

**Reports on the Assessment of the ReActing to the Past Pedagogy,
1999-2006**

Steven S. Stroessner
Department of Psychology
Barnard College, Columbia University

Contents

Summary Report on the Assessment of the ReActing to the Past Pedagogy: 2003-2006.....	1
Full Report on the Assessment of the “ReActing to the Past” Pedagogy (1999-2002):.....	7

Summary Report on the Assessment of the ReActing to the Past Pedagogy: 2003-2006

Steven S. Stroessner
Associate Professor of Psychology
Barnard College

March 2007

Abstract

An assessment of the ReActing to the Past pedagogy as implemented at affiliated campuses confirms that the method produces several reliable effects. Across campuses and over time, students who participated in a course utilizing the ReActing pedagogy demonstrate improvement in rhetorical skills and no difference in writing skills compared with controls. They also showed increases in self-esteem, decreased belief that they control events in their lives, and higher levels of emotional empathy for others. The ReActing pedagogy tends to be enjoyed to a greater degree by individuals high in self-esteem, mastery, and confidence in speaking and low in emotional empathy and social avoidance, but these individual differences did not predict performance.

An Assessment of ReActing to the Past: 2003-2006

The evaluation project during the years 2003-2006 examined the impact of the ReActing pedagogy in two different ways, depending on how the pedagogy was utilized on the different campuses involved. On some campuses (i.e., Barnard, Trinity, Smith, and Pace), the ReActing pedagogy was used in elective classes that students were (more or less) randomly assigned, allowing comparison between students who were enrolled in the course and students who took courses that used other, typically more traditional teaching methods. The assessment at the campuses addressed the question of replication: did the results obtained earlier at Barnard emerge at other colleges and universities, given different instructors, curricula, and local norms? On one other campus involved in the assessment, Loras College, the ReActing pedagogy was used differently. Rather than offering such as course as an elective, all first-year students were required to take a course using the ReActing pedagogy. This different implementation of the ReActing model meant that there would be no control group available for comparative purposes, and causality would be difficult to infer. Accordingly, the assessment method was modified to address a different question: what stable individual differences predict whether students will enjoy the pedagogy and do these same differences predict who will be successful in it? The results of each assessment will be briefly described here in turn, and a full report will be prepared and available in the coming year.

To assess replicability, data were collected at Barnard, Smith, and Trinity Colleges by student researchers (Alexis Whittaker, Jessica Keating, and Marissa Wong, respectively) who were trained in experimental design and implementation. Unfortunately, data collection at Pace University was not successful as there was difficulty identifying and retaining a student research assistant at this location. At the other three colleges, data were collected twice per semester, once at the beginning and once at the end of the semester, allowing comparisons between students who completed a course using the ReActing to the Past pedagogy versus more traditional First-Year Seminar formats. As was the case in the first assessment at Barnard from 1999-2002, two different types of data were collected. First, students' skill development was assessed by having them generate spontaneous writing and rhetoric samples (via tape recording). The former task involved them writing a short essay on an assigned topic (e.g., "If you had unlimited funds and freedom, what would you do with a day?") and the latter involved making

an persuasive argument to another person regarding a controversial topic where the students were provided with arguments supportive of a position on a particular issue (e.g., “there should be increased gun control legislation”). The quality of these writing and rhetorical samples were judged by evaluators who were unaware of the time of the assessment (beginning or end of the semester) and of whether the student was enrolled in a ReActing or alternate seminar. In other words, the study utilized double blind procedures. Second, students completed a set of standard and validated psychometric scales designed to assess psychosocial variables that have either been associated with success in academic environments (self-esteem, locus of control, confidence as a speaker, implicit person theory, mastery) or variables – both positive and negative – that focus groups had suggested might be affected by the pedagogy (emotional empathy, Machiavellianism, social avoidance and distress). The data were analyzed using ANOVA procedures with pedagogy (ReActing vs. Controls), campus (Barnard vs. Smith vs. Trinity) and semester (Spring 04 vs. Fall 04 vs. Spring 05) as between-participant independent variables and time (beginning vs. end of semester) as a within-participant factor.

The results show that most of the findings obtained originally at Barnard also emerged in this second wave of data collection, replicating across different campuses, different specific classes and instructors, and over time. With regard to the skill measures, the findings showed that ReActing and control students had statistically equivalent spontaneous writing and rhetorical skills at the beginning of the semester. Rhetorical speaking improved significantly for students in the ReActing courses over the semester whereas they did not for the Controls.

Speaking Skills (n = 82) (*M (SE)*)

	Beginning of semester	End of semester
ReActing	16.2 (.55)	17.7 (.44)
Controls	16.7 (.38)	17.0 (.31)

Writing skills improved significantly over the semester, but that effect did not vary as a function of any of the independent variables of interest.

Writing Skills (n = 112) (*M (SE)*)

	Beginning of semester	End of semester
ReActing	3.9 (.20)	4.2 (.15)
Controls	4.2 (.14)	4.4 (.10)

Consistent with the earlier assessment conducted only at Barnard, the ReActing pedagogy appears to enhance rhetorical skills but does not negatively affect students' spontaneous writing abilities. These effects were not qualified by campus or by semester in which the course was taken.

Data for the psychosocial measures also produced several results consistent with those that emerged in the Barnard assessment. In that study, it was found that participation in the ReActing pedagogy produced significantly higher self-esteem, greater emotional empathy, a more external locus of control, and a greater belief that individuals' traits and characteristics are malleable. In the current assessment, similar effects were obtained for of these variables, and these effects were not qualified by campus or semester:

Self esteem increased for students who took ReActing but did not change for Control Students.

Self-Esteem (n = 112) (*M (SE)*)

	Beginning of semester	End of semester
ReActing	3.17 (.11)	3.37 (.11)
Controls	3.22 (.07)	3.24 (.07)

Participants in ReActing courses reported feeling that they had less control over their own outcomes (i.e., a more external locus of control) whereas Control participants reported believing they had more control over events in their lives.

Locus of control (n = 112) (*M (SE)*)

	Beginning of semester	End of semester
ReActing	1.51 (.04)	1.57 (.05)
Controls	1.56 (.03)	1.52 (.03)

Emotional empathy also produced a ReActing x Time interaction, but the pattern reflected an unexpected *decrease* in emotional empathy by Control participants. ReActing participants did not show this decrease and, in fact, showed some evidence of increase in empathy (though not a standard levels of statistical significance).

