Journal of Personality and Social Psychology
1985, Vol. 49, No. 1, 287-293

0022-3514/85/$00.75

Improving the Performance of College Freshmen With Attributional Techniques

Timothy D. Wilson
University of Virginia

Patricia W. Linville
Yale University

In an earlier study (Wilson & Linville, 1982), college freshmen were given information suggesting that the causes of low grades are unstable. Compared with a control group, these students did better on both short-term and long-term performance measures. The long-term results, however, tended to be weak or open to alternative explanations. Two replication studies are reported here. Considered together, the three studies (original plus two replications) found that our attributional interventions improved the performance of college freshmen on both short-term and long-term measures. Presenting freshmen with information indicating that the causes of low grades in the first year are temporary led to (a) improvement on sample items from the Graduate Record Exam and (b) increases in actual grades in the semester after the studies were conducted. These results were stronger for males than for females. Possible reasons for this sex difference are discussed.

Wilson and Linville (1982) attempted to improve the performance of first-year college students by changing their attributions about the causes of their performance. Freshmen who were worried about their grades and who had not performed well their first semester received information suggesting that the causes of their poor performance were unstable. This information consisted of statistics indicating that grades typically increase after the first year, plus videotaped interviews with four upperclass students who reported that their grades were low their freshman year but improved thereafter. We predicted that this information would lead students to attribute their own low grades to unstable causes, which, by reducing anxiety about academic performance and increasing expectations about future grades, would improve actual performance (cf. Storms & McCaul, 1976; Weiner, 1974, 1980).

Subjects in the treatment condition—those who received the information about unstable causes—performed better on both short-term and long-term measures than subjects in the control condition. The short-term measure consisted of sample items from the Graduate Record Exam (GRE), administered right after the manipulation and 1 week later. The long-term measures consisted of actual grades earned in semesters after the study was conducted, plus the proportion of students in each condition who left college in the year following the study. Subjects in the treatment condition were less apt to leave college in the ensuing year and improved their grades more one year later as compared with subjects in the control condition.

Recently Block and Lanning (1984) questioned both the validity of Wilson and Linville's measures and the strength of their long-term results. Many of Block and Lanning's (1984) arguments were discussed in our original paper, such as the fact that some of the results were weak (dropout rates) or open to alternative explanations (grade increases). Rather than engaging in polemics on these issues, however, a more fruitful approach is to test the generalizability of the findings by attempting to replicate them. Two such replications have been performed.

Method

Replication 1

Wilson and Linville's (1982) study was replicated exactly, except for the following changes:

Bella DePaulo, Gregory Fischer, and Judy Harackiewicz provided valuable comments on an earlier draft of this manuscript. Some of the research described was supported by National Science Foundation Grant BNS79-21155.

Requests for reprints should be sent to Timothy D. Wilson, Department of Psychology, Gilmer Hall, University of Virginia, Charlottesville, Virginia 22901.

1. Subjects were students at the University of Virginia rather than Duke University and were selected on the basis of their answers to two instead of three questions on an initial survey. Freshmen were eligible for the study if their reported grade point average (GPA) for the first semester was below the median in our survey (*Mdn* = 3.00) and if their response to a question about how much they considered themselves to be a worrier about their academic performance was 5 or above on a 7-point scale, where 1 = *not at all a worrier* and 7 = *very much a worrier* (*Mdn* = 5.4). Forty-five percent of the freshmen in our survey met these criteria. Subjects were contacted by phone several weeks later and asked to participate in a freshmen survey in return for course credit. Fourteen subjects were randomly assigned to the control condition and 25 to the treatment condition, with the restriction that the ratio of males to females and blacks to whites be approximately equal in both conditions.

2. Two experimenters were used: One administered the independent variable, and the other, blind to condition, administered the dependent variables. The dependent measures were given at the first session only; that is, unlike Wilson and Linville (1982), subjects did not return 1 week later and complete the measures again.

