

# Response Shift Bias in Student Self Report Assessments

Frederick R. Rohs<sup>1</sup> and Chris A. Langone<sup>2</sup>  
Department of Extension Education, The University of Georgia  
Athens, GA 30602-4356

## Abstract

This study sought to determine the degree of response shift (change in level of understanding of leadership skills) by undergraduate students in a college leadership course. While the vast majority of students rated themselves as having "high" leadership skills at the end of the course, significant differences were found between their self-report ratings using the pretest/posttest approach and the then/posttest approach. The degree of response shift (pre/then comparison) was also significant. The findings from this study together with other studies cited suggest that when employing self-report measures, the post/then approach provides a less conservative and more accurate means of assessing a student's knowledge and understanding of the subject than would the traditional pretest/posttest approach. Suggestions for reducing the effects of response shift are also offered.

## Introduction

Business, industry and society are telling colleges and universities that there is a grave need for leadership and human resource preparation for today's students to succeed in the work place (Brown and Fritz, 1993). This has led to the proliferation of a number of leadership development courses that employ some form of self-report assessment measure. These introspective measures may vary from a listing of skills learned (Russell and Jones, 1995) or keeping a leadership journal (Dormody, 1996) to employing standardized assessment measures (Brungardt and Crawford, 1996) or similar rating scale. Youth organizations such as FFA and 4-H also have sought to measure their impact on a youth's leadership development. When such introspective measures are employed in the classroom, the conventional pretest/posttest method of evaluation is often used. In these instances differences between pretest and posttest ratings may appear to be non-existent when actually significant differences exist. The faculty need a more accurate measurement of behavioral change than the conventional pretest/posttest method.

<sup>1</sup> Professor and Head

<sup>2</sup> Associate Professor

One consequence of most leadership development courses is changing a person's awareness or understanding of the leadership skill being measured. For example, a class participant might feel at the beginning of a course that they are an "average" leader with "average" leadership skills. The course changes their understanding of leadership skills; after the course they understand their level of functioning was below average at the beginning of a course. Whenever such shift in understanding occur, conventional self-report pretest/posttest designs are unable to accurately gauge the impact of instructional programs.

Several studies (Howard and Dailey, 1979; Howard et al., 1979; Pohl, 1982; Sprangers and Hoogstraten, 1988; Rockwell and Kohn, 1989) have documented the "response-shift bias" phenomenon as a source of contamination of self-report measures that result in inaccurate pretest ratings and seriously compromise any assumption of internal validity.

Evidence of response shift biases have been found in college classrooms dealing with knowledge of subject matter and the learning of basic helping skills (Howard et al., 1979; Pohl, 1982). Extensive literature reviews by Pohl indicate that often when self-report measures are used, there is a lack of findings of significant differences between pre and posttest measurements.

To correct this problem, Howard et al., (1979) recommends that at the posttest session participants are asked to respond twice to each item on the self-report measure. The first asks participants to report their behavior/understanding as a result of the program(post). The second asks participants to report their behavior before the program(then). Because "then" ratings and post ratings are made in close proximity, it is more likely that both ratings will be made from the same perspective and thus be free of response-shift bias.

The purpose of this study was to determine the degree of response shift in the self report ratings of leadership skills development by undergraduate students enrolled in a leadership development course.

## Methods

Data were obtained from 28 students enrolled in AGR 300--an undergraduate course in agricultural leadership skills at the University of Georgia for students who want to

learn more about leadership and decision making skills. The 10-week course covers such topics as leadership theory, stages of group development, group maintenance skills, listening and feedback skills, conflict management and several techniques relating to group decision making and consensus building. Throughout the course students participate in various exercises that allow them to practice various skills being discussed in the class.

The Youth Leadership Life Skills Development Scale (YLLSDS), developed by Dormody et al., (1993), was used to measure students leadership skill development. The YLLSDS is a 30 item paper and pencil instrument which asks individuals to indicate on a four point scale (0=none, 3=a lot) the degree to which they possess each skill or characteristic. Total scale values can range from 0 to 90. For descriptive purposes Dormody et al., (1993) suggest scale values of 0 to 30 as "no to slight leadership skills development," from 31 to 60 "moderate development," and from 61 to 90 "high development."

According to Dormody et al., (1993) the YLLSDS was assessed for face and content validity by a panel of faculty from New Mexico State University and field tested with a stratified random sample of 262 New Mexico senior 4-H and FFA members. The Cronbach's alpha reliability coefficient for the scale was .98.

The YLLSDS was administered on the first day of class asking students how they would currently rate themselves on each of the 30 items (PRE). The YLLSDS was again administered on the last day of class asking students to respond twice to each item. First they are to report how they perceived themselves to be at the present (POST). Immediately after answering each item in this manner, they were asked to answer the same item again, this time in reference to how they perceive themselves at the beginning of the course (THEN).

