

# Response-Shift Bias in Critical Thinking<sup>1</sup>

**Brandi K. Phillips<sup>2</sup> and Brian E. Myers<sup>3</sup>**  
**University of Florida**  
**Gainesville, FL**



## Abstract

Students' responses to the EMI Critical Thinking Test were examined for response-shift bias, a phenomenon found in previous studies using tests of other constructs in which participants provided inconsistent responses in pre-tests compared to then-tests. Pre-test scores of a sample of 75 students enrolled in animal science courses at the University of Florida were compared to the students' then-test scores, which were obtained upon completion of the course and consisted of self-reports of students' prior critical thinking skills. Comparison of the pre-test scores and then-test scores in this study did not provide evidence of a response-shift bias. The influence of demographic variables including gender and ethnicity was also examined and results indicated that the appearance of response-shift bias was not impacted by either variable. The results of this study were not consistent with limited previous research and future studies should further investigate the phenomenon of response-shift bias with respect to the EMI Critical Thinking Test as well as other self-report tests.

## Introduction

Frequently in educational research, it is necessary to evaluate perceptions, knowledge, attitudes and behaviors of participants as they relate to a treatment. Self-reports of these constructs are often provided using a pre-test-post-test research design. Comparisons can then be made between the respondents' perceptions at the start of treatment and upon completion, allowing researchers to determine the effect of the treatment on the participants.

In some instances, however, obtaining a pre-test from participants may not be practical or feasible. Additionally, concerns have been expressed regarding the ability of participants to accurately self-report prior to a treatment due to their lack of knowledge surrounding the subject of interest (Rockwell and Kohn, 1989). The testing effect may also pose a threat in pre-

test-post-test designs, as research has shown that a pre-test can improve learning which is reflected in the post-test (McDaniel et al., 2007). Ary et al. (2010) have described pre-test sensitization as a threat to validity for attitude and personality inventories, resulting in students carefully considering their responses and changing their answers based on self-reflection and not necessarily on the effect of the treatment. Such instances may call for a post-then design, in which participants provide their self-report of pre-treatment knowledge or perceptions (then) at the same time as their post-treatment knowledge or perceptions (post).

Response-shift bias has been identified as a potential threat to the validity of pre-test-post-test research designs. Howard and Dailey (1979, p. 145) defined response-shift as "the difference between pre and then self-report ratings." Several studies have noted a response-shift in participants' responses (Howard and Dailey, 1979; Rohs, 1999). As a result, researchers have recommended that post-then data be collected in addition to pre-test data for all studies using self-rating measurement methods (Howard and Dailey, 1979; Rohs, 1999) before and after treatments.

One such study, conducted by Howard and Dailey (1979), tested for response-shift bias using a seven-item questionnaire to evaluate interviewer skills before and after a five day workshop. Twenty-one individuals participated in the study and completed a pre-test as well as a post-then-test. In addition, the researchers taped first and last practice interviews of each of the participants and trained judges rated the behavior of each on a 9-point scale. A response shift was discovered in the participants' self-reports on four of the seven items. Further, it was noted that the then-test reports were more closely aligned with the ratings assigned by judges as opposed to the pre-test reports. While a cause for the response shift was not investigated in this study, the shift was observed. The then-test scores were found

<sup>1</sup>This research was supported by the AG-STEM Education Research Lab in the Department of Agricultural Education and Communication, University of Florida. The mission of the AG-STEM Education Research Lab is to discover ways to improve student learning of STEM concepts in agricultural and life sciences through collaborative research in teaching and learning in formal and informal settings.

<sup>2</sup>Graduate Assistant and Master's Student, Agricultural Education and Communication; Email: brandikphillips@ufl.edu

<sup>3</sup>Associate Professor and Associate Chair, Agricultural Education and Communication; Email: bmyers@ufl.edu

to be more accurate representations of interviewer skills than pre-test scores (Howard and Dailey, 1979).

