on precisely the kind of research question that is precluded in repeated-measures ANOVA.

Unfortunately, the strengths of repeated-measures ANOVA are the weaknesses of this type of structural equation model. The first weakness is that structural equation models of this form preclude research questions about developmental trends. The regression coefficient *a* and *c* reflect the influence of earlier competence beliefs on later values, while *b* and *c* reflect the stability of values over time. As Eq. 2 illustrates, each wave of outcome data is treated as a distinct outcome variable, and comparisons between the means of those variables play no role in the analysis. Information about development is effectively discarded. Thus, this approach fails to address the one type of research question for which repeated-measures ANOVA is well suited.

The second weakness of this structural equation approach is that estimates of longitudinal relationships are not restricted to within-individual change. This statistical model is not founded on the clear separation of within-person versus between-person variance found in repeated-measures ANOVA. Instead, Eq. 2 adjusts for the earlier measure only to the degree it is correlated with the later measure. This adjustment does not remove all stable individual differences from the analysis, and therefore it remains possible that such factors could spuriously contribute to the estimated longitudinal effects (Cronbach & Furby, 1970; Rogosa, Brandt & Zimowski, 1982).

Pooled-Wave Regression Analysis

To take full advantage of the data generated by longitudinal panel studies of motivation and achievement, we need a statistical framework that combines the strengths of repeated-measures ANOVA with those of structural equation models, as shown in Fig. 1. This framework must be capable of estimating developmental trends as well as assessing longitudinal relationships between uncontrolled variables. Furthermore, to capitalize on the strength of repeated assessments for the same sample of individuals, it should allow us to isolate within-individual change over time, eliminating the contribution of stable individual differences to estimates of longitudinal relationships.

In the following sections of this chapter we will explain how to accomplish those aims through multi-level regression models that pool several waves of data. Figure 2 illustrates how data are organized for this purpose. The upper portion of this figure shows the layout of the data for estimating the structural equation model in Fig. 1. For that analysis, each case is a person, and there is a separate variable for each time that any construct is measured. Thus, perceived competence and achievement value are each repeated as three separate variables, one for every wave of data that was collected. The lower portion of Fig. 2 shows how to reorganize the data for pooled-wave regression analysis. There are several entries for each individual, one for each wave of data. Perceived self-competence and achievement value each appear as single variables in this layout. Age is now a measured variable as well, since it varies across observations for each person. This layout is a multi-level data structure because it has a hierarchical organization in which several observations are nested within each person.

Pooled-wave regression models relate a time-varying outcome variable, such as achievement value, to both explanatory variables that vary over time, such as perceived competence, and to explanatory variables that do not, such as gender. This approach brings the full flexibility of multiple regression models (e.g.

Standard Data Layout

Respondent	Gender	Perceived Competence at age 10	Perceived Competence at age 12	Perceived Competence at age 14	Value on Achievement at age 10	Value on Achievement at age 12	Value on Achievement at age 14
1	1	4	5	6	5	5	7
2	2	7	6	5	7	5	6
3	2	2	3	4	3	5	4
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
N	1	3	5	3	4	5	5

Pooled-Wave Data Layout

Respondent	Age	Gender	Perceived Competence	Value on Achievement
1	10	1	4	5
1	12	1	5	5
1	14	1	6	7
2	10	2	7	7
2	12	2	6	5
2	14	2	5	6
3	10	2	2	3
3	12	2	3	5
3	14	2	4	4
. . .	. . .	. . .	. . .	. . .
N	10	1	3	4
N	12	1	5	5
N	14	1	3	5

Fig. 2. Organizing Data for Pooled-Wave Regression Analysis.

Judd & McClelland, 1989; McClendon, 1995) to the analysis of panel data. Longitudinal regression models of this type are widely used in a number of fields including economics (Hsaio, 1986), sociology (Peterson, 1993), political science (Stimson, 1985), and family studies (Johnson, 1995). The growth curve models that have become popular in psychology and education (Bryk & Raudenbush, 1987) fall into this class as well.

Next we discuss how regression models of this sort can be used to address the main research questions of interest in longitudinal studies of motivation and achievement, and then we turn to the task of focusing the analysis on within-individual change. Specialized multi-level regression models are a third topic we must consider because they address the inevitable dependence among the several waves of data for each person. We finish this section with an illustrative analysis.

Three Types of Models

The pooled-wave regression approach brings the full flexibility of multiple regression to the analysis of longitudinal panel data. We will next consider how this approach can be used to address three general types of research questions by highlighting changes in each model that are needed to address the specific question.

Mean-Level Change

The first and most basic model addresses the overall pattern of developmental change in mean levels of the outcome variable across values of age or time. For example, research questions related to trajectories of achievement beliefs can be answered with this model (e.g. How does self-competence change with increasing age?). Equation 3 captures developmental trends of this sort with a model that allows for a quadratic, curvilinear relationship of age to achievement value:

$$Value_{it} = \beta_0 + \beta_1 Age_{it} + \beta_2 Age_{it}^2 + \varepsilon_{it} \tag{3}$$

The subscript i is for the respondent and the subscript t is for the time or wave of the data. The βs are regression coefficients that reflect the mean difference on the outcome measure associated with a unit of difference on the explanatory variable. The model includes a residual or error term, ε_{it}, which equals the difference between the observed value on the outcome variable and the fitted or predicted value for the remainder of the equation.

The top panel of Fig. 3 shows a hypothetical quadratic trend for achievement value that could be generated by this model. Here, value decreases over time,

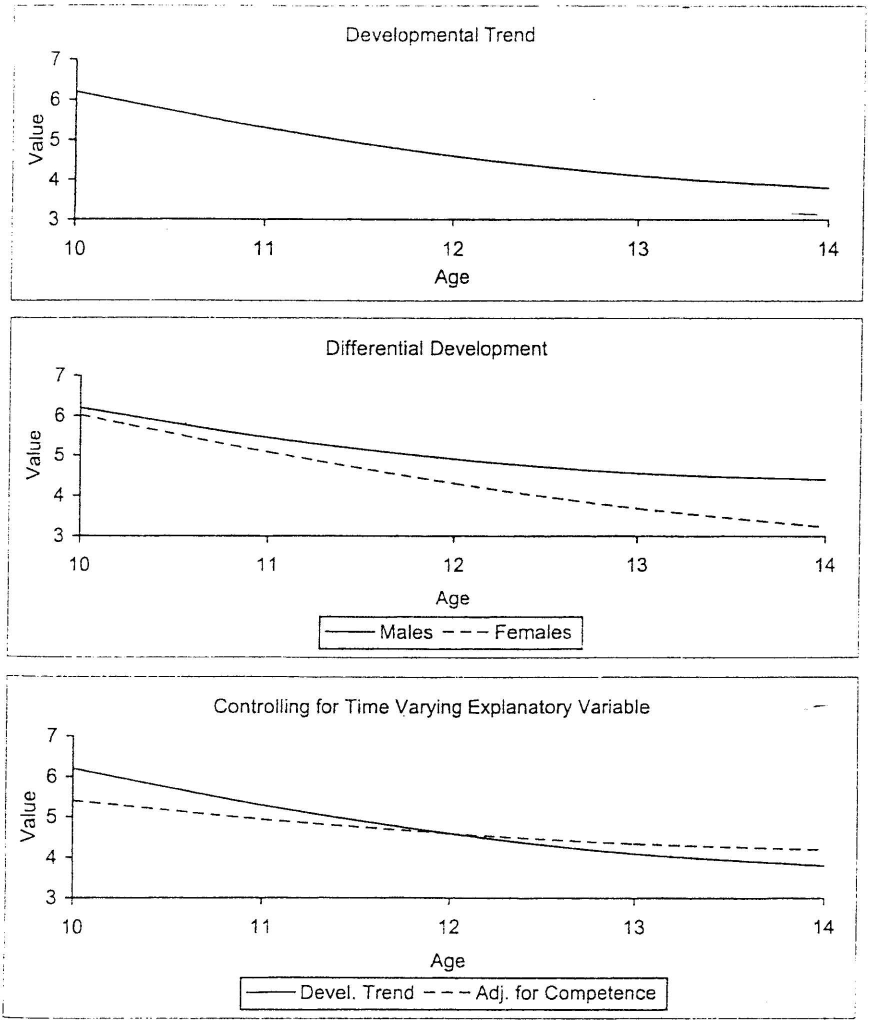


Fig. 3. Three Pooled-Wave Regression Models.

but the rate of decrease is declining and the trend leveling out toward the end of this period. Polynomial functions of this sort are a flexible and efficient means of summarizing non-linear trends, and standard regression texts explain the use of such functions to capture trends of varying levels of complexity (e.g.

