Restraining Golem and Harnessing Pygmalion in the Classroom: A Laboratory Study of Managerial Expectations and Task Design

DENNIS REYNOLDS Washington State University

In seeking to enhance teaching effectiveness, educators and trainers are demonstrating a growing interest in understanding positive expectations and the resulting Pygmalion effect. Unfortunately, the true impact of the Golem effect—Pygmalion in reverse—has gone unproven and the potential negative effects are not wholly understood. Furthermore, the literature largely fails to address the extent to which Pygmalion-related effects may differ based on task design. This study examines the effect of an instructor's verbalized expectations—both negative and positive—on the performance of 351 business-school undergraduate students. Analyses using pre- and posttreatment data collected during controlled-laboratory experiments indicated, most notably, that negative expectations of students' performance on cognitively based tasks tend to degrade that performance. The effects on noncognitively based tasks were, however, positive. Positive expectations had the opposite effect. Implications for management education research and related applications are discussed.

When teachers show that they expect students to perform well, students do perform well; when teachers project no such expectations, students do not attain the same level of performance. Research supporting this Pygmalion effect in the classroom is powerfully convincing, with studies spanning several decades (e.g., Rosenthal & Jacobson, 1968; Cooper, 1979; Dusak & Joseph, 1985; Babad, 1995). The implications of such research are evident in the management literature, too, with several studies underscoring the importance of managers' expressed positive expectations in generating superior performance from subordinates (Eden, 1990a; McNatt, 2000; Sutton & Woodman, 1989).

Stemming from Merton's (1948) self-fulfilling-

prophecy theory, the bulk of expectations-related research has, not surprisingly, focused on the Pygmalion effect. As illustrated in the movie *My Fair Lady* (based on George Bernard Shaw's play titled *Pygmalion*), the effect can be defined as improved subordinate performance resulting from expressed, positive supervisory expectations. But what about the reverse effects?

Theory and related research surrounding the consequences of negative Pygmalion effects coined Golem effects by Babad, Inbar, and Rosenthal (1982)—is even more provocative given suggestions that a superior's low expectations, made apparent through his or her behavior, can negatively impact subordinates' performance. This phenomenon has proved more elusive as a research topic than positive Pygmalion effects because researchers have had difficulty in directly testing the effects of negative supervisory expectations due to ethical constraints and operational challenges.

These constraints are particularly critical in a

An earlier version of this manuscript was named "Best Paper in Management Education and Development" at the 2005 Academy of Management meeting. I thank the meeting-paper reviewers, and especially J. B. Arbaugh, 2004-05 Program Chair, MED Division, as well as the James Bailey, AMLE Editor, and his reviewers, for their advice and counsel.

classroom setting given the potential implications for those subjected to the associated treatment. Some researchers have used existing teacher biases as a proxy for negative expectations, since such biases carry with them lower expectations for the students at whom the bias is directed (e.g., Baron, Tom, & Cooper, 1985). Meta-analyses demonstrate adequate effect magnitude for positive and negative expectations, but do not directly settle whether expectations—such as those explicitly verbalized by an instructor—have similar trending effects on subordinates' performance (cf., Brophy, 1983).

To further complicate the study of such effects, some researchers have posited that Pygmalionrelated effects vary with task design. For example, Lundberg (1975) found that positive-expectation effects affected team productivity and satisfaction differently based on the scope of the involved tasks. That is, output was greater under positiveexpectation conditions when tasks were more involved. Locke and Latham (2002) provided theoretical support for considering task design as a moderator in the goal-performance linkage. Other researchers have not, however, found empirical evidence that expectancy-induced effects vary significantly with task design (e.g., Kierein & Gold, 2000). My purpose in this study, then, is to test the Pygmalion and Golem effects in a classroom setting through an experiment with university business-school students as participants. In particular, the investigation is targeted at discerning whether students randomly assigned to treatment groupswherein an instructor provides positive, negative, or no expectations regarding performance, respectively—will perform differently. Thus, the primary contribution is to provide empirical evidence for largely untested Golem-related theory. Moreover, we explore the related causal conclusions by examining these potential effects on the basis of task design, thereby facilitating a broader understanding of Pygmalion and Golem in management education.

