

BOUNDARY CONDITIONS OF THE GALATEA EFFECT: A FIELD EXPERIMENT AND CONSTRUCTIVE REPLICATION

D. BRIAN McNATT
University of Georgia

TIMOTHY A. JUDGE
University of Florida

In a longitudinal field experiment, we examined the generalizability of the “Galatea effect” to businesses. This constructive replication provides the first test of the effect of a nonfictitious, nontraining intervention on actual job performance of new and established professionals. To examine duration, we measured the effect on performance until it disappeared for the sample as a whole. The Galatea intervention bolstered self-efficacy, motivation, and performance, though this latter effect was temporary. We also partially confirmed the “self-fulfilling prophecy at work model” and suggest possible extensions.

Almost 40 years ago, Rosenthal and Jacobson (1966) conducted the first experimental demonstration of the “Pygmalion effect,” which occurs when a person’s high expectations of another person result in high performance for that person. A related self-fulfilling prophecy phenomenon, *the Galatea effect*, occurs when high self-expectations lead to high performance. Since that time, hundreds of studies and scores of meta-analyses have shown that self-fulfilling prophecy interventions shape expectations, which, in turn, influence outcomes (Rosenthal, 2002). With such overwhelming evidence, why have businesses not embraced self-fulfilling prophecy methods to increase employee performance? The key reasons are likely several critical boundary conditions or limitations of previous research. Therefore, researchers have advised that those conducting future self-fulfilling prophecy research seek to understand these boundary conditions, their effects, and how they might be overcome (Eden et al., 2000; White & Locke, 2000).

One potential boundary condition is that experimental interventions have relied on false information. Participants have been told that they possessed more capability than others had when they did not, or that their exam scores identified them as having high potential when in fact the scores did not do so (e.g., Eden & Kinnar, 1991). It is true that

the people to or about whom high expectations were communicated tended to achieve greater gains. Still, to gain acceptance as a management tool, self-fulfilling prophecy interventions need to be shown to be effective without use of deception (Eden, 1990; White & Locke, 2000).

A second boundary condition of interest is that no study to date has tested the effectiveness of nontraining Galatea interventions on individuals with experience in a given study setting. Relatedly, Pygmalion studies have shown that expectation effects diminish or disappear when prior relationships exist between individuals and their instructors (Raudenbush, 1984). Since experience with tasks and prior contact among people are the norm within organizations, this is a severe limitation if self-fulfilling prophecy effects only occur in the absence of prior relationships (White & Locke, 2000).

A third boundary condition is that the vast majority of expectation-changing research has been conducted in clinical, educational, or military settings and not within businesses (Gist & Mitchell, 1992; Saks, 1995). To show generalizability, research must be conducted within businesses. Fourth, such research has also been confined mostly to training sessions that measure learning as the dependent variable (e.g., Eden & Ravid, 1982; Gist, 1989). Because learning is distinct from job performance, researchers have called for studies to test self-fulfilling prophecy effects on the latter (e.g., Gist & Mitchell, 1992; McNatt, 2000). Fifth, management researchers have indicated that effects may wear off over time, and so it is important to discover how long effects might last (Mitchell & James, 2001). Unfortunately, very few of the studies

This article is based on Brian McNatt’s dissertation. A previous version was presented at the 2002 Society for Industrial and Organizational Psychology conference. We acknowledge and are extremely grateful to the accounting firm that participated and to the associate editor and the three anonymous reviewers for their insightful and helpful recommendations.

testing self-fulfilling prophecy interventions have been longitudinal; an exception is Eden and Ravid (1982), which lasted seven weeks. To understand the possible duration of self-fulfilling prophecy intervention effects in, and thus the possible value to, businesses, researchers have called for more longitudinal research (e.g., Saks, 1995).

To address some of these limitations, Eden and colleagues conducted a group of seven studies to test the effectiveness of training in the "Pygmalion leadership style" (Eden, 1992) in producing Pygmalion effects (Eden et al., 2000). Interestingly, the training did not improve leaders' expectations for or behaviors toward their followers, or the followers' self-expectations or performance. In his *self-fulfilling prophecy at work model*, Eden (1990) proposed that followers' self-expectations and motivation mediate the impact of leader expectations and behaviors on follower performance. Thus, given these results and the limitations of previous research indicated above, a key direction of self-fulfilling prophecy research should be to learn how managers might enhance the self-expectations of employees in an attempt to create the Galatea effect (White & Locke, 2000). Likewise, Gist and Mitchell (1992) identified the malleability of self-expectations as one of the two most important topics for future self-expectation research. In addition, such research could help explain how self-fulfilling prophecy effects occur, especially since Eden's (1990) self-fulfilling prophecy model has received sparse testing. The only test using Baron and Kenny's (1986) procedure was published just recently, by Davidson and Eden (2000). They deemed their results inconclusive and called for additional exploration of the model. Particularly pressing is the need to study motivation as a mediator of the effect of self-expectations on performance.

Accordingly, our purposes for the present experimental study were (1) to conduct a constructive replication of the Galatea effect, (2) to simultaneously examine five previously untested boundary conditions regarding the generalizability and potential limitations of self-fulfilling prophecy interventions, and (3) to analyze the potential mediating impact of motivation as prescribed in the self-fulfilling prophecy at work model. Thus, our intent was not to test a new theory, but to conduct a novel replication. Eden (2002) argued the value and importance of constructive replications (replications that are as dissimilar as possible from previous studies). As indicated, our experiment is dissimilar to and extends previous research by examining several previously untested crucial boundary conditions. Specifically, we tested a nonfictitious, nontraining Galatea intervention; our sample included both new and established professionals; we used a business setting for

a Galatea intervention; we examined the Galatea intervention's impact on actual job performance; and we measured the effect on performance until it disappeared for the sample as a whole. We also provided needed testing of the self-fulfilling prophecy at work model.

THEORY, CONSTRUCTS, AND HYPOTHESES

The Galatea Effect

When high self-expectations lead to high performance, the Galatea effect is said to occur. There are several conceptual formulations of self-expectations of performance; these include specific self-efficacy, self-confidence, and performance expectancy. Although somewhat distinct conceptually, such constructs are highly correlated and have similar correlations with other constructs (Eden, 1990; Locke, Frederick, Lee, & Bobko, 1984). For a discussion of the similarities and differences between specific self-efficacy and expectancy, see Eden (1988) or Gist (1987). We chose to conceptualize self-expectations as specific self-efficacy because it is consistent with the self-fulfilling prophecy at work model and is the construct self-fulfilling prophecy researchers most often use (Eden, 1990; Eden & Kinnar, 1991). *Specific self-efficacy (SSE)* has been defined as "people's judgments of their abilities . . . to attain designated types of performances" (Bandura, 1986: 391). The two sources of specific self-efficacy that are most easily manipulated in nontraining interventions are vicarious experience/modeling and verbal persuasion (Bandura, 1997). Although related theories contain no distinction between accurate and deceptive information, what seems most important is that a given intervention be credible and powerful. As such, researchers have suggested ways that these sources can be most effective in raising specific self-efficacy (e.g., Bandura, 1997; Gist & Mitchell, 1992). We based our nondeceptive, nontraining intervention on these suggestions, seeking to raise participants' specific self-efficacy.