Judd & McClelland, 1989; McClendon, 1995). Trends over time need not be limited to smooth, continuous patterns of change, however. The pooled-wave regression framework is also amenable to capturing discontinuous change, as might be expected for a transition from elementary to middle school or the implementation of an instructional or treatment program (see Osgood & Smith, 1995).

Group Differences in Rates of Change

The second type of model allows for differential patterns of development or change by different groups of respondents. This type of model is effective for answering questions about topics such as differences in the rates of change in achievement beliefs between girls and boys or between high and low achievers. This model elaborates the first by adding terms for a grouping or individual difference variable (i.e. sex) and its interactions with age or time:

$$Value_{it} = \beta_0 + \beta_1 Age_{it} + \beta_2 Age_{it}^2 + \beta_3 Sex_i + \beta_4 Age_{it} Sex_i + \beta_5 Age_{it}^2 Sex_i + \varepsilon_{it} \qquad (4)$$

Note that *Sex* does not have the subscript t, which indicates that it is an attribute of the person that does not vary over time.

The middle panel of Fig. 3 shows a hypothetical example of differential development for males versus females. In this example, there is little difference between males and females in achievement value at the beginning of the study, and the value decreases with age for both sexes. The rate of decline is considerably sharper for females than males, however, so there comes to be a substantial gender gap by age 14.

The coefficients β_4 and β_5 reflect differences between the sexes in the pattern of change, while the coefficient β_3 will indicate the difference between the sexes at the age coded as zero. The data analyst can choose the age at which the difference is evaluated by simply subtracting it from the original measure of age, at which point that age becomes zero on the recoded measure. This process is called "centering" (Bryk & Raudenbush, 1992, pp. 25–29), and it must be done before computing the polynomial terms and interaction terms.

Time-varying Explanatory Variable

The third type of model incorporates time-varying explanatory variables that, unlike age, are not direct functions of time. For example, this model could be used to examine the way in which changes in one set of achievement beliefs affect changes in another set of beliefs or in achievement outcomes over time. This type of model is especially useful because it allows researchers to address the various causal processes that theorists hypothesize as underlying the development of motivation and achievement. Equation 5 illustrates this type of model with perceived competence as the time varying variable:

$$Value_{it} = \beta_0 + \beta_1 Age_{it} + \beta_2 Age_{it}^2 + \beta_3 Comp_{it} + \varepsilon_{it} \tag{5}$$

The coefficient β_3 will indicate the strength of relationship between perceived competence and achievement value, expressed in terms of the mean difference in *Value* per unit of change in *Comp*, controlling for the quadratic age trends in both.

Equation 5 also provides a means for addressing a less common research question that will be of special interest to scholars interested in the development of motivation and achievement. The comparison between Eqs 3 and 5 reveals the extent to which perceived competence accounts for developmental trends in achievement value. If perceived competence is strongly related to achievement values, and if the pattern of change is similar for the two, then the addition of *Comp* to the model has the potential to account for the age trend, as indicated by the reduction of coefficients β_1 and β_2 (Hirschi & Gottfredson, 1985). This is simply the application of the standard logic of direct and indirect effects from path analysis, with self-competence serving as the mediator between age and value on achievement (McClendon, 1995). The bottom panel of Fig. 3 shows a hypothetical illustration of this pattern. After controlling for self-competence, there is much less developmental change in achievement value.

Limiting the Analysis to Within-Individual Change

Pooled-wave regression models are not inherently analyses of within-individual change. In Eq. 5, the regression coefficient for self-competence, β_3, is based on all of the observed variance of that variable. Some of that variance corresponds to stable individual differences, which arise because some respondents generally, over time, feel more competent in this domain than do other people. The remainder of the variance in perceived competence will be within-individual change over time, and this variance coincides with people feeling more competent at some points in their lives than others. If the analysis of longitudinal data is not limited to within-individual change, it has little advantage over a cross-sectional analysis. Only by focusing on within-individual change does an analysis rule out the possibility that results are attributable to unmeasured individual difference factors, such as demographic factors or stable personality traits.

Fortunately, the solution to this problem is simply to elaborate the regression model in a way that incorporates the distinction between these two types of relationship (Bryk & Raudenbush, 1992, pp. 117–123):

$$Value_{it} = \beta_0 + \beta_{betw} Comp_{i\cdot} + \beta_{within} Comp_{it} + \varepsilon_{it} \tag{6}$$

This model includes both the original score on self-competence that is specific to time t, $Comp_{it}$, and the individual's mean over time on that explanatory variable, $Comp_{i\cdot}$. By the basic principles of regression analysis, the regression coefficient for any explanatory variable is entirely determined by the portion of its variance that is not shared with any other explanatory variable (Judd & McClelland, 1989). The only variance in the time-specific scores that is not shared with the individual mean is within-individual variance over time. Therefore, b_{within} will reflect the relationship of within-individual change in self competence to within-individual change in value for achievement. To obtain an analysis that is restricted to within-individual change over time, all that is needed is to control for the individual mean on that variable.

How does this distinction of within-individual versus between-individual relationships apply to age or time, variables that are in a sense controlled by the longitudinal panel research design? In a research design corresponding to Fig. 2, in which every respondent provides data at the same three ages, the mean age across time is identical for all of them. Thus, there is no between-individual variance on age so the regression coefficient for the time specific measure will automatically be restricted to within-individual change, even without including the individual mean for age or time. This would not be the case, however, if individuals differ in the mean ages for their sets of observations. Differences in mean age often do arise, either from sample attrition or as a by-product of the research design. When there is attrition, the average age of respondents who participate at all waves will be different from the average age of those who drop out of the study. In a multiple-cohort panel study, average age will correspond to initial differences in age at the beginning of the study. In both cases, the shared estimate of the relationship reflects not just within-individual change, so it is important to control for individuals' mean ages, as in Eq. 6.

Specialized Methods for Conducting Multi-level Regression Analyses

Thus far we have shown: (1) that pooled-wave regression models provide a flexible framework for answering important research questions about development and causal processes in motivation and achievement, and (2) that a simple elaboration of the basic regression model limits the analysis to within-individual change, thereby eliminating confounding with any and all stable individual differences. Next, we turn to a technical issue that will require the use of specialized statistical software for pooled-wave regression analyses. We want to emphasize, however, that this is not a substantive matter, and is has no consequence for what variables should be included in the regression equations or how the coefficients should be interpreted.

The technical problem arising for pooled-wave regression analysis is that data of this sort are almost certain to violate the standard regression assumption that all of the residuals, ε_{it}, in the entire data set are independent from one another. This assumption requires that there be no systematic pattern in which the residuals for some cases are more similar to one another than to the residuals for other cases. In longitudinal panel data, it is inevitable that we fail to explain a substantial share of the between-individual variation, which means that there will be considerable similarity among the several residuals for each person. Furthermore, there is a ubiquitous tendency for the residuals of observations closer together in time to be more highly correlated than those that are farther apart. Violating the assumption of independence does not systematically bias the estimates of the regression coefficients, but it does invalidate their significance tests, making it appear that one has more statistical power than is actually the case.

Multi-level regression models have the express purpose of addressing dependence in research designs with nested data, and they retain the full substantive flexibility of multiple regression. The key feature of these models is the addition of extra residual or error terms that capture the pattern of dependence among the residuals. Among the mostly widely known versions of these multi-level models are Bryk and Raudenbush's hierarchical linear models (HLM, 1992) and Goldstein's multi-level models (1995). Latent growth curve models (Willet & Sayer, 1994) express some of the same features within a covariance structure modeling framework.