THE PYGMALION EFFECT

Rosenthal and Jacobson (1968) provided the first empirical evidence of the Pygmalion effect. In a field experiment using an educational setting, these researchers confirmed their hypothesis that students for whom teachers held higher expectations would perform better, despite the random assignment of students to what teachers were told was the overachieving group. Later studies (e.g., Babad, 1995; Rosenthal & Rubin, 1978) firmly established the existence of positive-expectations effects and demonstrated the magnitude of the phenomenon in education.

In 1969, Livingston theorized that raising supervisory expectations regarding workplace productivity would serve as a determinant of subordinate performance. In the decades since Livingston's Pygmalion-in-management article, hundreds of studies have been conducted in nonbusiness settings (Rosenthal, 1994). However, as demonstrated in McNatt's (2000) meta-analysis, there have been relatively few Pygmalion studies in managementrelated contexts; the recent meta-analysis included only 17 studies. Underscoring the importance of the phenomenon in management, McNatt reported an overall corrected estimate of the average effect of d = 1.13 (58 effect sizes, n = 2,784). Similarly, Kierein and Gold (2000) reported an overall effect size of d = 0.81 (13 effect sizes, n = 2,853) in the nine studies included in their meta-analysis. Both effect sizes fall into Cohen's (1988) categorization of "large."

In 1974, King tested the effect on a somewhat different sample in which managers' expectations regarding the productivity of four plants were artificially raised. Productivity improved in all the plants, but significantly greater productivity improvements were found in those plants where managers' expectations were artificially elevated. As already mentioned, Lundberg (1975) replicated King's study of plant productivity in an experimental-laboratory setting using 108 college students. The purpose of the replication using a laboratory setting was to control for extraneous factors thought to have potentially influenced King's results. The author reported that positive expectations had a modest effect (p < .10) on output as compared with the output of the control group.

Still others have tested Pygmalion in the workplace. Sutton and Woodman (1989) tested the Pygmalion effect in two department stores, but their findings were inconclusive—most likely owing to experimental-design flaws (Eden, 1990a). Sutton and Woodman noted in their summary, too, that one possible reason for the inconclusive findings was that supervisors did not have enough time to interact differently with treatment employees than with their control counterparts.

Dating back to Rosenthal's (1969) early work, the common approach has been to randomly designate some group members as high performers and to plant the idea of such a designation as a seed in the supervisor or teacher; researchers then compare output from these perceived high performers with those who are treated as controls. Eden (1990b) suggested that such an approach does not control for interpersonal contrast effects that result when the referent against which improved performance is assessed—the control group—is part of the group of individuals under study. To test this assertion, Eden studied 29 platoons in the Israel Defense Forces (IDF) by randomly assigning entire platoons to either Pygmalion or control groups. Analyses suggested that nearly a fifth of the variance in mean platoon performance was determined by the experimental treatment.

While considerable empirical evidence exists to support the Pygmalion effect, the number of studies exploring Pygmalion effects in both educational and workplace settings are modest given the potential impact on performance and productivity. There are three primary reasons for this. First there are inherent difficulties associated with instilling "treatment" expectations. Second, as noted by Eden (1990a), it may be that this aspect of self-fulfilling prophecy is so ubiquitous and so embedded in the social fabric that it transcends everyday scrutiny. Moreover, teachers and managers might anticipate that the effects of their expectations will be overtly manifested in performance, when in reality, the effects are often difficult to enumerate unless both the target function of the expectation and the associated output can be objectively separated and quantified.

THE GOLEM EFFECT

The Golem effect is the negative or dark version of Pygmalion: Behavior reflecting low or negative supervisory expectations generates negative results in subordinates' performance. Babad et al. (1982), drawing from Hebrew slang, where the word means oaf or fool, used the term Golem to describe the negative version of Pygmalion effects. The term originates from a Jewish legend in which a robotlike being was created and brought to life to eradicate evil but ultimately became a monster owing to the increasingly corrupting influence of its power (Collins & Pinch, 1998).

The outcome of the Golem effect can emerge either as a net decline in the quality of subordinates' performance or simply as lower-than-otherwise-attainable levels of performance. With few exceptions, recent empirical evidence is anecdotal, drawing on extrapolations from Pygmalion experiments. The reason for this is again likely to be the troubling ethics of experimentally affecting an individual's performance by artificially lowering his or her supervisor's expectations. Even Babad et al. (1982) opted not to apply a treatment of artificially lowered teacher expectations toward students; rather, the researchers experimentally raised teachers' expectations toward some students and compared the performance results with data on students for whom teachers naturally had either high or low expectations. Of regret is that the teachers' familiarity with low-expectations students likely confounded the results.