According to self-fulfilling prophecy theory (Merton, 1948), people's behaviors are consistent with their expectations, and those behaviors in turn influence outcomes (that is, people's actions help fulfill their own prophecies/expectations). Similarly, self-efficacy theory (Bandura, 1982) states that people gain information from a variety of experiences and sources to formulate judgments of their ability to perform tasks at given levels. Such judgments are said to then influence the amount of effort and persistence people expend, which ultimately impacts their obtained levels of perfor-

mance. This description of motivation is most similar to those in cognitive choice theories (Kanfer, 1990), wherein motivation is conceptualized in terms of choices to expend given amounts of effort, to allocate time and energy, and to persist (Campbell & Pritchard, 1976; Naylor, Pritchard, & Ilgen, 1980). Thus, we defined motivation as the amount of effort and persistence individuals are willing to expend. Again, it is rational that if people believe they can accomplish something, their actions will follow those beliefs, and these actions will increase the likelihood that the beliefs will be realized. Specifically, then, people can become the prophets of their own performance, owing to their self-expectations or specific self-efficacy. In our experiment, we were interested in job performance, defined as the extent to which employees successfully com-

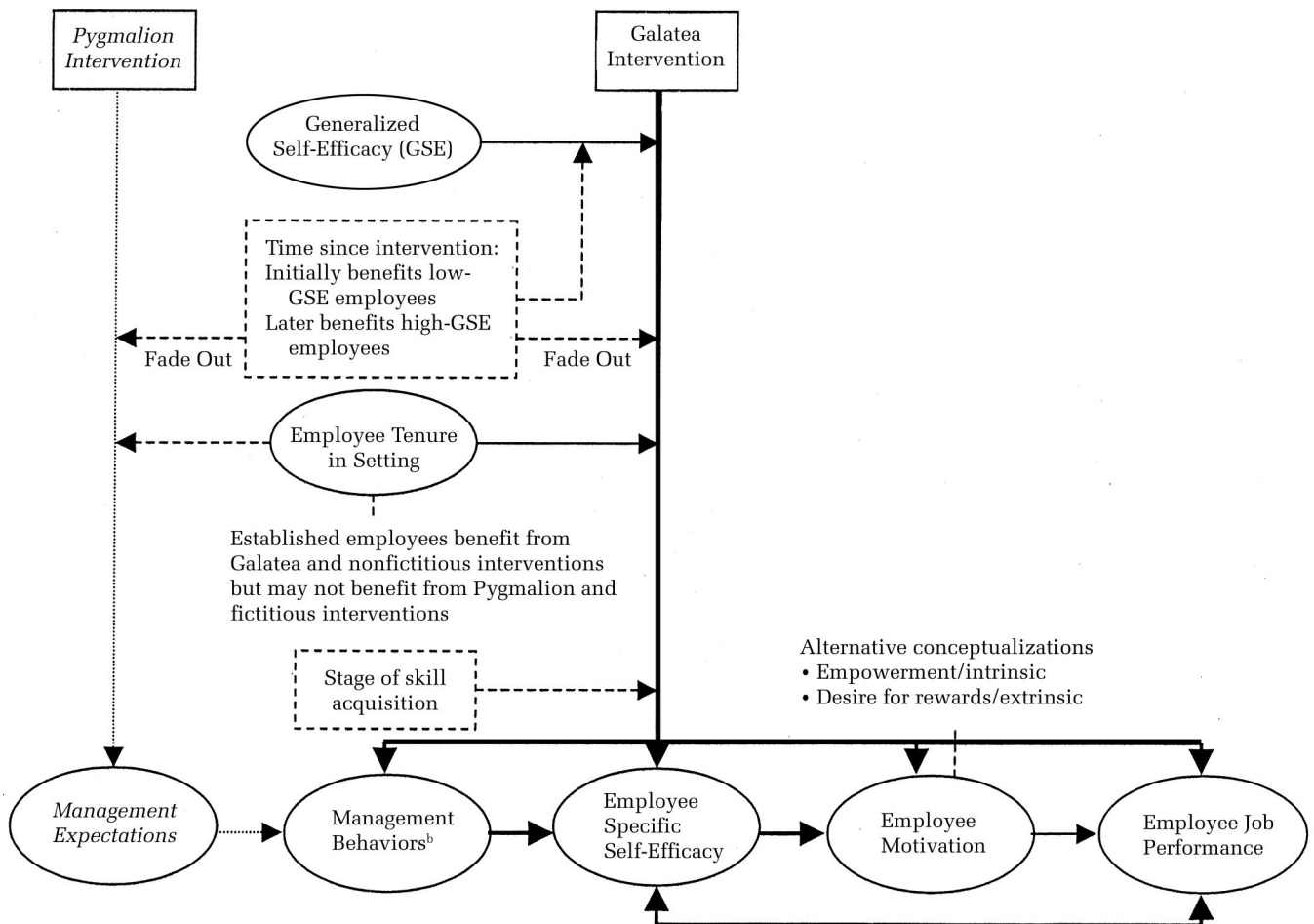
plete expected aspects of their jobs (McNatt, 2000). To test the Galatea effect, we sought to raise individuals' specific self-efficacy to examine whether such an increase would be related to higher motivation and performance. Our hypotheses were:

Hypothesis 1. Boosting specific self-efficacy improves performance.

Hypothesis 2. Boosting specific self-efficacy improves motivation.

In testing Hypothesis 1, we sought to replicate Eden and Zuk (1995). In Figure 1, a model of self-fulfilling prophecy relationships, we use solid lines to graphically depict the first two hypotheses as well as the other relationships tested.

FIGURE 1
Model of Self-Fulfilling Prophecy Relationships^a



^a Solid lines indicate tested relationships. Bold lines indicate statistical significance. Dotted lines indicate relationships proposed on the basis of results of this experiment. Italicized text and dotted lines indicate the parts of the self-fulfilling prophecy model that were not tested.

^b In this experiment, management behaviors were not measured but were manipulated as part of the Galatea intervention. Managers sent letters to enhance the specific self-efficacy of the experimental group participants.

Other Hypotheses

The self-fulfilling prophecy at work model. The self-fulfilling prophecy at work model is based on the self-fulfilling prophecy and self-efficacy theories explained above and indicates how managers' and employees' beliefs may be realized (Eden, 1990). Specifically, when managers and others communicate high expectations to employees, this communication becomes part of the information the employees use to estimate their performance expectations (or specific self-efficacy). With their expectations thus raised, they believe that they can achieve at higher levels, so they are motivated to exert greater effort and to persist. Higher levels of effort and persistence are then posited to result in greater job performance. Thus, in keeping with the self-fulfilling prophecy at work model, we hypothesized:

Hypothesis 3. Motivation mediates the effect of specific self-efficacy on performance.

Generalized self-efficacy. Researchers have indicated that traits likely play a key role in moderating how individuals respond to stimuli (e.g., Weiss & Adler, 1984). This moderation of response occurs because predispositions influence how a person perceives and thus acts upon experiences. *Generalized self-efficacy (GSE)* is a traitlike dimension defined as "individuals' perception of their ability to perform across a variety of different situations" (Judge, Erez, & Bono, 1998: 170). Drawing on Brockner's (1988) behavioral plasticity theory, researchers have examined generalized self-efficacy as a potential moderator of interventions involving specific self-efficacy (e.g., Eden & Kinnar, 1991). Brockner contended that people with low self-esteem are more susceptible to external influence (from experimenters and managers, for example) than those with high self-esteem. Although conceptually distinct, generalized self-efficacy has been strongly related to self-esteem (Judge, Bono, & Locke, 2000) and thus "can serve as a proxy . . . to test the plasticity hypothesis" (Eden & Aviram, 1993: 353). The rationale is that since persons with low generalized self-efficacy are less sure and less confident, they are more open to and influenced by the ideas and opinions of others—especially those in positions of authority or expertise. Those with low generalized self-efficacy may discount their ability to formulate accurate opinions of their capabilities. Thus, they may see external sources as more credible and rely on such sources to inform them. Those with higher generalized self-efficacy are more likely to be self-assured and independent and to look inward for confirmation, and they may

be more likely to discount external sources, particularly if the messages are inconsistent with their own opinions. If those lower in generalized self-efficacy are more open to influence, they could be affected more by experimental "treatments" that boost specific self-efficacy. Researchers have obtained empirical results indicating behavioral plasticity, although these have not always been statistically significant (Eden & Aviram, 1993; Eden & Zuk, 1995). Several factors that might have contributed to the inconsistent results include the small number of studies, poor generalized self-efficacy measures, and possibly small true effect sizes. Given those limits, and Eden and Kinnar's (1991) partial support of the plasticity theory, we hypothesized:

Hypothesis 4. Generalized self-efficacy moderates the effectiveness of specific self-efficacy treatments in that they are more effective among people low in generalized self-efficacy.