Emotional Empathy (n = 112) (*M (SE)*)

	Beginning of semester	End of semester
ReActing	5.11 (.18)	5.18 (.17)
Controls	5.31 (.12)	5.14 (.11)

Given that all students at Loras College completed a course using the ReActing pedagogy, comparisons with other FYS were precluded. Instead, we focused on that research on identifying individual differences in liking of the course (reflected in course evaluations) and performance in the course (reflected in earned grade). To do so, psychosocial scales were administered at the beginning of the semester, and the responses to these scales were used to predict student satisfaction with and performance in the course. These analyses indicated that students (n = 119) higher in self-esteem ($r = +.27$), mastery ($r = +.24$), and confidence as a speaker ($r = +.19$), and low in emotional empathy ($r = -.21$) and social avoidance ($r = -.15$) were most likely to rate the course positively. Despite the finding that some students were more likely to enjoy the course, actual performance was not predicted by any measured variables. All students could perform well in the course, regardless of whether they enjoyed its format.

**An Assessment of the “ReActing to the Past” Pedagogy (1999-2002):
Consequences for Belief, Competency, and Academic Performance**

Steven S. Stroessner
Department of Psychology
Barnard College

Laurie Susser
Barnard College

February 20, 2002

Abstract

Consequences of participating in a First-Year Seminar (FYS) utilizing a novel pedagogy were assessed over three years. The pedagogy, “ReActing to the Past,” involves collaborative role playing in three different “games” addressing classic conflicts and moments of historical importance. Reactions of student participants in focus group interviews confirmed the popularity of the method compared with more traditional pedagogies, although several concerns about the course were expressed. These reactions formed the basis of a more systematic analysis of the effects of participation on psychosocial variables, skill development, and subsequent academic performance. Compared with controls enrolled in other FYSs, students from “ReActing to the Past” showed elevated self-esteem, higher levels of empathy, a more external locus of control, and greater endorsement of the belief that human characteristics are malleable. Rhetorical skills were enhanced, but writing skills were not affected. Preliminary analyses of transcripts suggest a potential benefit in the four semesters following FYS for students enrolled in “ReActing to the Past.” These results suggest a number of benefits and generally assuage concerns about potential negative consequences of the pedagogy. Implications of these findings and future directions for research are discussed.

An Assessment of the “ReActing to the Past” Pedagogy: Consequences for Belief, Competency, and Academic Performance

Over the last three years, the Fund for the Improvement of Postsecondary Education (FIPSE) has supported the continued development of a novel pedagogy originally conceived by Professor of History Mark C. Carnes at Barnard College. This pedagogy, termed “ReActing to the Past,” relies exclusively on collaborative role-playing “games” in First-Year Seminar (FYS) to facilitate a student’s exploration of cultures, controversies, and historical moments of importance. This pedagogy attempts to increase student engagement by allowing inquiry under conditions where student concerns about self-presentation and evaluation are reduced, where the Instructor as the primary source of information and the arbiter of “truth” is eliminated, and where basic issues of truth and justice are examined within dramatic and psychologically involving contexts. The expectation was that students would show great enthusiasm for this method of learning and that the method would produce systematic changes in student beliefs and skills. These expectations were assessed over three years using varied by convergent methods and measures, and the results of that assessment are the focus of this report.

Description of the Pedagogy

Before discussing the assessment in detail, we briefly describe the pedagogy. Three games are played in each FYS, and different combinations of games are used within the different Seminars. To date, six different games have been used: “The Trial of Socrates,” “Succession Crisis in Ming China,” “The Trial of Anne Hutchinson,” “French Revolution: The Constitution of 1791,” “Jung vs. Freud: Whither the Psychoanalytical Institute,” and “India in the 1940s.” Numerous instructors have used the pedagogy since its inception. During the period of this assessment, 10 faculty supervised “ReActing to the Past” FYSs.

For any game, a student is randomly assigned a role and prepares for the game by studying prepared materials, reading primary source materials, and conducting additional research. Oral and written work is required from each student, regardless of the role assigned. Typically, roles are allocated to create factions of individuals with similar interests or agendas within a game (e.g., “Prosecutors” in the trial of Socrates) – these factions typically provide a student with research partners outside of class and allies within the playing of the game. A number of students are given roles as “indeterminists” who do not have a specific agenda in the

game (e.g., “People of Athens”). These individuals are the focus of the principle actors in any game, for their votes at the end of each game determine the winner. The winning team receives a slight bonus for the oral component of their course grade. After the game is completed, a debriefing procedure is used to review the game and to compare the events within the game against actual historical events.

Overview of the Assessment

Given the complex and multi-faceted potential consequences of “ReActing to the Past,” it was clear that a variety of methods and measures would be needed to perform an appropriate assessment. The variety of approaches was also necessitated by the initial ambiguity regarding the scope of possible effects. Although certain effects were predicted, it was quite possible that other consequences would arise. Reliance on a predetermined, rigid methodology was bound to lead to such oversights. Accordingly, the assessment began with open-ended interviews with students who had participated in a seminar using the pedagogy. The comments expressed during these interviews, conducted in the Spring of 1999 with students who participated in FYS in the Fall 1998 semester, served as the basis for designing a more systematic, detailed assessment methodology. Assessments using these quantitative methods were then conducted over the next five semesters (Fall 1999-Fall 2001). Finally, to assess academic performance, we collected transcripts in the Spring semester of 2002 of all current juniors and seniors who had taken “ReActing to the Past” and compared GPA over time with a control group who had taken traditional FYSs. If “ReActing to the Past” impacted students’ beliefs, competencies, and academic performance as hypothesized, the varied methods used

STUDY 1: PARTICIPANT REACTIONS

We began our assessment by gathering reactions of students who had completed a “ReActing to the Past” FYS in the relatively anonymous and unconstrained environment of a focus group. To do so, we asked students who had completed such a seminar to attend a discussion session about the course and its methods. The students were assured that neither faculty who taught the course nor college administrators would know who attended the session or the source of any comment. In addition, focus groups were led by a same-age female who had no affiliation with the college or faculty. The discussion leader used a script developed to assure that a variety of issues were discussed, but students were free to respond to inquiries in any way they wished and were free to introduce topics for discussion that were not incorporated in the

script. The goal of these focus group sessions was to solicit feedback, both positive and negative, about “ReActing to the Past” from students who had recently completed it.

Method

Participants. Three groups of students, each from a different FYS with a different Instructor, participated in focus group session. Each group included 8-12 students who had taken the same FYS in the previous semester. Students were compensated \$10 for participating in the focus group.

Procedure. Approximately one week before the first scheduled focus group session, students who had completed “ReActing to the Past” in the previous semester were contacted by email and invited to attend a discussion about the course. Students were informed that the discussion would be recorded, but they were also assured that their individual responses would be kept confidential from their FYS instructor. Of the students contacted, 62% attended a session.