To address the ethics of operationalizing negative expectations while maintaining reasonable ecological validity, Beez (1971) randomly applied an expectation manipulation through a fictitious psychological report given to students' tutors; the reports included low intelligence-test scores for artificially labeled "low-ability" children and unusually high scores for randomly assigned "highability" children. Results indicated that the treatments produced the hypothesized outcomes; the "low-ability" children performed significantly worse than the "high-ability" children. Feldman and Prohaska (1979) conducted an experiment in which confederates acting as students evinced either positive or negative expectations toward 40 participants acting as their teachers. Results indicated there was a general, significant impact on the participants' behavior resulting from the treatments corresponding to the corresponding treatment directionality. Feldman and Theiss's (1982) study of the joint effects of teachers' expectations of students and students' expectations of teachers on the performance and attitudes of both groups produced similar results.

Using a management setting, two studies go beyond offering anecdotal evidence of the Golem effect while addressing the aforementioned challenges associated with Golem research. Oz and Eden (1994) randomly led treatment-assigned squad leaders (n = 17) in a military unit to believe that low scores on physical fitness tests were not indicative of subordinates' ineptitude, while control squad leaders (n = 17) were not told how to interpret test scores. Tests indicated that low-scoring individuals in the experimental squads improved more than those in the control squads. While the researchers employed a respectable research design and were cautious to abide by ethical standards, the sample was extremely small. Of greater relevance, the researchers failed to introduce lower supervisory expectations or to compare the results with those of a control group, thus failing to follow the methodology applied in the majority of work designed to test the Pygmalion effect (e.g., Rosenthal & Jacobson, 1968). It is therefore arguable that the researchers failed to assess the Golem effect directly. The study is important, however, because it was the first to attempt identification of Golem effects in a managementlike context.

Davidson and Eden (2000) replicated the Oz and

Eden (1994) approach by using a treatment designed to prevent low expectations from forming on the part of instructors of trainees of disadvantaged women in the IDF. The researchers randomly instructed leaders from half of the platoons that the recruits possessed substantially higher-thanusual abilities for special recruits. Results indicated that the Golem effect was evident in the lower performance achieved by the control squads. It is important to note, however, that in both the Oz and Eden and Davidson and Eden (2000) studies, the researchers did not introduce a treatment consisting of lowering supervisory expectations toward the subordinates. Hence the findings, while provocative, fail to test the Golem effect explicitly.

Vrugt (1990) adopted a more targeted design and conducted an experiment that, while not central to the management or educational domain, investigated whether artificially induced negative expectations on the part of therapists might be conveyed by nonverbal behavior toward clients. Beginning male psychotherapists (n = 18), serving as interviewers, and male psychology students (n = 18), serving as interviewees, were randomly assigned to either treatment or control conditions. The treatment involved informing the respective interviewers that the interviewees were undergoing treatment for psychological problems. The significant findings confirmed the hypothesis that the therapists' negative expectations did affect clients in a negative manner.

TASK DESIGN

Weick (1984) introduced the notion that positive expectations are quickly reinforced by what he termed "small wins"—simple tasks that reinforce the self-expectations created by the superior. Eden (1988) explained how these small wins can lead to success in more complex tasks and that "Managers should...make these successes salient to workers through feedback that molds the right attributions of success and failure" (p. 649). But do such small wins or losses, when linked to expectations, lead individuals to perform differently on tasks that vary in design?

Individuals apply a greater variety of strategies to complex tasks than to easy tasks. Supporting Weick's (1979) related notion of requisite variety, Chesney and Locke (1991) found that the effects of expectations may vary with task assignments because individuals receiving expectancy information may link related success or failure expectations to some strategies but not to others and that larger potential gains warrant consideration of more strategies. Moreover, meta-analyses (Wood, Mento, & Locke, 1987) with goal difficulty effect sizes (d) of .48 for the most complex tasks versus .67 for the least complex tasks, suggest that individuals assess the simplicity of a task more readily.