Amount of tenure within context. Researchers have indicated that expectation effects may occur only with people who are young, new to a situation, or engaged in novel tasks (Eden & Zuk, 1995; Saks, 1995). Saks reasoned that such individuals may be more sensitive to training or other intervention experiences, and their specific self-efficacy may be more malleable than that of older or more experienced workers. On the other hand, established workers have more data upon which to base their estimates of self-efficacy, owing to performance feedback and greater knowledge of work tasks, and thus their specific self-efficacy is likely more fixed. Further, inasmuch as impressions or relationships have formed, others' perceptions of established workers may also be more rigid. For example, supervisors may have already formed impressions and thus be impervious to disconfirming interventions or employee behaviors. Thus, as people gain tenure in a given setting, it may become difficult for self-fulfilling prophecy interventions to have an impact. Therefore, we hypothesized:

Hypothesis 5. Tenure in an organization moderates the effectiveness of specific self-efficacy treatments in that they are more effective among people with less tenure.

Gender. Finally, self-fulfilling prophecy experiments have highlighted gender as a possible moderator of the Pygmalion effect, as some studies have found an effect with women participants and others have not. Dvir, Eden, and Banjo (1995) theorized that men and women may have different leadership styles that may then account for differences in self-fulfilling prophecy effects. McNatt's (2000)

meta-analysis of Pygmalion experiments found an effect for both genders; however, the effect was stronger in studies with all men than in those with all women—although the difference was not significant. Given these results and that such analyses have not been done with Galatea studies, we examined gender as a potential moderator. Since this was a side issue, we made no formal hypothesis.

METHODS

Sample, Design, and Measures

The participants were 72 auditors within three offices of a Big Four accounting firm. Forty-five were newcomers (beginning auditors just hired), and 27 were in their second year. Women comprised 52 percent of the sample, and the mean sample age was 24 years. The study was a true field experiment involving three major data collection points—prestudy (time 1); postintervention, at the end of the same day (time 2); and postintervention, three months later (time 3)—as well as work performance measures collected throughout the three months. Participants were randomly assigned to treatment and control “conditions.”

Unless otherwise indicated, a seven-point response scale ranging from 1, “strongly disagree,” to 7, “strongly agree,” was used to measure self-reported items in the experiment. Saks (1995) designed a *task-specific self-efficacy* measure of the dimensions of a staff-level financial auditor’s job. We eliminated items referring to review engagements, an aspect of auditing that did not occur in this experiment, obtaining a 37-item measure. A ten-point response scale (Eden & Kinnar, 1991) ranging from 1, “I completely lack the required ability,” to 10, “I have the ability to do extremely well,” measured participants’ estimates of their ability to complete tasks such as “perform and document procedures to audit cash” ($\alpha = .93$, time 1; $.96$, time 2; and $.97$, time 3). We assessed *generalized self-efficacy* only at time 1, since it is considered a trait, using 12 items from Sherer et al.’s (1982) GSE scale. A sample item is “When I set important goals for myself, I rarely achieve them” ($\alpha = .84$). To keep the survey length manageable, we eliminated the 5 items for which Sherer and colleagues reported the lowest factor “loadings.” *Motivation* was measured at all three times with a 7-item scale based on Davidson and Eden’s (2000) 16-item measure ($\alpha = .94$). A sample item is “I am willing to put in a great deal of effort beyond that normally expected” ($\alpha = .81$, time 1; $.73$, time 2; and $.78$, time 3). Additionally, management pro-

vided us with auditors’ start dates, which we used to calculate *tenure with firm*.

To measure *job performance*, we obtained copies of the evaluations completed at the end of audit engagements—which vary in duration from a week to several months—for the experimental period. These were six item/category general evaluations used with auditors at all levels. Second, we designed a seven-item appraisal relevant to staff auditors and to be completed more frequently. An example item is “This employee completed all audit or work tasks in a very effective manner.” These evaluations were e-mailed to supervisors weekly. Because of the intermittent timing of the standard evaluations and weekly variation in response rate, we combined and averaged the evaluation information from the two sources for each month and over the three months of the experiment (average $\alpha = .93$).

This study also included a measure of *cognitive ability*. A critical point of self-fulfilling prophecy and self-efficacy theories is that specific self-efficacy can have an impact on performance that goes beyond the impacts of ability or skill. However, many studies testing these theories have increased their participants’ specific self-efficacy with skills training (e.g., Gist, Schwoerer, & Rosen, 1989). When this is done, it is unclear whether the resultant performance is a consequence of the increased specific self-efficacy or the increased skill. Thus, we used a nontraining intervention and included American College Testing (ACT) or Scholastic Aptitude Test (SAT) scores as a measure of cognitive ability. We obtained these scores and grade point averages (GPAs) by permission from the participants’ universities. These tests are reasonable measures of cognitive or general mental ability (Schmidt, 1988), are commonly used (e.g., Gully, Payne, Coles, & Whiteman, 2002), and show stability (Stumpf & Stanley, 1998). To have a common metric, we used the ACT conversion table with SAT scores. The test scores had a correlation of $.32$ ($p \leq .05$) with participants’ GPAs.

Experimental Intervention

The nature and content of the intervention for the present experiment were based upon the theory, past empirical work, and recommendations in the self-fulfilling prophecy literature (Bandura, 1997; Eden, 1990; Eden & Zuk, 1995; Gist & Mitchell, 1992). For example, White and Locke (2000) recommended that interventions consist of specific manager behaviors as well as experimenters or other third parties communication of intervention information. Thus, in an attempt to raise auditors’

specific self-efficacy, we used scripted verbal persuasion in the initial intervention—an interview with the first author—combined with modeling in each of three “boosters” in the form of researcher-drafted letters to participants sent by and under the signatures of our contact manager and partner in each office of the firm. So that these communications would be based upon true information, we obtained copies of the auditors’ résumés to identify strengths and experiences and interviewed both the firm’s human resource director and one of the managers and partners noted above. Interviews concerned the firm’s hiring procedures and criteria, factors found to contribute to success at this firm, and suggestions for their employees.

Drawing on this information and past researchers’ theory and suggestions, we stressed the following in our interview text and letters: (1) informing employees of the competitive selection process and how they were hired from among the many who applied because they were extremely well qualified; (2) indicating that they possessed the qualifications and skills to be successful, that experience had shown that individuals of their caliber had excelled within their firm, and that management had the same expectation (belief) for them; (3) reminding them of their past successes, framing these as due to their ability and effort, and indicating how those experiences would help them be successful in their present jobs; (4) indicating the importance of their taking charge of their careers and that though the jobs would require hard work, they had the ability, and tasks were more doable than they might think; and (5) informing them of management’s support, commitment, and desire to see them succeed. In the letters, we framed the managers as models by stressing their similarity with the auditors and by generally referring to their challenges and successes. To evaluate the content and delivery of the intervention for sincerity and persuasiveness, we engaged in practice interviews with a colleague and had our contact partner and academic peers review the letters. We based iterative changes to each part of the intervention on this feedback.