A typical multi-level model addressing both types of dependence would be:

$$Value_{it} = \beta_0 + \beta_1 Age_{it} + \beta_2 Age_{it}^2 + (u_{0i} + u_{1i} Age_{it} + u_{2i} Age_{it}^2 + r_{it}) \qquad (7)$$

This equation divides the residual into four components, three that apply to all observations for person i (u_{0i}, u_{1i}, and u_{2i}) and one that is specific to a single wave of data t (r_{it}). This is a more elaborate version of the separation of within-individual and between-individual residual terms in repeated-measures ANOVA. u_{0i} applies to all observations for an individual, so it captures unexplained stable individual differences. The additional residual terms, u_{1i} and u_{2i}, apply to the linear and quadratic terms for age, and thus they correspond to the difference between this individual's pattern of change over time and the overall pattern of change captured by β_1 and β_2. Bryk and Raudenbush (1987) show that these additional residual terms allow for the possibility of greater correlation between observations that are closer together in time. The standard assumptions would be that the u_i terms have a multivariate normal distribution across individuals and that r_{it} is normally and independently distributed across all observations. The analysis provides estimates of the regression coefficients, the variances of both types of residuals, and the covariances among the u_{i}.

One often sees multi-level regression models presented in a different format that separates the time varying, or Level 1, aspects of the model from the between-individual, or Level 2, features. In this format, the Level 1 model for Eq. 7 would be:

$$Value_{it} = \pi_{0i} + \pi_{1i}Age_{it} + \pi_{2i}Age_{it}^2 + r_{it}$$

The Level 2 model for Eq. 7 is:

$$\pi_{0i} = \beta_0 + u_{0i}$$

$$\pi_{1i} = \beta_1 + u_{1i}$$

$$\pi_{2i} = \beta_2 + u_{2i}$$

Note that the π parameters of the Level 1 equation carry the subscript i, so they are the coefficients for a single individual rather than the entire sample. As the Level 2 model demonstrates, those individual coefficients are equal to the coefficients for the entire sample, β, plus the residual term reflecting the individual's deviation from the average relationship. Thus, the difference in presentation is simply a heuristic device, and this set of Level 1 and 2 equations specifies the same model as Eq. 7.

The format of separate Level 1 and 2 equations can be useful for specifying interactions between the two levels of analysis, while making clear what Level 2 residual terms would be appropriate. For instance, adding the variable sex to each of the Level 2 equations would allow for the possibility that boys and girls would differ in their pattern of change over time, comparable to Eq. 4:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + \beta_{01}Sex_i + u_{0i} \\ \pi_{1i} &= \beta_{10} + \beta_{11}Sex_i + u_{1i} \\ \pi_{2i} &= \beta_{20} + \beta_{21}Sex_i + u_{2i} \end{aligned} \tag{8}$$

All of the substantive examples of pooled-wave regression are amenable to these multi-level methods, including analyses focusing on causal processes over time, such as the influence of perceived competence on achievement value in Eq. 5. Again, multi-level models add a complex error structure that addresses dependence among observations; they bring no change or limitations to the substantive model. In using these methods, it is important not to make the mistake of interpreting the separate Level 1 and Level 2 equations as implying that these models automatically make a strict separation of within-individual relationships from between-individual level relationships. Coefficients for Level 1, time-varying, explanatory variables will be shared estimates that combine the two, as we explained in the preceding section. To limit multi-level regression coefficients to within-individual relationships, which control for all stable individual differences, one must add the individual mean of the explanatory variable to the model.

Next we present two empirical examples to illustrate these methods. The first will focus on within-individual change over time and the second will focus on contextual effects.

An Example Analysis of Within-Individual Change Over Time

This example is taken from a recent study documenting gender differences in age-related trends across grades one through twelve in perceived self-competence and activity values in three achievement domains (Jacobs et al., in press).[1] All of the equations and exemplars presented in our description of multi-level analysis of individual change have used these variables. Thus, the earlier description of the analysis techniques should be readily transferred to this example. The example illustrates all three types of models described earlier – mean level change, differential rates of change between groups, and the use of a time-varying explanatory variable.

This study used a cohort-sequential design to examine within-individual changes in beliefs across childhood and adolescence, focusing on differences between girls and boys in their rates of change in task values. Based on the previous literature and our own work, we predicted that subjective task values would be highest in the first grade, with decreases across grades. Earlier work has shown that gender differences in values exist as early as first grade (Marsh, 1989; Wigfield et al., 1997), and previous reports of increases in gender differences in self beliefs in middle childhood and in adolescence (Eccles et al., 1993; Huston, 1983; Ruble & Martin, 1998) led us to expect increases in the differences between the beliefs of males and females over time. Finally, when the time-varying explanatory variable of perceived competence was added to the model, we expected declines in competence beliefs with age to explain some of the decline in values, resulting in flatter trajectories for subjective task values after controlling for competence beliefs.

In this example, we used Hierarchical Linear Modeling to implement a growth curve analysis to test our hypotheses (Bryk & Raudenbush, 1987, 1992). Although changes in mean levels of self beliefs have been tested in short-term longitudinal studies (e.g. Wigfield et al., 1997) by using repeated measures ANOVA, such an approach has the limitations that we described earlier and becomes quite unwieldy for six waves of data covering a span of twelve grade levels. Our HLM growth model improves on that approach by efficiently capturing developmental trajectories with only a few parameters and by allowing us to include all respondents, even if they did not provide data for the full set of six observations. Using HLM, we were able to implement an analysis that is strictly limited to within-individual change, controlling for all stable individual differences, while addressing the possible non-independence of residuals due to repeated measures. This approach also allowed us to take full advantage of our cohort-sequential design by determining whether individual characteristics, such as gender, are related to initial status or to change and by analyzing time-varying factors that might explain change over time in the outcome.

Sample and Measures

These data were gathered as part of the Childhood and Beyond (CAB) longitudinal project investigating the development of children's self-perceptions, task values, and activity choices. The analyses presented here will be one of the outcome variables from the Jacobs et al. (in press) study: Values for Language Arts. Data were collected between 1989–1999 from children attending 10 elementary schools in four middle class, primarily European-American school districts in the suburbs of a large Midwestern city. A cross-sequential design was employed, in which three cohorts of children were followed longitudinally across the elementary, middle, and high school years. Children were in the second, third, and fifth grades during Wave 1 of the data used in this analysis, and they were assessed again one, four, five, and six years later for a total of five waves spanning grades two to twelve. The results reported here are based on 761 students who were present at the first wave and who provided data for both sex and grade. As is necessary for pooled-wave regression analysis, the same scales were used to measure Perceived Competence and Values at each grade level. The scale for Language Arts Competence Beliefs contained five items (e.g. How good are you at reading?) and Language Arts Values were measured with a four-item scale (e.g. How useful will reading be to you in the future?). Alphas varied for each grade level, ranging between 0.73 and 0.93. The Slossen Intelligence Test – Revised (1991 edition), given to all children when they joined the study, was used as a measure of cognitive ability.

The Multi-level Model for the Analysis

The multi-level modeling framework is especially useful for our purposes because there is no assumption that the number and spacing of observations will be consistent across individuals or across time. Thus, HLM is quite compatible with the complexities of our research design, including multiple cohorts sampled in different years of school, a three-year gap in data collection, and sample attrition.

We present the model and results for this first analysis in the two level format of HLM. The Level 1 equation for our initial analysis is the following quadratic growth curve model:

$$Value_{it} = \pi_{0i} + \pi_{1i}Grade_{it} + \pi_{2i}Grade_{it}^2 + r_{it}$$

This equation is identical to several pooled-wave regression models discussed above, with grade as the time index instead of age. We "centered" grade at six by subtracting this value from the original codes for grade (Bryk & Raudenbush, 1992, pp. 25–29). We do this to take advantage of the definition of the constant as the fitted value when all explanatory variables equal zero. *Grade* and *Grade*^{2} will not equal zero for the sixth grade, so the constant, π_{0i}, will characterize subjective values in the sixth grade. The rate of change in values varies over time because of the inclusion of *Grade*^{2}. As a result of centering grade, however, π_{1i} will reflect the "instantaneous" slope at the sixth grade.