With respect to the present study, it is important to note that tasks that are similar in design to those for which individuals have received expectancy information may be consonantly affected. Latham and Seijts (1999), for example, found that goals were better achieved when participants received expectations that were distinct and immediately related to the task outcomes. As Dorner (1991) noted, individuals perform better in achieving complex tasks when success with related goals has been realized or-equally critical-linked to explicit expectations that are correspondingly complex. Finally, Frese and Zapf (1994) demonstrated that expectations are more readily manifested in correlative performance outcomes when the linkage between task and expectation is unequivocal.

Such research provides a foundation for understanding linkages between expectations and related tasks. Suppose, however, that tasks are introduced to which the expectations do not relate? How will students or subordinates respond? More to the point, will performance on tasks of different design (e.g., intellective vs. physical) be different for distinct treatment groups?

Kavanagh (1972), addressing a topic that is tangential to the present one, discerned that respondents, when given a choice, tended toward choosing and completing tasks for which they perceived no negative expectations. Conversely, individuals who perceived positive expectations with respect to a specific task gravitated more strongly toward completing such a task. More recently, Bolt, Killough, and Koh (2001) found that positive expectations were more strongly linked with positive outcomes when task design was intellectually more difficult. When task complexity was low, the relationship was apparent but was less pronounced. Unfortunately, these and other researchers studying task design have failed to explore the implications of negative expectations and the resulting outcomes. Nonetheless, the implications suggest that expectation effects-both positive and negative—will carry over to unrelated tasks.

Hypotheses

Relating to the primary objective of this research, which is to test the existence of both Pygmalion and Golem effects resulting from an instructor's expressions of positive and negative expectations, respectively, of a business student's performance, the first set of hypotheses is:

- Hypothesis 1a: Positive expectations verbalized by the instructor will positively influence students' performance.
- Hypothesis 1b: Negative expectations verbalized by the instructor will negatively influence students' performance.
- Hypothesis lc: The absence of verbalized expectations will neither positively nor negatively influence students' performance.

Pertaining to expectancy effects on task design, the related hypotheses are:

- Hypothesis 2a: Positive expectations verbalized by the instructor will result in a significant positive effect on students' performance of an unrelated task.
- Hypothesis 2b: Negative expectations verbalized by the instructor will result in a significant negative effect on students' performance of an unrelated task.

METHODS

Design and Procedure

As Babad (1993) noted, "Numerous obstacles prevent effective application of expectancy research" (p. 145). Traditionally, Pygmalion studies have introduced treatments in which the teacher's expectations have been artificially manipulated, therein providing evidence supporting self-fulfillingprophecy theory. Moreover, directional effects associated with positive and negative expectations have not been successfully assessed due to the ethical and operational challenges.

Embracing what Archer (1993) describes as "imaginative epistemology," then, we employed a research design in which the treatments were applied directly to students by way of support instructors. The course instructor first administered what was explained as a management-acumen assessment, which was simply a test comprised of 15 logic puzzles, commonly considered brainteasers ($\alpha = .91$). The following day, students were randomly divided (although neither the subjects nor support instructors knew the segregation was random), allegedly based on scores earned on the first test.

Next, the supporting instructors were supplied with information about students. Specifically, support instructor A was told she had been assigned the high-performing group; that is, the students had performed well on the previously administered test. She was also told that based on such performance, this group would likely perform well on the subsequent tests she was to administer during the next meeting. Support instructor B was informed of her group's dismal performance, along with the knowledge that such performance portended equally poor outcomes on the upcoming tests. As Easterly-Smith, Thorpe, and Lowe (1994) suggested, the third support instructor was given no information about the students in the respective group but was directed to give the two subsequent tests during the same class period. It is important to note that—as suggested by Eden (1990a)—expectations are more likely to be perceived as accurate by the receiver when there is a basis for the expectation, such as past achievement, thus the reference to the pretest aided in the treatment application and expression of expectations. Finally, the outline of how the posttests were applied was supplied to the support instructors to ensure the information delivery was kept constant across treatments, thus supporting the experimental design to test the effects of expectations rather than those associated with feedback or goal setting.

As a means of exploring task design and in a manner consistent with that outlined by Dawn and Latham (1996), the first of the two posttests was very similar in subject and scope to the pretest given to the students. The second test, which was described to the students as "completely different from the previous test" and an exercise of "noncognitive" skills, required students to create origami cups. Students were given verbal instructions along with a handout describing the folding procedure. Following a practice period, they were then given 5 minutes to fold as many cups as possible (with the explanation that only completed cups would be counted).