Procedures

One week before the experimental treatment, we sent packets to the auditors’ homes containing the first survey and letters from our contact partner and from us describing our research as gathering data about personality and job attitudes. Our letter framed the interviews as a way to get to know them better and to obtain answers to open-ended questions. Seventy auditors responded (97%) and were

then randomly assigned to conditions. The firm sent a different letter to the homes of supervisor auditors, informing them of the firm’s participation in a research project related to employee evaluations and of the weekly e-mail evaluations that they would complete. The interviews were conducted at the firm’s offices. Working from a word-for-word script, the interviewer (the first author) introduced himself and established credibility (that is, accounting degrees, CPA certification, and auditor experience). He then asked them about (1) important achievements, attributing these to their ability and effort, and (2) what they worried about related to their jobs, underscoring that they had the needed qualifications and aptitudes. He then communicated the script summarized above. In contrast, control group participants received no feedback but were asked additional questions about themselves and their current jobs. All interviews lasted approximately 15 minutes. Participants completed the postintervention (time 2) survey later that day. Seventy-one participants returned the second survey (99%). As staff auditors, all of the participants worked at different client locations apart from other participants. Contact partners from two offices of the firm decided that pulling all involved employees from their jobs would be too difficult. Therefore, we sent the second survey and a follow-up letter to control group participants from these offices informing them that, because of scheduling difficulties, they would not need to come into the office to meet with us. We found no significant difference between control group participants receiving interviews and those who did not; details are in “Results.”

Auditors in the experimental group later received a letter and an e-mail from the recruiting manager (in weeks 3 and 9) and a letter from the participating partner (in week 6). Although we did not draft letters to be sent to the control group, all employees including the control group did receive “informational mail” from management during the time of the experiment. The time 3 survey was mailed to participants three months after the interviews. Sixty-two of these were completed (86%). Over the course of the study, we took steps to maintain the integrity of the experimental design. Only the contact partner and manager at each office were informed of the nature of the project—and they were completely removed from supervisors and the participants; we never used the terms “experiment,” “treatment group,” or “control group” with anyone; communications to participants and supervisors masked the intent of the project; all letters were sent to the participant’s residences; there was very little opportunity for contact among

participants—they mostly worked in different locations from one another; and the newcomers (63%) essentially did not know other participants. Finally, our communications with the firm throughout and at the end of the experiment indicated that the integrity of our design had been maintained.

RESULTS

To assure the discriminant validity among the self-report constructs, we specified a confirmatory factor analysis model using three “parcels” per construct (Schallow, 2000). This factor structure fitted the data reasonably well ($\chi^2 = 65.16, df = 24$; root-mean-square residual [RMSR] = .07; comparative fit index [CFI] = .91; incremental fit index [IFI] = .91). Furthermore, this model fitted the data much better than alternative models that combined any two of the constructs ($\chi^2 = 118.63, df = 26$; RMSR = .12; CFI = .79; IFI = .79). Thus, results suggested that these were indeed separate constructs. Next, to assure preexperimental similarity of treatment and control participants, we regressed all of the time 1 variables on the treatment. None of the effects was significant. Table 1 provides descriptive statistics, reliabilities, and correlations among the primary study variables, and Table 2 lists the means and standard deviations by experimental and control conditions for the manipulation checks and dependent variables.

Manipulation Checks and Hypothesis Tests

For manipulation checks, we first regressed the independent variable, specific self-efficacy, on the treatment. The regression results—see Table 3—

TABLE 2
Means and Standard Deviations for Galatea Effect Variables^{a, b}

Variables	Experimental		Control	
	Mean	s.d.	Mean	s.d.
Specific self-efficacy, time 2	8.17	1.01	7.73	1.03
Specific self-efficacy, time 3	8.25	1.22	7.81	0.97
Motivation, time 2	6.14	0.67	5.80	0.67
Motivation, time 3	6.02	0.72	5.60	0.85
Overall job performance	5.45	0.54	5.30	0.77
Job performance, month 1	5.57	0.65	5.20	0.88
Job performance, month 2	5.39	0.73	5.34	0.87
Job performance, month 3	5.47	0.78	5.44	0.98

^a Specific self-efficacy was measured with a ten-point response scale; the other variables were measured with a seven-point response scale.

^b $n = 29-32$ for the experimental group auditors and $31-39$ for the control group auditors.

confirmed that the intervention raised the auditors’ specific self-efficacy as intended ($\beta_{t2} = .21, p \leq .05$) and that this effect lasted at least three months ($\beta_{t3} = .19, p \leq .05$). These are close to medium-sized effects ($d_{t2} = .43; d_{t3} = .40$). We then tested whether the treatment influenced specific self-efficacy differently, depending on participants’ generalized self-efficacy or tenure at the firm. Table 3 shows that the treatment did not interact with generalized self-efficacy or tenure at time 2; however, we did find significant interactions at time 3 for generalized self-efficacy ($\beta = .18, p \leq .05$) and tenure ($\beta = .18, p \leq .05$). To detect the nature of these interactions, we analyzed the impact of the treatment for participants low and high (± 1 s.d.) on the proposed moderators (Aiken & West, 1991). The

TABLE 1
Descriptive Statistics, Scale Reliabilities, and Correlations^a

Variables	Mean	s.d.	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Treatment	0.44	0.50													
2. Generalized self-efficacy	5.97	0.61	-.04	.84											
3. Ability ^b	25.35	3.32	.05	-.37	.92										
4. Tenure with firm	3.97	4.79	.04	-.18	.19										
5. Initial self-efficacy	7.74	1.20	.00	.25	.07	.23	.93								
6. Specific self-efficacy, time 2	7.92	1.02	.21	.14	.13	.24	.83	.96							
7. Specific self-efficacy, time 3	8.01	1.10	.20	.27	.07	.23	.67	.71	.97						
8. Motivation, time 2	5.95	0.67	.25	.42	-.18	-.15	.32	.39	.39	.73					
9. Motivation, time 3	5.79	0.79	.25	.40	-.24	-.14	.37	.27	.45	.53	.78				
10. Overall job performance	5.37	0.68	.11	-.06	.15	.23	.07	.16	.32	-.02	.15	.93			
11. Job performance, month 1	5.37	0.80	.23	-.04	.08	.30	.24	.37	.43	.10	.15	.79	.93		
12. Job performance, month 2	5.36	0.80	.03	.01	-.01	.14	-.03	.02	.19	-.01	.21	.88	.42	.91	
13. Job performance, month 3	5.46	0.88	.02	-.08	.23	.18	-.06	-.01	-.02	-.10	-.09	.70	.27	.70	.95

^a Coefficient alpha reliability estimates are in bold on the diagonal. $n = 57-72$. For correlations greater than or equal to .20, $p \leq .05$. For correlations greater than or equal to .29, $p \leq .01$.

^b Measured as American College Testing (ACT) scores. For these correlations, $n = 37-44$.