The Level 2 equations for this analysis are:

$$\pi_{0i} = \beta_{00} + \beta_{01}Sex_i + \beta_{02}Ability_i + \beta_{03}Grade_{i\cdot} + \beta_{04}Grade_{i\cdot}^2 + u_{0i}$$
$$\pi_{1i} = \beta_{10} + \beta_{11}Sex_i + \beta_{12}Ability_i + u_{1i}$$
$$\pi_{2i} = \beta_{20} + \beta_{21}Sex_i + \beta_{22}Ability_i$$

Here we have elaborated the basic model for differential development (Eq. 8) in two ways. First, controlling for ability insures that results for gender differences in development are not attributable to ability differences. Second, as we recommended above, the equation for the intercept, π_{0i}, also includes individual means over time for the Level 1 explanatory variables, *Grade* and *Grade*^{2}. Doing so insures that our results for trends across grades will reflect within-individual change over time rather than preexisting differences between the cohorts or between respondents who completed the study versus those who were lost in earlier waves.

We have centered all of the Level 2 explanatory variables at their sample means (i.e. subtracted the mean from the original scores) so that the Level 2 constant terms (β_{00}, β_{10}, and β_{20}) characterize the growth curves averaged across boys and girls and for students of average ability.

The coefficients for sex, β_{01}, β_{11}, and β_{21}, will indicate the difference in the growth curves for males and females, with β_{01} reflecting the mean sex difference in the sixth grade, β_{11} the difference in slope in the sixth grade, and β_{21} the difference in curvature. Because sex is a dummy variable with zero assigned to males and one to females, positive values indicate higher means, slopes, and more convex curvature (i.e. positive change in slope over time) for females than for males.[1]

To allow for dependence among the several observations for each respondent, these equations include residual terms for the Level 1 intercept, u_{0i}, and for linear change, u_{1i}. Preliminary analyses indicated that the quadratic trend, π_{2i}, did not vary significantly across respondents.

Growth Curve Results

The results for this analysis appear both as Model 1 in Table 1, which lists the regression coefficients, their statistical significance, and their standard errors, and as the solid lines in Fig. 4, which expresses their meaning by graphing the fitted values that correspond to these coefficients.

We expected subjective task values for Language Arts to decline across grades. This expectation was upheld. As can be seen in Fig. 4, initial subjective task values were high, 5.6 for boys and 5.9 for girls out of a possible score of 7.0. Though the growth curve was curvilinear, the figure shows that decline is the dominant trend for task value beliefs. In Table 1 the intercept term for linear change indicates that, at grade 6 and averaged across boys and girls, values were declining at a highly significant rate of -.12 units per year. The intercept term for grade-squared is positive. Correspondingly, Fig. 4 shows that average subjective values for language arts decline most rapidly during the elementary years, and the rate of decline slows over time.

Prior research led us to expect that females would have higher values for language arts (e.g. Eccles et al., 1993), and we expected these differences to grow larger over time due to gender role socialization. As can be seen in Table 1, tests of the Level 1 intercept for sex indicate that males and females hold significantly different task values in language arts at grade 6; females hold higher values for language arts than males. Although gender differences in the Level 1 intercept were found in language arts, no significant gender differences were found in the linear or quadratic *rate of change*. A joint test of the coefficients for sex on linear and quadratic change indicate that the overall sex differences in the pattern of change in language arts is only a statistical trend ($p=0.12$). If it were significant, it would suggest the following: Girls initially place higher value on language arts than boys in second grade. Because girls' values decline more rapidly than boys, however, the gap narrows by late elementary school. The gap then increases again during high school as girls' value for language arts increases and boys' value levels off (see Fig. 4).

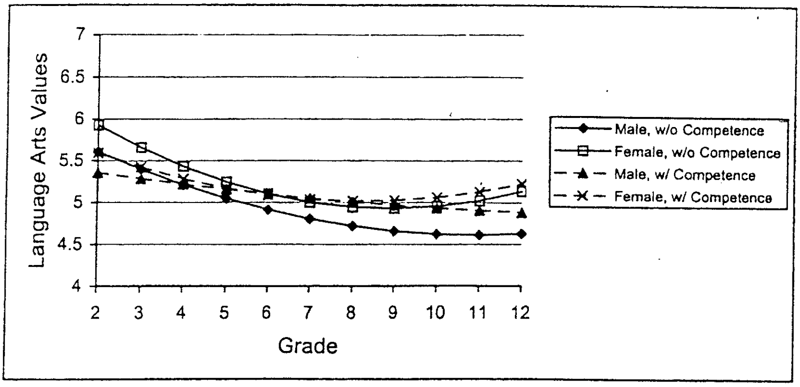
The Contribution of Perceived Competence

After describing the growth trajectories for subjective task values, we were interested in examining the potential of a time varying explanatory variable, perceived competence, to account for trajectories of change in subjective task values. As discussed above, we can determine this by adding competence beliefs as an explanatory variable in the HLM model for task values. To the

Table 1. Growth Models for Language Arts Values, With and Without Controlling for Perceived Competence.

	Model 1		Model 2	
	Coeff.	Std. Err.	Coeff.	Std. Err.
For Intercept				
Intercept	5.008*	0.042	5.095*	0.036
Sex	0.189*	0.080	−0.011	0.068
Ability Test	−0.002	0.003	−0.008*	0.002
Average Grade	0.066*	0.022	0.040*	0.019
Average Grade^2	−0.008	0.010	−0.008	0.008
Ave. Competence Beliefs			0.068	0.039
For Linear Change				
Intercept	−0.122*	0.010	−0.060*	0.009
Sex	0.002	0.018	−0.013	0.016
Ability Test	−0.0002	0.0006	−0.0006	0.0005
For Quadratic Change				
Intercept	0.017*	0.003	0.009*	0.003
Sex	0.009	0.006	0.012*	0.005
Ability Test	0.0003	0.0002	0.0004*	0.0002
For Competence Beliefs				
Intercept			0.433*	0.029
	Var.		Var.	
Residual Variance				
For Intercept	0.320		0.155	
For Linear Change	0.011		0.006	
For Competence Beliefs			0.083	
Within Individual	0.802		0.594	

* $p < 0.05$.



[1]Adapted from Jacobs, et al., in press.

Fig. 4. Multi-Level Model of Growth Curves for Language Arts Values with and without Controls for Language Arts Perceived Competence.[1]

extent that competence beliefs are related to trajectories of change in subjective task values, introducing that variable to the model will reduce coefficients for change over time and the trajectories will become more flat, showing less change. Also, to the extent that perceived competence is related to gender differences, the coefficients for gender will be reduced, the trajectories for males and females will be more similar in slope and shape, and their trajectories will lie closer together.

For this analysis, we modify our original HLM model by adding the time-specific measure of perceived competence to the Level 1 equation of our HLM model and the individual mean for perceived competence to the Level 2 equation for the Level 1 intercept. By including both, we insure that the Level 1 relationship is restricted to within-individual change. These results appear as Model 2 in Table 1 and as the dashed lines in Fig. 4.

Table 1 shows that respondents are much more likely to value language arts when they feel competent in that domain. The coefficient for perceived competence is both large (0.433) and highly significant. The strength of the relationship is also reflected in the variance explained by perceived competence for all of the residual terms in the analysis. Perceptions of competence explained 52% of the previously unexplained variance in stable individual differences (i.e. residual variance for intercept). Perceived Competence also explained substantial portions of the variance in change over time, in the form of both individual differences in slope (46%) and within-individual variation around the growth curves (26%). In sum, students' perceptions of competence in language arts are strongly associated with how much they value related tasks.

A time-varying explanatory variable like perceived competence has the potential to explain not only between-individual variation in trajectories of task values, but also the amount and pattern of change in the average trajectory. For this to occur, competence beliefs must be strongly associated with subjective task values, and we have just seen that they are. Furthermore, the pattern of change in perceived competence must be similar to that of task values. The extent to which perceived competence accounts for the overall trajectory will be reflected in the impact of controlling for competence beliefs upon the intercept terms for linear and quadratic change in Table 1 and in the difference in trajectories before and after controlling for perceived competence, as shown in Fig. 4.

Perceived competence is closely tied to the trajectory of values for language arts. Coefficients for both linear and quadratic change are reduced over 40% by controlling competence beliefs. Accordingly, Fig. 4 shows that the trajectories for both boys and girls are considerably flattened by taking competence beliefs into account. Other analyses (Jacobs et al., in press) indicated that self-competence in language arts declines rapidly during the elementary school years. Accordingly, Fig. 4 shows that competence beliefs account for much of the decline in language arts values during that period for both boys and girls.