Data were summarized and analyzed using SAS 608. Statistical tests (t-test for matched groups) were employed to determine if differences existed between the sets of scores testing evidence of response shift.

## Results and Discussion

Ninety six percent of the students rated themselves as "high" in leadership skills on the posttest (Table 1). However, their pretest self-report ratings also revealed they felt they were "high" in leadership skills development. The students "then" ratings reveal a different story. Students "then" ratings indicate that only 14% rated their leadership skills as high with 75% falling into the moderate category and 11% into the low level category.

No significant differences were found between the pretest and posttest means (Table 2). The posttest score (

=70.0) and the "then" score ( =52.0) were significantly different. Thus, students felt their leadership skills had improved since the beginning of the course with the post/then method. To determine the response shift in students self-report ratings, pretest means were compared with their "then" means. A significant difference existed between the "pretest" and "then" means. The difference between the two means or response shift was 18.0 points in students perception of their leadership skills development.

This study provides evidence of the impact of response shifts on self-report ratings of leadership skills by students enrolled in the leadership class. The then/post procedure provided radically different results with which to evaluate the leadership class compared to the pre/post procedure. The response-shift effects, differences between the "then" pretest and the pretest, are treatment dependent. While the lack of a control group may limit this study, it should be noted that the danger of such an instrumentation effect cannot be eliminated by the use of a control group. The score on a given scale may have a different meaning for the "treatment" group than for those in the "control" group (Rohs and Langone, 1997). Response-shift theory provides a plausible explanation for these findings. An increase in the students' understanding of the phenomenon under consideration or an increased appreciation of their initial level of functioning on that dimension could have caused them to report leadership 'then' scores which were lower than their pretest scores. However, other explanations are also possible. For example, these same results might have occurred if (1) students remembered their pretest rating and level of functioning and consciously overrepresented their posttest level rating or underrated their pre course level on the retrospective/then pretest to report a positive experience or (2) biased their reports to provide the instructors with more favorable results. However, the time period between the administration of the pretest and posttest/then procedure (10 weeks) would not enhance the students' memory. Students were also asked on their posttest to record what they thought was their pretest scale score. No accurate readings occurred. The students were also told at both administrations that their responses were confidential and would not be taken into account when class grades were computed. Studies by Howard et al. (1979) also refute these alternative explanations.

## Summary

The Then/Post analysis provided a drastically different set of conclusions regarding the effectiveness of the leadership class from the Pre/Post approach. The Then/Post data revealed the course produced major changes in the leadership skills of students verses a "no change"

Table 1. Pretest, then, and posttest levels of leadership skill development scores.

YLLSDS SCORES	Pretest <sup>x</sup>		Then <sup>y</sup>		Posttest <sup>z</sup>	
	n	%	n	%	n	%
Low 0-30	0	.00	3	.12	0	.00
Moderate 31-60	5	.18	21	.75	1	.04
High 61-90	23	.82	4	.14	27	.96
Totals	28	100	28	100	28	100

<sup>x</sup> =  $\bar{X}$  =70.0, SD=10.8

<sup>y</sup> =  $\bar{X}$  =52.0, SD=17.5

<sup>z</sup> =  $\bar{X}$  =70.0, SD=8.2

Table 2. Means, standard deviations and test of significance of self-report leadership skill scores by condition

Condition	Pre		Then		Post		T
	Mean	SD	Mean	SD	Mean	SD	
Pre/Posttest	70	10.8			70	8.2	NS
Then/Posttest			52	17.5	70	8.2	6.61***
Then/Pretest			52	17.5	70	10.8	-4.49***

\*\*\* = .001 significance level

conclusion using Pre/Post data. Furthermore, the Pre/Then data indicate a "response shift" or change in the level of understanding of leadership skills by students took place during the course. Studies of college courses by Howard (1980), Bray and Howard (1980), and Pohl (1982) have produced similar results.

Findings from this study suggest that the Then/Post approach provides a more accurate estimate of measuring change. Pre/Post methods remain popular, therefore further research is needed to assess the conditions under which a Pre/Post method would be more appropriate. Additionally, research is needed to identify and clarify the various causal determinants of the response shift. One factor may be the level of information students have at the pretest regarding the dimension, in this case leadership skills, on which they are asked to self-report.

To lessen response-shift bias "informed pretests" may be employed where a thorough description of the variable being measured is provided to the student prior to the administration of the self-report pretest. As with all research the adequacy of the measures used affects the

quality of the findings. While this study employed the YLLSDS scale as a valid and reliable measure, our experience with leadership skills assessment has been that most self-report measures/studies do not. Integrating self-report, objective and behavioral measures, if possible, may help to provide a more complete assessment of change. Use of pretest, posttest and retrospective/then pretest self-report data will provide a more sensitive assessment of a students perspective of personal changes and skill development.