TABLE 3
Results of Regression Analysis of Specific Self-Efficacy on the Treatment^a

Variables	Generalized Self-Efficacy ^b		Tenure with Firm ^b	
	Step 1	Step 2	Step 1	Step 2
	Main Effect	Interaction	Main Effect	Interaction
Specific self-efficacy, time 2				
Initial self-efficacy	.85**		.82**	
Treatment	.21*		.22*	
Generalized self-efficacy	-.06			
Tenure with firm			.05	
Generalized self-efficacy × treatment		.01		
Tenure with firm × treatment				-.09
R^2	.72	.72	.71	.73
ΔR^2		.00		.02
F	54.57**	40.32**	54.13**	41.50**
Specific self-efficacy, time 3				
Initial self-efficacy	.63**		.66**	
Treatment	.19*		.20*	
Generalized self-efficacy	.12			
Tenure with firm			.07	
Generalized self-efficacy × treatment		.18*		
Tenure with firm × treatment				.18*
R^2	.50	.53	.49	.52
ΔR^2		.03		.03
F	18.50**	15.35**	17.98**	14.98**

^a $n = 60-69$. Standardized regression coefficients are shown.

^b Proposed moderator.

* $p \leq .05$

** $p \leq .01$

results indicated that the treatment had a slightly stronger impact on the time 3 specific self-efficacy of auditors who were *higher* in generalized self-efficacy and had *more* tenure; this was in the opposite direction to that we hypothesized.

Next we tested for the presence of the Galatea effect. The main effect of the treatment on job performance over the three months was small ($d = 0.22$) and not statistically significant ($\beta = .19$, n.s.). Table 5 shows these results. Still, an estimated 55 percent of the auditors receiving the intervention performed above the median, compared with 45 percent of the control group; these percentages constitute a binomial effect size display (BESD; Rosenthal & Rubin, 1982). Experimental auditors did have significantly better job performance during the first month ($\beta = .30$, $p \leq .05$); their performance was one-half of a standard deviation higher than their peers' ($d = 0.47$). However, during the next two months, the effect essentially disappeared ($\beta = .11$, n.s., and $\beta = .07$, n.s.; $d = 0.05$ and $d = 0.03$, respectively). Thus, results supported the ex-

istence of a Galatea effect (Hypothesis 1), though the effect weakened over time. In addition, as predicted, the treatment also raised auditors' motivation, and this effect persisted for three months ($\beta_{t2} = .24$, $p \leq .05$; $\beta_{t3} = .25$, $p \leq .05$), thus supporting Hypothesis 2; see Table 4. This impact was equivalent to half of a standard deviation ($d_{t2} = .50$; $d_{t3} = .52$) or an increase in the success rate on motivation from 37 to 63 percent (BESD). Next, using Cohen's (1988) Table 2.3.2, we calculated that the average power for the supported main effects was .59. Given that the sample sizes were similar for the nonsupported main effects, if their true effect sizes were also similar in magnitude, power would not be a reason for the lack of statistical significance. Finally, to examine whether there was a difference between control participants having informational interviews and those who did not, we computed Z-scores for each set of correlations between the treatment and the independent and dependent variables (for instance, for specific self-efficacy, time 2, r 's = .25 and .20; for motiva-

TABLE 4
Results of Regression Analysis of Motivation on the Treatment^a

Variables	Generalized Self-Efficacy ^b		Tenure with Firm ^b	
	Step 1	Step 2	Step 1	Step 2
	Main Effect	Interaction	Main Effect	Interaction
Motivation, time 2				
Treatment	.24*		.24*	
Generalized self-efficacy	.43**			
Tenure with firm			-.15	
Generalized self-efficacy × treatment		-.19*		
Tenure with firm × treatment				-.34**
<i>R</i> ²	.17	.21	.06	.18
ΔR^2		.04		.12
<i>F</i>	7.38**	5.76**	2.17	4.67**
Motivation, time 3				
Treatment	.25*		.26*	
Generalized self-efficacy	.41**			
Tenure with firm			-.15	
Generalized self-efficacy × treatment		-.21*		
Tenure with firm × treatment				.31**
<i>R</i> ²	.15	.19	.06	.11
ΔR^2		.04		.05
<i>F</i>	5.84**	4.50**	2.02	2.48*

^a $n = 61-70$. Standardized regression coefficients are shown.

^b Proposed moderator.

* $p \leq .05$

** $p \leq .01$

tion, time 2, r 's = .20 and .27; and for job performance, month 1, r 's = .29 and .22). There were no significant differences.

Next we tested the mediation hypothesis stating that motivation explains the impact of specific self-efficacy on job performance. The first two conditions were met in that specific self-efficacy was significantly related to motivation ($r_{t2} = .39$, $p \leq .001$) and to overall and month 1 job performance ($\beta = .34$, $p \leq .01$; $\beta = .53$, $p \leq .001$). Next, although motivation was marginally related to overall job performance ($\beta = .19$, $p \leq .10$), it was not significantly related to month 1 performance ($\beta = .12$, n.s.), and in neither case did it reduce the impact of specific self-efficacy. Thus, the results did not support motivation as a mediator of specific self-efficacy's effect on job performance (Hypothesis 3).

We then tested whether the treatment differently influenced the auditors' motivation and job performance because of their level of generalized self-efficacy. We found interaction effects with motivation at times 2 and 3 ($\beta_{t2} = -.19$, $p \leq .05$; $\beta_{t3} = -.21$, $p \leq .05$), and with job performance for the average of the three months ($\beta = .33$, $p \leq .05$) as well as month 1 ($\beta = .33$, $p \leq .05$) and month 2 ($\beta =$

.28, $p \leq .05$) specifically; see Tables 4 and 5 for results. Thus, the significant interaction effect on performance faded before the end of the third month. Analysis of the interactions revealed that the treatment produced greater motivation with those lower in generalized self-efficacy, but more favorably affected the job performance of auditors with higher generalized self-efficacy. Thus, these results, combined with the interaction manipulation check, indicated that the plasticity hypothesis (Hypothesis 4) was mostly unsupported. We also examined whether the experimental treatment benefited those with less tenure with the firm differently. Essentially, the results revealed that they did not. There was no performance difference and, although we did find significant interactions with motivation at times 2 and 3 ($\beta_{t2} = -.34$, $p \leq .01$; $\beta_{t3} = .31$, $p \leq .01$), only the time 2 measure was in the predicted direction—that is, favoring new auditors. Thus, Hypothesis 5 was mostly unsupported, as the treatment did not have a stronger effect for new auditors. Finally, the gender by treatment interaction was not significant for specific self-efficacy, motivation, or job performance. The beta coefficients ranged from $-.14$ to $.16$. Thus, the

TABLE 5
Results of Regression Analysis of Job Performance on the Treatment^a

Variables	Moderator Tests				Mediator Test	
	Generalized Self-Efficacy		Tenure with Firm		Motivation	
	Main Effect	Interaction	Main Effect	Interaction	Step 1	Step 2
Overall job performance						
Ability	.14		.09		.13	.07
Treatment	.19		.11			
Generalized self-efficacy		-.05				
Tenure with firm			.31*			
Generalized self-efficacy × treatment		.33*				
Tenure with firm × treatment				-.09		
Self-efficacy, time 2					.34**	.34**
Motivation, time 2						.19
<i>R</i> ²	.06	.17	.13	.14	.13	.19
ΔR^2		.11		.01		.06
<i>F</i>	.77	1.85	1.95	1.52	3.07*	2.91*
Job performance, month 1						
Ability	.09		.02		.05	.01
Treatment	.30*		.22*			
Generalized self-efficacy		-.08				
Tenure with firm			.29*			
Generalized self-efficacy × treatment		.33*				
Tenure with firm × treatment				-.11		
Self-efficacy, time 2					.53**	.55**
Motivation, time 2						.12
<i>R</i> ²	.09	.20	.13	.15	.28	.34
ΔR^2		.11		.02		.06
<i>F</i>	1.21	2.12	1.93	1.55	7.42**	5.88**
Job performance, month 2						
Ability	.02		.05			
Treatment	.11		.05			
Generalized self-efficacy		.11				
Tenure with firm			.25			
Generalized self-efficacy × treatment		.28*				
Tenure with firm × treatment				-.08		
<i>R</i> ²	.02	.10	.06	.07		
ΔR^2		.08		.01		
<i>F</i>	.31	.96	.83	.66		
Job performance, month 3						
Ability	.19		.20			
Treatment	.07		.08			
Generalized self-efficacy		-.09				
Tenure with firm			.15			
Generalized self-efficacy × treatment		.17				
Tenure with firm × treatment				.01		
<i>R</i> ²	.06	.10	.08	.08		
ΔR^2		.04		.00		
<i>F</i>	.75	.81	.92	.67		

^a *n* = 36–41. The lower *n* is the result of having American College Testing (ACT) scores for only 44 participants. Regressions run with all participants and without ACT scores gave similar results. Standardized regression coefficients are shown. The mediator analysis examined the extent to which motivation mediated the effect of self-efficacy on job performance.