Analyses reported elsewhere indicate that girls consistently feel more competent in language arts than boys do after first grade, and boys' perceived competence beliefs decline at a faster rate than those of girls (Eccles et al., 1993; Jacobs et al., under review). Figure 4 shows that perceived competence explains much of the gender difference during the middle grades, while a moderate difference remains during the earlier and later grades. Averaging across grades, controlling for competence beliefs reduces the sex difference in values by 62%, from 0.28 to 0.11. Thus, the higher perceived competence that girls experience in language arts may be an important source of the higher value they place in this domain.

ANALYSIS OF CONTEXT EFFECTS

As discussed earlier, motivation researchers all acknowledge that achievement is embedded within the context of schools, classrooms, or neighborhoods. In recent years, many models have emphasized differences in both processes and outcomes due to context or to individual differences. Despite the acknowledged importance of considering the role of context in the development of achievement motivation, researchers have seldom included true tests of context effects. We begin this section by laying out two major methodological issues that arise in studying context effects: units of analysis and separating context effects from aggregated individual effects. Fortunately, both of these issues are successfully resolved by the same multi-level regression models that we discussed above. Then, as an empirical example, we present an analysis that examines attitudes toward school achievement in relation to the contextual factor of attending a school with high or low average levels of parent education.

Methodological Issues

Unit of Analysis

In many studies of motivation and achievement, individual respondents are selected as members of larger units such as classrooms, schools, or neighborhoods. For many years researchers struggled with the question of whether individuals or aggregate units were more appropriate for analyzing such data. Suppose that 20 different teachers instructed 300 female students in a study of the effect of math teachers' gender on female students' interest in the subject. Those 300 students do not represent fully independent observations for this research question. Girls who have the same teacher share many potential influences that go well beyond their teacher's gender, such as the teacher's instructional style and skill and the quality of the classroom interaction that particular year. Therefore the data set has less information about the effect of teachers' gender than if the 300 students all had different teachers. In other words, when analyzed at the individual level, these data violate the standard statistical assumption of independence, with the consequence that significance tests are biased toward results appearing significant when they should not (Bryk & Raudenbush, 1992).

Before multi-level statistical models were available, the standard recommendation to overcome this problem of dependent observations was to conduct the analysis at the aggregate level rather than the individual level. To accomplish this, one would compute the mean of student responses for each teacher and then conduct the analysis on the 20 means rather than the 300 individual scores. This analysis solves the problem of dependence (assuming that the teachers are not further clustered, such as by school), so this analysis is legitimate. Unfortunately, the solution comes with costs. The accuracy of these means depends on the number of students in each class, so the analysis requires weighted least squares rather than ordinary regression or ANOVA (McClendon, 1995). Basing the analysis on aggregated means makes it difficult to-control for individual level factors (such as gender, race, and prior achievement) that might be confounded with the aggregate variable of interest. Furthermore, limiting the analysis to means precludes the separation of individual and context effects that we discuss below.

Multi-level statistical models overcome the level of analysis problem by making use of information at both individual and aggregate levels of analysis. Here is a multi-level model for the analysis of the impact of teacher gender on students' interest in math:

$$Interest_{ij} = \beta_0 + \beta_1 TGender_j + \beta_2 Perf_{ij} + \beta_3 Race_{ij} + (u_j + r_{ij})$$

The subscript i differentiates individual student respondents and the subscript j the classrooms of which they are members. The outcome variable is interest in math, $Interest_{ij}$, which is assessed for each student. The coefficient β_1 will reflect the relationship between that outcome and the teachers' gender, $TGender_j$. Note that this variable carries the subscript j but not i because it is constant for all students in the same classroom. This model also controls for the two individual level variables of prior performance, $Perf_{ij}$, and race (such as a dummy variable for minority versus majority ethnic group). The residual term u_j allows for unexplained mean differences between classrooms, which is the main source of dependence among the observations. Thus, a multi-level regression model can simultaneously examine the relationship of an outcome to both individual level and aggregate level explanatory variables, while adjusting for any redundancy among the responses in each context.

It is also possible that the strength of an individual level relationship would vary across contexts. For instance, some teachers might be more effective than others at encouraging the interest of students who had done poorly at math in the past. If so, then the relationship of prior performance to current interest would be weaker in some classrooms than in others. To avoid an erroneous significance test, the analysis must take this variation into account. A multi-level regression model does so by adding a residual term that applies to the relevant individual level variable, such as u_{1j} in this regression equation:

$$Interest_{ij} = \beta_0 + \beta_1 TGender_j + \beta_2 Perf_{ij} + \beta_3 Race_{ij} + (u_{0j} + u_{1j} Perf_{ij} + r_{ij}) \quad (10)$$

The term u_{1j} captures the difference between the effect of prior performance in classroom j and the overall effect of classroom performance, averaged across classrooms. Texts on multi-level regression analyses (e.g. Bryk and Raudenbush, 1992; Goldstein, 1995; Kreft and deLeeuw, 1998) explain procedures for determining whether or not such a residual term is necessary for any given individual level explanatory variable.

Multi-level regression analyses can also investigate the reasons that individual level relationships vary across contexts. For instance, a researcher might want to know whether experienced teachers are more successful at eliciting greater interest among students who have done poorly at math in the past. We can capture this in a multi-level regression model by adding a variable that is the product of a measure of teacher experience and the measure of prior performance. The format of Bryk and Raudenbush's HLM program makes it quite easy to specify such interaction; one simply adds a Level 2 variable to the equation for the Level 1 variable of interest.

Significance tests for these cross-level interaction terms are sensitive to variation in the strength of the individual level relationship across contexts, just as was the case for tests of the "main effects" of Level 2 variables (described above). Fortunately, incorporating a residual term such as u_{1j} in Eq. 9 not only insures accurate significance tests of the overall relationship between prior performance and interest in math, but also insures accuracy for any interactions of prior performance with aggregate level variables, like teacher experience.

Separating Context Effects from Individual Effects

One of the most interesting reasons for studying individuals within larger units such as classrooms or neighborhoods is the possibility that the aggregate properties of these units may have emergent effects that are distinct from individual level relationships. For instance, Pong (1998) argued that student achievement is higher at schools with a larger proportion of two parent families because all students benefit from the stronger network of social relationships among adults made possible by the greater social capital of two parent families. In Pong's research, the contextual effect of the aggregate characteristic amplified the positive individual level relationship between achievement and having a two parent family. It is also possible, however, that a context effect would be in the opposite direction of an individual level effect. For example, although school achievement is positively associated with self esteem at the individual level, Felson and Reed (1986) predicted and found that the contextual effect of average school achievement is in the opposite direction. They reasoned that a high average level of achievement provides a stricter standard of social comparison that leads students to evaluate their own performance more poorly than if the average achievement were lower.

Contextual effects of this sort cannot be inferred from aggregate level analysis alone, but rather they require multi-level analysis of both individual and contextual relationships (Bryk and Raudenbush, 1992). A strictly aggregate analysis confounds individual and contextual effects. Consider the relationship of parents' level of education to students' attitudes about school, which is the research question we investigate in our empirical example of context effects below. It is likely that students will have more favorable attitudes about school achievement if their parents obtained higher levels of education. This individual level process alone is sufficient to generate more favorable mean attitudes in those schools where mean parental education is higher. This type of aggregated individual level relationship is often called a compositional or selection effect. It is conceptually distinct from a true contextual effect, in which average parents' education would impact all students, regardless of their own parents' education. This is an emergent effect of the school context rather than an individual level relationship that holds regardless of which school the student attends.

If both individual and aggregate data are available, the two types of relationship are readily separated in a multi-level regression model in this form:

$$AttSch_{ij} = \beta_0 + \beta_1 ParEd._j + \beta_2 ParEd_{ij} + (u_j + r_{ij}) \tag{9}$$

The equation specifies that the outcome variable of attitudes toward school, $AttSch_{ij}$, is a function of both the education of the individual respondents' parents, $ParEd_{ij}$, and of the mean education of all students at the school, $ParEd._j$. The coefficient β_2 will capture the individual level relationship between each student's outcome and their own parents' education. The coefficient β_1 will reflect the context effect of mean parental education on all children at the school, above and beyond the effect of their own parents' education, β_2. Thus, multi-level regression modeling provides a straightforward means of implementing this important conceptual distinction between effects at the two levels of analysis.