This phenomenon was investigated later by Rohs (1999). Students in an undergraduate agricultural leadership course participated in a similar study using the Youth Leadership Life Skills Development Scale in pre-post and post-then comparison (Rohs, 1999). A group of 30 students participated in a pre-post-test and 28 completed a post-then-test. The data appeared to indicate a response shift, as post-then students reported greater changes compared to the pre-post participants (Rohs, 1999).

In some cases, however, response-shift bias may not pose a threat. Sprangers and Hoogstraten (1988) tested the effects of a bogus-pipeline induction on response-shift bias in testing first aid knowledge of psychology students before and after a first aid film. Results from this research showed no response-shift in the bogus-pipeline experiment, fitting with the researchers' hypothesis. An unexpected finding was that response-shift had also not occurred in the non-bogus-pipeline component (Sprangers and Hoogstraten, 1988). This indicates that there may be certain circumstances under which response-shift bias is not a threat to validity for pre-test-post-test designs.

Although several studies have been conducted to test for response-shift bias (Howard and Dailey, 1979; Sprangers and Hoogstraten, 1988; Rohs, 1999), this phenomenon may not occur under all pre-test-post-test circumstances (Sprangers and Hoogstraten, 1988). Previous studies have looked at student groups as a whole, without providing any data on possible relationships between response-shift and student characteristics. This information may provide valuable insight into response-shift bias. This study investigated response-shift bias using the Engagement, Cognitive Maturity and Innovativeness (EMI) critical thinking test, considering demographic variables which included gender and ethnicity.

A pre-test-post-test analysis of EMI critical thinking test scores of students at the University of Florida was used to determine whether participation in animal science courses and activities impacted critical thinking (Miller et al., 2011). Results of this analysis demonstrated that as a result of participation in animal science courses and activities, students demonstrated improvement on the Innovation and Engagement scales. Then-test data were also collected from these students, but had not been analyzed in the study conducted by Miller et al. (2011). By analyzing the then-test data of these students, this study attempted to validate the results of the former study.

## Methods

The purpose of this study was to determine if a response shift existed between then-test responses and pre-test responses of participants providing a self-evaluation using the EMI critical thinking test. Given this information, researchers may be more able to appropriately determine the accuracy of self-reports evaluated in both pre-then-post as well as post-then-pre designs.

The following objectives were used to guide this study:

1. Evaluate the difference between pre-test scores and then-test scores of the EMI instrument for students enrolled in classes at the University of Florida.
2. Evaluate the difference between pre-test scores and then-test scores of the EMI instrument based on demographics.

The population for this study consisted of students enrolled in the Introduction to Animal Sciences course ( $n = 66$ ), as well as those enrolled in the Meat Selection and Grading ( $n = 3$ ) and Live Animal Evaluation ( $n = 6$ ) courses, at the University of Florida during the 2009-2010 academic school year. Each of the courses provided students with both lecture and laboratory instruction.

Participating students were asked to evaluate their critical thinking skills before and after one semester of participation in the courses. Ricketts and Rudd (2005) developed the EMI test to measure critical thinking disposition in a 26 item response test, consisting of 11 questions measuring engagement (defined as "students' predisposition... to use reasoning" p. 33), eight questions measuring cognitive maturity ("awareness... of their own and others' biases and predispositions" p.33) and seven questions measuring innovativeness (students' predisposition to seek truth). Cronbach's alpha scores of .79, .75 and .89 were given for Innovativeness, Cognitive Maturity and Engagement, respectively (Ricketts and Rudd, 2005). Students were administered the test at the beginning of the programs (pre-test); upon completion of the program, students were asked to fill out the instrument again, including their responses after the course or team activities (post-test). Following the post-test, the participating students were asked to evaluate their responses previous to enrollment or participation (then-test).

Data were then analyzed using SPSS® for Windows™ software. A paired t-test was used to compare pre-treatment responses given prior to participation (pre-test) with pre-treatment responses given after participation (then-test) for totaled values for the following constructs: engagement, cognitive maturity and innovativeness. The total values for the combined constructs were also compared using a paired t-test analysis. *A priori*, a

## Response-Shift Bias

significance level of  $p < .05$  was set. Responses of each construct and the totals were also compared based on gender and ethnicity to determine what trends, if any, may have existed based on demographic information provided by the participants.