* *p* ≤ .05

** *p* ≤ .01

men and women in this sample did not seem to respond differently to the Galatea intervention.

DISCUSSION

Our constructive replication makes contributions within two general categories. First, we shed important light on five previously untested boundary conditions of producing self-fulfilling prophecy effects. Specifically, we examined whether and for how long a nonfictitious, nontraining Galatea intervention could raise new and established professionals' day-to-day work performance within a business organization. Results indicated that we produced a Galatea effect, although the effect was temporary for performance. Thus, whereas previous research has relied upon fictitious interventions to improve learning performance within military or academic settings, our experiment provides important external validity to the generalizability of self-fulfilling prophecy phenomena with these "new" characteristics and a new setting. Second, we conducted needed testing of the self-fulfilling prophecy model and constructs proposed to moderate it. Our results provided general support for self-fulfilling prophecy theory, self-efficacy theory, the self-fulfilling prophecy model, and the Galatea effect. For example, our results supported the spiral effect explained in these theories—the reciprocal relationships between specific self-efficacy and performance. Our experimental design clarifies that the treatment did cause a difference in experimental auditors' specific self-efficacy that was related to a difference in subsequent performance. In turn, performance (at time 1) was then strongly related to subsequent (time 3) self-reports of specific self-efficacy. Thus, it appears that initial specific self-efficacy caused subsequent performance, which along with the treatment, in turn influenced later specific self-efficacy estimates. We now elaborate theoretical implications, including possible extensions of current theory, and note practical implications of our findings.

Theoretical Implications

Inasmuch as our experiment lasted three months, the results may provide additional insights into self-fulfilling prophecy phenomena. In fact, in management research, time often has been a neglected variable in developing or evaluating theories (Ancona, Goodman, Lawrence, & Tushman, 2001). First, no previous research has continued to take measurements until after the self-fulfilling prophecy effect disappeared in order to get a sense of when that might be. In the present experiment,

the effect on job performance for all auditors lasted approximately one month. Thus, our results suggest that self-fulfilling prophecy interventions may fade within a relatively short time (see Figure 1). Perhaps over time information that is more proximal or salient dilutes the treatments.

One possible factor that may delay decay is individuals' generalized self-efficacy. According to plasticity theory, people with lower self-esteem/generalized self-efficacy are more susceptible to external influence; however, the self-fulfilling prophecy studies showing this finding have measured impact on performance relatively soon after interventions have taken place—from right away to two and a half weeks later (e.g., Eden & Aviram, 1993¹; Eden & Kinnar, 1991). However, our results showed that auditors with high generalized self-efficacy had higher specific self-efficacy at three months and were the only ones with higher job performance throughout the three months. Therefore, the impact of communicating high expectations may not last as long for those lower in generalized self-efficacy. They may have strong initial reactions to messages reassuring them of their competence, but then the pressures and ambiguities produced by a challenging professional environment, combined with their natural, lower self-perceptions, may cause them to fall back to previous self-doubts. On the other hand, our results suggest that inasmuch as messages of self-competence are more consistent with the dispositions of employees with high generalized self-efficacy, the effects of periodic reminders of what they already generally hold to be true about themselves may have more staying power (see Figure 1).

Next, previous self-fulfilling prophecy research has only included individuals new to a particular study setting or else has found no performance effect for those with time in a particular setting (e.g., Raudenbush, 1984). Our results essentially showed that the treatment benefited both newcomers and established auditors regardless of their tenures with the firm. Differences between our experiment and past research suggest three implications of these findings. First, results showing no effect for participants with tenure in a particular setting are from Pygmalion studies, whereas ours was a Galatea study. Self-serving bias may thus explain the different outcomes. People tend to attribute their own past successes to internal factors and their own past failures to external factors, and they tend

¹ These authors also found a plasticity effect at two months, but it was for influencing generalized self-efficacy, not specific self-efficacy.

to do the opposite for other people (Bradley, 1978). Thus, in a Pygmalion setting, when participants' experience with people causes them to have initially low expectations, they may attribute others' failures to internal factors and thus not believe disconfirming high-expectation interventions about those people. However, inasmuch as Galatea interventions attempt to raise expectations about *oneself*, people may attribute their own past failures to external factors and believe interventions indicating that their past successes are predictive of future success. Thus, although Pygmalion interventions have not worked with established employees, our results suggest that Galatea interventions may (see Figure 1).

Second, interventions in previous self-fulfilling prophecy research used fictitious information, whereas ours did not. With fictitious interventions, established participants may have contradictory prior experience, and so the intervention may be discounted outright. However, interventions based on factual data are consistent with actual experience and thus they may be seen as credible and have a chance to influence expectations. Thus, whereas interventions based on fictitious information may not work with established employees, our findings show that nonfictitious interventions may (see Figure 1). Third, research has indicated that boosting specific self-efficacy may work better during early skill-acquisition trials than in later trials, after tasks have been mastered (Mitchell, Hopper, Daniels, George-Falvy, & James, 1994). This argument suggests that established employees will be less susceptible to specific self-efficacy interventions, since they will likely have mastered their tasks. However, the premise is that expectation solidification comes through task mastery or lack of novelty versus just time. This formulation is consistent with our finding similar performance effects with established auditors who were beginning new roles as "in-charge auditors" and thus still learning. Therefore, our results suggest that specific self-efficacy interventions may work with new or established employees if they are still engaged in learning novel tasks (see Figure 1).

Finally, our results suggested that motivation did not mediate the effect of specific self-efficacy on job performance. Although inconsistent with the self-fulfilling prophecy model, this finding is consistent with the mixed support found in similar studies (e.g., Davidson & Eden, 2000; Dvir, Eden, & Banjo, 1995). Our findings, together with past research, suggest that motivation may need to be conceptualized differently (see Figure 1). First, Thomas and Velthouse (1990) conceptualized motivation as empowerment and "operationalized" it as intrinsic

task motivation. They argued that to the degree workers have choices of tasks that they deem to be meaningful, believe they have sufficient competence, and believe that their behaviors will have an impact, they will choose to engage and persist in tasks. This conceptualization has found some support in expectation research (Chen & Klimoski, 2003). Second, motivation might be conceptualized extrinsically, as employees' level of desire for potential rewards and outcomes.