3. We included two different treatment conditions. One was identical to Wilson and Linville's (1982), where students received information indicating that grades are low in the freshmen year but improve thereafter. Subjects in a second treatment condition received the same information indicating that grades are low in the freshmen year, but were told nothing about whether or not grades improve in the upperclass years. Our intent was to separate consensus information about poor performance in the first year from stability information about increases in grades after the first year. As it happened, however, subjects in both treatment groups responded very similarly on the dependent measures and the manipulation checks. In retrospect, the interviews of upperclass students reporting that their grades as freshmen were low implicitly communicated stability information. That is, subjects seem to have inferred, even without being explicitly told, that the students' grades must have improved from what they were. This was borne out on questions about what percentage of students subjects thought improve their grades from the first to the upperclass years. Subjects who received only the consensus information were as likely to report improvement as subjects who received consensus and stability information. Therefore, the two treatment conditions were combined and considered as one group who received information that grades are low at first but often improve thereafter. This combined treatment group (10 males, 15 females) was compared with a control group (6 males, 8 females) who received filler statistics and saw videotaped interviews where there was no mention of grades.

Replication 2

Unlike the previous studies, Replication 2 was conducted in the first (as opposed to the second) semester of the academic year. A questionnaire was handed out in introductory psychology classes at the University of Virginia midway through the semester, after students had had midterms in their classes. Subjects were asked what their average grade was in each of their classes so far that semester, as well as our standard question about how much they considered themselves to be a worrier, this time on a 9-point scale (9 = *very much a worrier*). Students were eligible for the study if they were freshmen, if their reported performance in their courses was below the median (*Mdn* = 3.02 for freshmen), and if they responded to the worry question with a 6 or higher (*Mdn* = 7.01). Thirty-eight percent of the freshmen met these criteria. Twenty subjects (5 males, 15 females) were randomly assigned to the control condition, and 21 (5 males, 16 females) were randomly assigned to the treatment condition. (It should be noted that in both replications there were more females than males in our pool of potential subjects. Eligibility for the study was not differentially related to sex of subject).

Though the treatment manipulation was very similar to the ones in the previous studies, the cover story was quite different. Subjects were seen in groups of between four and six people, ostensibly as part of a large-scale project designed to give high school students more information about what to expect their first year of college. Information about the first year was being gathered from University of Virginia students, they were told, which would be distributed to high school students around the state. Subjects were then given statistics and shown a videotape, ostensibly to show them the type of data we would be giving to the high school students. Subjects in the control condition received filler statistics and heard upperclass students on the videotape respond to filler questions, with no mention of grades. Subjects in the treatment condition received the same information, with the addition of responses on the videotape and statistics indicating that grades are often low in the freshman year but increase thereafter.

Subjects then filled out a questionnaire that would supposedly be shown to high school students. This questionnaire contained several filler items consistent with the cover story, for example, how different the social life was at college versus high school. To strengthen the treatment manipulation, subjects in the treatment condition were asked, after completing the questionnaire, to write an essay explaining why their grades were lower than they would have liked them to be. Subjects were told to answer this question as honestly as possible. The experimenter explained, however, that she was trying to get a variety of explanations to show the high school students and would appreciate it if people could incorporate one of the following reasons for poor performance in their essays: Not knowing how to take college tests or exams, not knowing how to choose college courses, and unpleasant living conditions. Each of these causes had been rated by 13 pretest students as being very unstable, using Russell's (1982) Causal Dimension Scale. Subjects in the control condition did not write an essay.

All subjects then were taken to a separate room to take part in what was ostensibly an unrelated study. A second experimenter, blind to condition, explained that she was conducting a survey on college students' attitudes. The same dependent measures used in the previous study were administered, with one exception. The sample items from the Graduate Record Exam were expanded to increase the reliability of this measure. Subjects read two paragraphs and answered a total of 13 comprehension

questions, as opposed to six questions about one paragraph in Wilson and Linville (1982) and Replication 1.

Results

For ease of comparison, the results of all three studies (Wilson & Linville plus two replications) will be presented, as well as the results averaging across studies. Wilson and Linville's (1982) short-term measures were assessed twice, once right after the manipulation and again 1 week later. To allow more meaningful comparisons to the replications, only their results right after the manipulation will be presented.

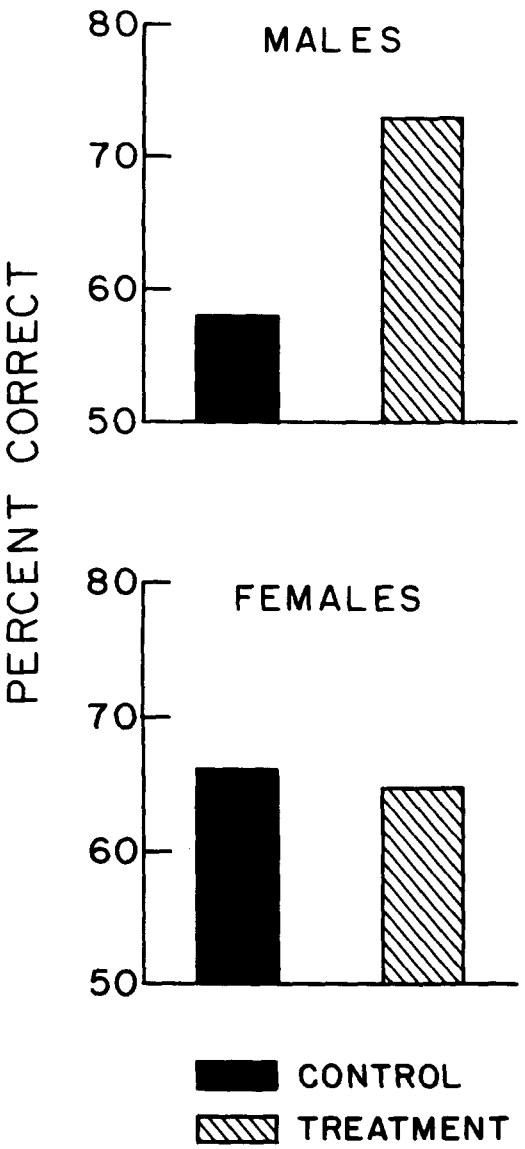

Figure 1. Short-term performance: Percentage of sample items from the Graduate Record Exam answered correctly.

Table 1
Performance on Sample Items From the Graduate Record Exam (Percentage Correct)

Study	Males		Females	
	Control	Treatment	Control	Treatment
Wilson and Linville (1982)	57	83	59	70
Replication 1	50	62	75	60
Replication 2	68	74	64	64

Short-Term Performance

Subjects were given sample items from the Graduate Record Exam. Figure 1 displays the percentage of items answered correctly by condition and by sex averaged across studies (see Table 1 for means from the individual studies). It can be seen that in all three studies, males in the treatment condition performed better than males in the control condition. On average, however, females in the two conditions performed at very similar levels. Averaging across studies, using the method of adding *z*s (Mosteller & Bush, 1954; Rosenthal, 1978), the Condition × Sex interaction was significant, $z = 2.09$, $p = .04$.[1] The treatment manipulation was significant for males ($z = 2.21$, $p = .03$), but not for females ($z < 1$, *ns*). The size of the effect for males, using Cohen's (1977) *d* statistic (averaged across studies), was .45.[2]

Long-Term Performance

Grade point averages were obtained from the transcripts of students who gave us per-

[1] Rosenthal (1978) discusses several methods of combining the results of independent studies. The method of adding *z*s was chosen because it is conservative relative to other methods (e.g., adding *p*s). Other possible methods were tested (adding weighted *z*s and adding *t*s), yielding nearly identical results.

[2] Cohen's *d* statistic is an estimate of effect size in standard deviation units. He recommends that *d*s of .20, .50, and .80 be considered small, medium, and large effects, respectively.

mission to examine their records and who remained in college after the study. As of one semester after the studies were completed these *n*s were 31 of 40, 37 of 39, and 36 of 41 in the Wilson and Linville study and the two replications, respectively.