Sessions were conducted in a research lab located on campus. The lab contained a two-way mirror, allowing observation and recording of the discussion sessions. Focus groups were led by a 22-year old female interviewer, an employee of a local market research firm and recent graduate of a large northeastern public university. To begin the discussion, the leader introduced herself and explained that she had been recruited by a member of the faculty who was conducting research on “ReActing to the Past.” The research was being conducted to identify both the strengths and weaknesses of the pedagogy, and students were told that they should feel free to express both positive and negative reactions. Students again were assured that their responses would remain confidential. The discussion leader then began by asking a number of questions developed in collaboration with the researchers. Although all questions were eventually posed to each discussion group, the discussion leader allowed the discussion to flow naturally to topics that had not been anticipated, pursued intriguing comments, and sought to clarify reactions when they were ambiguous. Focus group sessions lasted two hours, after which students were thanked and compensated.

Results and Discussion

There was general consensus on the majority of issues discussed. Students were generally very positive about the course. The majority of students felt that the class met or exceeded their expectations. This was true even for students who had not requested placement in

the “ReActing to the Past” course and those who expressed having a fear of public speaking. Several students stated that the seminar was their best educational experience to date.

Students consistently cited several consequences they believed were unique to the pedagogy. First, students suggested that the game-playing method provoked a deep understanding of historical and philosophical issues. Students claimed that much research was required to play a role effectively, and the debates in the classroom were complex because of the need to simultaneously consider the arguments and agendas of numerous parties. When the discussion leader questioned whether their understanding might perhaps be biased by the role played in a game, the majority of students emphasized that success depends on the ability to understand differing points of view. Second, students agreed that there tended to be a reduction in self-consciousness as roles were developed and games progressed. This reduction in self-consciousness was suggested to reduce fear of speaking and debating, but it also led some to embrace arguments and positions radically different from their own. When asked whether the roles actually changed their attitudes, most students reported that they continued to hold their former beliefs but could better appreciate varying points of view. Several students also mentioned that they found themselves at times behaving toward others in ways they might not normally find acceptable. Third, there was consensus that the debates tended to be intense and involving. Although most students reported that the intensity of debates was invigorating and motivated additional research, several students also mentioned that they found themselves behaving toward others in ways they might not normally find acceptable. These students tended to report that they responded by making an extra effort to reach out positively to the other students.

There was disagreement on several issues. First, some students expressed concern that writing skills suffered in the class. These students suggested that the emphasis in the class was on oratorical skills, and that feedback on written assignments was more sporadic and inconsistent across instructors. Other students felt that there were sufficient writing assignments and feedback. They suggested that their writing skills had improved, and that the emphasis on oratory simply only appeared to de-emphasize writing. Second, there was some discussion regarding the lingering effects that might result from the intensity of the games. As described above, some students felt they had successfully compensated for any conflict that arose in the games, but others expressed surprise that some negative feelings toward others persisted after the

class was completed. Third, there was some discussion about whether the pedagogy was equally successful across instructors. Some felt that the games were so involving that variability in instruction was unimportant. Others felt that instructors played crucial roles in the quality of the games, even if their contribution was often undetectable. It was difficult to gauge the accuracy of these opinions, given that each session was made up of students who had the same instructor.

The reactions of students helped determine the nature and scope of the quantitative, controlled research to come. It was apparent, at least based on the students in attendance, that the pedagogy was very popular. Therefore, it seemed imprudent merely to quantify the level of student enthusiasm. Instead, student comments had identified a set of benefits that might flow from the pedagogy, including increased self-confidence and empathy. Students also suggested a number of possible deleterious consequences, including impairment of writing skills and an increase in anti-social or manipulative behavior. Finally, student comments also raised a number of questions that warranted further investigation such as the importance of variability in instruction and whether the time and energy dedicated to the seminar was detracting from studies in other courses. These questions were investigated in the subsequent studies.

STUDY 2: EFFECTS ON PSYCHOSOCIAL VARIABLES

Student reactions and theorizing about the nature of role playing helped identify a set of possible consequences of the “ReActing to the Past” pedagogy. Given the demonstrated unreliability of self-report regarding cognitive and affective processes (Nisbett & Wilson, 1977), it was clear that these consequences could be firmly established only through the use of methodologies more conducive to establishing the reliability of effects and allowing inferences about causality. Accordingly, we chose to approach the FYS experience as a “manipulated independent variable,” in that students were assigned, more or less randomly, either to a seminar that employed the “ReActing to the Past” or an alternate (i.e., more traditional) pedagogy. Students from seminars using both pedagogies then would be asked to complete a set of established instruments and tasks that are designed to assess beliefs and competencies that might be impacted by participating in “ReActing to the Past.” Comparisons between the two groups would allow identification of reliable consequences of each pedagogy.

There were two primary benefits to this approach. First, this approach allowed a systematic test regarding the validity of student reactions to the pedagogy. Although students had firm beliefs about the effects associated with role-playing (and, it must be acknowledged,

these beliefs might have important consequences regardless of their validity), people are often inaccurate in judging both the nature and the extent of the impact of experience. It also would be impossible for students to accurately compare the impact of two pedagogies used in FYS since they had experienced only one method of instruction. Second, by using instruments that are widely accepted in the research literature, any effects identified might be more easily integrated with existing knowledge. The impact of role-playing (although previously used only on a short-term basis) has long been of interest to researchers in education, sociology, philosophy, and psychology. Findings obtained by systematic, reliable, and accepted methods could easily complement existing literatures.

Despite these benefits, several weaknesses must be recognized. First, the treatment of FYS as an independent variable is potentially misleading. Ideally, differing levels of an independent variable differ in only one respect. Any effects associated with the independent variable, therefore, could be reasonably viewed as caused by the variable. FYSs, regardless of the general pedagogy they employ, differ in a multitude of ways. Although any effects associated with a pedagogy can be viewed as “caused” by it, the exact causal mechanism cannot be identified. In addition, the aggregation of data collected from courses using role-playing and those that do not should not be taken to imply that any other sources of variation in content or quality are unimportant. Nothing could be further from the truth. It is undoubtedly true that some FYSs using more “traditional” methods are exceptional, and it is certainly possible that some seminars utilizing role-playing could be abject failures. Because of the need to control for time, maturity, impactful events and other factors, however, data from the “ReActing to the Past” courses must be compared with data from courses not using this method, even at the risk of overgeneralizing within each group. It should be clear, however, that this is essentially a conservative strategy. To the degree that FYSs using pedagogies not based on role-playing are exceptional, this should undermine the ability to identify benefits of the “ReActing to the Past” pedagogy.