### Literature Cited

- Bray, J.H. and G.S.Howard. 1980. Methodological considerations in the evaluation of a teacher-training program. *Jour. of Educational Psychology* 72(1): 62-70.
- Brown, W.F. and S.M. Fritz. 1993. Determining the breath of leadership and human resource management development offerings in post secondary Departments of Agricultural Education. *NACTA Jour.* 37 (3):11.

- Brungardt, C. and C.B. Crawford. 1996. A comprehensive approach to assessing leadership students and programs: preliminary findings. *Jour. of Leadership Studies* 3 (1):37-48.
- Dormody, T. 1996. 30 question leadership journal. *Jour. of Leadership Studies* 3 (2):75-81.
- Dormody, T., B.S. Seevers, and D.L. Clason. 1993. The youth leadership life skills development scale: an evaluation and research tool for youth organizations. *N.M.S.U. Agr. Expt. Sta. Bul.* 672.
- Howard, G.S. 1980. Response shift bias--a problem in evaluating interventions with pre/post self-reports. *Evaluation Review* 4 (1):93-106.
- Howard, G.S. and P.R. Dailey. 1979. Response-shift bias: a source of contamination in self report measures. *Jour. of Applied Psychology* 64 (2):144-150.
- Howard, G. S., K.M. Ralph, N.A. Gulanick, S.E. Maxwell, D.W. Nance, and S.K. Gerber. 1979. Internal invalidity in pretest/posttest self report evaluations and a re-evaluation of retrospective pre-tests. *Applied Psychological Measurement* 3:1-23.
- Pohl, N.F. 1982. Using retrospective pre-ratings to counteract response-shift confounding. *Jour. of Experimental Education* 50(4):211-214.
- Rockwell, S.K. and H. Kohn. 1989. Post-then pre evaluation. *Jour. of Extension* 27 (2):19-21.
- Rohs, F.R. and C.A. Langone. 1997. Increased accuracy in measuring leadership impacts. *Jour. of Leadership Studies* 4 (1):150-159.
- Russell, M.A. and H.W. Jones. 1995. A leadership development course for animal industry careers. *NACTA Jour.* 39(4):30-33.
- Sprangers, M. and J. Hoogstraten. 1988. Response-style effects, response-shift bias, and a bogus pipeline: a replication. *Psychological Reports* 62:11-16.

---

## Cooperative Learning: Group Activity Projects in Reproductive Biology Instruction<sup>1</sup>

Darrel J. Kesler<sup>2</sup>, Department of Animal Sciences,  
University of Illinois, Urbana, IL 61801

### Abstract

Course and instructor evaluations from a reproductive biology course that included group activity projects were analyzed. Group activity projects consisted of four to eight students responsible for an objective such as establishing pregnancy in a ewe by embryo transfer. At the conclusion of the project students gave a class presentation. Although the majority of students believed they learned the most from the lecture section, the majority believed the group activity projects were the most challenging and enjoyable. Ninety-four percent of the students believed the group activity projects complemented and heightened interest in the other sections and 97% of students believed the time required for the group activity projects was worthwhile. Group activity projects 1) stimulated student interest in the subject, 2) taught teamwork, 3) encouraged competition, 4) taught responsibility, 5) encouraged development of public speaking skills, and 6) motivated students by involving and challenging them in an enjoyable peer-learning environment.

---

<sup>1</sup> The author acknowledges and appreciates advice from P.J. Dzuik in the development and instruction of the group activity projects.

<sup>2</sup> Address correspondence to 1207 W. Gregory Dr.

### Introduction

One objective of higher education is to provide students with the ability to critically evaluate ideas and information from many sources, creatively solve problems, and/or discover new concepts. Accomplishing this objective, however, requires more than just dispensing knowledge to students. Students must be involved in the processes described by the instructor. In other words, they must get their hands dirty. One method of allowing extensive student involvement is through activity/special projects, either on an individual or group basis (Kauffman et al., 1989).

Group activity projects that stimulate analysis and thought have been successfully incorporated into higher education courses (Buhr and Knauff, 1984; Hall, 1989). Individual and group activity projects that involve practical problems and procedures stimulate higher-level thought and improve student understanding and are generally regarded by students as valuable learning experiences (Howe and Durr, 1982; Hall, 1989). Student activity projects stimulate student interest by encouraging them to take responsibility for obtaining information and creating ideas (Schaefer and Kauffman, 1975).

This article describes a course that incorporated