## Results and Discussion

### Objective One - Evaluate the Difference between Pre-Test Scores and Then-Test Scores of the EMI Instrument for Students Enrolled in Classes at the University of Florida.

The average score of participants' responses to the engagement portion of the EMI critical thinking test was  $M = 43.83$  at pre-test and  $M = 43.85$  at then-test. A  $p$  value of 0.96 indicated no significant difference between pre and then responses for this construct. Participants' measures of cognitive maturity were reported as  $M = 30.73$  at pre-test and  $M = 31.03$  at then-test. No significant difference between pre and then responses existed ( $p$  value = 0.44). Average values of  $M = 27.75$  at pre-test and  $M = 28.01$  at then-test were reported for innovativeness. A  $p$  value of 0.47 indicated no significant difference in response scores.

**Table 1. Mean Pre and Then Scores of Critical Thinking Constructs**

Item	Mean - Pre	Mean - Then	T	p
Engagement	43.83	43.85	-0.06	0.96
Cognitive Maturity	30.73	31.03	-0.77	0.44
Innovativeness	27.75	28.01	-0.72	0.47
Total	102.31	102.89	-0.61	0.54

### Objective 2 - Evaluate the Difference between Pre-Test Scores and Then- Test Scores of the EMI Instrument Based on Demographics.

Male respondents' ( $n = 19$ ) average score for the total EMI critical thinking test was  $M = 104.26$  at pre-test and  $M = 102.53$  at then-test. No significant difference between the pre and then-tests was determined based on a  $p$  value of 0.31. The average score of female respondents ( $n = 56$ ) for the total EMI critical thinking test was  $M = 101.64$  at pre-test and  $M = 103.02$ . A  $p$  value of 0.24 indicated no significant difference between pre and then-test scores.

The majority of participants were White ( $n = 64$ ), with total average scores of  $M = 101.75$  at pre-test and  $M = 102.33$  at then-test. A  $p$  value was calculated at 0.57, so no significant difference existed between the pre and then-tests. The Non-White participants ( $n = 11$ ) had similar results. Average scores were 105.55 at pre-test and 106.18 at then-test. The  $p$  value of 0.83 indicated that no significant change occurred in this group of participants either.

No significant differences were found between the pre-test scores and the then-test scores reported by participating students with respect to any of the constructs

**Table 2. Mean Pre and Then Scores of Critical Thinking Constructs of Male Participants**

Males	n	Mean - Pre	Mean - Then	t	p
Engagement	19	31.26	30.84	0.64	0.53
Cognitive Maturity	19	44.89	43.63	1.41	0.18
Innovativeness	19	28.11	28.05	0.08	0.94
Total	19	104.26	102.53	1.06	0.31

**Table 3. Mean Pre and Then Scores of Critical Thinking Constructs of Female Participants**

Females	n	Mean - Pre	Mean - Then	t	p
Engagement	56	30.55	31.09	-1.177	0.244
Cognitive Maturity	56	43.46	43.93	-0.821	0.415
Innovativeness	56	27.63	28.00	-0.846	0.401
Total	56	101.64	103.02	-1.194	0.238

**Table 4. Mean Pre and Then Scores of Critical Thinking Constructs of White Participants**

White	n	Mean - Pre	Mean - Then	t	p
Engagement	64	30.52	30.81	-0.715	0.477
Cognitive Maturity	64	43.63	43.69	-0.120	0.905
Innovativeness	64	27.61	27.83	-0.557	0.579
Total	64	101.75	102.33	-0.556	0.573

**Table 5. Comparison of Mean Pre and Then Scores of Critical Thinking Constructs Non-White Participants**

Non-White	n	Mean - Pre	Mean - Then	t	p
Engagement	11	32.00	32.27	-0.280	0.785
Cognitive Maturity	11	45.00	44.82	0.132	0.898
Innovativeness	11	28.55	29.09	-0.493	0.633
Total	11	105.55	106.18	-0.217	0.832

measured by the EMI test. Total scores likewise yielded no significant difference between pre and then scores. Average scores for the EMI critical thinking test in total at the time of pre-test was  $M = 102.31$  and  $M = 102.89$  at the time of then-test.