Grades earned the semester after each study was conducted were compared with grades earned before the studies were conducted. In Replication 2 prior grades were reported midterm grades for the first semester, because the study was conducted at the end of the first semester. For the other two studies prior grades were those earned in the entire first semester, because these studies were conducted in the second semester of the freshman year.

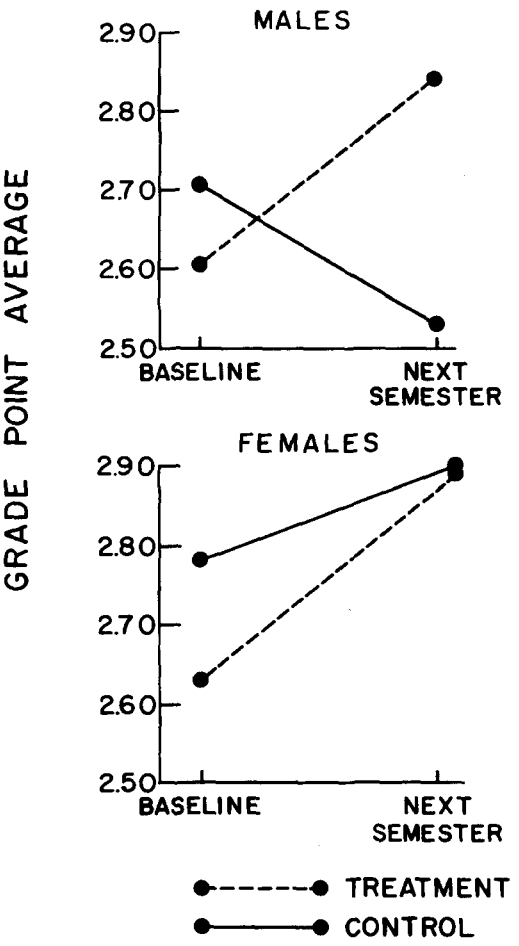

Figure 2. Long-term performance: Actual grades achieved before versus after the studies were conducted (on a 4-point scale, where 4.0 = A, 3.7 = A−, 3.3 = B+, 3.0 = B, etc.).

Table 2
Mean Actual Grades Obtained Before and After Each Study

Study	Males		Females	
	Control	Treatment	Control	Treatment
Wilson and Linville (1982)				
Baseline	2.79	2.41	2.89	2.67
Next semester	2.86	2.97	3.14	2.89
Replication 1				
Baseline	2.76	2.63	2.83	2.65
Next semester	2.56	2.71	2.78	2.86
Replication 2				
Baseline	2.59	2.80	2.61	2.57
Next semester	2.17	2.83	2.79	2.91

Note. All grades are on a 4-point scale, with A = 4, A− = 3.7, B+ = 3.3, B = 3.0, etc. Baseline grades are those obtained before the study was conducted; next semester grades are those obtained the semester after the study was conducted.

Subjects in the treatment conditions improved their grades more than did subjects in the control conditions (cf. Figure 2; means for the individual studies are displayed in Table 2). Averaging across studies, 2 × 2 × 2 (Control vs. Treatment × Males vs. Females × Later Grades vs. Previous Grades) between-within analyses of variance revealed a significant Condition × Time interaction, $z = 2.34$, $p = .02$, $d = .27$. Once again the effect was larger for males ($z = 2.17$, $p = .03$, $d = .47$) than for females ($z = 1.03$, *ns*, $d = .21$), though the Condition × Sex × Time interaction was not significant, $z = 1.24$, $p = .33$.[3]

[3] This analysis of grade improvement excluded final grades from the semesters in which the studies were conducted, because it is unclear whether these grades should be considered as baseline or posttreatment performance. All three studies were conducted toward the end of the semester, when course grades were largely determined; thus if these grades are to be used, it seems logical to consider them as part of the baseline. When this is done, similar results are obtained. The Condition × Time interaction is significant ($z = 2.49$, $p = .01$, $d = .38$), reflecting greater improvement in the treatment condition. The increase was greater for males ($z = 2.49$, $p = .01$, $d = .54$) than for females ($z = .95$, *ns*, $d = .18$). The Condition × Sex × Time interaction was marginally significant, $z = 1.74$, $p = .08$.