Limitations

One limitation of our study is that not all control participants received informational interviews. However, the informational interviews were not essential as the intervention was consistent with our research question, nor did their absence have an impact since we found no difference in the effect of the treatment between control participants who received an informational interview and those who did not. As such, in two offices one group received the treatment (SSE interview) and one group did not (no SSE interview). In keeping with past research in this area (e.g., Eden et al., 2000), we were not testing whether various kinds of information (for instance, specific self-efficacy-enhancing versus informational) were differentially effective. Second, there is always the potential for some sensitive information leakage among participants during a study. As indicated above, however, we believe the study design and the truly unique characteristics of the site ensured that leakage was extremely unlikely. Third, since the job performance evaluations were completed at the end of audit engagements, which varied in length from a week to several months, it was not possible to synchronize the timing of these evaluations with the collection of the self-reported measures. Thus, we were not able to employ multilevel modeling techniques that would have allowed us to investigate within-individual variance in the motivational states.

Next, having a researcher conduct the interviews raises the question of potential experimenter bias. Several considerations mitigated this concern. First, such a bias could never be fully controlled. Bias could be transmitted directly to the participants or indirectly through whomever the researcher instructed to conduct the intervention. Second, in self-fulfilling prophecy research, transferring expectations is an essential part of the treatment by design; the interviewer wants to transfer high expectations—this is what is being tested. Third, the letters and e-mail came from others. Fourth, the interviewer memorized a word-for-

word script for each condition that was repeated verbatim in all interviews to create a uniform verbal delivery and reduce the chance for any variation in treatment. Granted, it is possible that the interviewer's nonverbal behaviors, such as facial expressions and body position, could have transmitted different expectations to members of the experimental and control groups. Finally, researchers conducting the treatment is a generally accepted practice in this research area (e.g., Davidson & Eden, 2000; Eden & Ravid, 1982).

Next, Chen, Gully, and Eden (2001) have questioned the validity of the Sherer GSE scale and have constructed and validated a new scale measuring generalized self-efficacy (the NGSE); however, at the time of our experiment the NGSE was not available. This might be an alternative explanation for our lack of plasticity results. We also did not explicitly test the opposite sides of some boundary conditions (for instance, we did not conduct experiments in both a nonbusiness *and* a business setting) and thus cannot specify which conditions accounted for the difference in our findings from previous research. We felt sufficient research had addressed nonbusiness settings, and we were seeking instead to discover whether the five boundaries could be hurdled. Finally, the age and tenure of the participants was restricted. We still feel that older people with more tenure in organizations (five or ten years, for example) and more task experience may respond differently to specific self-efficacy interventions and that lack of support for the related hypothesis in our experiment may be a result of this range restriction.

Practical Implications

The boundary conditions addressed in this experiment have been literal boundaries on business adoption of self-fulfilling prophecy practices as a viable management tool for nearly 40 years. Our findings imply this need not be the case. The present results demonstrated that communications based upon factual personal and organizational information can strengthen employees' specific self-efficacy beliefs, motivation, and job performance. These findings should help disabuse managers of the notion that they must use fictitious information to create positive self-fulfilling prophecy effects. Our results indicate that there is much true information at managers' disposal that can be used to boost specific self-efficacy; management need only learn to recognize such information. Possible sources include employees' résumés, past and current successes at work, and beneficial characteristics or behaviors that they observe. In addition,

managers can share accounts of how they experienced common difficulties and then model how they overcame them. One challenge is how to use nonfictitious information with employees whose performance is below average. What specific self-efficacy-enhancing truths do managers communicate to the lower quartile? First, managers could find positive contributions these employees are making and focus on these rather than on the negative aspects of performance—even poor employees do some things well. Second, managers could recognize and laud improvement rather than the absolute level of performance. Managers could also reinforce to employees their potential. Managers must have seen something good in such employees when they hired them; that potential is likely still there but has not been tapped. Of course, managers must examine their own core beliefs about poorly performing employees. Employees can sense the duplicity in a manager who communicates high expectations but does not believe them, and such attempts have little chance for success.

In fact, manager expectations are the key to producing Pygmalion effects (high manager expectations of employees leading to high employee performance). For example, to attempt to produce a Pygmalion effect, managers might tack mini high-expectation interventions onto everything that they do. All changes, assignments, announcements, and so forth could be communicated with the expectation of success. To clarify, managers may also be able to create Galatea effects (employees' high self-expectations leading to high performance) by directly raising employees' self-expectations. In fact, management and supervisory personnel are likely a better source of Galatea interventions than a third party is, since they interact with employees every day. Whereas our intervention and boosters were infrequent and related to auditors' jobs in general, supervisors could embed their efficacy-enhancing attempts within employees' daily activities and problems and could do so consistently. For example, inasmuch as specific self-efficacy is increased through successful performance, management could also structure successful repetitious and progressively more difficult mastery experiences for employees, so that they *learn* that they can "do it."

Our study also provides managers with direct experimental evidence that self-fulfilling prophecy effects can be applied more broadly than had previously been shown. Specifically, these effects may be produced within businesses, with both newcomer and second-year employees, and they may affect ongoing job performance in addition to learning performance within a training setting. Of course, inasmuch as the effects on performance that

we found were small to medium-sized and were temporary, questions arise as to whether the effects were large enough and long enough to be of practical importance. Our results are consistent with evidence indicating that self-fulfilling prophecy effects within businesses may just be smaller than in other settings. McNatt's (2000) meta-analytic review of Pygmalion studies estimated a smaller effect within civilian settings ($d = 0.42$) than in military settings ($d = 1.07$). Likewise, a meta-analysis of five Galatea studies with military and school settings yielded an estimated effect of 1.64 (McNatt, 2003)—compared to our effect within a business of 0.47. In terms of BESD equivalents, our treatment raised auditors' likelihood of higher than average job performance from 38 percent to 62 percent for one month and from 45 percent to 55 percent for at least three months. Is this increase in performance large enough or long enough? Perhaps supervisors might look for situations in which such a temporary increase in performance could make a meaningful difference—situations such as special projects, critical periods, beginnings of new assignments or jobs, changes requiring employee commitment, and performance slumps. Thus, although the effects found were not large and were temporary, we still believe such interventions have a practical benefit to business organizations. Additional ways to implement self-fulfilling prophecy principles and expectation effects along with aid for learning to do so have been discussed elsewhere (e.g., Eden 1990; McNatt, 2001).

Future Research

The results and limitations of our experiment suggest several important areas that future research might target. First, research should test whether managers alone can effectively produce Galatea effects. Next, studies should examine effects with older and more experienced people—those with five to ten years in an organization or industry. Research could also address the propositions made in the theoretical implications section of this article. Next, although we found no moderator effect for gender, self-fulfilling prophecy effects may differ for men and women in ways that future research can help elucidate. Further, research is needed to test the source, content, and frequency of interventions to learn more about the longevity of the effects they produce and how to delay fade-out; specific ideas appear in the practical implications section above. Finally, since these empirical results are based on the first tests of several important boundary conditions, there is a need to continue to examine these boundaries. Self-fulfilling prophecy ef-

fects seem to have great potential benefits for organizations, but researchers need to continue to learn how, and how much, management might harness their power.