To see whether “ReActing to the Past” produced some of the expected benefits, students in the second study completed a series of scales embedded in a questionnaire. These scales were selected because they measure some of the variables expected to be affected by the role-playing method. All scales also had desirable psychometric properties including high reliability and validity. Participants in the study were recruited during the last week of the semester from two

sources: approximately half of the students were enrolled in one of the “ReActing to the Past” FYSs and half were recruited from FYSs not using the role-playing. The data collected from these groups were then compared to assess how the role-playing pedagogy affected the beliefs and attitudes of students.

Method

Participants. Students currently enrolled in “ReActing to the Past” or other concurrent FYSs were invited by email to participate in a confidential, one-hour research project to take place in the last week of the semester. Students were told that the project focused on the experiences of students in FYS and that they would be paid \$10 for participating. All students in “ReActing to the Past” classes during a two-year period were invited ($n = 184$), and students were randomly chosen to be invited from the rosters for the other contemporaneous FYSs ($n = 368$) until sample sizes were approximately equal in any semester. Of those invited to participate, 87 students from “ReActing to the Past” and 73 students from concurrent FYSs consented to participate. Therefore, the participation rate for “ReActing” students was modest (47 %) and quite low for the students from other FYSs (20 %).

The different rate of responding could suggest differences between these two groups independent of pedagogical variation. To assess this possibility, a subset of these students ($n = 99$) was invited to attend two sessions, one at the beginning and one at the end of the semester. Data were collected from these students to test the assumption that students from both types of seminars would respond similarly to the scale questions at the beginning of the semester, before the effects of FYS could be manifest. The absence of differences at this point would mitigate the likelihood that differential responding alone could account for any findings obtained at the end of the semester.

Procedure. Students who responded to the email invitation were asked to attend one of several one-hour testing sessions during the coming week. Students were again told that the focus of the research was on their FYS experience, that their responses would be collected without identifying information, and that their data would be kept confidential.

At the scheduled time, students arrived at a campus research lab and were greeted by a female, undergraduate research assistant. A brief description of the project was provided, and students were informed of their rights as a participant in research. All students then signed a form providing their informed consent to participate. The research assistant then asked each

student which course she was taking for FYS and noted this information for data analytic purposes.

Students were then given a questionnaire packet containing items from 9 different established scales. There were a total of 136 items, and items from different scales appeared in a mixed, random order (although scales that used an identical response scale were grouped together). The scales were included either because they were designed to measure some of the variables that might be impacted by participation in a role-playing pedagogy or for exploratory purposes. All measured variables have all been shown to have varied and important consequences for social behavior. The 9 scales included in the questionnaire were designed to assess self-esteem (Rosenberg, 1965), confidence as a speaker (Paul, 1966), emotional empathy (Mehrabian, 1996), orientation to life (Antonovsky, 1987), Machiavellianism (Christie & Geis, 1970), mastery (Pearlin et al., 1981), locus of control (Rotter, 1966), social avoidance and distress (Watson & Friend, 1969) and implicit person theories (Levy, Stroessner, & Dweck, 1998). Sample items from the 9 scales are presented in Table 1. (Items from the Mehrabian's (1996) Balanced Emotional Empathy Scale – designed to assess “one's vicarious experience of another's emotional experiences” – are not listed at the author's request). Following the scales was a final questionnaire that required students to rank order the importance to them of 15 different social values (e.g., equality, wisdom) (Rokeach, 1973).

Students were told they would have an hour to complete the items, to respond to all items, but not to write their name or any identifying information on the questionnaire. After students completed all the items, they then completed two tasks that were the focus of Study 3. After finishing all the tasks, students were debriefed about the nature of the research, allowed to ask any questions they had, and were paid \$10. For those participants who completed the procedure at the beginning and end of the semester, they were debriefed only after the second session was completed.

Results and Discussion

Preliminary analyses. Three preliminary analyses were conducted before examining the main theoretical questions of interest. First, responses to scale items were analyzed to determine the reliability of each scale and to assess whether elimination of any problematic items was warranted. Cronbach's α was computed for each scale. Reliability was acceptable for all scales (range .71-.91) except for Orientation to Life (Antonovsky, 1987) ($\alpha = .20$). The low reliability

of this scale precluded further analysis. Therefore, all but one of the scales had sufficient reliability to allow further analysis and conclusions.

Second, data from the subset of 99 participants who completed questionnaires at both the beginning and end of the semester were examined. If first-year students were randomly assigned to FYS courses, equivalency of the groups could be assumed. However, students' indicate their preferences for FYSs based on brief descriptions in application forms, allowing the possibility that the "ReActing to the Past" seminars might be preferred by students who differ systematically from students who do not. If these students were different before enrolling in the course, then any conclusions about the causal role of the pedagogy would be invalid. Therefore, the responses of students in "ReActing to the Past" and other FYSs collected in the first week of the semester were compared. There were no statistically significant differences on any measures (all $ps > .10$). Therefore, there was no basis for questioning the assumption of equivalency in groups.

Third, an analysis was conducted to determine the degree of variation in effects across instructors. On the one hand, it was not our intent to quantify the effects of any single instructor. This would be impossible given the small amount of data collected from courses led by any one instructor, and we doubted the wisdom of pursuing such a strategy given that conclusions about causality would be problematic at best. On the other hand, it was imperative to show that any obtained effects were not limited to any one instructor. Specifically, a question was raised near the end of the period of data collection about whether any effects we might obtain relied on the involvement of Prof. Carnes, the individual who developed the pedagogy. To empirically address this issue, we took advantage of the fact that Prof. Carnes was not one of the "ReActing to the Past" instructors in the Fall, 2001 semester. Any effects that emerged from that semester could not be attributed to his involvement that semester. Therefore, the data were coded as to whether Prof. Carnes could have taught the course, and any interaction effects involving this variable would suggest the need to qualify any findings. This analysis, however, produced no effects on any of the measures (all $ps > .10$), allowing us to collapse across this variable in the main analyses. More important, this analysis suggests that the effects reported were not dependent on the involvement of any particular individual.

Primary analyses. One-way Analyses of Variance (ANOVAs) were conducted on each of the measures. Data collected at the beginning of the semester from the subset of students were

included in the analyses to provide a comparison of changes over time. Therefore, the averages of three “groups” (Beginning, ReActing, NonReActing) were compared in any one analysis.