Every effort was made to create what Aronson and Carlson (1968) termed *mundane reality* in applying the treatments. To accomplish this, the same course instructor applied each of the treatments to the respective support instructors (which were new each term, supporting the treatment application's legitimacy) throughout the 4-year datacollection period. In addition, the "tests" were administered in a serious fashion to persuade students that they would be graded and that the "management-assessment" scores would contribute to their course grades. Furthermore, the students were told that the origami cups were needed for an event to be held on campus later that day and that only cups of good quality could be used.

To ensure the legitimacy of the experimental design, support instructors were new to the course each year and had no knowledge of similar experiments performed during previous semesters. The

December

assignments to the "smart" or "loser" groups were explicitly made, creating a very real psychological expectations schema among support instructors. Furthermore, we tested the manipulation by querying the support instructors at the conclusion of the experiment as to how their respective students had performed during the posttest; each responded as expected. That is, the support instructor with the Golem group reported that her students had not performed well (even though she did not know how the other group did), while the instructor with the Pygmalion group reported that this group had performed admirably (again without a performance referent).

At the conclusion of the experiment, the students assembled for a formal debriefing given by the course instructor. The entire experiment was discussed with particular focus on the effects. To this end, the data were entered for a cursory analysis and the results, which demonstrated the related effects, were explained, underscoring the importance for managers to understand the phenomenon. Finally, students were presented with the option to pull their respective data from the research study.

Sample

Drawing from undergraduate students in required, 2nd-semester introductory management courses during a 4-year period (2001–2004), the sample included 374 full-time students at a large university in the northeastern United States. Owing to absenteeism during the pre- or posttest application, the final sample for which "cup" data were available totaled 351. Furthermore, on one occasion the posttest was not distributed; thus, the sample producing change in pretest/posttest scores was available for 275 participants. To guard against cohort effects, the students were queried during the first class meeting about what they expected and about what they had heard would be involved in the class. No reports or references were made to suggest students had heard of the experiment described herein. Similarly, safeguards were taken to ensure the support instructors were not aware of the manipulations.

RESULTS

The means and standard deviations for pretest, posttest, resulting change, and number of cups are shown in Table 1. The information regarding pretest/posttest change and cups deserves more attention because these reflect treatment effects more explicitly. To this end, the table also includes change-value means and standard deviations by treatment.

Of the 351 students participating in the study, 51% were female and 12% were international students. These characteristics were similar across the four classes. Neither gender nor international status significantly differentiated the treatment effects. Similarly, no differences were found on the basis of which year the data were gathered (supporting the earlier check against cohort effects).

To ensure that the randomization scheme was adequate and that resulting differences among treatment effects were not the result of selection anomalies, we analyzed pretest scores on the basis of treatment using analysis of variance. No significant difference was found, indicating adequacy of randomization in creating pre-experimental equivalence among treatment groups.

Next, differences between pretest and posttest scores were compared on the basis of each treatment group. Results of the corresponding t tests demonstrated that participants reported significantly higher scores under positive treatment conditions and significantly lower scores under negative treatment conditions. As expected, there were no significant differences between pretest and posttest scores for the control group. Furthermore, an additional analysis of variance demonstrated significant differences among the treatment groups on the basis of the posttest scores. Distinguishing the supporting analyses by treatment group underscored the differences between the positive treatment effect and the control, as well as between the negative treatment effect and the control.

The change in pretest/posttest scores was regressed on the treatments to assess the effect of the positive and negative expectations, as stated in H1. The results indicated that, indeed, both treatments were significant and that the main effects offered considerable explanatory power ($R^2 =$.16; F = 8.38; df = 2, 272; and p < .001). Subsequent t tests demonstrated highly significant differences between treatment effects, confirming each of the related hypotheses.

To test the second set of hypotheses, we regressed the number of cups on the treatments. Again, the main effects were significant ($R^2 = .11$; F = 19.75; df = 2,348; and p < .001). Further analysis, however, did not support the hypotheses. As demonstrated by the means in Table 1, which were significantly different in comparing each pairing of values, those who received positive expectations regarding their managerial acumen produced a lesser number of cups than did those who received corresponding negative expectations.

