

# The Attitude Polarization Phenomenon: Role of Response Measure, Attitude Extremity, and Behavioral Consequences of Reported Attitude Change

Arthur G. Miller, John W. McHoskey, Cynthia M. Bane, and Timothy G. Dowd

Four studies examined the generality of attitude polarization (C. G. Lord, L. Ross, & M. R. Lepper, 1979). Biased assimilation of essays on 2 controversial issues was substantial and correlated with reported attitude change. Polarization was observed for reported attitude change on capital punishment and generally stronger in Ss with extreme than moderate attitudes. Polarization was not indicated in a pre-post measurement design. For affirmative action, reported polarization was not observed. The hypothesis that Ss reporting polarization would subsequently write particularly strong essays was not supported, although those reporting depolarization wrote relatively weak essays. The results suggest the relevance of individual differences in reported attitude change but do not confirm the powerful inferences frequently drawn regarding the pervasive, undesirable consequences of self-reported attitude polarization.

An influential, frequently cited experiment in the social cognition literature is that by Lord, Ross, and Lepper (1979). In this study, proponents and opponents of capital punishment were shown an identical set of alleged research findings containing evidence both in favor of and against capital punishment. Two important phenomena were observed. First, subjects evaluated more positively the research that agreed with their preexisting position, a result termed *biased assimilation*. Thus, subjects who were in favor of capital punishment rated method X as superior to method Y if the former yielded support for the deterrent effect of capital punishment, but rated the identical method as inferior if it refuted the efficacy of capital punishment. The converse was true for subjects opposed to capital punishment.

The second, particularly intriguing finding was that subjects reported that their attitudes had become more extreme in the direction of their initial point of view after evaluating the research evidence, an effect termed *attitude polarization*. Thus, showing proponents and opponents of capital punishment an identical body of mixed evidence appeared "to increase further the gap between their views" (Lord et al., 1979, p. 2105).

Lord et al. (1979) suggested that in the subjects' biased processing of the research information, they had emerged with an inflated sense of the amount of evidence supporting their initial belief. Correlational analyses involving the relative degree of biased assimilation and the magnitude of reported attitude change (polarization) supported this account. That subjects gave substantially higher evaluations to research that supported

their initial position was not construed, in and of itself, as a particularly unacceptable bias, given the adaptive, often rational strategy of interpreting new information "in the light of past knowledge and experience" (p. 2107). It was the polarization effect, in particular its linkage to biased assimilation, that was viewed as a particularly unsettling departure from optimal judgment:

Their sin lay in their readiness to use evidence already processed in a biased manner to bolster the very theory or belief that initially "justified" the processing bias. In so doing, subjects exposed themselves to the familiar risk of making their hypotheses unfalsifiable—a serious risk in a domain where it is clear that at least one party in a dispute holds a false hypothesis—and allowing themselves to be encouraged by patterns of data that they ought to have found troubling. (Lord et al., 1979, p. 2107).

The Lord et al. (1979) study has received wide acclaim from social psychologists (e.g., Myers, 1990; