

In Search of the Powerful Self-Fulfilling Prophecy

Stephanie Madon and Lee Jussim
Rutgers—The State University of New Jersey

Jacquelynne Eccles
University of Michigan

This research examined moderators of naturally occurring self-fulfilling prophecies. The authors assessed whether positive or negative self-fulfilling prophecies were more powerful and whether some targets were more susceptible to self-fulfilling prophecies because of their self-concepts in a particular achievement domain and previous academic records. Participants were 98 teachers and 1,539 students in sixth-grade public school math classes. Results yielded a strong pattern showing that teacher perceptions predicted achievement more strongly for low achievers than for high achievers. Results also yielded a much weaker pattern showing that teacher overestimates predicted achievement more strongly than teacher underestimates. Implications for social perceptual accuracy, self-enhancement theory, and understanding when self-fulfilling prophecies are stronger are discussed.

When do expectations create social reality, and how much do expectations influence social reality? The power of interpersonal expectations to influence others' behavior is a stronghold of social psychology that dates back to Merton's (1948) introduction of the self-fulfilling prophecy. Merton introduced the term *self-fulfilling prophecy* to refer to situations in which initially false beliefs become true. He argued that self-fulfilling prophecies explain a host of social problems, ranging from bank failures to discrimination against Jews and African Americans. However, it was not until Rosenthal and Jacobson's (1968a) classic empirical demonstration that teacher expectations could influence student achievement that self-fulfilling prophecies became a major area of social psychological research (see, e.g., reviews by Brophy, 1983; Darley & Fazio, 1980; R. A. Jones, 1977; Jussim, 1986).

Stephanie Madon and Lee Jussim, Department of Psychology, Rutgers—The State University of New Jersey; Jacquelynne Eccles, Department of Psychology, University of Michigan.

This research was supported by National Institute of Child Health and Development (NICHD) Grant 1 R29 HD28401-01A1 and by grants from the National Institute of Mental Health, NICHD, and the National Science Foundation. This article reports research based on Stephanie Madon's master's thesis, conducted under the direction of Lee Jussim. This thesis won the 1994 Best Master's Thesis award from the New Jersey Psychological Association. An earlier version of this article was presented at the 1995 conference of the Society of Experimental Social Psychology.

Special thanks are extended to Max Guyll, for his assistance with the calculation of effect sizes, and to Dave Kenny and Jon Krosnick, for their invaluable help with complex statistical issues. Thanks are also extended to Richard Contrada for helpful comments on an earlier version of this article. We also thank the following people for their aid in collecting and processing the data: Bonnie L. Barber, Christy Miller Buchanan, Harriet Feldlaufer, Connie Flanagan, Janis Jacobs, Dave Klingel, Doug MacIver, Carol Midgley, David Reuman, and Allan Wigfield. Finally, we thank the personnel in all of the participating schools for their help in data collection.

Correspondence concerning this article should be addressed to Stephanie Madon, Department of Psychology, Tillet Hall, Livingston Campus, Rutgers—The State University of New Jersey, New Brunswick, New Jersey 08903.

Through the 1980s many researchers believed that the self-fulfilling prophecy was a powerful and pervasive phenomenon (e.g., Fiske & Taylor, 1984; E. E. Jones, 1986; Snyder, 1984). However, neither meta-analyses of the experimental research (e.g., Raudenbush, 1984; Rosenthal & Rubin, 1978) nor naturalistic studies (see Brophy, 1983; Jussim, 1991; Jussim & Eccles, 1995a, for reviews) supported this conclusion. Under naturalistic conditions, perceiver expectations (in terms of standardized regression coefficients) typically have effects of about .1 to .2 on targets' behavior and achievement (see Jussim, 1991; Jussim & Eccles, 1995a, for reviews).

Does this mean that perceiver expectations never have powerful effects on targets? Not necessarily. Effect sizes are averages. Under certain conditions, and for certain types of targets, perceiver expectations may be more powerful than average effect sizes indicate (e.g., Jussim, Eccles, & Madon, 1996; Snyder, 1992). In the present study we examined some of the conditions under which self-fulfilling prophecies might be more powerful. We examined whether positive or negative expectations produced more powerful self-fulfilling prophecies and whether students with low self-concepts in a particular achievement domain and poor records of academic performance were more susceptible to self-fulfilling prophecies than were students with high self-concepts in a particular achievement domain and good records of academic performance. We addressed these issues in a study that involved nearly 100 teachers and more than 1,500 students in sixth-grade public school math classes.

Self-Fulfilling Prophecies

Experimental Studies

Experiments have demonstrated the possible existence of self-fulfilling prophecies. In a typical experiment, researchers induce false expectations in perceivers and then assess differences in target behavior (e.g., Rosenthal & Jacobson, 1968a; Snyder, Tanke, & Berscheid, 1977). Although the early research on teacher expectations was controversial (e.g., Elashoff & Snow, 1971), the self-fulfilling prophecy was later firmly established as a real phenomenon (Darley & Fazio, 1980; Jussim, 1986;

Rosenthal & Rubin, 1978). Nonetheless, experimental studies show only that perceiver expectations *can* lead to self-fulfilling prophecies (Jussim, 1993). Whether perceivers typically develop false expectations naturally, and whether self-fulfilling prophecies typically follow, are questions that cannot be addressed by the experimental paradigm.

Naturalistic Studies

The limitations of experimental studies led many researchers to examine self-fulfilling prophecies under naturalistic conditions. Consistent with experimental results, naturalistic studies have repeatedly supported the existence of self-fulfilling prophecies (see reviews by Brophy & Good, 1974; Eccles & Wigfield, 1985; Jussim & Eccles, 1995a). For example, teachers do indeed develop erroneous expectations for their students (Brophy, 1983; Brophy & Good, 1974; Cooper, 1979; Cooper & Good, 1983; Jussim, 1986; Jussim & Eccles, 1995a), and these expectations predict student motivation and achievement even after statistically controlling for students' previous motivation and achievement (e.g., Brattesani, Weinstein, & Marshall, 1984; Jussim, 1989; Jussim & Eccles, 1992; Eccles-Parsons, Kaczala, & Meece, 1982; West & Anderson, 1976). Also consistent with the experimental research (see, e.g., meta-analyses by Raudenbush, 1984; Rosenthal & Rubin, 1978), such effects are typically small, ranging from about .1 to .2 (see Brophy, 1983; Jussim, 1991; Jussim & Eccles, 1995a, for reviews).

The Search for Moderators

There may be, however, conditions under which self-fulfilling prophecies are more powerful. Goals, personality characteristics, power relationships, and situational factors all moderate self-fulfilling prophecies (Harris, 1989; see meta-analyses by Cooper & Hazelrigg, 1988; Raudenbush, 1984; see reviews by Neuberger, 1994; Snyder, 1992). Because nearly all of this research has involved experimental laboratory studies, however, little is known about moderators of naturally occurring self-fulfilling prophecies.

We know of only two studies of moderators of naturally occurring self-fulfilling prophecies. Teacher expectations produced more powerful self-fulfilling prophecies when students believed that teachers treated high- and low-achieving students differently (Brattesani et al., 1984). Our own research has shown that self-fulfilling prophecies are more powerful among students who are either African American or from lower socioeconomic backgrounds (Jussim et al., 1996).

In the present study we further examined moderators of naturally occurring self-fulfilling prophecies by addressing four questions: (a) Do positive or negative expectations produce more powerful self-fulfilling prophecies? (b) Do expectations produce more powerful self-fulfilling prophecies when they match students' self-concepts in a particular achievement domain? (c) Are students with low self-concepts in a particular achievement domain more susceptible to self-fulfilling prophecies than students with high self-concepts in a particular achievement domain? and (d) Are students with poor records of academic performance more susceptible to self-fulfilling prophecies than students with good records of academic perfor-

mance? We addressed these questions by testing several different hypotheses that were derived from theory and empirical research.

Hypotheses Tested

Golem Hypothesis: Do Negative Expectations Produce More Powerful Self-Fulfilling Prophecies?

Rosenthal and Jacobson's (1968a, 1968b) study demonstrated the power of positive expectations. Inducing negative expectations in a real classroom would have been unethical. However, several reviews have speculated that under naturalistic conditions expectations lead to more powerful self-fulfilling prophecies when they are negative than when they are positive (e.g., Brophy, 1983; Brophy & Good, 1974; Eccles & Wigfield, 1985). This pattern has been referred to as *golem effects*. The term *golem* originates from Eastern European Jewish mythology in which the "Golem" was a mechanical creature who was created to be a servant but who turns into a monster (Babad, Inbar, & Rosenthal, 1982).

There are several reasons why negative expectations may be more powerful than positive expectations. People often consider negative information more useful than positive information (e.g., Kanouse & Hanson, 1971; Skowronski & Carlston, 1989; Taylor, 1991). People react more strongly to negative feedback than to positive feedback (Coleman, Jussim, & Abraham, 1987), and they often weigh costs more heavily than rewards when making decisions (Kahneman & Miller, 1986). The tendency for individuals to focus on, and to be more influenced by, negative versus positive information suggests that negative expectations may create more powerful self-fulfilling prophecies than do positive expectations.

Self-Enhancement Hypothesis: Do Positive Expectations Produce More Powerful Self-Fulfilling Prophecies?

Other research, however, suggests that positive expectations may be more powerful than negative expectations. Self-enhancement theory proposes that people strive to think well of themselves (Jussim, Yen, & Aiello, 1995; Sedikides, 1993; Swann, Pelham, & Krull, 1989; for reviews, see S. C. Jones, 1973; Shrauger, 1975). This suggests that people should respond more to positive evaluations than to negative evaluations (Shrauger, 1975). Extending this premise to research on self-fulfilling prophecies suggests that positive self-fulfilling prophecies may be more powerful than negative self-fulfilling prophecies. This pattern may arise if targets, motivated to view themselves favorably, focus their attention on and believe in the validity of positive expectations and overlook or discount negative expectations.

Self-Consistency Hypothesis: Do Expectations That Match Targets' Self-Concepts in a Particular Achievement Domain Produce More Powerful Self-Fulfilling Prophecies?

Alternatively, expectations may be more powerful when they match targets' self-concepts. Self-consistency theory proposes that people desire self-relevant information that matches their

preexisting self-concepts (Jussim et al., 1995; Sedikides, 1993; Shrauger, 1975; Swann et al., 1989; Swann, Griffin, Predmore, & Gaines, 1987). Research has shown that people retain feedback more accurately, view it as more credible, and believe it to reflect their ability more when it matches their own self-concepts than when it does not (Jussim et al., 1995; Shrauger, 1975; Swann et al., 1987, 1989). Perceiver expectations that match targets' self-concepts, therefore, may produce more powerful self-fulfilling prophecies than perceiver expectations that do not match targets' self-concepts. We examined whether negative expectations produce more powerful self-fulfilling prophecies for students with low self-concepts in a particular achievement domain than for students with high self-concepts in a particular achievement domain and whether positive expectations produce more powerful self-fulfilling prophecies for students with high self-concepts in a particular achievement domain than for students with low self-concepts in a particular achievement domain.

Susceptibility Hypothesis 1: Are Students With Low Self-Concepts in a Particular Achievement Domain More Susceptible to Self-Fulfilling Prophecies Than Students With High Self-Concepts in a Particular Achievement Domain?

It also is possible that perceiver expectations produce more powerful self-fulfilling prophecies for targets with low self-concepts than for targets with high self-concepts, regardless of whether those expectations are positive or negative. Persuasive messages influence people who are low in self-esteem more than they influence people who are high in self-esteem (e.g., Abelson & Lesser, 1966; A. R. Cohen, 1966; Janis, 1954; Leventhal & Perloe, 1962). People with low self-esteem are also less clear and less confident about their self-concepts (Campbell & Lavelle, 1993; Pelham, 1991). This lack of clarity and confidence may render them more susceptible than people high in self-esteem to all sorts of social influence (Brockner, 1983; Campbell & Lavelle, 1993; Jussim, 1986, 1990; Swann & Ely, 1984). This suggests that perceivers' expectations may produce more powerful self-fulfilling prophecies among targets with low self-concepts than among targets with high self-concepts. We therefore examined whether teacher expectations produced stronger self-fulfilling prophecies among students with low self-concepts in a particular achievement domain than among students with high self-concepts in a particular achievement domain.

Susceptibility Hypothesis 2: Are Low Achievers More Susceptible Than High Achievers to Self-Fulfilling Prophecies?

Students with poor records of academic performance may also be especially susceptible to positive and negative self-fulfilling prophecies. Lows' greater susceptibility to negative self-fulfilling prophecies may arise because lows might simply have less ability, on average, than do highs. To improve academically, therefore, lows may need to compensate for their lower ability with hard work. However, hard work requires motivation, something that lows may have in short supply because of the higher

frequency with which they, in comparison to highs, experience negative feedback (Deci & Ryan, 1985; Eccles & Wigfield, 1985; Jussim, 1986; Jussim, Soffin, Brown, Ley, & Kohlhepp, 1992). This suggests that negative expectations may undermine the motivation that lows need to compensate for their lower ability.

In addition, students who find school consistently difficult and unpleasant (as lows might if they are struggling academically and receiving frequent negative feedback from their teachers) may devalue the importance that they place on school (Eccles & Midgley, 1989; Eccles, Arberton, et al., 1993; Jussim, 1986; Midgley, Feldlaufer, & Eccles, 1989; Steele, 1992; Steele & Aronson, 1995). These students may give up trying to succeed in school, allow their achievement to decline, and thus ultimately fulfill their teachers' negative expectations. In contrast, because of highs' possibly greater ability, and because they experience negative feedback less frequently than do lows, highs may find school a much more pleasant place to be. Highs, therefore, may be much less likely than lows to devalue school when they do find it difficult. They may also be able to maintain a relatively high level of achievement even when faced with negative teacher expectations. This may explain why highs may be less susceptible to negative teacher expectations than are lows.

Lows may also be more susceptible than highs to positive expectations. When low achievers find themselves faced with an encouraging and challenging teacher, their motivation may rise dramatically, leading them to work especially hard in school. If lows sustain this level of hard work over the course of the school year, their performance should improve. Of course, highs may also be more motivated by an encouraging and challenging teacher. However, because highs are already performing well, their level of performance may not have as much room for growth as does lows' performance. Positive expectations, therefore, may increase lows' performance more than they increase highs' performance.

Lows' greater susceptibility to positive and negative expectations may also stem from processes similar to that discussed for targets with low self-concepts. Low achievers may be less certain than highs about how well they can perform in school. Although lows typically hope to improve, they may not be sure that they can (Jussim, 1986; Swann, 1987). Therefore, low achievers, much like targets with low self-concepts, may seek certainty by internalizing their teachers' expectations. This may be especially likely among low achievers who have multiple vulnerabilities. Imagine a student who has low ability, frequently receives negative feedback, has fewer past academic successes, and is uncertain about future academic successes. Such a student may have fewer emotional and psychological resources with which to resist internalizing his or her teachers' expectations (see Jussim et al., 1996, for research on other types of multiple vulnerabilities to self-fulfilling prophecies). Highs, in contrast, may have many more emotional and psychological resources than do lows with which to resist teacher expectations. They may find school easier, receive positive feedback more frequently, have more past successes, and be more certain about future successes. For these reasons highs may be better able than lows to resist internalizing their teachers' expectations. These processes, too, may render low achievers more susceptible

Table 1
Summary of Hypotheses

Hypothesis	Prediction
Golem	Negative expectations produce more powerful self-fulfilling prophecies than do positive expectations.
Self-enhancement	Positive expectations produce more powerful self-fulfilling prophecies than do negative expectations.
Self-consistency	Expectations produce more powerful self-fulfilling prophecies when they match students' self-concepts in a particular achievement domain. Therefore, positive expectations should produce more powerful self-fulfilling prophecies for students with high self-concepts in a particular achievement domain; negative expectations should produce more powerful self-fulfilling prophecies for students with low self-concepts in a particular achievement domain.
Susceptibility	Expectations produce more powerful self-fulfilling prophecies for students with low self-concepts in a particular achievement domain than for students with high self-concepts in a particular achievement domain. Expectations produce more powerful self-fulfilling prophecies for low achievers than for high achievers.

than high achievers to both positive and negative expectations. A summary of the hypotheses is presented in Table 1.

Conceptual Model

Background Variables

Figure 1 presents the conceptual model underlying this research. The model assumes that student background variables

can influence teacher perceptions and students' future achievement. Previous research using this data set has clearly identified how these background variables serve as a basis for teacher perceptions (represented by the thin arrow; Jussim, 1989; Jussim & Eccles, 1992). They are not the focus of this study, however, and are not discussed further.

Self-Fulfilling Prophecies

The model also assumes that teacher perceptions can influence students' future achievement (represented by the thick horizontal arrow). Conceptually, this arrow represents self-fulfilling prophecies—influences of teacher expectations on student achievement. However, because in this study we used naturalistic data, in the Results section operational hypotheses are phrased in terms of some variables "predicting" rather than "causing" other variables.

Self-fulfilling prophecies, by definition, start with inaccurate expectations. Although an in-depth analysis of the accuracy of expectations is beyond the scope of this research (see Jussim, 1989, 1991, 1993; Jussim & Eccles, 1992, 1995a; Jussim et al., 1996), we present a brief overview here of how accuracy and inaccuracy in expectations are represented in this model. This model, related conceptual models (e.g., Jussim, 1989, 1991, 1993; Jussim & Eccles, 1992, 1995a, Jussim et al., 1996), and nearly all existing formal models of accuracy in social perception (e.g., Funder, 1995; Judd & Park, 1993; Kenny, 1994) define accuracy quantitatively, rather than qualitatively. In other words, teacher expectations can be completely accurate, completely inaccurate, or anything in between. For example, the expectation that a student will score at the 95th percentile on a standardized test is completely accurate if the student does perform at the 95th percentile, is slightly less accurate if the student scores at the 90th percentile, and is considerably less accurate if the student scores at the 50th percentile. Therefore, even mostly accurate teacher expectations may be inaccurate to some degree. It

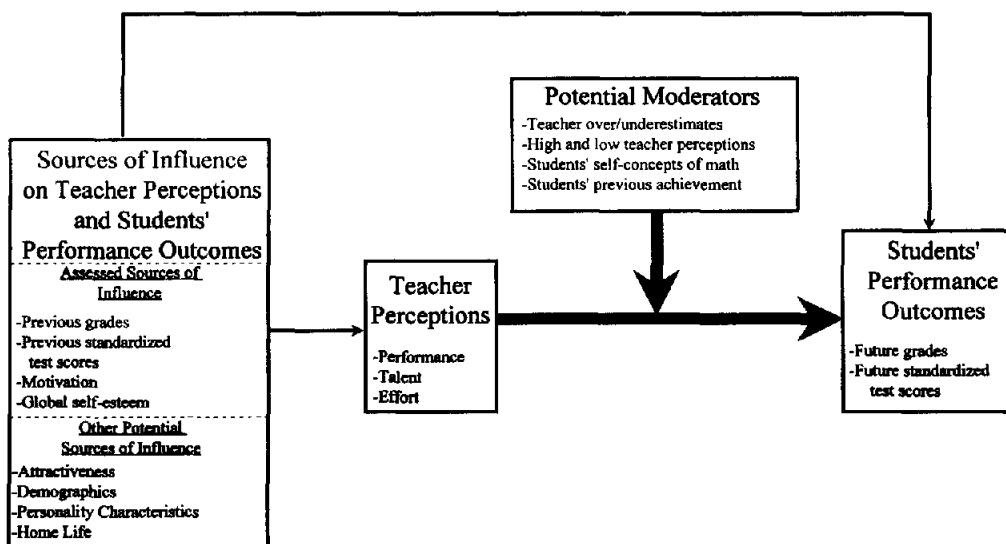


Figure 1. Conceptual model of relations among teacher perceptions and student achievement.

is this inaccuracy that has the potential to create self-fulfilling prophecies.

In this model, inaccuracy is the component of teacher perceptions that is not based on valid predictors of future achievement. Relations between teacher perceptions and students' future achievement are assessed in the context of this model, which controls for major predictors of future achievement. Conceptually, therefore, the path relating teacher perceptions to student achievement represents the influence of the inaccurate portion of teacher perceptions on students' future achievement. Operationally, however, the extent to which such paths represent self-fulfilling prophecies depends on the quality of the model, measures, and background variables included as controls, issues to which we return in the Results and Discussion sections (see also Jussim, 1989, 1991, 1993; Jussim & Eccles, 1995a, 1995b; Jussim et al., 1996, for detailed discussions of theoretical and empirical issues involved in identifying accuracy, inaccuracy, and self-fulfilling prophecies under naturalistic conditions).

Moderation

The thick vertical arrow represents the idea that various proposed moderators may increase or decrease the self-fulfilling influence of teacher perceptions on student achievement. The relationships represented by this arrow are the main focus of the present research.

Method

Participants

This research was based on data collected for the Michigan Study of Adolescent Life Transitions (Eccles, 1988), which focused on adolescents' transition from elementary to junior high school in 12 school districts in southeastern Michigan.¹ Analyses for the present study were based on 98 teachers and 1,539 students. There were 1,396 White students, 60 African American students, 40 other, and 43 race unknown. There were 823 girls and 716 boys. Each student had the same math teacher over the course of the entire school year.

Questionnaires

All teacher perception variables were assessed in early October of sixth grade. All student variables were assessed in late September or early October, shortly before the assessment of teacher perceptions. For a complete discussion of all of the teacher and student questionnaire items included in this study, see Eccles et al., 1989; Eccles (Parsons) et al., 1983; and Eccles-Parsons et al., 1982.

Teacher perceptions. Teacher questionnaires assessed teachers' beliefs about a variety of issues related to teaching (e.g., student achievement, teaching efficacy, and obedience issues). Included in the present study were three questions about teachers' perceptions of their students' achievement. Specifically, teachers evaluated the performance, natural math talent, and effort of each student in their math class.

Global self-esteem. Global self-esteem was assessed with the General Self-Worth subscale of the Perceived Competence Scale for Children (Harter, 1982). Initial analyses provided no evidence that global self-esteem moderated self-fulfilling prophecy effects. Therefore, these effects are not discussed further. Global self-esteem was, however, included as a background/control variable in all of the reported analyses.

Self-concept of math ability. Questionnaires also assessed students' self-concept of math ability. *Self-concept of math ability* refers to stu-

dents' perceived skill and competence in the specific domain of math, and thus it differs from global self-esteem, which assessed students' perceived self-worth in general. Self-concept of math ability was assessed with two questions. Students rated their overall math ability, in general, and their math ability in comparison to all other students in their math class (also see Eccles et al., 1989; Eccles (Parsons) et al., 1983; Eccles-Parsons et al., 1982).

Student motivation and effort in math. Students indicated the amount of effort they exerted in math, the time they spent on math homework, and the value they placed on math. Two aspects of math value were assessed: interest value and utility value. Interest value in math was assessed with two questions. First, students indicated their degree of interest in math. Second, students indicated how much they liked math. Utility value of math was assessed with three questions. Students first indicated the general usefulness of the math they were currently learning. Second, students indicated how useful the math they were currently learning would be after graduating from high school. Third, students indicated how useful the math that they would learn in high school would be after graduating from high school. These scales were developed out of a comprehensive expectancy value model of achievement. For a complete discussion of the theoretical rationale for these value items and confirmatory factor work on the distinction between these two components of value, see Eccles and Wigfield (1985) and Eccles (Parsons) et al. (1983).

Measures of Students' Achievement

Previous achievement. Students' previous achievement consisted of final marks in fifth-grade math and percentile ranks on one of four standardized math tests taken during the fifth or beginning of the sixth grade.

Future achievement. Students' future achievement consisted of scores on the math section of the Michigan Educational Assessment Program (MEAP), a standardized test taken in October by all seventh-grade Michigan students.

Results

Preliminary Analyses

We used Cronbach's alpha to assess the reliability of three scales that measured students' interest value of math, utility value of math, and self-concept of math ability. Alpha equaled .84 for interest value of math, .72 for utility value of math, and .81 for self-concept of math ability. Descriptive statistics are reported in Table 2; correlations are reported in Table 3.

Overview of Main Analyses

Analyses examined four questions: (a) Do positive or negative teacher perceptions more strongly predict student achieve-

¹ Using this same data set, Jussim (1989) and Jussim and Eccles (1992) addressed methods for teasing apart self-fulfilling prophecies, perceptual biases, and accuracy. Jussim et al. (1996) examined demographic moderators of self-fulfilling prophecies and tested whether self-fulfilling prophecies were more powerful among girls or boys, among students of low or high socioeconomic status, and among African American or White students. The present study differs from this work in that it examines the power of positive and negative self-fulfilling prophecies and whether previous achievement and students' self-concepts in a particular achievement domain moderate the self-fulfilling effects of teacher perceptions.

Table 2
Descriptive Statistics

Variable	Min-max value	<i>M</i>	<i>SD</i>
Teacher perceptions			
Performance	1-5	3.49	1.10
Talent	1-7	4.92	1.31
Effort	1-7	5.14	1.36
Students' previous achievement			
Test scores in percentile ranks	1-99 ^a	62.70 ^a	25.04 ^a
Final marks in fifth-grade math	3-16 ^b	11.56	2.44
Students' future achievement:			
MEAP scores	4-28 ^c	23.11	4.55
Interest value of math	2-14 ^d	9.52	3.39
Utility value of math	3-21 ^d	18.41	3.09
Global self-esteem	5-20 ^e	13.91	3.25
Self-concept of math ability	2-14 ^e	10.22	2.42

Note. *N* = 1,539. All variables except test scores are in raw score units. *Test scores* refers to previous standardized test scores in percentile ranks. Min = minimum; max = maximum; MEAP = Michigan Educational Assessment Program.

^a Values reflect percentile ranks. ^b Higher values reflect higher grades. ^c Higher values reflect higher MEAP scores. ^d Higher values reflect greater value placed on math. ^e Higher values reflect more favorable self-concepts of math ability.

ment? (b) Do teacher perceptions predict achievement more strongly when they match students' self-concepts of math ability than when they do not? (c) Do teacher perceptions predict achievement more strongly among students with low self-concepts of math ability than among students with high self-concepts of math ability? and (d) Do teacher perceptions predict achievement more strongly among low achievers than among high achievers?

Assessing moderation. These questions all address issues

of moderation—identifying conditions under which, or groups among whom, self-fulfilling prophecies are more powerful. Multiple regression analyses tested for moderation in two steps. First, we assessed a base model that assumed that the control variables (students' final marks in fifth-grade math, scores on a standardized math test, self-concept of math ability, global self-esteem, effort spent on math, time spent on math homework, and interest and utility value placed on math) and the three teacher perception variables (performance, talent, and effort) predicted students' future achievement (see Table 4 for results regarding the base model).

Next we assessed the moderation hypotheses (see Baron & Kenny, 1986) by estimating a new model that added three product terms to the base model. We created these terms by multiplying the hypothesized moderator with each of the three teacher perception variables (performance, talent, and effort). For example, to test moderation by student self-concept of math ability, we created three new variables: performance × self-concept, talent × self-concept, and effort × self-concept (where *performance*, *talent*, and *effort* all refer to teacher perceptions of students, and *self-concept* refers to students' self-concept of math ability).

A different type of multiplicative product is a quadratic or squared term used to test curvilinear relationships. A quadratic term is virtually identical to a moderator in that the variable moderates itself (see Judd & McClelland, 1989, for a more detailed discussion). When we tested curvilinear relationships, values representing teacher perceptions were squared. For example, to examine whether positive or negative teacher perceptions were more powerful, we created three quadratic terms: performance × performance, talent × talent, and effort × effort (where *performance*, *talent*, and *effort* all refer to teacher perceptions of students).

In all second steps, interest value of math, utility value of

Table 3
Correlations

Variable	1	2	3	4	5	6	7	8	9	10	11	12
Teacher perceptions												
1. Performance	—	.80	.72	.60	.53	.54	.26	.21	.15	-.16	.19	.48
2. Talent		—	.58	.63	.51	.52	.22	.19	.10	-.19	.17	.46
3. Effort			—	.45	.44	.39	.25	.17	.18	-.11	.17	.36
Achievement												
4. Test scores				—	.56	.65	.23	.17	.05	-.17	.17	.43
5. Final marks in fifth-grade math					—	.49	.25	.15	.10	-.15	.17	.39
6. MEAP scores						—	.23	.17	.08	-.13	.14	.35
Motivation for math												
7. Interest value of math							—	.33	.30	-.04	.22	.43
8. Utility value of math								—	.26	.07	.17	.25
9. Effort exerted in math									—	.11	.13	.27
10. Time spent on math homework										—	-.07	-.11
Self-conceptions												
11. Global self-esteem											—	.31
12. Self-concept of math ability												—

Note. *Performance*, *talent*, and *effort* refer to mean deviated teacher perceptions of performance, talent, and effort. *Test scores* refers to previous standardized test scores. All correlations are independent of students' classrooms. These correlations were calculated by (a) obtaining residuals for each variable by regressing each variable onto dummy variables that represented students' classrooms and (b) correlating the residuals. Correlations $\geq |.05|$ were significant at $p \leq .05$. MEAP = Michigan Educational Assessment Program.

Table 4
Base Models

Predictor variable	Model based on teacher perceptions that over- and underestimated students ($R^2 = .604$)			Model based on high and low teacher perceptions ($R^2 = .604$)		
	<i>B</i>	β	<i>t</i>	<i>B</i>	β	<i>t</i>
Performance	.77	.14	5.14***	.77	.19	5.14***
Talent	.08	.02	0.66	.08	.02	0.66
Effort	-.11	-.03	1.15	-.11	-.03	1.15
Test scores	.10	.54	21.48***	.09	.47	16.95***
Final marks in fifth-grade math	.29	.16	6.57***	.22	.12	4.76***
Interest value of math	.06	.04	1.94	.06	.04	1.98*
Utility value of math	.04	.03	1.45	.03	.02	0.95
Time spent on math homework	-.10	-.02	0.99	-.04	-.01	0.38
Effort exerted in math	.04	.01	0.58	.03	.01	0.43
Global self-esteem	-.01	-.00	0.27	-.00	-.00	0.08
Self-concept of math ability	.05	.03	1.13	-.03	-.02	0.70

Note. Dependent variable: Michigan Educational Assessment Program Scores. Both regression analyses controlled for classroom, thus removing sources of nonindependence between classrooms. *Performance*, *talent*, and *effort* refer to teacher perceptions of performance residuals, talent residuals, and effort residuals for teacher over- and underestimates and refer to mean deviated teacher perceptions of performance, talent, and effort for the model based on high and low teacher perceptions. *Test scores* refers to previous standardized test scores. *B* refers to the unstandardized regression coefficient, and β refers to the standardized regression coefficient. Error degrees of freedom for all analyses ranged from 1,405 to 1,413.

* $p \leq .05$. *** $p \leq .0001$.

math, time spent on math homework, effort exerted in math, global self-esteem, and self-concept of math ability did not significantly predict MEAP scores. Therefore, they are omitted from all tables and are not discussed further (Tables 5 and 6). All analyses are based on a sample that included 1,539 students from 114 classrooms. Error degrees of freedom for all analyses ranged from 1,405 to 1,413. These degrees of freedom reflect the number of students (1,539) minus the number of predictors, which ranged from 12 to 20, minus the number of classroom dummy variables, which was always 113 (number of classrooms minus 1; see Cohen & Cohen, 1983).

Collinearity. The hypothesis that a particular variable moderated the self-fulfilling effects of teacher perceptions could be confirmed if either the block of three product terms, or any single product term, significantly predicted students' future achievement. However, because the three product terms were highly correlated with each other,² testing all three simultaneously could artificially reduce the size and significance of their corresponding regression coefficients (e.g., Gordon, 1968), thereby substantially underestimating the role of any one moderator.³

Therefore, we adopted the following procedures to reduce underestimation of moderator effects due to collinearity among the product terms. We examined the individual product terms in more detail if the block of three product terms significantly increased R^2 at $p < .15$ or if one of the individual product terms was significant at $p \leq .05$. If either of these conditions was met, we used a two-step procedure to examine the individual product terms in more detail. First, we examined a model that added only the product term that most strongly predicted future achievement to a model that included the base model variables and all lower order product terms. The product term

significantly predicted future achievement in each of the analyses that we performed. Next, we added the other two product terms to the model. In each of the analyses, adding the two product terms as a block never led to even a marginally significant increase in R^2 (all $ps \geq .15$). Thus, all models yielding statistically significant evidence of moderation included only a single product term. We also transformed all variables from raw scores into mean deviated variables by subtracting the scale mean from each variable (see the discussion on "centering" in Chapter 6 of Cohen & Cohen, 1983, for an explanation about why this method reduces problems associated with collinearity).

Nonindependence of teacher perceptions. All analyses reported below are multiple regression analyses in which the stu-

² The three product terms were correlated with each other for two reasons: (a) Each teacher perception variable was multiplied by the same potential moderator, and (b) the teacher perception variables themselves were moderately to highly intercorrelated (about .5-.8).

³ This is "artificial" because each individual coefficient is estimated after controlling for all other variables in the model. This includes not only the real control variables but also the other teacher perception-moderator product terms—in essence, potentially controlling "out" much of the moderational relation we are attempting to assess. We considered eliminating this problem by simply summing the three teacher perception variables, thereby providing a single index of teacher perceptions. However, LISREL analyses using this data set have shown that the assumption that the three teacher perception variables reflect a single, underlying factor is not viable. Models making this assumption have major theoretical weaknesses, are clearly misspecified, and do not fit the data (Jussim, 1989; Jussim & Eccles, 1992). Therefore, the analyses in the present article do not combine the three teacher perception variables to form a single scale.

Table 5
Teacher Perception Residuals Predicting MEAP Scores

Predictor variable	Golem versus self-enhancement ($R^2 = .605$)			Golem versus self-enhancement controlling for achievement moderation ($R^2 = .608$)			Susceptibility Hypothesis 2 achievement moderation ($R^2 = .608$)		
	<i>B</i>	β	<i>t</i>	<i>B</i>	β	<i>t</i>	<i>B</i>	β	<i>t</i>
Performance residuals squared	.20	.04	2.14*	.13	.03	1.37	—	—	—
Performance Residuals \times Test Scores	—	—	—	-.01	-.06	3.37**	-.02	-.07	3.75**
Performance residuals	.82	.15	5.40***	.74	.13	4.87***	.71	.13	4.70***
Talent residuals	.06	.01	0.49	.06	.01	0.52	.07	.02	0.64
Effort residuals	-.11	-.03	1.13	-.08	-.02	0.88	-.08	-.02	0.86
Test scores	.10	.54	21.54***	.10	.54	21.51***	.10	.54	21.48***
Final marks in fifth-grade math	.29	.16	6.57***	.29	.16	6.65***	.29	.16	6.66***

Note. Performance, talent, and effort refer to teacher perceptions of performance residuals, talent residuals, and effort residuals. Test scores refers to previous standardized test scores. All regression analyses controlled for classroom, thus removing sources of nonindependence between classrooms. All results can be considered examples of the effects of the overall set of teacher perception residuals. All regression analyses included as controls interest value of math, utility value of math, time spent on math homework, effort exerted in math, global self-esteem, and self-concept of math. However, because none were significant predictors of MEAP scores, they are not reported here. *B* refers to the unstandardized regression coefficient, and β refers to the standardized regression coefficient. Error degrees of freedom for all analyses ranged from 1,405 to 1,413. Dashes indicate variables that were excluded from an analysis. MEAP = Michigan Educational Assessment Program.

* $p \leq .05$. ** $p \leq .001$. *** $p \leq .0001$.

dent was the unit of analysis. Because teachers rated all of the students in their classrooms, teacher perceptions are not independent of one another. However, all analyses included classroom (coded as dummy variables) as predictors of future achievement, thereby removing the basic nonindependence between classrooms⁴ (for an examination of teacher expectations for whole classes, see Vannoy, Smith, Jussim, & Eccles, 1996).

Teacher over- and underestimates and high and low teacher perceptions. We performed two sets of regression analyses. The first focused on moderators of the relation of teacher over- and underestimates to students' future achievement. Teacher over- and underestimates reflect teacher perceptions relative to each student's own background. That is, they indicate the extent to which teachers over- or underestimated students' future achievement in comparison to what was predicted by each student's own background variables. We used teacher over- and underestimates to test the power of positive and negative self-fulfilling prophecies.

We empirically identified teacher over- and underestimates by regressing the three teacher perception variables on the student background variables (i.e., final marks in fifth-grade math, scores on a standardized math test, self-concept of math ability, global self-esteem, effort spent on math, time spent on math homework, and interest and utility value placed on math). Next, we created three new variables by saving the residuals from these regression analyses. This yielded three residual teacher perception variables for each student: one for performance, one for talent, and one for effort. Each residual represents the portion of teacher perceptions that was not predicted by students' background variables. Positive residuals indicate that a teacher had more positive perceptions of a student than was predicted by that student's background variables. Negative residuals indicate that a teacher had more negative perceptions than was predicted by that student's background variables.

Our second set of analyses focused on moderators of the

relation of high and low teacher perceptions to students' future achievement. High and low teacher perceptions reflect how favorably teachers viewed each student relative to other students. However, high teacher perceptions do not necessarily overestimate students, and low teacher perceptions do not necessarily underestimate students. For example, the high teacher perception that a student is performing "above the middle of the class" will underestimate a particular student who is actually "one of the best in the class." Therefore, this set of analyses cannot test the power of positive and negative self-fulfilling prophecies. It can, however, test the power of objectively favorable and unfavorable teacher perceptions to create self-fulfilling prophecies.

Because of the important conceptual differences between over- and underestimates and between high and low teacher perceptions, the results reported below have been divided into two sections. The first section focuses exclusively on teacher over- and underestimates. The second section focuses exclusively on high and low teacher perceptions. Both sections test the golem, self-enhancement, self-consistency, and susceptibility hypotheses.

Depiction of significant moderator effects. We depicted significant moderator effects by plotting the predicted relations separately for different levels of each significant moderator (see, e.g., Judd & McClelland, 1989). For example, quadratic relationships show the relationship between teacher perceptions and students' future achievement for each level of teacher percep-

⁴ Controlling for classroom does not remove the possibility that there may be interactions between students' classrooms and the predictor variables. Hierarchical linear modeling (HLM) would, therefore, have been the most appropriate method of analysis. However, because of the small number of students in many of the classrooms and the complexity of our hypotheses, HLM was not a practical solution. Thus, the conclusions drawn from this study should be taken with some caution.

Table 6
High and Low Teacher Perceptions Predicting MEAP Scores

Predictor variable	Golem versus self-enhancement (R ² = .608)			Susceptibility hypotheses			Golem versus self-enhancement, achievement moderation controlled (R ² = .617)		
				Moderation by achievement (R ² = .617)			Moderation by self-concept (R ² = .609)		
	B	β	t	B	β	t	B	β	t
Talent squared	-.15	-.08	3.95***	-.02	-.13	7.06***	-.02	-.01	0.56
Performance × Test Scores							-.02	-.12	5.85***
Performance × Self-Concepts	.73	.18	4.84***	.67	.16	4.49***	.66	.16	4.47***
Performance	.01	.00	0.07	.09	.03	0.80	.08	.02	0.68
Talent	-.05	-.02	0.54	-.05	-.02	0.55	-.04	-.01	0.48
Effort	.09	.47	16.92***	.08	.44	15.76***	.08	.44	15.76***
Test scores	.23	.12	5.07***	.23	.12	5.06***	.23	.12	5.09***
Final marks in fifth-grade math									

Note. Performance, talent, and effort refer to mean deviated teacher perceptions of performance, talent, and effort. Test scores refers to previous standardized test scores. Self-concepts refers to self-concepts of math. All regression analyses controlled for classroom, thus removing sources of nonindependence between classrooms. All results can be considered examples of the effects of the overall set of high and low teacher perceptions. All regression analyses included as controls interest value of math, utility value of math, time spent on math homework, effort exerted in math, global self-esteem, and self-concept of math. Because none were significant predictors of MEAP scores, they are not reported here. The following coefficients for self-concept of math are, however, needed to calculate predicted values and effect sizes (see footnote 4 for effect size calculations) for the analysis testing moderation by self-concept: B = -.07; β = -.04. In the above table, B refers to the unstandardized regression coefficient, and β refers to the standardized regression coefficient. Error degrees of freedom for all analyses ranged from 1,405 to 1,413. Dashes indicate variables that were excluded from an analysis. MEAP = Michigan Educational Assessment Program. *** p ≤ .0001.

tions. Linear relationships show the relationship between teacher perceptions and students' future achievement separately for highs and lows (i.e., high and low achievers). Highs and lows were represented by scores 1 SD above and below the mean, respectively.

Teacher Over- and Underestimates

Do teacher perception over- or underestimates predict achievement more strongly? The golem and self-enhancement hypotheses. The first set of analyses tested the competing predictions of the golem and self-enhancement hypotheses. The golem hypothesis is that teacher underestimates predict achievement more strongly than teacher overestimates. The self-enhancement hypothesis is that teacher overestimates predict achievement more strongly than teacher underestimates. Because the only difference between the golem and self-enhancement hypotheses is in the direction of the predicted effect, the same analyses simultaneously tested both hypotheses.

Both hypotheses predict curvilinear relationships between teacher perception residuals and student achievement. Therefore, we squared each teacher perception residual to create three quadratic terms: teacher perceptions of performance residuals squared, teacher perceptions of talent residuals squared, and teacher perceptions of effort residuals squared. Step 1 of the initial analyses included the base model variables (see Table 4) as predictors of MEAP scores. Step 2 added the three squared terms to the base model.

Results showed that the set of three teacher perception residuals squared marginally predicted MEAP scores, F(3, 1,411) = 1.96, p = .12. Teacher perceptions of performance residuals squared was the strongest moderator in this model. As an example of the overall set's effect, we re-estimated the model using only the base model variables plus teacher perceptions of performance residuals squared as predictors of MEAP scores. In this model, teacher perceptions of performance residuals squared significantly predicted MEAP scores (β = .04, p = .03). Table 5 presents the results from this final model. Figure 2 depicts the direction of the effect and shows that teacher perceptions that overestimated students' performance predicted MEAP scores more strongly than did teacher perceptions that underestimated students' performance. This result is consistent with the self-enhancement hypothesis and not with the golem hypothesis.

We also calculated effect sizes for the lowest and highest teacher perception of performance residuals (-2.97 and 2.27) and for each integer between the lowest and highest (i.e., -2 through 2).⁵ Their effects were -.09, .00, .10, .20, .29, .39,

⁵ The following equation was used to determine effect sizes for all quadratic moderator effects (see, e.g., Judd & McClelland, 1989, for more details): achievement = [B₁ + (B₂ × 2 × TP)] × (SD_{teacher perceptions} / SD_{dependent variable}). B₁ is the unstandardized regression coefficient from the final model (see Tables 5 and 6). For example, if teacher perceptions of performance residuals squared significantly predicted MEAP scores, then B₁ refers to the unstandardized regression coefficient associated with teacher perceptions of performance residuals in the final model. B₂ is the unstandardized regression coefficient associated with the significant squared teacher perception variable (e.g., the unstandardized coefficient associated with teacher perceptions of performance residuals squared). We obtained the numeral 2 by taking the first derivative of the significant quadratic teacher perception variable (see, e.g., Judd &

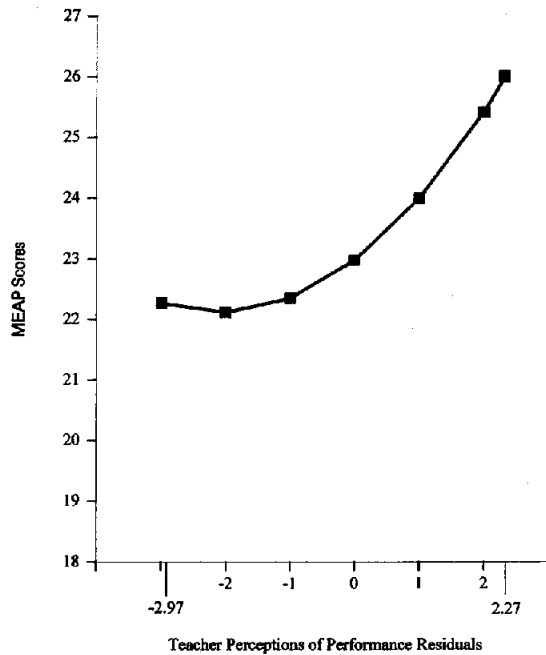


Figure 2. Teacher perceptions that overestimated students predicted MEAP scores more powerfully than teacher perceptions that underestimated students. Predicted MEAP scores are shown for students who were at the mean on all variables except teacher perceptions of performance residuals. The depicted relation represents an example of the effects of the overall set of teacher perception residuals. MEAP = Michigan Educational Assessment Program.

and .42 for performance residuals ranging from -2.97 through 2.27 , respectively. Teacher perceptions that overestimated students' performance, therefore, predicted MEAP scores more strongly than teacher perceptions that underestimated students' performance.

Do teacher perception residuals predict achievement more strongly when they match students' self-concepts of math ability? The self-consistency hypothesis. We also tested the self-consistency hypothesis that teacher perceptions that overestimate students predict achievement more strongly for students with high self-concepts of math ability, whereas teacher perceptions that underestimate students predict achievement more strongly for students with low self-concepts of math ability. A quadratic interaction was necessary to test this hypothesis, because teacher perceptions that over- and underestimated students were expected to predict achievement more strongly (i.e., teacher perception residuals squared) when their valence matched students' self-concepts of math ability (i.e., teacher perception residuals squared \times students' self-concepts of math ability). Despite having substantial power in the analysis, however, the results failed to provide any support for the self-consis-

tency hypothesis. The quadratic interaction product terms did not even approach statistical significance, $F(3, 1,405) = 0.82$, $p = .48$, and are not discussed further.

Do teacher perception residuals predict achievement more strongly among students with low self-concepts of math ability than among students with high self-concepts of math ability? The susceptibility hypothesis 1. The first susceptibility hypothesis is that teacher perceptions predict achievement more strongly for students with low self-concepts of math ability than for students with high self-concepts of math ability. To test this prediction we created three product terms: teacher perceptions of performance residuals \times students' self-concepts of math ability, teacher perceptions of talent residuals \times students' self-concepts of math ability, and teacher perceptions of effort residuals \times students' self-concepts of math ability. Step 1 of the initial analysis included the base model variables (see Table 4) as predictors of MEAP scores. Step 2 added the three product terms to the base model.

In spite of substantial power in the analysis, results indicated that, when tested as a block, the three product terms did not significantly predict MEAP scores, $F(3, 1,411) = 0.73$, $p = .53$. This means that teacher perceptions that overestimated or underestimated students predicted achievement about just as strongly for students with low self-concepts of math ability as they did for students with high self-concepts of math ability. This result fails to support the first susceptibility hypothesis, which stated that teacher over- and underestimates predict achievement more strongly for students with low self-concepts of math ability than for students with high self-concepts of math ability.

Do teacher perception residuals predict achievement more strongly among low achievers than among high achievers? The susceptibility hypothesis 2. The second susceptibility hypothesis is that teacher perceptions predict achievement more strongly for low achievers than for high achievers. We tested this hypothesis with students' final marks in fifth-grade math and their scores on a previous standardized test. However, because only students' previous standardized test scores moderated the relationship between teacher perception residuals and students' future achievement, moderation by previous grades is not discussed.

To examine the moderating influence of previous standardized test scores on teacher perceptions that over- and underestimated students, we created three product terms: teacher perceptions of performance residuals \times previous standardized test scores, teacher perceptions of talent residuals \times previous standardized test scores, teacher perceptions of effort residuals \times previous standardized test scores. Step 1 of the initial analysis included the base model variables (see Table 4) as predictors of MEAP scores. Step 2 added the three product terms to the base model. Results showed that when tested as a block, the three product terms significantly predicted MEAP scores, $F(3, 1,411) = 4.94$, $p < .01$. Teacher perceptions of performance residuals \times previous standardized test scores was the strongest moderator in this model. As an example of the overall set's effect, we re-estimated this model including the base model variables (Table 4) plus the teacher perception of performance residuals \times previous standardized test scores product term as predictors of MEAP scores, $\beta = -.07$, $p < .01$. Table 5 presents the results from this final model.

McClelland, 1989, for more details). TP refers to each level of teacher perceptions (e.g., teacher perception residuals ranging from one extreme residual to the other extreme residual). Reported effect sizes are interpretable as standardized regression coefficients.

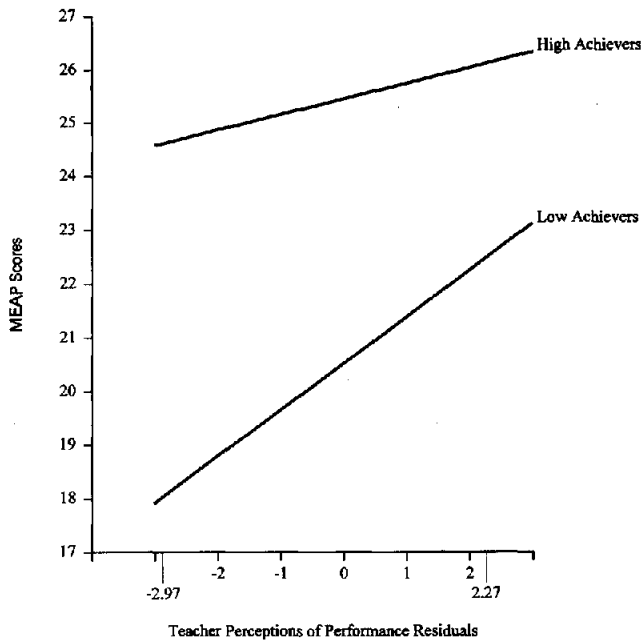


Figure 3. Teacher perceptions that over- and underestimated students predicted MEAP scores more powerfully for low achievers than for high achievers. *Low achievers* refers to students with previous standardized test scores 1 *SD* below the mean; *high achievers* refers to students with previous standardized test scores 1 *SD* above the mean. Predicted MEAP scores are shown for students who were at the mean on all variables except teacher perceptions of performance residuals and previous achievement. The depicted relation represents an example of the effects of the overall set of teacher perception residuals for low and high achievers. MEAP = Michigan Educational Assessment Program.

Figure 3 depicts the relation between teacher perceptions of performance residuals and MEAP scores for low and high achievers separately. Teacher perceptions of performance residuals predicted low achievers' MEAP scores more strongly than they predicted high achievers' MEAP scores. Effect sizes were .26 for low achievers and .08 for high achievers.⁶ This pattern supports the second susceptibility hypothesis, that teacher perceptions predict achievement more strongly among low achievers than among high achievers.

Was the greater power of positive versus negative self-fulfilling prophecies independent of moderation by achievement? Analyses indicate two patterns. First, teacher perceptions that overestimated students predicted achievement more strongly than did teacher perceptions that underestimated students. Second, teacher perceptions predicted achievement more strongly among low achievers than among high achievers. However, these patterns need not be independent of one another. To probe the independence of the self-enhancement and susceptibility effects, we estimated a model that included the base model variables (Table 4), teacher perceptions of performance residuals squared, and the product term between teacher perceptions of performance residuals and students' previous standardized test scores as predictors of MEAP scores. If positive self-fulfilling prophecies are more powerful among students in general, rather than just among lows in particular, then teacher perceptions of perfor-

mance residuals squared should remain a significant predictor of MEAP scores even when moderation by achievement is controlled. This did not occur, however. The results indicated that only the product term between teacher perceptions of performance residuals and students' previous standardized test scores was a significant predictor of MEAP scores ($\beta = -.06$, $p < .01$).

Nonetheless, there was still a small nonsignificant tendency for teacher perceptions that overestimated students' performance to predict achievement more strongly than teacher perceptions that underestimated students' performance ($\beta = .03$, $p = .17$). This suggests that there may be a very weak tendency for positive self-fulfilling prophecies to be more powerful than negative self-fulfilling prophecies among students in general, even after controlling for moderation by achievement.

Therefore, we plotted the relationship between teacher perceptions of performance residuals squared for high and low achievers separately (Figure 4). This figure sheds considerable light on the relation between teacher perceptions and student achievement. It shows that: (a) high achievers were almost completely invulnerable to teacher perceptions that underestimated them (the slope from teacher perception residuals of -2.97 to 0 is nearly flat for high achievers); (b) teacher overestimates predicted increased achievement among high achievers (the slope from 0 to 2.27 increases for high achievers); (c) the entire spectrum of teacher perception residuals predicted achievement among low achievers, so that underestimates predicted lower achievement and overestimates predicted higher achievement; (d) when comparing the curves, it is clear that, at every point, teacher perception residuals predicted achievement more strongly among low achievers than among high achievers; and (e) among low achievers, the increases in achievement predicted by teacher overestimates were greater than the decreases predicted by teacher underestimates (the slope above 0 is steeper than the slope below 0).

We also calculated effect sizes for the lowest and highest teacher perception of performance residuals (i.e., -2.97 and 2.27) and for each integer between the lowest and highest (i.e., -2 through 2). We did this separately for low and high achievers. Effect sizes for low achievers were .08, .14, .20, .26,

⁶ The following equation was used to determine effect sizes for the significant relations between teacher perceptions and students' previous achievement (see, e.g., Judd & McClelland, 1989, for more details about his equation): $\text{achievement} = [B_1 + (B_2 \times \text{significant moderator})] \times (SD_{\text{teacher perceptions}} / SD_{\text{dependent variable}})$. B_1 is the unstandardized regression coefficient from the final model (see Tables 5 and 6). For example, if students' previous standardized test scores significantly moderated the relation between teacher perceptions of performance residuals and MEAP scores, then B_1 refers to the unstandardized regression coefficient associated with teacher perceptions of performance residuals in the final model. B_2 is the unstandardized regression coefficient associated with the significant product term (i.e., the unstandardized regression coefficient associated with teacher perceptions of performance residuals \times students' previous standardized test scores). *Significant moderator* refers to the variable that moderated the relationship between teacher perceptions and MEAP scores (e.g., students' previous standardized test scores). The values used for *significant moderator* corresponded to scores ± 1 *SD* from the mean. Reported effect sizes are interpretable as standardized regression coefficients.

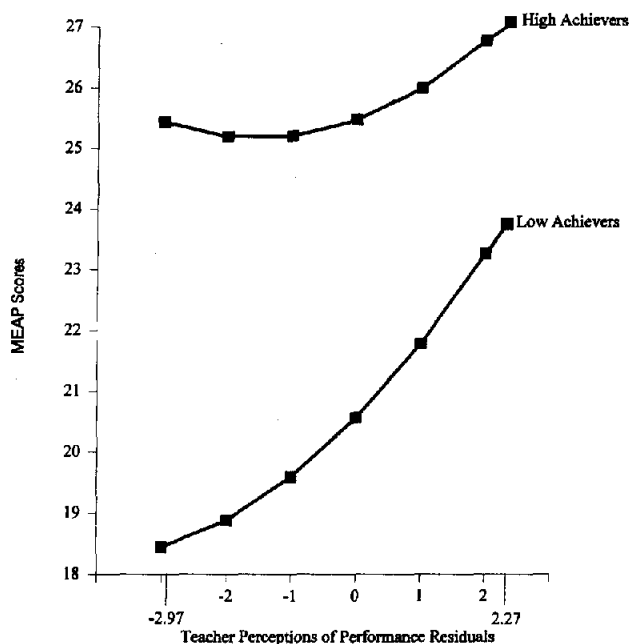


Figure 4. Among both low and high achievers, teacher perceptions that overestimated students predicted achievement more strongly than teacher perceptions that underestimated students. *Low achievers* refers to students with previous standardized test scores 1 *SD* below the mean; *high achievers* refers to students with previous standardized test scores 1 *SD* above the mean. Predicted MEAP scores are shown for students who were at the mean on all variables except teacher perceptions of performance residuals and previous achievement. The depicted relation represents an example of the effects of the overall set of teacher perception residuals for low and high achievers. MEAP = Michigan Educational Assessment Program.

.33, .39, and .41 for teacher perceptions of performance residuals ranging from -2.97 through 2.27 , respectively. Effect sizes for high achievers were $-.09$, $-.03$, $.03$, $.10$, $.16$, $.22$, and $.24$ for teacher perceptions of performance residuals ranging from -2.97 through 2.27 , respectively (see footnote 4 for calculations).

These results provide strong support for the susceptibility hypothesis that low achievers are more susceptible to self-fulfilling prophecies than are high achievers. They also suggest the possibility that self-enhancement effects are operating among students in general. However, because the self-enhancement effects we observed were weak at best, especially after controlling for moderation by achievement, strong conclusions about the power of positive self-fulfilling prophecies is premature. Nonetheless, even very weak support in favor of the self-enhancement hypothesis strongly argues against the golem hypothesis. Negative self-fulfilling prophecies were clearly not more powerful than positive self-fulfilling prophecies.

Was the greater susceptibility to self-fulfilling prophecies of lows versus highs due to floor and ceiling effects? Might our results, showing that lows are more susceptible than highs to self-fulfilling prophecies, be due to floor and ceiling effects rather than to a greater susceptibility among lows per se? A floor-and-ceiling explanation predicts that lows, who are closer

to the achievement floor and therefore have more room to improve academically, would be more susceptible to positive self-fulfilling prophecies. It also predicts that highs, who are closer to the achievement ceiling and therefore have more room to decline academically, would be more susceptible to negative self-fulfilling prophecies.

We empirically tested the viability of the floor-and-ceiling explanation with a single regression analysis. The analysis included the base model variables (see Table 4), teacher perceptions of performance residuals squared, teacher perceptions of talent residuals squared, teacher perceptions of effort residuals squared, and three quadratic interaction product terms as predictors of MEAP scores. We created the quadratic interaction product terms by multiplying each squared term by students' previous standardized test scores. The quadratic interaction product terms directly test the power of positive and negative self-fulfilling prophecies among low and high achievers. Specifically, they examine whether teacher overestimates predict MEAP scores more strongly than do teacher underestimates for low achievers and whether teacher underestimates predict MEAP scores more strongly than do teacher overestimates for high achievers.

Results indicated, however, that the quadratic interaction product terms were not significant predictors of MEAP scores, $F(3, 1,405) = 0.92, p = .43$. This null finding argues against floor and ceiling effects by showing that the shape of the curves relating teacher over- and underestimates to MEAP scores did not differ significantly for low and high achievers. Our previous susceptibility and self-enhancement results also argue against floor and ceiling effects. Our susceptibility results showed that lows were more susceptible to both positive and negative self-fulfilling prophecies. This pattern is inconsistent with floor effects: Lows were more susceptible to negative self-fulfilling prophecies than were highs. Our self-enhancement results showed that there was a very weak tendency for positive self-fulfilling prophecies to be more powerful than negative self-fulfilling prophecies among students in general, including highs. Regardless of how much stock one puts in the tendency for positive self-fulfilling prophecies to be more powerful than negative self-fulfilling prophecies, this finding nonetheless strongly argues against a ceiling effects explanation: Highs were *not* more susceptible to negative self-fulfilling prophecies than they were to positive self-fulfilling prophecies.⁷

⁷ The MEAP was negatively skewed (a long tail at the low end of scores). In negatively skewed distributions there is more variation among low scores than among high scores. This may have led teacher perception residuals to predict MEAP scores more strongly at the low end (where low achievers tend to score) and to predict MEAP scores less strongly at the high end (where high achievers tend to score). If so, then teacher perception residuals may have predicted MEAP scores more strongly among low achievers than among high achievers because of the greater variation at the low end of the MEAP rather than because of lows' greater susceptibility per se. To examine this possibility we squared students' MEAP scores. This created a distribution close to normal. We then tested the final model again to determine whether students' past achievement still moderated relations between teacher perception residuals and MEAP scores squared. This model included the base model variables (Table 4) and the product term between teacher perceptions of performance residuals \times students' previous standardized test scores as predictors of MEAP squared. Results indicated that the teacher perceptions of performance residuals \times previous standardized test scores

High and Low Teacher Perceptions

We also examined the golem, self-enhancement, self-consistency, and susceptibility hypotheses with high and low teacher perceptions. These analyses were identical to the residual analyses described previously, with the exception that high and low teacher perceptions replaced teacher perception residuals. Initial analyses examined whether the block of three product terms (i.e., teacher perceptions of performance squared, talent squared, and effort squared) were significant predictors of MEAP scores. If the block was significant at $p < .15$, then we performed a final model in which we included only the base model variables (Table 4) and the strongest moderator as predictors of MEAP scores. We report only the final models here. Because we found no support whatsoever for the self-consistency hypothesis, we do not present those results.

Do high or low teacher perceptions predict achievement more strongly? The golem and self-enhancement hypotheses. Analyses indicated that teacher perceptions of talent squared most strongly predicted MEAP scores, $\beta = -.08$, $p < .01$. Table 6 presents the results from this final model as an example of the overall set's effect. Effect sizes were .35, .26, .17, .09, .00, $-.09$, $-.17$ for teacher perceptions of talent equaling 1 through 7, respectively (see footnote 4 for calculations). Consistent with the golem hypothesis, but not with the self-enhancement hypothesis, therefore, low teacher perceptions predicted achievement more strongly than did high teacher perceptions.

Do high and low teacher perceptions predict achievement more strongly among students with low self-concepts of math ability than among students with high self-concepts of math ability? The susceptibility hypothesis 1. Analyses indicated that the product term between teacher perceptions of performance and students' self-concepts of math ability most strongly predicted MEAP scores, $\beta = -.08$, $p < .01$. Table 6 presents the results from this final model as an example of the overall set's effect. Effect sizes of teacher perceptions of performance were .24 for students with low self-concepts of math ability and .10 for students with high self-concepts of math ability (see footnote 5 for computations). These results are consistent with the hypothesis that students with low self-concepts of math ability are more susceptible to self-fulfilling prophecies than are students with high self-concepts of math ability.

Do teacher perceptions predict achievement more strongly among low achievers than among high achievers? The susceptibility hypothesis 2. Analyses indicated that the product term between teacher perceptions of performance and previous standardized test scores most strongly predicted MEAP scores, $\beta = -.13$, $p < .01$. Table 6 presents the results from this final model as an example of the overall set's effect. Effect sizes of teacher perceptions of performance were .04 for high achievers and .28 for low achievers (see footnote 5 for computations). These analyses support the hypothesis that teacher perceptions

predict achievement more strongly for low achievers than for high achievers.

Were the self-concept-of-math-ability and achievement moderators independent? Both self-concept of math ability and previous achievement significantly moderated relations between high and low teacher perceptions and MEAP scores. However, because self-concept of math ability correlates with previous achievement (see Table 3), these moderator effects may not be independent. We examined their independence with a single regression analysis that included the base model variables (Table 4) and the product terms for both moderation by self-concept of math ability and moderation by achievement. Moderation by achievement significantly predicted MEAP scores ($\beta = -.12$, $p < .01$), but moderation by self-concept of math ability did not ($\beta = -.02$, $p = .28$). This means that self-concept of math ability had no moderating effects over and above that of previous achievement.

Why were low teacher perceptions more powerful? Low teacher perceptions predicted achievement more strongly than did high teacher perceptions. This result, consistent with the golem hypothesis, seems to suggest that low teacher perceptions are inherently more powerful than high teacher perceptions. However, low teacher perceptions may have predicted achievement more strongly than did high teacher perceptions because the typical targets of low perceptions (i.e., low achievers) were more susceptible to self-fulfilling prophecies.

We examined this alternative to the golem hypothesis with a single regression analysis that included the base model variables (Table 4), the quadratic teacher perception term, and the product term for moderation by achievement as predictors of MEAP scores. If low teacher perceptions were more powerful, then the squared term would have remained significant in this analysis. This did not happen—the squared teacher perception of talent term was not even close to significance ($\beta = -.01$, $p = .57$). However, the product term for moderation by achievement remained a significant predictor of MEAP scores in this analysis ($\beta = -.12$, $p < .01$). This means that low teacher perceptions predicted achievement more strongly than did high teacher perceptions because those low perceptions were held for low achievers (see Table 6). This analysis supports the susceptibility hypothesis and provides no support for the golem hypothesis.⁸

⁸ We again examined whether the pattern of achievement moderation was due to floor and ceiling effects (just as we did for the residual analyses). According to a floor-and-ceiling-effects analysis, high teacher perceptions should predict MEAP scores more strongly for low achievers than for high achievers, whereas low teacher perceptions should predict MEAP scores more strongly for high achievers than for low achievers. We examined this with a model that included the base model variables (Table 4), the teacher perceptions of performance \times previous standardized test scores product term, teacher perceptions of performance squared, talent squared, effort squared, and three quadratic interaction product terms as predictors of MEAP scores. We created the quadratic interaction product terms by multiplying each squared term by students' previous standardized test scores. Results indicated that the quadratic interaction product terms were not significant predictors of MEAP scores, $F(3, 1,405) = 1.71$, $p > .15$. This indicates that high teacher perceptions did not predict MEAP scores more strongly for low achievers than for high achievers and that low teacher perceptions did not predict MEAP scores more strongly for high achievers than for low achievers. Thus, this analysis provides no support for a floor-and-ceiling interpreta-

product term significantly predicted MEAP squared in this model, $\beta = -.05$, $p < .01$. The pattern of moderation was virtually identical to those of the findings reported in the text. Teacher perceptions of performance residuals predicted students' future achievement more strongly for low achievers than for high achievers.

Discussion

This research examined conditions under which perceivers' expectations produce more powerful self-fulfilling prophecies. Results yielded one strong pattern and one much weaker pattern. The strong pattern was that teacher perceptions predicted achievement more strongly for low achievers than for high achievers. There was also a much weaker tendency for teacher perceptions that overestimated students to predict achievement more strongly than teacher perceptions that underestimated students. Before interpreting these findings, we discuss general issues involved in the interpretation of results from naturalistic studies and the relevance of those issues to the present study.

Interpreting Results From Naturalistic Studies

Correlation design. Several limitations to our research qualify the insights it has provided regarding the conditions under which self-fulfilling prophecies are more powerful. Most important, naturalistic studies do not provide as strong a basis for causal inferences as do experiments. Typically, with correlational designs one cannot identify whether the predictor(s) caused the dependent variable, the dependent variable caused the predictor(s), or whether both were caused by a third factor. However, with longitudinal designs one can rule out the possibility that the dependent variable influenced the predictor(s). For example, we can be certain that achievement at the beginning of the seventh grade did not cause teacher perceptions at the beginning of sixth grade.

The omitted-variable problem. A second limitation to our research is the omitted-variable problem—the possibility that a relevant predictor was excluded from the analyses. Might our results showing that teacher perceptions predicted student achievement (even after controlling for students' previous achievement and motivation) reflect a spurious relation of both teacher perceptions and student achievement to some omitted variable? They might. In fact, the omitted-variable problem characterizes all nonexperimental studies. No matter how many control variables are included in a naturalistic study, it is always possible that a relevant one was omitted (see, e.g., Judd & McClelland, 1989; Pedhazur, 1982).

The accuracy alternative. With respect to understanding naturally occurring relations between perceiver expectations and target behavior, omitted-variable explanations are actually predictive accuracy explanations. *Predictive accuracy* is predictive validity without (self-fulfilling) influence (see Jussim, 1989,

1991, 1993; Jussim & Eccles, 1992, 1995a). Perceivers' expectations will be accurate when they predict targets' behavior because both those expectations and targets' behavior are "spuriously" related to some third variable. If such a third variable were omitted from our analyses, we might obtain significant coefficients linking teacher perceptions to students' future achievement—not because of self-fulfilling prophecies but because teachers may have been more accurate than we had given them credit for being. Thus, an omitted-variable explanation for the relations we observed between teacher perceptions and student achievement is an accuracy alternative to our self-fulfilling prophecy explanation.

General evaluation of the accuracy explanation. Several general aspects of our (and others') research argue against an accuracy interpretation of our results. First, a self-fulfilling prophecy interpretation of our findings is consistent with a long history of experimental findings demonstrating that perceiver expectations can influence target behavior (e.g., see reviews by Darley & Fazio, 1980; Miller & Turnbull, 1986; Snyder, 1984, 1992). Of course, just because experiments find self-fulfilling prophecies does not prove that our results reflect self-fulfilling prophecies. Nonetheless, confidence in the validity of a general conclusion (e.g., that perceiver expectations sometimes produce self-fulfilling prophecies) increases when naturalistic and experimental studies yield converging evidence (see also Jussim & Fleming, 1996).

Second, our study included more control variables than most other naturalistic investigations of teacher expectations (see Jussim & Eccles, 1995a, for a review). Rarely have studies controlled for student grades and standardized test scores plus student motivation. Of those that have controlled student motivation (e.g., Brattesani et al., 1984), few have operationalized student motivation as broadly as we did in this study. The extensive controls that we used reduce the likelihood that a relevant predictor was omitted from the regression analyses, thereby decreasing the probability that relations between teacher perceptions and students' future achievement were caused by uncontrolled factors.

Furthermore, because we controlled for previous achievement, significant coefficients relating teacher perceptions to students' future achievement mean that teacher perceptions predicted *changes* in student achievement (Cook & Campbell, 1979; Jussim, 1991). For an accuracy interpretation to be viable, our analyses would have had to omit some variable that not only related to student achievement but also predicted both teacher perceptions and changes in student achievement.

What might be some of the most likely contenders for such predictors? One possibility could be students' demographic group memberships. Abundant evidence attests to the existence of race, sex, and social class differences in achievement (e.g., Eccles [Parsons], Adler, & Meece, 1984; Eccles, Arberton, et al., 1993; Jussim et al., 1996; Marger, 1994). However, demographic group membership cannot account for our results. According to an accuracy explanation, students' demographic group membership must predict both teacher perceptions and students' future achievement independent of the control variables. Research based on the same data set (Jussim et al., 1996; Jussim & Eccles, 1995b) has shown that neither of these conditions were met. After controlling for the same set of background

tion of our main results. We also examined the possibility that the pattern of achievement moderation was due to greater variability in MEAP scores at the low end of the distribution. We performed an analysis that included the base model variables (Table 4) and the teacher perceptions of performance \times students' previous standardized test scores product term as predictors of MEAP squared (squaring the MEAP scores yielded a distribution close to normal). Results indicated that the teacher perceptions of performance \times previous standardized test scores product term still significantly predicted MEAP squared in this model, $\beta = -.08$, $p < .01$. The pattern of results was consistent with our previous findings. Teacher perceptions of performance predicted students' future achievement more strongly for low achievers than for high achievers.

variables that were included in the current study, race/ethnicity, sex, and social class had little or no influence on teacher perceptions and also did not significantly predict students' future achievement.

Moreover, additional regression analyses directly examined whether the inclusion of race/ethnicity, sex, and social class as controls changed the significant moderator effects that we observed. These analyses always included the base model variables (Table 4), students' race/ethnicity, sex, and social class, and the significant moderator effect (e.g., teacher perceptions of performance residuals squared) as predictors of MEAP scores. Results provided no evidence that race/ethnicity, sex, and social class changed the significant moderator effects we had initially found. Teacher overestimates still predicted MEAP scores significantly more strongly than teacher underestimates. Teacher over- and underestimates still predicted low achievers' MEAP scores significantly more strongly than high achievers' MEAP scores. High and low teacher perceptions still predicted low achievers' MEAP scores significantly more strongly than high achievers' MEAP scores. We conclude, therefore, that students' race/ethnicity, sex, and social class are not viable sources of accuracy that could explain our results. Other alternative explanations that might account for our specific patterns of results are considered next.

Other Alternative Explanations

Thus far, we have considered general arguments for and against an accuracy interpretation of our results. Next, we consider specific accuracy interpretations for our results showing that teacher perceptions predicted greater achievement changes among low achievers than among high achievers and that teacher overestimates predicted slightly greater achievement changes than did teacher underestimates. First, an accuracy interpretation would mean that one or more omitted variables led teachers to develop more accurate perceptions of low achievers than of high achievers. Second, it would mean that omitted variables led teachers to develop more accurate perceptions when those perceptions overestimated students than when they underestimated students. We consider both of these accuracy interpretations next.

Do teachers perceive low achievers more accurately? Perhaps teachers held more accurate perceptions of low achievers than they did of high achievers. Although possible, the empirical evidence points to the opposite conclusion. Research shows that, in comparison to lows, teachers interact more with highs, are friendlier to highs, prepare more for teaching highs, and provide highs with greater opportunities to learn and display knowledge (see reviews by Brophy, 1983; Brophy & Good, 1974; Jussim, 1986; Jussim, Madon, & Chatman, 1994; Rosenthal, 1973, 1989). There also is evidence that the more perceivers get to know targets, the more accurate perceivers become, and the less likely they are to create self-fulfilling prophecies (Fiske & Neuberg, 1990; Kenny, 1994; Raudenbush, 1984; Swann & Ely, 1984). This evidence strongly suggests that teachers most likely develop more accurate expectations for highs than they do for lows.

The alternative—that teachers develop more accurate expectations for low achievers (with whom they interact the least)—

seems highly implausible. This conclusion will warrant revision when new research empirically demonstrates that teachers actually spend more time with, and lavish more attention on, lows than on highs. Until then, we conclude that: (a) an accuracy interpretation of results showing that teacher perceptions predicted greater achievement changes among lows than among highs is not very plausible, and (b) the most viable interpretation of these results is that low achievers are more susceptible to self-fulfilling prophecies than are high achievers.

Are "overestimates" more accurate? Our other finding was the very weak tendency for teacher perceptions that overestimated students to predict greater achievement changes in comparison to teacher perceptions that underestimated students. Might this pattern reflect a tendency for overestimates to be more accurate than underestimates? Perhaps. Teachers may have been sensitive to some factor(s) that when present increased achievement more than could be predicted by students' backgrounds but when absent did not decrease achievement more than could be predicted by students' backgrounds.

However, to our knowledge, no empirical or theoretical article in any area of psychology or education has ever identified an attribute (intellectual, motivational, or familial) that when present increases achievement but when absent does not decrease achievement (consider, e.g., nutrition, intrinsic motivation, confidence, effort, studying, class attendance, supportive family environment, etc.). The only exception to this is the present study's finding regarding teacher expectations! Although we do not deny the possibility that some other such variable may someday be discovered, we conclude that until that time, the most plausible explanation for our results is that teacher overestimates had a very weak tendency to produce greater self-fulfilling prophecies than did teacher underestimates.

Lows' Susceptibility to Self-Fulfilling Prophecies

Analyses showed that teacher perceptions predicted achievement changes more strongly among low achievers than among high achievers. Moreover, this pattern was independent of any tendency for teacher overestimates to predict greater achievement changes than did teacher underestimates. Our results, therefore, are consistent with the second susceptibility hypothesis—that low achievers are more susceptible to self-fulfilling prophecies than are high achievers. The strongest self-fulfilling prophecy effects obtained in this study, and one of the strongest self-fulfilling prophecy effects ever obtained in naturalistic research ($\beta = .41$), occurred when teachers substantially overestimated low achievers.

Why This Greater Susceptibility?

Why might low achievers be more susceptible than high achievers to self-fulfilling prophecies? Although we have no direct empirical data bearing on this question, we next speculate on some possible reasons by addressing two related but separate subquestions: (a) Why might lows be more susceptible to negative self-fulfilling prophecies? and (b) Why might lows be more susceptible to positive self-fulfilling prophecies?

Why might lows be more susceptible to negative self-fulfilling

prophecies? School may be a less friendly place for lows than it is for highs. Low achievers may find schoolwork more difficult than do highs. Also, of course, low achievers are more likely than high achievers to be the targets of negative expectations. When school is consistently difficult and unpleasant, students may devalue the importance they place on it (Eccles & Midgley, 1989; Eccles, Arberton, et al., 1993; Jussim, 1986; Midgley et al., 1989; Steele, 1992; Steele & Aronson, 1995). Teachers often act on their negative expectations in ways that undermine students' motivation (see reviews by Cooper, 1979; Jussim, 1986). Therefore, by communicating their negative expectations to low achievers, teachers may undermine those students' motivation and achievement (see Jussim, 1986; Rosenthal, 1973, for reviews; see Harris & Rosenthal, 1985, for a meta-analysis). Highs, in contrast, by virtue of their ability to draw on a wealth of successful past experiences, and perhaps their genuinely greater ability, may have greater psychological resources available to combat negative expectations.

Why might low achievers be more susceptible than high achievers to positive self-fulfilling prophecies? Perhaps when teachers communicate positive expectations to students who have a history of negative school experiences, it feels like a breath of fresh air. Such a teacher may inspire lows to improve their school performance considerably. Steele (1992) described one such situation in his analysis of the sources of African American underachievement (also see Eccles & Midgley, 1989; Eccles, Arberton, et al., 1993; Midgley et al., 1989). Low-achieving minority students who were encouraged to perform difficult honors-level work came to perform as highly as their White classmates. Thus, lows' motivation may be dramatically enhanced by positive expectations. Although highs, too, may be motivated more by positive expectations than by negative expectations, they may not have as much room for improvement as do lows, possibly because they are already achieving at or close to their potential. Because our analysis here is speculative, however, an important area for future research will be to more precisely identify the factors that render low achievers, in comparison to high achievers, more susceptible to self-fulfilling prophecies.

Power of Positive Self-Fulfilling Prophecies

There was also a slight tendency for teacher overestimates to increase achievement more than teacher underestimates tended to decrease achievement among students in general. This pattern was statistically significant in the analysis that did not control for moderation by achievement but was only a nonsignificant trend in the analysis that did control for moderation by achievement (see Figure 4). These results indicate that there was a very weak tendency for positive self-fulfilling prophecies to be more powerful than negative self-fulfilling prophecies among students in general. Despite its weakness this pattern nonetheless has theoretical and practical importance.

Theoretical importance. First, this pattern supports predictions derived from self-enhancement theory. Self-enhancement theory proposes that people are motivated to view themselves favorably (e.g., S. C. Jones, 1973; Jussim et al., 1995; Sedikides, 1993; Shrauger, 1975). Self-enhancement theory suggests that targets, motivated to view themselves favorably, may attend

more closely to perceiver behaviors that convey positive expectations, and pay less attention to, and more readily discount, perceiver behaviors that convey negative expectations. Future research should directly assess some of the social and psychological processes that may explain why overestimates might produce slightly more powerful self-fulfilling prophecies than do underestimates.

Second, even a very weak tendency for positive self-fulfilling prophecies to be more powerful than negative self-fulfilling prophecies resoundingly disconfirms the golem hypothesis. This finding is itself important given that much of the concern about self-fulfilling prophecies as a social problem involves their potential to undermine targets' achievement (e.g., Brophy, 1983; Eccles & Wigfield, 1985; Jussim, 1990). The brief summary of Rosenthal and Jacobson's (1968b) classic Pygmalion study that appeared in *Scientific American* was titled "Teacher Expectations for the Disadvantaged," even though Rosenthal and Jacobson induced only positive expectations. The implicit assumption seemed to be, even back then, that negative expectations are at least as damaging as positive expectations are helpful. However, the present study has provided the first clear empirical test of this proposition and completely failed to support it. There was no evidence that teacher underestimates were more damaging than teacher overestimates were helpful.

Practical implications. This optimistic finding also has practical implications. First, it raises the possibility that, even when inaccurate, teacher expectations may more often increase than decrease student achievement. This suggests that teacher expectations may be more likely to contribute solutions to problems associated with scholastic underachievement than they are to create those problems in the first place. If such findings are obtained with different age groups, in different geographic areas, and in different contexts, it will mean that some of the concerns often expressed about the damaging effects of teachers' expectations for targets (e.g., Devine, 1995; E. E. Jones, 1986; Miller & Turnbull, 1986; Snyder, 1984) may be overstated. Instead, teachers may be more likely to uplift struggling students than to make matters worse by damaging them further.

Of course, this does not mean that self-fulfilling prophecies are never negative or never damaging. To the contrary, our results showed that negative self-fulfilling prophecies were indeed more powerful among lows than among highs. Thus, when expectations are negative they may have their most powerful effects among those who can least afford it.

Why Did Self-Concept in a Particular Achievement Domain Not Moderate Self-Fulfilling Prophecies?

Our results show that after controlling for moderation by achievement there was no evidence that self-concept in a particular achievement domain moderated self-fulfilling prophecies. Nonetheless, moderation by self-concept in a particular achievement domain may occur in ways that we did not examine. Self-verification theory (Swann, 1987; Swann & Ely, 1984) suggests that self-fulfilling prophecies may be stronger for targets who hold unclear self-concepts than they are for targets who hold clear self-concepts. Our measure of students' self-concepts of math ability, however, did not assess clarity. Therefore, self-concept in a particular achievement domain may moderate self-

fulfilling prophecies, but only among students whose self-concepts are unclear.

Self-Fulfilling Prophecies Among Stigmatized Groups

This study demonstrated that teachers' expectations led to powerful self-fulfilling prophecies among low achievers. Elsewhere (Jussim et al., 1996), we have demonstrated powerful self-fulfilling prophecies among students from stigmatized groups. Thus, there is mounting evidence that although self-fulfilling prophecies are small on average (Jussim, 1991; Jussim & Eccles, 1992), they can be especially powerful among certain targets. Sometimes a greater susceptibility to self-fulfilling prophecies can have beneficial consequences for targets. Expectations that substantially overestimated low achievers led to large gains in achievement.

However, we also suspect that some targets may encounter negative expectations in disproportionate amounts. For example, prejudice and stereotypes may lead perceivers to hold negative expectations for members of stigmatized groups far more often than positive expectations. Thus, even if members of stigmatized groups can benefit greatly from positive self-fulfilling prophecies, they may only rarely have the chance to reap those benefits. One way to increase the chance that members of stigmatized groups will benefit from self-fulfilling prophecies is by instituting policy changes that encourage perceivers (e.g., employers, teachers) to hold realistic, but high, expectations for targets (Eden, 1986). Our research suggests that under these conditions, members of stigmatized groups will be able to benefit considerably from the self-fulfilling effects of perceivers' expectations.

Conclusion

When do perceiver expectations create more powerful self-fulfilling prophecies? This study indicated that self-fulfilling prophecies were more powerful for low achievers. However, the self-fulfilling effects of teacher expectations on lows were usually small, typically below .30. This pattern of relatively small effect sizes is consistent with most naturalistic studies on self-fulfilling prophecies. Although throughout the 1980s and early 1990s social psychology abounded with testimonies to the power of expectations to create social reality (Fiske & Taylor, 1991; Hamilton, Sherman, & Ruvolo, 1990; E. E. Jones, 1986, 1990; Snyder, 1984; see Jussim, 1991, for a review), naturalistic studies have consistently failed to support such claims. Naturally occurring expectancy effects rarely exceed .1 to .2 in terms of standardized regression coefficients (see Jussim, 1991; Jussim & Eccles, 1995a, for reviews).

Nonetheless, this does not mean that naturally occurring self-fulfilling prophecies must be, or are always, small. Under certain conditions, and for certain types of targets, self-fulfilling prophecies may be more powerful than average effect sizes suggest. Consistent with this, we identified one condition under which self-fulfilling prophecies were considerably more powerful than usual: when teachers overestimated low achievers. Elsewhere we have demonstrated powerful self-fulfilling prophecies among African American students and students of lower socioeconomic status (Jussim et al., 1996). Thus, our research indicates that

although self-fulfilling prophecies are often small, there are certain conditions under which they are quite powerful.

References

- Abelson, R. P., & Lesser, G. S. (1966). The measurement of persuasibility in children. In C. I. Hovland & I. L. Janis (Eds.), *Personality and persuasibility* (pp. 141–166). New Haven, CT: Yale University Press.
- Babad, E. Y., Inbar, J., & Rosenthal, R. (1982). Pygmalion, galatea, and the golem: Investigations of biased and unbiased teachers. *Journal of Educational Psychology, 71*, 459–474.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.
- Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology, 76*, 236–247.
- Brockner, J. (1983). Low self-esteem and behavioral plasticity: Some implications. In L. Wheeler & P. Shaver (Eds.), *Review of personality and social psychology* (Vol. 4, pp. 237–271). Beverly Hills, CA: Sage.
- Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 70*, 631–661.
- Brophy, J., & Good, T. (1974). *Teacher–student relationships: Causes and consequences*. New York: Holt, Rinehart & Winston.
- Campbell, J. D., & Lavalley, L. (1993). Who am I? The role of self-concept confusion in understanding the behavior of people with low self-esteem. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 3–20). New York: Plenum.
- Cohen, A. R. (1966). Some implications of self-esteem for social influence. In C. I. Hovland & I. L. Janis (Eds.), *Personality and persuasibility* (pp. 102–120). New Haven, CT: Yale University Press.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Coleman, L. M., Jussim, L., & Abraham, J. (1987). Students' reactions to teacher evaluations: The unique impact of negative feedback. *Journal of Applied Social Psychology, 17*, 1051–1070.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton-Mifflin.
- Cooper, H. (1979). Pygmalion grows up: A model for teacher expectation communication, and performance influence. *Review of Educational Research, 49*, 389–410.
- Cooper, H., & Good, T. (1983). *Pygmalion grows up: Studies in the expectation communication process*. New York: Longman.
- Cooper, H., & Hazelrigg, P. (1988). Personality moderators of interpersonal expectancy effects: An integrative research review. *Journal of Personality and Social Psychology, 6*, 937–949.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35*, 867–881.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Devine, P. G. (1995). Prejudice and out-group perception. In A. Tesser (Ed.), *Advanced social psychology* (pp. 467–524). New York: McGraw-Hill.
- Eccles, J. S. (1988). *Achievement beliefs and the environment* (Report No. NICHD-PS-019736). Bethesda, MD: National Institute of Child Health and Development (National Institute of Health). (ERIC Documentation Reproduction Service No. ED 332 830)
- Eccles (Parsons), J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: W. H. Freeman.

- Eccles (Parsons), J., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, 46, 26–43.
- Eccles, J. S., Arberton, A., Buchanan, C. M., Jacobs, J., Flanagan, C., Harold, R., MacIver, D., Midgley, C., Reuman, D., & Wigfield, A. (1993). School and family effects on the ontogeny of children's interests, self-perceptions, and activity choices. In R. Dienstbier & J. E. Jacobs (Eds.), *Nebraska Symposium on Motivation: Developmental perspectives on motivation* (Vol. 40, pp. 145–208). Lincoln: University of Nebraska Press.
- Eccles, J. S., & Midgley, C. (1989). Stage/environment fit: Developmentally appropriate classrooms for early adolescents. In R. E. Ames & C. Ames (Eds.), *Research on motivation in education: Goals and cognitions* (Vol. 3, pp. 139–186). New York: Academic Press.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage–environment fit on young adolescents' experiences in schools and families. *American Psychologist*, 48, 90–101.
- Eccles, J. S., & Wigfield, A. (1985). Teacher expectations and student motivation. In J. Dusek (Ed.), *Teacher expectancies* (pp. 185–226). Hillsdale, NJ: Erlbaum.
- Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D., & Yee, D. (1989). Self-concepts, domain values, and self-esteem: Relations and changes at early adolescence. *Journal of Personality*, 57, 283–310.
- Eccles-Parsons, J. S., Kaczala, C. M., & Meece, J. L. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development*, 53, 322–339.
- Eden, D. (1986). OD and self-fulfilling prophecy: Boosting productivity by raising expectations. *Journal of Applied Behavioral Science*, 22, 1–13.
- Elashoff, J. D., & Snow, R. E. (1971). *Pygmalion reconsidered*. Worthington, OH: Charles A. Jones.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.
- Fiske, S. T., & Taylor, S. E. (1984). *Social cognition*. Reading, MA: Addison-Wesley.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652–670.
- Gordon, R. A. (1968). Issues in multiple regression. *American Journal of Sociology*, 73, 592–616.
- Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, 46, 35–60.
- Harris, M. J. (1989). Personality moderators of expectancy effects: Replication of Harris and Rosenthal (1986). *Journal of Research in Personality*, 23, 381–387.
- Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, 97, 363–386.
- Harter, S. (1982). The perceived competence scale for children. *Child Development*, 53, 87–97.
- Janis, I. L. (1954). Personality correlates of susceptibility to persuasion. *Journal of Personality*, 22, 504–518.
- Jones, E. E. (1986, October 6). Interpreting interpersonal behavior: The effects of expectancies. *Science*, 234, 41–46.
- Jones, E. E. (1990). *Interpersonal perception*. New York: W. H. Freeman.
- Jones, R. A. (1977). *Self-fulfilling prophecies: Social, psychological, and physiological effects of expectancies*. Hillsdale, NJ: Erlbaum.
- Jones, S. C. (1973). Self and interpersonal evaluations: Esteem theories versus consistency theories. *Psychological Bulletin*, 79, 185–199.
- Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model-comparison approach*. Orlando, FL: Harcourt Brace Jovanovich.
- Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100, 109–128.
- Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review*, 93, 429–445.
- Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, 57, 469–480.
- Jussim, L. (1990). Social reality and social problems: The role of expectancies. *Journal of Social Issues*, 46, 9–34.
- Jussim, L. (1991). Social perception and social reality: A reflection–construction model. *Psychological Review*, 98, 54–73.
- Jussim, L. (1993). Accuracy in interpersonal expectations: A reflection–construction analysis of current and classic research. *Journal of Personality*, 61, 637–668.
- Jussim, L., & Eccles, J. (1992). Teacher expectations II: Construction and reflection of student achievement. *Journal of Personality and Social Psychology*, 63, 947–961.
- Jussim, L., & Eccles, J. (1995a). Naturally occurring interpersonal expectancies. *Review of Personality and Social Psychology*, 15, 74–108.
- Jussim, L., & Eccles, J. (1995b). Are teacher expectations biased by students' gender, social class, or ethnicity? In Y. T. Lee, L. Jussim, & C. R. McCauley (Eds.), *Stereotype accuracy: Toward appreciating group differences* (pp. 245–271). Washington, DC: American Psychological Association.
- Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 281–388). San Diego, CA: Academic Press.
- Jussim, L., & Fleming, C. (1996). Self-fulfilling prophecies and the maintenance of social stereotypes: The role of dyadic interactions and social forces. In N. C. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 161–192). New York: Guilford Press.
- Jussim, L., Madon, S., & Chatman, C. (1994). Teacher expectations and student achievement: Self-fulfilling prophecies, biases and accuracy. In L. Heath, R. S. Tindale, J. Edwards, E. J. Posavac, F. B. Bryant, E. Henderson-King, Y. Suarez-Balcazar, & J. Myers (Eds.), *Applications of heuristics and biases to social issues* (pp. 303–334). New York: Plenum.
- Jussim, L., Soffin, S., Brown, R., Ley, J., & Kolhepp, K. (1992). Understanding reactions to feedback by integrating ideas from symbolic interactionism and cognitive evaluation theory. *Journal of Personality and Social Psychology*, 62, 402–421.
- Jussim, L., Yen, H., & Aiello, J. (1995). Self-consistency, self-enhancement, and accuracy in reactions to feedback. *Journal of Experimental and Social Psychology*, 31, 322–356.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136–153.
- Kanouse, D. E., & Hanson, R., Jr. (1971). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 41–62). Morristown, NJ: General Learning Press.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford Press.
- Leventhal, H., & Perloe, S. (1962). A relationship between self-esteem

- and persuasibility. *Journal of Abnormal and Social Psychology*, 64, 385–388.
- Marger, M. N. (1994). *Race and ethnic relations* (3rd ed.). Belmont, CA: Wadsworth.
- Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review*, 8, 193–210.
- Midgley, C. M., Feldlaufer, H., & Eccles, J. S. (1989). Changes in teacher efficacy and student self- and task-related beliefs during the transition to junior high school. *Journal of Educational Psychology*, 81, 247–258.
- Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 37, pp. 233–256). Stanford, CA: Annual Review.
- Neuberg, S. L. (1994). Expectancy-confirmation processes in stereotype-tinged social encounters: The moderating role of social goals. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario Symposium* (Vol. 7, pp. 103–130). Hillsdale, NJ: Erlbaum.
- Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.
- Pelham, B. W. (1991). On confidence and consequence: The certainty and importance of self-knowledge. *Journal of Personality and Social Psychology*, 60, 518–530.
- Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy inductions: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, 76, 85–97.
- Rosenthal, R. (1973). *On the social psychology of the self-fulfilling prophecy: Further evidence for Pygmalion effects and their mediating mechanisms*. New York: MSS Modular.
- Rosenthal, R. (1989, August). *Experimenter expectancy, covert communication, and meta-analytic methods*. Invited address at the 97th Annual Convention of the American Psychological Association, New Orleans, LA.
- Rosenthal, R., & Jacobson, L. (1968a). *Pygmalion in the classroom: Teacher expectations and student intellectual development*. New York: Holt, Rinehart & Winston.
- Rosenthal, R., & Jacobson, L. (1968b). Teacher expectations for the disadvantaged. *Scientific American*, 218, 19–23.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, 3, 377–386.
- Sedikides, C. (1993). Assessment, enhancement, and verification determinants of the self-evaluation process. *Journal of Personality and Social Psychology*, 65, 317–338.
- Shrauger, S. J. (1975). Responses to evaluation as a function of initial self-perceptions. *Psychological Bulletin*, 82, 581–596.
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105, 131–142.
- Snyder, M. (1984). When belief creates reality. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 247–305). Orlando, FL: Academic Press.
- Snyder, M. (1992). Motivational foundations of behavioral confirmation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 67–114). San Diego, CA: Academic Press.
- Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 35, 656–666.
- Steele, C. M. (1992, April). Race and the schooling of Black Americans. *Atlantic Monthly*, 68–78.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811.
- Swann, W. B., Jr. (1987). Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology*, 53, 1038–1051.
- Swann, W. B., Jr., & Ely, R. J. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology*, 46, 1287–1302.
- Swann, W. B., Jr., Griffin, J. J., Jr., Predmore, S. C., & Gaines, B. (1987). The cognitive affective crossfire: When self-consistency confronts self-enhancement. *Journal of Personality and Social Psychology*, 52, 881–889.
- Swann, W. B., Jr., Pelham, B. W., & Krull, D. S. (1989). Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *Journal of Personality and Social Psychology*, 57, 782–791.
- Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization–minimization hypothesis. *Psychological Bulletin*, 110, 67–85.
- Vannoy, M., Smith, A., Jussim, L., & Eccles, J. (1996, June). *Teacher expectations and student achievement: Ability grouping and class level dynamics*. Poster presented at the Convention of the American Psychological Society, San Francisco.
- West, C., & Anderson, T. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research*, 46, 613–630.

Received February 7, 1996

Revision received November 14, 1996

Accepted November 21, 1996 ■