	(+)	(–)	(c)
	Treatment	Treatment	Treatment
Pretest ($N = 275$)	M:10.78	M:10.83	M:10.81
Posttest ($N = 275$)	M:10.42	M:8.94	M:9.76
Change Value ($N = 275$)	SD:2.70	<i>SD</i> :2.97	<i>SD</i> :2.24
	M:36	<i>M</i> :-1.89	<i>M</i> :1.05
Cups ($N = 351$)	SD:2.67	SD:3.17	SD:1.82
	M:17.87	M:20.82	M:23.70
	SD:6.00	SD:8.04	SD:7.02

TABLE 1 Means, Standard Deviations, Change Values, and Cups by Treatment

Moreover, the control group produced significantly more cups that either of the treatment groups.

DISCUSSION

The linkage between verbally expressed expectations—both positive and negative—and resulting performance on related tasks was confirmed. The effect of positive expectations is not surprising given both the ubiquitous relationship between positive expectations and outcomes and the evidence provided in the literature linking positive supervisory expectations and subordinate performance in both classroom and workplace settings (e.g., Crawford, Thomas, & Fink, 1980; Eden & Ravid, 1982; King, 1974; Rosenthal, 1994).

The findings pertaining to the Golem effect represent a unique contribution toward the study of interpersonal expectations in that they demonstrate empirically—for the first time—that a supervisor's verbally expressed negative expectations have a direct impact on subordinates' performance on related tasks. While previous studies have attempted to confirm the Golem effect (e.g., Babad et al., 1982; Davidson & Eden, 2000; Oz & Eden, 1994), none of these studies actually applied a treatment of negative expectations. Only Vrugt (1990) directly tested the role of negative expectations in his study involving therapy interviews; that study did not, however, involve real patients, was not within the management or education domains, and focused only on nonverbal-communication behaviors.

The Golem effect produced in this study is also important because it was induced while circumventing the ethical issues cited earlier. To date, researchers have attributed their inability to directly test the Golem effect in a manner similar to the method used in Pygmalion experiments because of these ethical concerns. Finally, this finding is noteworthy because the research involved only a single exposure to a treatment involving an instructor's brief comments regarding his expectations to students. The consequences resulting from repeated exposure would likely have been even more dramatic.

The findings regarding the unrelated task (origami cup folding) are both interesting and provocative. It appears that students who were told they were inferior in a cognitive task (logic puzzles put forward as a test of managerial acumen) may have compensated when presented with a lessskilled task. Similarly, it would seem that students in the "smart" group might have felt that applying a significant effort to such a menial task was unwarranted.

The control group, which produced more origami cups than either treatment group, offers a puzzle. Were these students simply bored and sought to excel as they might in any exercise? If so, why did they not perform better than the treatment group in the cognitive portion of the experiment? Clearly, such questions are food for future research on task design as it applies to expectancy effects.

These composite results also bring to bear the research on self-efficacy as a partial explanation. As Stajkovic and Luthans (1998) explained, one of the strongest antecedents of self-efficacy is past performance. Thus, students who believed they did well on the pretest may have felt more efficacious when taking the posttest. Reynolds (2001) explicated the likely effect of expectations in the selfefficacy/performance link and suggested, as we do here, that more research is needed to understand the related variables.

Limitations

Although the results of this study make important contributions to our understanding of Pygmalionrelated effects, there are several limitations. The participants were undergraduate students in a management course. It is conceivable that graduate students might respond differently. Similarly, we cannot discern whether these effects can be generalized to executive-education or management-development settings. Additionally, the experiments were performed at a top-ranking business school. Would similar effects be found in lesscompetitive academic settings? This remains an empirical question to be answered only by future research.

The experimental design of the study, particularly pertaining to the application of the expectations treatment, also may limit generalizability. While such an approach was necessary to circumvent the ethical obstacles to subjecting participants to low supervisory expectations, this arguably limits the extent to which we can generalize the findings. For example, it would be interesting to introduce lower expectations for instructors regarding students' abilities on a random basis (employing random assignment of students to treatment and control groups) and ensure these are communicated fully to the students; the results might expound the finding reported here.

It is also possible that some participants may have realized or suspected that they were part of an experiment, and therefore, did not respond as they might had no such suspicions been present. Thus, attempts to replicate this research should be conducted in a variety of settings and among different student groups.