REFERENCES

- Aiken, L. S., & West, S. G. 1991. *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Ancona, D. G., Goodman, P. S., Lawrence, B. S., & Tushman, M. L. 2001. Time: A research lens. *Academy of Management Review*, 26: 645–663.
- Bandura, A. 1982. Self-efficacy mechanism in human agency. *American Psychologist*, 3: 122–147.
- Bandura, A. 1986. *Social foundations of thought and action: A social-cognitive view*. Englewoods Cliffs, NJ: Prentice-Hall.
- Bandura, A. 1997. *Self-efficacy: The exercise of control*. New York: Freeman.
- 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.
- Bradley, G. W. 1978. Self-serving biases in the attribution process: A re-examination of the fact or fiction question. *Journal of Personality and Social Psychology*, 36: 56–71.
- Brockner, J. 1988. *Self-esteem at work: Research, theory, and practice*. Lexington, MA: Lexington Books.
- Campbell, J. P., & Pritchard, R. D. 1976. Motivation theory in industrial and organizational psychology. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*: 63–130. Chicago: Rand McNally.
- Chen, G., Gully, S., & Eden, D. 2001. Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4: 62–83.
- Chen, G., & Klimoski, R. J. 2003. The impact of expectations on newcomer performance in teams as mediated by work characteristics, social exchanges, and empowerment. *Academy of Management Journal*, 46: 591–607.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.
- Davidson, O., & Eden, D. 2000. Remedial self-fulfilling prophecy: Two field experiments to prevent golem effects among disadvantaged women. *Journal of Applied Psychology*, 85: 386–398.
- Dvir, T., Eden, D., & Banjo, M. L. 1995. Self-fulfilling prophecy and gender: Can women be Pygmalion and Galatea? *Journal of Applied Psychology*, 80: 253–270.
- Eden, D. 1988. Pygmalion, goal setting, and expectancy:

- Compatible ways to boost productivity. *Academy of Management Review*, 13: 639–652.
- Eden, D. 1990. *Pygmalion in management: Productivity as a self-fulfilling prophecy*. Lexington, MA: Lexington Books.
- Eden, D. 1992. Leadership and expectations: Pygmalion effects and other self-fulfilling prophecies in organizations. *Leadership Quarterly*, 3: 271–305.
- Eden, D. 2002. Replication, meta-analysis, scientific progress, and *AMJ's* publication policy. *Academy of Management Journal*, 45: 841–846.
- Eden, D., & Aviram, A. 1993. Self-efficacy training to speed reemployment: Helping people to help themselves. *Journal of Applied Psychology*, 78: 352–360.
- Eden, D., Geller, D., Gewirtz, A., Gordon-Terner, R., Inbar, I., Liberman, M., Pass, Y., Salomon-Segev, I., & Shalit, M. 2000. Implanting Pygmalion Leadership Style through workshop training: Seven field experiments. *Leadership Quarterly*, 11: 171–210.
- Eden, D., & Kinnar, J. 1991. Modeling Galatea: Boosting self-efficacy to increase volunteering. *Journal of Applied Psychology*, 76: 770–780.
- Eden, D., & Ravid, G. 1982. Pygmalion versus self-expectancy: Effects of instructor and self-expectancy on trainee performance. *Organizational Behavior and Human Performance*, 30: 351–364.
- Eden, D., & Zuk, Y. 1995. Seasickness as a self-fulfilling prophecy: Raising self-efficacy to boost performance at sea. *Journal of Applied Psychology*, 80: 628–635.
- Gist, M. E. 1987. Self-efficacy: Implications for organizational behavior and human resource management. *Academy of Management Review*, 12: 472–485.
- Gist, M. E. 1989. The influence of training method on self-efficacy and idea generation among managers. *Personnel Psychology*, 42: 787–805.
- Gist, M. E., & Mitchell, T. R. 1992. Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, 17: 183–211.
- Gist, M. E., Schwoerer, C., & Rosen B. 1989. The effects of alternate training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology*, 74: 884–891.
- Gully, S. M., Payne, S. C., Koles, K. L. K., & Whiteman, J. K. 2002. The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective. *Journal of Applied Psychology*, 87: 143–155.
- Judge, T. A., Bono, J. A., & Locke, E. A. 2000. Personality and job satisfaction: The mediating role of job characteristics. *Journal of Applied Psychology*, 85: 237–249.
- Judge, T. A., Erez, A., & Bono, J. E. 1998. The power of being positive: The relation between positive self-concept and job performance. *Human Performance*, 11: 167–187.
- Kanfer, R. 1990. Motivation theory and industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed.), vol. 1: 75–170. Palo Alto, CA: Consulting Psychologists Press.
- Locke, E. A., Frederick, E., Lee, C., & Bobko, P. 1984. Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology*, 69: 241–251.
- McNatt, D. B. 2000. Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, 85: 314–322.
- McNatt, D. B. 2001. *The Pygmalion effect: Managing the power of expectations*. Carlsbad, CA: CRM Learning.
- McNatt, D. B. 2003. *Building others' self-confidence: A meta-analysis of the Galatea effect*. Working paper no. 200310, University of Georgia, Athens.
- Merton, R. K. 1948. The self-fulfilling prophecy. *Antioch Review*, 8: 193–210.
- Mitchell, T. R., Hopper, H., Daniels, D., George-Falvy, J., & James, L. R. 1994. Predicting self-efficacy and performance during skill acquisition. *Journal of Applied Psychology*, 79: 506–517.
- Mitchell, T. R., & James, L. R. 2001. Building better theory: Time and the specification of when things happen. *Academy of Management Review*, 26: 530–547.
- Naylor, J. C., Pritchard, R. D., & Ilgen, D. R. 1980. *A theory of behavior in organizations*. New York: Academic Press.
- Raudenbush, S. W. 1984. Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, 76: 85–97.
- Rosenthal, R. 2002. Covert communications in classrooms, clinics, courtrooms, and cubicles. *American Psychologist*, 57: 839–849.
- Rosenthal, R., & Jacobson, L. 1966. Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports*, 19: 115–118.
- Rosenthal, R., & Rubin, D. B. 1982. A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 74: 166–169.
- Saks, A. M. 1995. Longitudinal field investigation of the moderating and mediating effects of self-efficacy on the relationship between training and newcomer adjustment. *Journal of Applied Psychology*, 80: 211–225.
- Schallow, J. R. 2000. A comparison of three approaches to constructing item parcels to improve subject-to-parameter ratios in confirmatory factor analysis. *Multivariate Experimental Clinical Research*, 12: 29–41.

- Sherer, M., Maddux, J., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. 1982. The Self-Efficacy Scale: Construction and validation. *Psychological Reports*, 51: 663–671.
- Schmidt, F. L. 1988. The problem of group differences ability test scores in employment selection. *Journal of Vocational Behavior*, 33: 272–292.
- Stumpf, H., & Stanley, J. C. 1998. Stability and change in gender-related differences on the College Board Advanced Placement and Achievement Tests. *Current Directions in Psychological Science*, 7: 192–196.
- Thomas, K. W., & Velthouse, B. A. 1990. Cognitive elements of empowerment: An “interpretive” model of intrinsic task motivation. *Academy of Management Review*, 15: 666–681.
- Weiss, H., & Adler, S. 1984. Personality and organizational behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior*, vol. 6: 1–50. Greenwich, CT: JAI Press.
- White, S. S., & Locke, E. A. 2000. Problems with the Pygmalion effect and some proposed solutions. *Leadership Quarterly*, 11: 389–415.



D. Brian McNatt (*bmcnatt@terry.uga.edu*) is an assistant professor in the Department of Management at the Terry College of Business, University of Georgia. He received his Ph.D. from the University of Iowa. His research interests are in the areas of expectations broadly defined (e.g., self-fulfilling prophecies, belief perseverance, self-efficacy), and academic-practitioner research relationships.

Timothy A. Judge (*tjudge@ufl.edu*) is a professor in the Department of Management, Warrington College of Business, University of Florida. He received his doctoral degree from the University of Illinois at Urbana-Champaign. His principal research interests are in the areas of personality and affect. He also does research on job attitudes, leadership, and careers.



Copyright of Academy of Management Journal is the property of Academy of Management and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.