Block and Lanning (1984) questioned our original conclusions about grade increases, emphasizing that the baseline grades in the treatment condition were lower than in the control condition of the Wilson and Linville (1982) study. It can be seen in Table 2 that baseline differences were not a problem in the two replications. It is thus important to note that averaging across the two replications only, the grade increases were significantly higher in the treatment conditions ($z = 1.97$, $p < .05$).

Another way of dealing with baseline differences is to perform an analysis of covariance on the posttreatment grades, using baseline grades as a covariate. Averaging across studies the effect of treatment was significant in this analysis ($p = .056$ using pretreatment grades as the covariate, $p = .014$ when the baseline measure discussed in Footnote 3 was the covariate).

Wilson and Linville (1982) reported an additional long-term behavioral measure, namely the number of subjects in each condition who left college in the year after the study. Drop out rates were very low in the two replications; only 1 out of 74 subjects in these two studies left college by the next semester. As we would predict, this one subject was in the control condition (of Replication 1). Considering the three data sets together, 10% (5 of 52) of the subjects in the control conditions left college by the semester following the studies, whereas 2% (1 of 62) left in the treatment conditions. This difference is significant at the .068 level, using Fisher's exact test.

Self-Report Results

The effect of the manipulations on self-reported moods, expectations, and attitudes toward the freshman year were, for the most part, similar to those reported by Wilson and Linville. Averaging across the three studies, the treatment manipulation had little effect on mood, $z = -.77$, *ns* (Condition × Sex interaction, $z = .61$, *ns*). The treatment manipulation appears to have focused attention on the difficulties of the freshman year, in that subjects in the treatment conditions reported being less pleased about their first year ($z = -2.25$, $p = .02$) and reported that it is more difficult to adjust to the first year ($z = -3.13$, $p < .005$) than did subjects in the control conditions.

Averaging across the three studies the treatment manipulation had little effect on subjects' reports about their expectations for their next semester grades ($z = -1.15$, *ns*) or expectations about their grade point average upon graduating ($z = .29$, *ns*). Unexpectedly, it decreased expectations about performance in the semesters the studies were performed ($z = -2.34$, $p = .02$), perhaps due to the fact that the manipulation focused attention on how difficult the freshman year can be.

Discussion

Communicating to college freshmen that the causes of low grades are temporary has been shown to have beneficial effects, on both short-term and long-term performance measures. These effects were stronger for males. Though the reason for the sex difference is not entirely clear, responses to manipulation checks suggest an answer: Females were more apt to find out on their own that the causes of poor grades are unstable. That is, our message about the causes of academic performance in the first year may have come as a surprise more to males than to females. In each study, subjects were asked to estimate the percentage of students at their university who improve their grades from their first year to their upperclass years. This question was included as a manipulation check. If our statistics and videotaped interviews had the intended impact, subjects in the treatment condition should have estimated that a large percentage of students improve their grades after the first year. Averaging across studies, the treatment manipulation increased these estimates only for males (Condition × Sex interaction, $z = 2.10$, $p = .04$). Females already believed that this percentage was relatively high, thus the treatment information may have had little impact (cf. Figure 3 and Table 3).

This result may clear up a puzzling aspect of Wilson and Linville's (1982) results. People experiencing problems are often given consensus information ("a lot of people have

that problem") and stability information ("don't worry, things will get better") by others. Why, then, did our treatment manipulation have a beneficial effect, when it communicated the same information that students may well have been getting from friends, advisers, and parents? Females have been shown to be more likely to self-disclose (Cozby, 1973) and to seek out help for their problems (DePaulo, 1982), thus the females in our studies may have been more likely to elicit information from others about the instability of academic problems in the freshman year. This would explain why our manipulation worked better for males: Females may have gotten the "treatment" information on their own.