This study showed no evidence of response-shift bias. Within this sample, pre-test and then-test scores of participants demonstrated no significant difference in self-reports on the EMI critical thinking test administered ( $p > .05$ ). No significant difference was reported in the individual components of the EMI critical thinking test, including engagement, cognitive maturity and innovativeness ( $p > .05$ ). Additionally, analysis revealed no significant difference between pre and then reported scores of males compared to females ( $p > .05$ ). Scores between pre and then reports of White students and Non-White students also showed no significant difference ( $p > .05$ ). Demographic variables investigated in this study appeared to have no effect on the likelihood of response-shift bias for the participants.

The findings of this study contradict those of Rohs (1999) and Howard and Dailey (1979). As the study conducted by Sprangers and Hoogstraten (1988) indicated, response-shift bias may not threaten the validity of all tests. This may include the EMI Critical Thinking Test or possibly measures of the critical thinking construct. A deeper understanding of response-shift bias is needed, as well as how to address response-shift bias if it is found to be present. Relatively few

studies have investigated this phenomenon; therefore, research is needed to test whether response-shift bias exists as a threat to validity in pre-test-post-test designs using the EMI instrument, as well as other self-report measures. Tests used to measure perceptions of individuals with regard to animal welfare issues, use of genetically modified agricultural products and other issues faced by the agriculture industry could benefit from further investigation of response-shift bias.

Studies should continue to collect pre-test-post-test data in conjunction with post-test-then-test designs to verify results. Future research may also include demographic variables to determine whether factors such as gender and ethnicity affect response-shift bias when such a phenomenon is discovered. The impact of participant variables such as age and experience should also be considered in future research.

## Summary

The purpose of this research was to determine if response shift occurred between participants' responses to the EMI critical thinking test before a treatment and a then-test following treatment. A total of seventy five students participated in the study. Participating students completed the EMI critical thinking pre-test at the beginning of the courses, as well as a then-test upon completion of the courses. The participants of this study were selected purposively and consisted of students enrolled in animal science courses at the University of Florida. Results therefore cannot be generalized outside of this population.

No significant differences were found between pre-test and then-test scores of participants selected for this study. Gender and ethnicity of the participants did not result in significant differences between pre-test and then-test scores. Response-shift bias was not a threat to the validity of the EMI Critical Thinking Test within the population selected for this study.

## Literature Cited

- Ary, D., L.C. Jacobs and C. Sorenson. 2010. Introduction to research in education. 8th ed. Belmont, CA: Wadsworth, Cengage Learning.
- Howard, G.S. and P.R. Dailey. 1979. Response-shift bias: A source of contamination of self-report measures. *Jour. of Applied Psychology* 64(2): 44-150.
- McDaniel, M.A., J.L. Anderson, M.H. Derbish and N. Morrisette. 2007. Testing the testing effect in the classroom. *European Jour. of Cognitive Psychology* 19(4-5):494-513.
- Miller, C.C., C.C. Carr, J.H. Brendemuhl, J.C. Ricketts, B.E. Myers and T.G. Roberts. 2011. Quantifying the critical thinking skills of students who receive instruction in meat-animal or meat product evaluation. *NACTA Jour.* 55(1):50-54.
- Ricketts, J.C. and R.D. Rudd. 2005. Critical thinking skills of selected youth leaders: The efficacy of critical thinking dispositions, leadership and academic performance. *Jour. of Agricultural Education* 46(1):32-43.
- Rockwell, S.K. and H. Kohn. 1989. Post-then-pre evaluation. *Jour. of Extension* 50(4):19-21.
- Rohs, F.R. 1999. Response-shift bias: A problem in evaluating leadership development with self-report pretest-post-test measures. *Jour. of Agricultural Education* 40(4):28-37.
- Sprangers, M. and J. Hoogstraten. 1988. Response-style effects, response-shift bias, and a bogus-pipeline: A replication. *Psychological Report* 62:11-16.