Data from these analyses are presented in Table 2. Results were not significant for three of the measures: Machiavellianism, Confidence as a Speaker, and Mastery. Although there was a concern that the role-playing games in “ReActing to the Past,” with their notion of “winning” and “losing” teams, might facilitate deception and manipulation (i.e., Machiavellianism). The lack of any effect for that variable suggests that these negative consequences do not necessarily flow from the game structure, even given its competitive aspects. The lack of an effect on students’ confidence in public speaking was surprising. Many students in the focus group discussions commented on having increased proficiency in the oratorical skills. One possible reason that the effect did not emerge might be that students in other FYs, in which debate and oratory are less emphasized, might have inaccurate, or even misplaced, confidence in their oratorical abilities. An effect would not emerge if students from the “ReActing” courses were accurately confident about their speaking skills whereas students from the other seminars were inaccurately but equally confident about their abilities. (Findings consistent with this interpretation will be presented in Study 3). Finally, the absence of an effect on Mastery was also surprising, although perhaps to a lesser degree. It was expected that students might develop a stronger sense of their capabilities by participating in role-playing. An examination of the items comprising this scale, however, suggests that the construct measures a blend of beliefs about personal efficacy and beliefs about controllability. It is quite likely that the role-playing pedagogy might affect these two beliefs in very different ways. Responses to the Locus of Control scale (reported below) should clarify whether beliefs about controllability were affected.

The pedagogy produced clear effects on the other variables. The analyses of responses to the Rosenberg (1965) Self-Esteem scale was significant, $F(2, 247) = 3.85, p < .05$. A post-hoc Scheffé test revealed that students who had completed “ReActing to the Past” had elevated self-esteem compared with students in the other FYs and compared with students who completed the measure at the beginning of the semester. Self-esteem is associated with a number of benefits, although excessive esteem has been shown to produce some types of anti-social behavior (Tesser, 2001). The average self-esteem of students obtained here is well within normal range, suggesting that beneficial consequences are likely to dominate.

The analysis of responses to Mehrabian's (1996) Balanced Emotional Empathy Scale was produced a significant effect, $F(2, 247) = 3.98, p < .05$, although the nature of that effect was more complex. Although neither the "ReActing" nor "NonReActing" group differed significantly in empathy from the students at the beginning of the semester, they differed significantly from each other. The pattern of data suggests that students in the "ReActing" seminars showed a slight increase in emotional empathy compared with the beginning of the semester. Students in the other FYs, in contrast, showed a slight decrease in empathy compared with the beginning of the semester. Although the former effect was expected, the latter effect was not.

A similar significant effect emerged on the measure assessing students' beliefs about the malleability of human characteristics (Levy, Stroessner, & Dweck, 1998), $F(2, 247) = 3.04, p = .05$. Neither means collected at the end of the semester were significantly different from the beginning of the semester, but they differed significantly from one another. Students who had completed a "ReActing" course were more likely than students in other seminars to endorse the view that human characteristics were malleable. This belief has been associated with a variety of benefits in both social and academic contexts (Dweck, 1999). For example, people who tend to believe that characteristics are malleable rather than stable are more likely to persist after failure and are less likely to endorse stereotypes.

A different but significant pattern emerged on responses to Rotter's (1966) Locus of Control scale, $F(2, 247) = 3.57, p < .05$. Responses to this scale, which measures beliefs about whether outcomes are determined by forces that are internal or external to the self, showed that "ReActing" had no effect. Instead, students in the other FYs were more likely to endorse the view that they controlled their own outcomes. An internal locus of control, so long as not too extreme, is generally beneficial for self-regulation and for responding to challenges. It is not particularly surprising, however, that students in "ReActing" did not develop a stronger internal locus of control, given the frequent occurrence of random and chance events in the games scenarios. In fact, what might be surprising is that these students did not develop a more external locus of control as the semester progressed.

A highly significant but puzzling effect emerged on the measure designed to assess social avoidance and distress, $F(2, 247) = 107.07, p < .001$. Students from both groups showed a substantial increase in avoidance and distress at the end compared with the beginning of the

semester. The most plausible account of this finding is that data were collected at the end of the semester, just before finals were to take place. It is very likely that all students, regardless of their FYS, were experiencing high levels of stress that might interfere with normal social functioning.

An analysis of the responses to Rokeach's (1973) Value Scale did not produce any significant effects. The values endorsed by students in both types of seminars were similar and did not vary from the beginning of the semester.

In sum, the analyses suggest that participating in "ReActing to the Past" was associated with a number of psychological benefits. Students enrolled in a FYS utilizing the role-playing method showed elevated self-esteem, greater empathy with the needs and feelings of others, and greater agreement with the belief that human characteristics are amenable to change across time and contexts. Research has shown that each of these effects can provide immense benefits in both academic and social interactions.

STUDY 3: EFFECTS ON SKILL DEVELOPMENT

Study 2 identified a number of possible salutatory psychological and social consequences of participating in "ReActing to Past." Although these effects may be important, it was also important to determine the impact on skills central to academic success. Students who had completed the course had taken note of the benefits for oratorical skills. Nonetheless, some students also expressed concern that writing skills were not developed in "ReActing" seminars to the same degree as in other FYSs.

Study 3 was designed to assess the development of writing and rhetorical skills under differing rubrics of the FYS. Data for Study 3 were collected at the same time as for Study 2. After students completed the questionnaires that were the focus of Study 2, they then completed two tasks designed to assess writing and oratorical skills. To test extemporaneous writing, students were asked to write a brief essay on a topic we provided. To test rhetorical skills, students were presented with information on a controversial topic, were asked to consider an argument consistent with a particular stance on the topic, and then verbally presented that argument for several minutes into a tape recorder as if they were trying to persuade another person. Researchers unaware of any student's FYS then coded Writing and rhetoric samples. Analysis of these ratings allowed a test of the potential benefits and detriments for skills associated with participating in "ReActing to the Past."

Method

Participants. Participants were the same students (n = 160) who completed Study 2. Two independent evaluators have coded samples from 110 students at this time. Therefore, analyses reported here focus on the writing and rhetorical samples of 70 students from “ReActing to the Past” and 51 students from other FYSS. As in Study 2, 99 of the participants also completed writing and rhetoric samples at the beginning of the semester. Rhetorical samples from 5 students were not collected because of tape recorder malfunctions.

Procedure. After completing the questionnaire packet described in Study 2, participants were told that they would complete two additional tasks. The first task involved extemporaneous writing. Students were given a sheet of paper with the following instruction:

“Assume that it is early spring and you have a week-long break from school. You have an unlimited amount of money and resources to do whatever you would like for the entire week. You are void from all your responsibilities and worries. What would you do during your Spring break? Think about it for a few minutes. Where would you go? Who would you spend time with? What would you do? What goals do you have for the week? Why? Please write a short essay for approximately 10 minutes.”