Note that Eq. 10 directly parallels Eq. 6, which addressed the analysis of within-individual change over time. Each of those equations includes both a Level 1 variable and its Level 2 mean. The reason for this similarity is that the two equations are designed to solve methodological problems by making a clear distinction between relationships at the two levels of analysis.

The analysis of context effects and of within-individual change are not parallel in one respect, however. The Level 1 estimate automatically controls for any and all Level 2 variables, whether measured or not. This is a considerable methodological advantage for the analysis of within-individual change because it rules out the broad class of alternative explanations having to do with stable individual differences. Unfortunately, the Level 2 coefficient only controls for the individual level factors that are included in the model. Therefore, it remains possible that any apparent context effect of mean parents' education in Eq. 10 could actually be due to other individual level factors, such as race, economic status, or family structure. Therefore, when studying context effects it is important to identify other relevant variables of this sort and include them in the analysis (Bryk & Raudenbush, 1992).

An Empirical Example

We illustrate the application of multi-level regression models to contextual data with analyses of students' attitudes about school in relation to the school context. The data for this example come from a study of 5,253 eighth grade students attending 36 different schools located in 10 different school districts widely dispersed across the U.S. The data were collected as part of an evaluation of a school-based gang prevention program (Esbensen and Osgood, 1999).[2] The sample is highly diverse, with only 44% of the respondents identifying themselves as white. It is especially useful for a contextual analysis that the schools vary widely in their population characteristics. In terms of ethnicity, these schools ranged from a low of 3% minority students to a high of 97%; in terms of family structure, they vary from a low of 9% of students living with two parents to a high of 91%; and in terms of parents' education, they varied from a low of only 39% graduating from high school or earning a GED to a high of 65% completing college.

The measure we use as our outcome variable is students' views about whether their educational opportunities are limited. This four item measure ($\alpha = 0.70$) includes questions indicating that the student feels he or she doesn't have much chance of going far in school. The explanatory variables included in the analyses are respondent's sex (1 = male, 2 = female), minority ethnic status (1 = African American or Latino, 0 = other), and parents' education (coded as the higher of two parents, 1 = less than high school, 2 = high school graduate, 3 = college graduate).

The focus of these analyses is on whether attending a school with higher or lower average parental education has a context effect on students' sense of limited educational opportunities. We would expect that individual students are more likely to see their educational prospects as limited if their parents attained a low level of education. This analysis will also address whether, in addition to this individual level relationship, students' views about these limitations are affected by attending a school where most other students come from families with higher or lower levels of parental education. In other words, is there an emergent effect whereby all students' views about educational prospects are affected by attending a school in which more or fewer families have attained higher levels of education.

To illustrate the use of multi-level regression for analyzing contextual data, we estimated three models using Raudenbush, Bryk, Cheong, and Congdon's

Table 2. Hierarchical Linear Models of Students Views that Their Educational Opportunities are Limited.

	Individual Level		School Level		Joint	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Individual Parents' Educ.	−0.230*	0.017	–		−0.220*	0.017
Mean Parents' Education			−0.567*	0.078	−0.359*	0.082
Female	−0.093*	0.020			−0.092*	0.020
Minority	0.041	0.032			−0.004	0.032
Constant	1.880*	0.019	1.914*	0.017*	1.902*	0.016
	Var.	Expl. Var.	Var.	Expl. Var.	Var.	Expl. Var.
Residual Variance	–					
School Mean	0.007	65.1%	0.005	73.2%	0.004	81.4%
Minority Slope	0.014				0.012	
Within School	0.500	5.5%	0.528	0.0%	0.499	5.5%

* $p < 0.05$.

(2000) HLM program. The first of these is an individual level model in the sense that all of the explanatory variables characterize the individual respondent and his or her family. As Table 2 shows, students whose parents had less education see their own educational prospects as considerably more limited. Given the coding of parents' education, students whose parents did not finish high school score an average of 0.46 lower than students whose parents graduated from college, a difference of 0.84 standard deviations. Girls saw their educational prospects as significantly less limited than boys, but there was no difference between minority and non-minority students (controlling for the other factors).

The individual level model includes three residual variance components. The first is the variance of the residual mean differences between schools (u_{0j} in the equations above), which allows for the dependence among the observations for students from the same school. There is also a residual variance component for the relationship of minority ethnicity to perceptions of limited educational opportunity, which varied significantly across schools.[3] The final residual variance component reflects differences among students within schools, and this term is most comparable to the residual variance in a standard ordinary least squares multiple regression analysis.

The variance of the residual term for school means provides a standard for gauging how well a model accounts for differences among schools in students' perceptions of limited educational opportunities. Though the first HLM model includes only individual level variables, those variables still have the potential to account for differences between schools, provided that the explanatory variable differs among schools as well. Indeed, the individual level model explains fully 65.1% of the between-school variance. This figure is computed by comparison to a "null model" which includes no explanatory variables, and it constitutes an R^2 value for the aggregate level of analysis. The proportion of the school-level variance that is explained is far larger than the 5.5% of individual level variance that is explained. It must be remembered, however, that there is far more within-school variance than between-school variance.

The second analysis in Table 2 is restricted to our school-level explanatory variable, the mean parents' education. The significant school-level relationship is more than twice as strong as the individual level relationship of parents' education to perceived limitations on educational opportunity. Does that imply that the average parents' education for an entire school has a larger impact on this outcome measure than does a student's own parents' education? The answer is "not necessarily" for two different reasons.

The larger school-level coefficient does imply that a one unit difference in a school's mean parental education makes more difference in the outcome for each student than does a one unit difference in their own parents' education. Yet the consequence of this larger coefficient is dampened by the limited variance of school means in parental education in comparison to individual level variance. Though these schools vary widely in mean parental education, a one unit difference between schools is rare while one unit difference between individuals is common. Accordingly, the school-level model explains a larger proportion of the variance between schools than does the individual level model. The school-level model can explain none of the within-school variance, however, because the explanatory variable does not differentiate among students at the same school.

Second, we cannot interpret the results of the school level model as indicating a strong effect of mean parental education on students' perceptions of limited opportunities because this coefficient does not distinguish between individual and contextual effects. Instead, it reflects the total association between means on the two variables, which combines the two sources. A model including both individual scores and school means, as in Eq. 10, is necessary to estimate the context effect.

The third HLM model in Table 2 is just such a joint model. In comparison to the previous models, the individual level relationship is essentially unchanged, while the mean level relationship is considerably smaller. In fact, the mean level relationship has been reduced by the magnitude of the individual level relationship. This is precisely what we seek as an estimate of the context effect: the relationship between the school mean for parental education level and students' views about limited educational opportunities, above and beyond the individual level relationship. Thus, the coefficient in the joint model reflects the dependence of the outcome on the overall level of parental education in a school, regardless of the student's own parents' educational levels.

In this analysis the context effect remains statistically significant and larger than the individual level effect. In combination, the individual and contextual effects account for the lion's share of variation across schools in students' views of limited educational opportunities (81.4%). As noted above, even strong context effects may account for little of the total variation in the outcome because between-school variance in the explanatory variable is typically far less than within-school variance. Even so, a finding such as this would have considerable theoretical importance for showing that students' expectations for educational success are dependent on the broader social context in which they live, rather than just their immediate family. Furthermore, such a result would point to especially pernicious effects of extreme concentration of economic disadvantage in particular neighborhoods and schools, as has been the focus of W. J. Wilson's writing (e.g. 1987). Of course, this analysis is only a simple example to illustrate these methods. The results for the context effect should not be treated as definitive without more thorough controls for other individual level factors that would provide alternative explanations (see preceding section).

CONCLUSIONS

We began this chapter by suggesting that theories and models of the development of achievement motivation rely on implicit assumptions regarding within-person change and the role of context. Researchers have not been able to test those suppositions easily in the past due to the limitations of statistical programs; thus, this chapter described potential uses of multi-level modeling to address three common issues in achievement motivation research that are based on these assumptions: (1) groups differ in their patterns of change over time; (2) changes in one motivational construct will be related to changes in another: and (3) context will affect achievement motivation.