Table 3
Reported Percentage Who Improve Grades From First Year to Upperclass Years

Study	Males		Females	
	Control	Treatment	Control	Treatment
Wilson and Linville (1982)	62	81	67	71
Replication 1	44	75	62	70
Replication 2	53	63	62	67

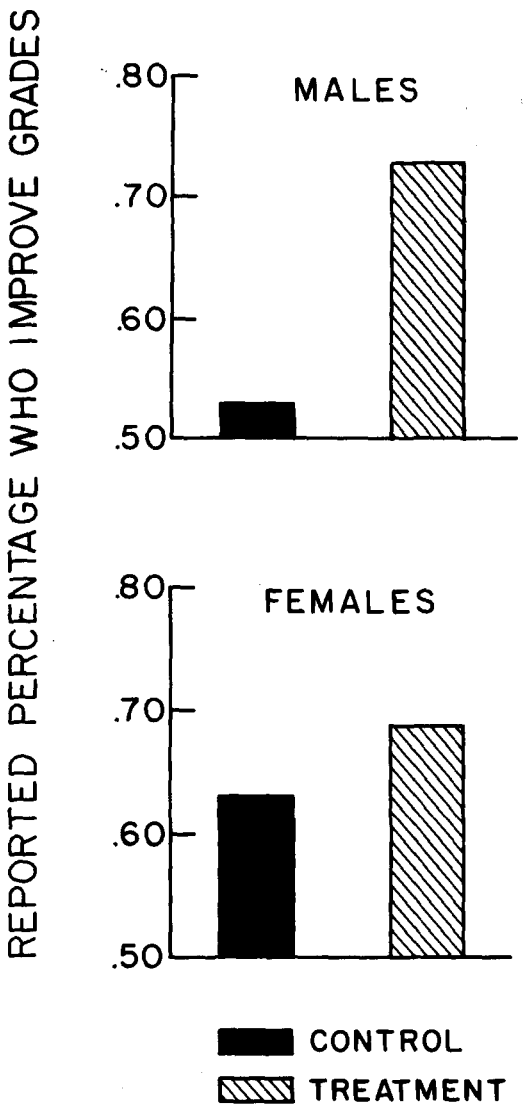

Figure 3. Manipulation check: Reported percentage of students who improve their grades from the freshman to the upperclass years.

Were the effects of our attributional interventions "dramatic," as we suggested earlier (Wilson and Linville, 1982), but questioned by Block and Lanning (1984)? Averaging across the three studies, the magnitudes of the effects of the treatment manipulations on short-term and long-term performance were not huge (for males, *d*s = .45 and .47, respectively). On the other hand, it is encouraging that a simple, one-time intervention led to long-term improvement, even if the size of the effect was modest. We can now conclude on the basis of three studies that an inexpensive, straightforward procedure has been found to help people (especially males) who are concerned about their academic performance.

References

Block, J., & Lanning, K. (1984). Attribution therapy requestioned: A secondary analysis of the Wilson–Linville study. *Journal of Personality and Social Psychology, 46,* 705–708.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin, 79,* 73–91.

DePaulo, B. M. (1982). Social-psychological processes in

informal help seeking. In T. A. Wills (Ed.), *Basic responses in helping relationships.* New York: Academic Press.

Mosteller, F. M., & Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology: Vol. 1. Theory and method* (pp. 289–334). Cambridge, MA: Addison-Wesley.

Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin, 85,* 185–193.

Russell, D. (1982). The Causal Dimension Scale: A measure of how individuals perceive causes. *Journal of Personality and Social Psychology, 42,* 1137–1145.

Storms, M. D., & McCaul, K. D. (1976). Attribution processes and emotional exacerbation of dysfunctional behavior. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attributional research* (Vol. 1, pp. 143–164). Hillsdale, NJ: Erlbaum.

Weiner, B. (1974). *Achievement motivation and attribution theory.* Morristown, NJ: General Learning Press.

Weiner, B. (1980). The role of affect in rational (attributional) approaches to human motivation. *Educational Researcher, 9,* 4–11.

Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology, 42,* 367–376.

Received November 8, 1983
Revision received February 21, 1984 ■