Students were given ten minutes to complete this task.

After completing the writing task, students were provided with materials for the task assessing rhetorical skills. They were told that they would be given information concerning a controversial topic, that they would be asked to formulate a persuasive message supporting a particular point of view, and that they would present their argument orally into a tape recorder. Students were then provided with the following task:

“Many issues are currently being debated in our society. Two particular issues of concern regard gun control and the decline of quality public education. Assume that a nationally televised debate will be planned for two political candidates to discuss their platforms, yet there will be only enough time allotted to fully discuss one of the two topic mentioned above. Please compose in your head a persuasive argument in favor of discussing gun control, and be sure to address why education should not be the foremost concern. Use information below as suggestions, and please feel free to bring in any new ideas you might have. You have five minutes to prepare. You will be asked to deliver orally a 3-5 minute speech on this topic.”

Students were then given 5 minutes to examine list of arguments favoring and apposing the discussion of gun control. After five minutes, students were prompted to begin recording their oral argument until completed.

Students who completed these tasks at both the beginning and end of the semester received different specific topics for their writing and rhetorical tasks, but the procedures were identical

Results and Discussion

Coding. Two undergraduate research assistants served as coders of the writing and rhetoric samples. The writing samples were coded using a 6-point scale developed by the College Board for use in scoring the SATII: Writing Test (available on-line at <http://cbweb6.collegeboard.org/writewellCB/student/sat/html/satsguid.html>). The rhetoric samples were coded using the Competent Speaker Speech Evaluation Form (Morreale, Moore, Taylor, Surges-Tatum, & Hulbert-Johnson, 1994). Both coders were blind regarding any student's FYs, and ratings were made independently. Correlations of the ratings were then computed. For the writing and rhetorical samples, coders ratings were weakly, but significantly correlated ($r_s = .32$ and $.28$, respectively, both $p_s < .01$). Because the correlations were low, average ratings of the two coders for each sample were computed for analytical purposes.

Preliminary analyses. As described in Study 2, analyses were conducted to assess differences in students at the beginning of the semester. Ratings of these writing and rhetoric samples from students in "ReActing to the Past" and other FYs were statistically undistinguishable (all $p_s > .10$). Accordingly, these data were averaged to serve a baseline against which data from the end of the semester could be compared.

Primary analyses. The average ratings from the beginning of the semester and the two ratings ("ReActing" and "NonReActing") were submitted to ANOVA. For the writing data, there was no significant effect, $F < 1$. Students' samples from the "ReActing" ($M = 4.12$, $SE = .11$) and "NonReActing" ($M = 4.00$, $SE = .14$) FYs were equivalent to each other and statistically identical with the samples from the beginning of the semester ($M = 4.00$, $SE = .09$). In contrast, the ANOVA for the ratings of the rhetorical samples was significant, $F(2, 201) = 6.41$, $p < .01$. Post-hoc analyses indicated that students from "ReActing to the Past" made more effective rhetorical presentations ($M = 16.2$, $SE = .32$) compared with students from "NonReActing" seminars ($M = 14.7$, $SE = .38$), and they improved from the beginning of the semester ($M = 14.8$, $SE = .28$).

These analyses suggest that writing ability, at least extemporaneous writing, was not negatively affected, but that rhetorical skills actually improved as a result of participating in

“ReActing to the Past.” Given the focus of the role-playing pedagogy on debate and discussion, it is not particularly surprising that students show improvement in these skills. At the same time, the lack of effects for writing tends to assuage concerns that students who participate in “ReActing to the Past” might lag behind others in the development of writing skills.

Two obvious limitations in this study should be recognized. First, the reliability between the two student raters was quite low. This might have occurred because of differing interpretations of the coding scheme or reactions to individual samples, or because of a restriction of range in the quality of samples analyzed. Because of the lack of a criterion for determining which ratings were more “accurate,” we simply averaged the ratings of the two coders. Given the high amount of disagreement between the raters, this might serve to undermine the ability to detect real differences in the quality of writing or rhetoric. The fact that actual differences were detected in rhetoric quality suggests is impressive given this potential problem. The ultimate solution, however, might be to include a third coder with greater expertise in these domains. The ratings of this coder might allow identification of the basis of the low reliability between coders and may be combined with or substituted for the student ratings. Second, any conclusions about the lack of effects for writing must be restricted to extemporaneous writing. In our assessment, students were asked to write spontaneously on a topic about which they presumably had given little thought. Although there are times that such writing is required, this task (we hope!) does not resemble typical academic writing. There could be differences in the quality of reflective, persuasive, or creative writing as a function of FYS that might not have been identified in our assessment in Study 3. Study 4 will revisit this issue.

STUDY 4: EFFECTS ON ACADEMIC PERFORMANCE

Study 4 involved an assessment of student transcripts with three purposes in mind. First, and most important, we wished to assess whether academic performance subsequent to completing FYS differed between those students who completed “ReActing to the Past” and other seminars. Although dramatic differences were not expected, it seemed possible that the psychological and skill benefits identified in the earlier studies, or effects not yet identified, might enhance performance in subsequent semesters. Transcripts of Juniors and Seniors, who had completed several semesters subsequent to their FYS, were the focus of this assessment. Second, we wanted at least one aspect of the assessment to include data from all the students

who participated in “ReActing to the Past.” Studies 1-3 relied on students’ voluntary participation, and the degree of participation varied across studies and never approached 100%. Third, we choose to focus in particular on performance in courses emphasizing writing. Although there were no effects of “ReActing to the Past” on writing quality in Study 3, we acknowledged that extemporaneous writing might bear little correspondence to the writing processes that the curriculum is designed to develop. Accordingly, we focused on performance in classes that Barnard has identified as “Writing Intensive Courses.” These courses are designated because they contain multiple writing requirements involving numerous drafts and revisions. These courses also have a “Writing Fellow,” a student assigned to assist others in the writing process. If there were any obvious weakness in writing, we hypothesized that it would be most obvious in these courses.

Method

Materials and Procedure. Transcripts for all students who completed “ReActing to the Past” as a FYS were obtained from the College Registrar. For each student, her original application file completed before arriving at Barnard was then obtained from the Office of the Dean of Studies. These application files contained two pieces of information that were necessary in controlling for extraneous sources of variance in academic performance. First, each student’s combined SAT score (Verbal + Quantitative) was recorded to control for differences in pre-existing academic aptitude. Second, each student’s expressed preference for FYS was noted. A student is allowed to express her top three preferences of six briefly described “clusters” of FYS, of which “ReActing to the Past” is one. Because there might be systematic differences between students interested in this pedagogy and those who were not, the expressed interest level (first preference, second preference, third preference, no preference) was recorded.