The strength of multi-level modeling to answer the first two questions lies in its ability to take full advantage of the data generated by longitudinal panel studies of motivation and achievement by combining, in a unified statistical framework, the strengths of repeated-measures ANOVA with those of structural equation models. This framework is capable of estimating developmental trends as well as assessing longitudinal relationships between uncontrolled variables. In addition, it allows us to capitalize on repeated assessments for the same sample of individuals by isolating within-individual change over time, eliminating the contribution of stable individual differences to estimates of longitudinal relationships. A second strength of multi-level modeling over previous methods is its ability to test context effects without sacrificing individual-level data. Methods that aggregate at the school or classroom level confound individual and contextual effects. Contextual effects of this sort cannot be inferred from aggregate level analyses alone, but require multi-level analysis of both individual and contextual relationships.

It should be noted that, throughout the chapter, we have limited our discussion to two level hierarchical data, but the principles generalize to research designs with three levels or more. A three level data structure would arise, for instance, if students at a sample of schools were assessed on several occasions over time. In this case, a multi-level regression analysis would allow the investigator to study potential context effects of school variables on the patterns of within-individual change over time, while adjusting for dependence among observations within schools and within individuals. Bryk and Raudenbush (1988) provide an example of such an analysis. The HLM software is capable of estimating three level models, and Goldstein's MLWin program will handle up to ten levels!

With more than two levels of analysis, it is possible that the levels are not hierarchical, but rather "cross-nested." This would arise, for instance, if one took into account classroom changes over time, so that there were multiple classrooms per individual as well as multiple individuals per classroom (Goldstein, 1994; Raudenbush, 1995). The MLWin software can be used to conduct analyses of cross-nested data.

The goal of this chapter was to whet the reader's appetite for learning more about multi-level modeling, an analysis tool that is especially well-suited for research in achievement motivation. The emphasis here has been on the value of multi-level regression models for addressing interesting research questions that help solve important methodological problems within research on achievement motivation. There are, of course, many issues in the proper use of these methods that we have not addressed or have only mentioned briefly, such as centering variables and deciding which residual variance components to include. Readers interested in applying these methods in their own research can find useful guidance on such matters in texts by Bryk and Raudenbush (1992), Goldstein (1995), and Kreft and deLeeuw (1998). Finally, although we have focused primarily on longitudinal change and context effects, we want to emphasize the versatility of this technique – the use of multi-leveling modeling is not limited to those research designs. For example, it may be used to consider such things as within-subject relationships across items rather than across time: similarities and differences between siblings; or achievement differences between schools, nested within districts or states. In sum, we believe that this methodology has the potential to enhance our understanding of the developmental and educational aspects of achievement motivation by providing a powerful tool to test a broad array of research questions.

NOTES

1. This research was supported by Grant HD17553 from the National Institute for Child Health and Human Development to Jacquelynne S. Eccles, Allan Wigfield, Phyllis Blumenfeld, and Rena Harold. We would like to thank the principals, teachers, students, and parents of the cooperating school districts for their participation in this project.

2. The full analysis of these data also included tests of whether the growth curves varied across the three cohorts. There were no significant differences between cohorts. We do not present that aspect of the analysis here to avoid unnecessary complexity.

3. The original dataset includes 45 schools, but these analyses are limited to 36 schools for which there were at least 35 respondents.

4. All variables in this analysis were "grand mean centered," which indicates that the sample means were subtracted to give all variables a mean of zero. The value of doing so in this analysis is that the variance component for school means has a consistent interpretation across the several models (Bryk and Raudenbush, 1992).

REFERENCES

Alexander, P. A., Kulikovich, J. M., & Jetton, T. L. (1994). The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Review of Educational Research, 64*, 201–252.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261–271.

Anderman, E. M., & Maehr, M. L. (1992).

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260–267.

Atkinson, J. W. (1957). Motivational determinants of risk taking behavior. *Psychological Review, 64*, 359–372.

Atkinson, J. W. (1966). Motivational determinants of risk taking behavior. In: J. W. Atkinson & N. T. Feather (Eds), *A Theory of Achievement Motivation* (pp. 11–31). New York: Wiley.

Bailey, D. E. (1971). *Probability and Statistics: Models for Research.* New York: Wiley.

Bandura, (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Rview, 84*, 191–215.

Bandura, (1989). Human agency in social cognitivie theory. *American Psychologist, 44*, 1175–1184.

Blumenfeld, P., Pintrich, P. R., Meece, J., & Wessels, K. (1982). The formation and role of self-perceptions of ability in elementary school classrooms. *Elementary School Journal, 82*, 401–420.

Borkowski, J. G., & Muthukrisna, N. (1995). Learning environments and skill generalization: How contexts facilitate regulatory processes and efficacy beliefs. In: F. Weinert & W. Schneider (Eds), *Recent Perspectives on Memory Development*. Hillsdale, NJ: Erlbaum.

Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*, 147–158.

Bryk, A. S., & Raudenbush, S. W. (1988). On heterogeneity of variance in experimental studies: A challenge to conventional interpretations. *Psychological Bulletin, 103*, 396–404.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models*. Newbury Park, CA: Sage.

Covington, M. W. (1984). The motive for self-worth. In: R. Ames & C. Ames (Eds), *Research on Motivation in Education* (Vol. 1, pp. 77–113), Academic Press, New York.

Cooper, H., & Dorr, N. (1995). Race comparisons on need for achievement: A meta-analytic alternative to Graham's narrative review. *Review of Educational Research, 65*, 483–508.

Cronbach, L. J., & Furby, L. (1970). How should we measure change – or should we? *Psychological Bulletin, 74*, 68–80.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In: M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds), *Optimal Experience* (pp. 15–35). Cambridge, MA: Cambridge University Press.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

Dauber, S. L., & Benbow, C. P. (1990). Aspects of personality and peer relations of extremely talented adolescents. *Gifted Child Quarterly, 34*, 10–15.

Dunkin, M., & Biddle, B. (1974). *The study of teaching*. Holt, Rinehart and Winston.

Dweck, C. S., & Goetz, T. E. (1978). Attributions and learned helplessness. In: J. H. Harvry, W. Ickes & R. F. Kidd (Eds), *New Directions in Attribution Research* (Vol. 2), Hillsdale, NJ: Erlbaum.

Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. In: P. H. Mussen (Ed.), *Handbook of Child Psychology* (3rd ed., Vol. 4, pp. 643–691). New York: Wiley.

Dweck, C. S. & Leggett, E. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.

Eccles (Parsons), J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values, and academic behaviors. In: J. T. Spence (Ed.), *Achievement and Achievement Motivation* (pp. 75–146). San Francisco: W. H. Freeman.

Eccles, J. S., Adler, T. F., & Meece, J. L. (1984). Sex differences in achievement: A test of alternative theories. *Journal of Personality and Social Psychology, 46*, 26–43.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135–172.

Eccles, J. S., & Midgley, C. (1989). Stage/environment fit: Developmentally appropriate classrooms for early adolescents. In: R. Ames & C. Ames (Eds), *Research on Motivation in Education* (Vol. 3, pp. 139–181). New York: Academic Press.

Eccles, J. S., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's achivement self-perceptions during the elementary school years. *Child Development, 64*, 830–847.

Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., Mac Iver, D., & Feldlaufer, H. (1993). Are traditional middle grades schools undermining the academic motivation of early adolescents? *Elementary School J., 93*, 553–574.

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly, 18*, 585–609.

Eccles, J. S. & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin, 21*, 215–225.

Esbensen, Finn-Aage, & Osgood, D. W. (1999). Gang Resistance Education and Training (G.R.E.A.T.): Results from the National Evaluation. *Journal of Research on Crime and Delinquency, 36*, 194–225.

Feather, N. T. (1982). Expectancy-value approaches: Present status and future directions. In: N. T. Feather (Ed.), *Expectations and actions: Expectancy-value models in psychology* (pp. 395–420). Hillsdale, NJ: Erlbaum.

Feather, N. T. (1988). Values, valences, and course enrollment: Testing the role of personal values within an expectancy-value framework. *Journal of Educational Psychology, 80*, 381–391.