CONCLUSION

In 1984, Eden wrote about harnessing Pygmalion. Now with years of related research behind us, we understand well the importance of Eden's assertions: When teachers expect students to perform well, they do. With the findings reported here, we now also have empirical evidence that, when teachers verbalize negative expectations, students' performance is negatively affected. Thus, just as Pygmalion must be harnessed, so Golem must be restrained.

REFERENCES

- Archer, D. 1993. The methodological imagination: Insoluble problems or investigable questions? In P. D. Blanck (Ed.), *Interpersonal expectations:* 337–349. Cambridge, England: Cambridge University Press.
- Aronson, E., & Carlson, V. M. 1968. Experiments in social psychology. In G. Lindzey, & E. Aronson (Eds.), *Handbook of* social psychology: 441-486. Reading, MA: Addison-Wesley.
- Babad, E. 1993. Pygmalion—25 years after interpersonal expectations. In P. D. Blanck (Ed.), *Interpersonal expectations:* 125–153. Cambridge, England: Cambridge University Press.
- Babad, E. 1995. The "teacher's pet" phenomenon: Students' perceptions of teachers' differential behavior and students' morale. *Journal of Educational Psychology*, 87, 361–368.
- Babad, E., Inbar, J., & Rosenthal, R. 1982. Pygmalion, Galatea, and the Golem: Investigations of biased and unbiased teachers. *Journal of Education Psychology*, 74, 459–474.
- Baron, R. M., Tom, D. Y., & Cooper, H. M. 1985. Social class, race, and teacher expectations. In J. B. Dusek (Ed.), *Teacher expectancies:* 251–268. Hillsdale, NJ: Erlbaum.
- Beez, W. V. 1971. Influence of biased psychological reports on "teacher" behavior and pupil performance. Unpublished doctoral dissertation, Indiana University.
- Bolt, M. A., Killough, L. N., & Koh, H. C. 2001. Testing the interaction effects of task complexity in computer training using the social cognitive model. *Decision Sciences*, 32, 1–20.

Brophy, J. E. 1983. Research on the self-fulfilling prophecy and

teacher expectations. *Journal of Educational Psychology* 75, 631–661.

- Chesney, A., & Locke, E. 1991. An examination of the relationship among goal difficulty, business strategies, and performance on a complex management simulation task. Academy of Management Journal, 34, 400–424.
- Cohen, J. 1988. Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Collins, H., & Pinch, T. 1998. *The Golem: What you should know about science* (3rd ed.). Cambridge, England: Cambridge University Press.
- Cooper, H. 1979. Pygmalion grows up: A model for teacher expectations, communications, and performance influence. *Review of Education Research*, 49, 389–410.
- Crawford, K. S., Thomas, E. E., & Fink, J. J. 1980. Pygmalion at sea: Improving the work effectiveness of low performers. *The Journal of Applied Behavioral Science*, 16, 482–503.
- Davidson, O., & Eden, D. 2000. Remedial self-fulfilling prophecy: Two field experiments to prevent Golem effects among disadvantaged women. *Journal of Applied Psychology*, 85(3): 386–398.
- Dawn, W., & Latham, G. P. 1996. The effect of learning versus outcome goals on a simple versus a complex task. Organization Management, 21(2): 236–251.
- Dorner, D. 1991. The investigation of action regulation in uncertain and complex situations. In J. Rasmussen & B. Brehmer (Eds.), Distributed decision making: Cognitive models for cooperative work: 349–354. Chichester, England: Wiley.
- Dusak, J., & Joseph, G. 1985. The bases of teacher expectancies. In J. B. Dusek (Ed.), *Teacher expectancies:* 229–250. Hillsdale, NJ: Erlbaum.
- Easterly-Smith, M., Thorpe, R., & Lowe, A. 1994. *Management* research. London: Sage.
- Eden, D. 1984. Self-fulfilling prophecy as a management tool: Harnessing Pygmalion. *Academy of Management Review*, 9(1): 64–73.
- Eden, D. 1988. Pygmalion, goal setting, and expectancy: Compatible ways to boost productivity. *Academy of Management Review*, 13(4): 639–652.
- Eden, D. 1990a. *Pygmalion in management: Productivity as a self-fulfilling prophecy.* Lexington, MA: D.C. Heath and Company.
- Eden, D. 1990b. Pygmalion without interpersonal contrast effects: Whole groups gain from raising manager expectations. *Journal of Applied Psychology*, 75(4): 394–398.
- Eden, D., & Ravid, G. 1982. Pygmalion vs. self-expectancy: Effects of instructor- and self-expectancy on trainee performance. Organizational Behavior and Human Performance, 30, 351–364.
- Feldman, R. S., & Prohaska, T. 1979. The student as Pygmalion: Effect of student expectation on the teacher. *Journal of Educational Psychology*, 71(4): 485–493.
- Feldman, R. S., & Theiss, A. J. 1982. The teacher and student as Pygmalions: Joint effects of teacher and student expectations. Journal of Educational Psychology, 74(2): 217–223.
- Frese, M., & Zapf, D. 1994. Action as the core of work psychology: A German approach. In H. C. Triandis & M. D. Dunnette (Eds.), Handbook of industrial and organizational psychol-