After the transcripts were collated with SAT scores and preference information, transcripts were then requested for a set of students who had taken one of the classes from the other five clusters of FYS to serve as a control. The specific transcripts requested were based on data gleaned from application files regarding class rank (Junior or Senior), SAT combined score, and expressed seminar preference. Therefore, a total of 241 transcripts from four categories of students were analyzed: 1) students who expressed a preference for “ReActing to the Past” who received the course ($n = 75$), 2) students who expressed a preference for “ReActing to the Past” who did not receive the course ($n = 108$), 3) students who did not express a preference for

“ReActing to the Past” but who nonetheless received the course ($n = 29$), and 4) students who did not express a preference for “ReActing to the Past” and who did not receive the course ($n = 29$).

After identifying information was removed from each transcript, a student research assistant recorded three different pieces of information: GPA in Writing-Intensive courses, total GPA, and GPA for each semester relative to enrolling in FYS. (For the semester in which students completed FYS, GPA was computed for all other courses without FYS so that differences in FYS grades would not bias results). These data were the focus of statistical analysis.

Results and Discussion

Preliminary analyses. A preliminary analysis was conducted to insure that the four different groups of students did not differ in SAT combined score. Accordingly, a 2 (ReActing vs. NonReActing) \times 2 (ReActing Preferred vs. ReActing Not Preferred) ANOVA was conducted with SAT combined as the dependent variable. This analysis yielded no effects for ReActing nor the interaction, both $ps > .10$. There was an intriguing main effect for Preference, $F(1, 236) = 13.55$, $p < .001$, indicating that students who indicated “ReActing to the Past” was one of their three preferred clusters had higher average SAT scores than those who did not ($M_s = 1332$ and 1280 , respectively), regardless of whether they actually received the course. This suggests that the role-playing pedagogy might hold greater appeal for those students with higher academic aptitude.

Because of this difference, it was imperative to include Preference as a variable in all analyses of GPA. Although analyses including this variable did produce some effects favoring students who preferred “ReActing to the Past” (regardless of FYS completed) the preference variable did not interact with the ReActing variable (i.e., the actual course students completed). Therefore, the primary analyses collapsed across this variable and all reported effects generalize across preference level.

Primary analyses. The three measures of academic performance were analyzed separately via one-way ANOVAs. To assess the issue about writing skills originally addressed in Study 3, the average GPA in “Writing Intensive” courses were first analyzed. Of the 241 students for whom transcripts were available, 106 (61 ReActing, 55 NonReActing) had taken at least one such course. The analysis of these GPAs was not significant, $F(1, 104) = .039$, $p = .84$. There was

virtually no difference in the GPA of students who completed “ReActing to the Past” ($M = 3.43$, $SE = .06$) and students who completed other FYSs ($M = 3.41$, $SE = .06$). The analysis of overall GPA also was not significant, $F(1, 236) = 1.43$, $p = .23$, although there was a slight trend favoring students who had completed “ReActing to the Past” compared with other seminars ($M_s = 3.36$ vs. 3.29 , respectively).

Despite the lack of effects for total GPA, we felt that a more detailed analysis was appropriate for several reasons. First, overall GPA includes performance from throughout the college career, some from courses taken before or along with FYS. It would not be expected that any effects of FYS should be evident on these data. Second, it is possible that any effects of FYS might occur for a limited period. An analysis that incorporates grades from all semesters, although important in its own right, might obscure more subtle and specific effects.

Therefore, a final set of analyses examined GPA for each semester relative to completion of FYS. Separate analyses were required for each semester because of the differing number of students in each comparison (e.g., whereas all 241 students completed one semester after FYS (FYS+1), only 30 students had completed a fifth semester after FYS (FYS+5) when this analysis was conducted). Average GPAs by semester can be seen in Figure 1. The separate analyses produced only one significant effect detailed below, but the overall pattern of GPA over time was striking. In the two semesters before FYS (FYS-2 and FYS-1), GPA was slightly but not significantly higher for students who were not eventually to take “ReActing” compared with those students who. In the FYS semester that pattern reverses, with there emerging a slight but insignificant advantage for students in “ReActing,” $F(1, 234) = 1.61$, $p = .21$. This slight advantage persists across the next several semesters, and is in fact statistically significant in the fourth semester after FYS (FYS+4), $F(1, 80) = 3.83$, $p = .05$. This advantage for students from “ReActing” is no longer evident in FYS+5, although, as noted above, the number of cases in this analysis was substantially lower.

This pattern of the data is suggestive of some overall benefit in academic performance for those students who completed “ReActing to the Past” as a FYS. Although performing somewhat more poorly than students in other seminars in the semester before FYS, students from “ReActing to the Past” performed marginally better in the four semesters following completion of the course, $F(1, 80) = 3.75$, $p = .06$, according to a repeated-measures ANOVA. This pattern is inconsistent with an interpretation that students who were assigned to a “ReActing” FYS

simply had higher academic aptitude. Instead, some aspect of the FYS course appears to have produced a slight benefit in performance in subsequent semesters.

The causal factor behind this benefit is admittedly unclear. There are several possibilities. One possibility is that students who take different types of FYSs tend to take different courses and to major in different disciplines. Given the clear differences in GPA across majors, this might account for the differences obtained in performance. That possibility is being assessed in analyses this semester. A second, more intriguing possibility, is that the psychosocial effects identified in Study 2 or the benefits regarding rhetorical skill identified in Study 3 might account for the results. This would not be particularly surprising, given the established benefits of self-esteem (Tesser, 2001) and belief in malleable characteristics (Dweck, 1999) already identified in the literature.

Unfortunately, the mediating role of these variables cannot be examined here. Such an analysis would require that the data from Studies 2 and 3 be collated with the GPA data assessed in Study 4. Students who participated in the first two studies, however, were told that no identifying information would be collected. Indeed, none was. Therefore, it is impossible to tie together the results on psychosocial and skill measures with the data concerning academic performance. That flaw should be redressed in future research.