Feather, N. T. (1992). Values, valences, expectations, and actions. *Journal of Social Issues, 48*, 109–124.

Felson, R. B., & Reed, M. D. (1986). Reference Groups and Self-Appraisals of Academic Ability and Performance. *Social Psychology Quarterly, 49*, 103–109.

Fraser, B. J., & Fisher, D. L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*, 498–518.

Fuligni, A. J., Eccles, J. S., & Barber, B. L. (1995). The long-term effects of seventh-grade ability grouping n mathematics. *Journal of Early Adolescence, 15*(1), 58–89.

Goldstein, H. (1994). Multilevel Cross-Classified Models. *Sociological Methods and Research, 22*, 364–375.

Goldstein, H. I. (1995). *Multilevel Statistical Models* (2nd ed.). London: Arnold.

Graham, S. (1992). Most of the subjects were European American and middle class: Trends in published research on African Americans in selected APA journals 1970–1989. *American Psychologist, 47*, 629–639.

Graham, S. (1994). Motivation in African Americans. *Review of Educational Research, 64*, 55–117.

Hare, B. R. (1985). Stability and change in self-perceptions and achievement among African American adolescents: A longitudinal study. *Journal of African American Psychology, 11*, 29–42.

Harter, S. (1982). The Perceived Competence Scale for Children. *Child Development, 53*, 87–97.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development, 55*, 1969–1982.

Hidi, S. (1990). Interest and its contributions as a mental resource for learning. *Review of Educational Research, 60*, 549–571.

Hirschi, T., & Gottfredson, M. (1985). All wise after the fact learning theory, again: reply to Baldwin. *American Journal of Sociology, 90*, 1330–1333.

Hsaio, C. (1986). *Analysis of panel data*. Cambridge: Cambridge University Press.

Jacobs, J. E. (1991). The influence of gender stereotypes on parent and child math attitudes: Differences across grade-levels. *Journal of Educational Psychology, 83*, 518–527.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (in press). Longitudinal changes in self beliefs between first and twelfth grades. *Child Development.*

Johnson, D. R. (1995). Alternative methods for the quantitative analysis of panel data in family research: pooled time-series methods. *Journal of Marriage and the Family, 57,* 1065–1077.

Judd, C. M., & McClelland, G. H. (1989). *Data Analysis: A Model Comparison Approach.* San Diego, CA: Harcourt Brace Jovanovich.

Kessler, R. C., & Greenberg, D. F. (1981). *Linear Panel Analysis: Models of Quantitative Change.* New York: Academic Press.

Kreft, Ita, & Jan de Leeuw, J. (1998). *Introducing Multilevel Modeling.* London: Sage.

Kukla, A. (1972). Foundations of an attributional theory of performance. *Psychological Review, 79,* 454–470.

Maehr, M. L. (1984). Meaning and motivation. In: R. Ames & C. Ames (Eds), *Research on motivation in education* (Vol. 1, pp. 115–144). New York: Academic Press.

Maehr, M. L., & Braskamp, L. (1986). *The motivation factor: Toward a theory of personal investment.* Lexington, MA: Lexington Books.

Maehr, M. L., & Midgley, C. (1996). *Transforming school cultures.* Boulder, CO: Westview Press.

Mac Iver, D. J. (1987). Classroom factors and student characteristics predicting students' use of achievement standards during ability self-assessment. *Child Development, 58,* 1258–1271.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81,* 417–430.

Marsh, H. W., & Hocevar, D. (1985). The application of confirmatory factory analyses to the study of self-concept: First and higher-order factor structures and their invariance across age groups. *Psychological Bulletin, 97,* 562–582.

McClendon, M. J. (1995). *Multiple Regression and Causal Analysis.* Itasca, IL: Peacock Publishers.

Meece, J. L., Eccles-Parsons, J., Kaczala, C. M., Goff, S. E., & Futterman, R. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91,* 324–348.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its consequences for young adolescents' course enrollment intentions and performances in mathematics. *Journal of Educational Psychology, 82,* 60–70.

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perceptions of academic attainment, and the understanding that difficult tasks require more ability. *Child Development, 49,* 800–814.

Nicholls, J. G. (1984). Conceptions of ability and achievement motivation. In: R. E. Ames & C. Ames (Eds), *Motivation in Education.* New York: Academic Press.

Nicholls, J. G., & Miller, A. T. (1984). The differentiation of the concepts of difficulty and ability. *Child Development, 54,* 951–959.

Nicholls, J. G., Patashnick, M., & Mettetal, G. (1986). Conceptions of ability and intelligence. *Child Development, 57,* 636–645.

Osgood, D. W., & Smith, G. (1995). Applying hierarchical linear modeling to extended longitudinal evaluations: the Boys Town follow-up study. *Evaluation Review, 19*(1), 3–38.

Pallas, A. M., Entwisle, D. R., Alexander, K. L., & Stluka, M. F. (1994). Ability-group effects: Instructional, social, or institutional? *Sociology of Education, 67,* 27–46.

Peterson, T. (1993). Recent advances in longitudinal methodology. *Annual Review of Sociology, 19,* 425–454.

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In: D. H. Schunk & J. L. Meece (Eds), *Student Perceptions in the Classroom* (pp. 149–183). Hillsdale, NJ: Erlbaum.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, *63*, 167–199.

Pong, S. (1998). The School Compositional Effect of Single Parenthood on 10th-Grade Achievement. *Sociology of Education*, *71*, 23–42.

Raudenbush, S. W. (1995). Hierarchical linear models to study the effects of social context on development. In: J. M. Gottman, (Ed.), *The Analysis of Change* (pp. 165–201). Mahwah, New Jersey: Lawrence Erlbaum.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T. Jr. (2000). *HLM5: Hierarchical Linear and Nonlinear Modeling*. Chicago: Scientific Software International.

Renninger, K. A., Hidi, S. & Krapp, A. (Eds) (1992). *The role of interest in learning and development*. Hillsdale, NJ: Erlbaum.

Renninger, K. A., & Wozniak, R. H. (1985). Effect of interest on attentional shift, recognition, and recall in young children. *Developmental Psychology*, *21*, 624–632.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, *92*, 726–748.

Rokeach, M. (1979). From individual to institutional values with special references to the values of science. In: M. Rokeach (Ed.), *Understanding Human Values* (pp. 47–70). New York: Free Press.

Rosenholtz, S. R., & Rosenholtz, S. J. (1981). Classroom organization and the perception of ability. *Sociology of Education*, *34*, 132–140.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*, 299–323.

Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology*, *21*, 3–18.

Schunk, D. H. (1983). Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. *Contemporary Educational Psychology*, *8*, 76–86.

Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. *Educational Psychologist*, *19*, 48–58.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, *25*, 71–86.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context.* Hawthorn, NY: Aldine de Gruyter.

Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology*, *32*, 22–32.

Stimson, J. A. (1985). Regression in space and time: a statistical essay. *American Journal of Political Science*, *29*, 914–947.

Stipek, D. J. (1996). Motivation and instruction. In: R. C. Calfee & D. C. Berliner (Eds), *Handbook of Educational Psychology*. New York: Macmillan.

Stipek, D. J., & Gralinski, H. J. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, *83*, 361–371.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, *64*, 37–54.

Trickett, E. J., & Moos, R. H. (1974). Personal correlates of contrasting environments: Student satisfaction in high school classrooms. *American Journal of Community Psychology*, *2*, 1–12.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, *71*, 3–25.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548–573.

Wentzel, K. (1989). Adolescent classroom grades, standards for performance, and academic achievement: An interactionist perspective. *Journal of Educational Psychology*, *81*, 131–142.

Wigfield, A., & Eccles, J. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, *12*, 265–310.

Wigfield, A., Eccles, J., Mac Iver, D. Reuman, D. & Midgley, C. (1991). Transitions at early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, *27*, 552–565.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., & Blumenfeld, P. C. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology*, *89*, 451–469.

Willett, J. B., & Sayer, A. G. (1994). Using Covariance Structure Analysis to Detect Correlates and Predictors of Individual Change. *Psychological Bulletin*, *116*, 363–381.

Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, The Underclass, and Public Policy*. Chicago: University of Chicago Press.