ogy (2nd ed., pp. 271–340). Palo Alto, CA: Consulting Psychologists Press.

- Kavanagh, M. J. 1972. Leadership behavior as a function of subordinate competence and task complexity. Administrative Science Quarterly, 17(4): 591–604.
- Kierein, N., & Gold, M. A. 2000. Pygmalion in work organizations: A meta-analysis. *Journal of Organizational Behavior*, 21, 913–928.
- King, A. S. 1974. Expectation effects in organizational change. Administrative Science Quarterly, 19, 221–230.
- Latham, G. P., & Seijts, G. H. 1999. The effects of proximal and distal goals on performance on a moderately complex task. *Journal of Organizational Behavior*, 20, 421–429.
- Livingston, J. S. 1969. Pygmalion in management. Harvard Business Review, 47(4): 81–89.
- Locke, E. A., & Latham, G. P. 2002. Building a practically useful theory of goal setting and task motivation. *American Psy*chologist, 57(9): 705–717.
- Lundberg, C. C. 1975. The effect of managerial expectation and task scope on team productivity and satisfaction: A laboratory study. *Journal of Business Research*, 3(3): 189–198.
- McNatt, D. B. 2000. Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, 85(2): 314–322.
- Merton, R. K. 1948. The self-fulfilling prophecy. Antioch Review, 8, 193–210.
- Oz, S., & Eden, D. 1994. Restraining the Golem: Boosting performance by changing the interpretation of low scores. *Journal* of Applied Psychology, 79, 744–754.
- Reynolds, D. 2001. Integrating "My Fair Lady" with foodservice management: A theoretical framework and research

agenda. Journal of Foodservice Business Research, 4(4), 263–302.

- Rosenthal, R. 1969. Interpersonal expectations: Effects of the experimenter's hypothesis. In R. Rosenthal, & R. L. Rosnow (Eds.), *Artifact in behavioral research*. New York: Academic Press.
- Rosenthal, R. 1994. Interpersonal-expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, 3, 176–179.
- Rosenthal, R., & Jacobson, L. 1968. *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development.* New York: Holt, Rinehart & Winston.
- Rosenthal, R., & Rubin, D. B. 1978. Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, 3, 377–415.
- Stajkovic, A. D., & Luthans, F. 1998. Self-efficacy and workrelated performance: A meta-analysis. *Psychological Bulletin*, 124(2): 240–261.
- Sutton, C. D., & Woodman, R. W. 1989. Pygmalion goes to work: The effects of supervisor expectations in a retail setting. *Journal of Applied Psychology*, 74, 943–950.
- Vrugt, A. 1990. Negative attitudes, nonverbal behavior, and selffulfilling prophecy in simulated therapy interviews. *Journal* of Nonverbal Behavior, 14, 77–86.
- Weick, K. 1979. Social psychology of organizing. New York: McGraw-Hill.
- Weick, K. 1984. Small wins: Redefining the scale of social problems. *American Psychologist*, 39, 40–49.
- Wood, R. E., Mento, A., & Locke, E. A. 1987. Task complexity as a moderator of goal effects. *Journal of Applied Pschology*, 17, 416–425.

Dennis Reynolds (der@wsu.edu), is the Ivar Haglund Distinguished Professor of Hospitality Management at the Washington State University College of Business in Pullman, Washington. Reynolds received his PhD in hospitality management from Cornell University where, prior to joining WSU, he served as the J. Thomas Clark Professor of Entrepreneurship and Personal Enterprise. His related research interests include relational exchanges as antecedents to operational efficiency. Copyright of Academy of Management Learning & Education 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.