General Discussion

These studies clearly point to a number of benefits associated with completion of a FYS utilizing the “ReActing to the Past” methodology. The course was clearly popular, even for students who did not request the course or were originally hesitant once learning more about the nature of the course. Students praised the pedagogy for producing a level of motivation and involvement in the course material they had not anticipated. According to student self-reports, the pedagogy also facilitated the development of an appreciation of multiple points of view on controversial topics. These effects were confirmed and extended in studies designed to compare students from “ReActing to the Past” seminars with students from other seminars and over time. These studies showed that the pedagogy produced higher self-esteem, empathy, and belief that people can change over time and across contexts. Rhetorical skills were developed through the seminar, but writing skills, measured both at the time of FYS and later, did not appear to be negatively impacted.

Despite these positive findings, a few cautions are in order. With one exception (Study 4), participation in these studies was voluntary and less than universal. This allows for the possibility that the obtained effects might simply reflect a bias in the sampling of students who agreed to complete the various tasks we used to assess the pedagogy. Arguing against this possibility, however, is that analyses comparing students from “ReActing to the Past” and the other FYS courses at the beginning of the semester never yielded significant differences. Nonetheless, future research should utilize other methods independent in addition to financial compensation to increase rates of participation.

It must also be emphasized that the specific cause of the effects is unclear. Role-playing differs from other pedagogical approaches in a variety of ways, and within these “other” approaches is equal or greater heterogeneity. Moreover, particular instructors who utilize the role-playing method presumably also use differing strategies and possess varying strengths. To tie the effects discussed here to any single aspect of the pedagogy would be an error. Nonetheless, it is clear that some aspect or aspects of the pedagogy produce systematic benefits, at least for some of the students.

The extent of these effects is also not settled based on the effects reported here. We selected our measures based on student comments solicited in focus group sessions and based on our theorizing about the potential impact of role-playing over a semester. It is entirely possible (and, indeed, likely) that other effects were not identified because appropriate measures were not included in our studies. These effects might point to other benefits or might instead point to detrimental consequences of the pedagogy that we did not consider.

Finally, it is unclear whether these effects will occur at other colleges and universities that decide to utilize the pedagogy. Barnard College describes itself as an “independent college of liberal arts and sciences for women, affiliated with Columbia University” that is “located in the cosmopolitan urban environment of New York City.” The students at Barnard tend to have high academic and professional potential, and they tend to come from the East Coast of the United States. Each of these variables (and others too numerous to mention) might play an important role in determining the success or failure of the “ReActing to the Past” pedagogy. Although the course has been largely successful across instructors, across classes, and over time, it has yet to be tried with students with substantially different backgrounds and characteristics.

The findings obtained in those different contexts will ultimately be the basis for judging the degree and nature of the consequences associated with the pedagogy.

Table 1

Sample Items from PsychoSocial Scales.

Title and Source	Sample Items (items with a * are reverse-scored)
Self-Esteem (Rosenberg, 1965)	“On the whole, I am satisfied with myself” “I certainly feel useless at times”*
Implicit Person Theory (Levy et al., 1998)	“People can substantially change the kind of person they are”* “People can do things differently, but the important parts of who they are can’t really be changed”
Locus of Control (Rotter, 1966)	“People’s misfortunes result from the mistakes they make” “I have often found that what will happen will happen”*
Social Avoidance and Distress (Watson & Friend, 1969)	“I often find social occasions upsetting” “It is easy for me to relax when I am with strangers”*
Machiavellianism (Christie & Geis, 1970)	“One should take action only when sure it is morally right”* “Never tell the real reason you did something unless it is useful to do so”
Mastery (Pearlin et al., 1981)	“I have little control over the things that happen to me”* “What happens in the future mostly depends on me”
Confidence as a Speaker (Paul, 1966)	“I look forward to an opportunity to speak in public” “I perspire and tremble just before getting up to speak”*
Orientation to Life (Antonovsky, 1987)	“Doing the things you do every day is a source of deep pleasure and satisfaction” “Has it happened that people whom you counted on disappointed you?”*

Table 2

Effects of ReActing to the Past on PsychoSocial Variables.

	Beginning	NonReActing	ReActing
Self-Esteem	3.14 _a (.047)	3.16 _a (.059)	3.32 _b (.049)
Balanced Emotional Empathy	5.23 (.082)	5.15 _a (.089)	5.37 _b (.069)
Implicit Person Theory	3.62 (.134)	3.92 _a (.128)	3.42 _b (.137)
Locus of Control	1.47 _a (.019)	1.55 _b (.025)	1.49 (.023)
Social Avoidance & Distress	1.29 _a (.024)	1.74 _b (.030)	1.74 _b (.025)
Machiavellianism	3.40 (.086)	3.51 (.089)	3.48 (.110)
Mastery	3.02 (.039)	3.05 (.050)	3.03 (.049)
Confidence as a Speaker	1.54 (.025)	1.46 (.033)	1.50 (.030)

Note: Standard error of the mean in parentheses. Means with different subscripts differ at the $p < .05$ level.

References

- Antonovsky, A. (1987). Unraveling the mystery of health. San Francisco: Jossey-Bass.
- Christie, R., & Geis, F. L. (1970). Studies in Machiavellianism. New York: Academic Press.
- Dweck, C. S. (1999). Self-theories: Their role in motivation, personality, and development. Philadelphia, PA: Psychology Press.
- Levy, S. R., Stroessner, S. J., & Dweck C. S. (1998). Stereotype formation and endorsement: The Role of implicit theories. Journal of Personality and Social Psychology, 74, 1421-1436.
- Mehrabian, A. (1996). Manual for the Balanced Emotional Empathy Scale (BEES). (Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA 93940).
- Morreale, S. P., Moore, M. R., Taylor, K. P., Surges-Tatum, D., & Hulbert-Johnson, R. (1994). "The competent speaker" speech evaluation form. Annandale, VA: Speech Communication Association.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.
- Paul, G. L. (1966). Insight vs desensitization in psychotherapy. Stanford, CA: Stanford University Press.
- Pearlin, L., Lieberman, M., Menaghan, E., & Mullan, J. (1981). The stress process. Journal of Health and Social Behavior, 22, 337-356.
- Quianthy, R. L. (1990). Communication is life: Essential college sophomore speaking and listening competencies. Annandale, VA: National Communication Association.
- Rokeach, M. (1973). The nature of human values. New York: Free Press.
- Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs (General & Applied), 80, 1-28.
- Tesser, A. (2001). Self-esteem. In A. Tesser & N. Schwarz (Eds.), Blackwell handbook of social psychology: Intraindividual processes (pp. 479-498). Oxford, UK: Blackwell.
- Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. Journal of Consulting and Clinical Psychology, 33, 448-457.

Figure Captions

Figure 1. Mean GPA for students from “ReActing to the Past” vs. alternate seminars as a function of semesters relative to completing FYS, Study 4. Bars denote one standard error.

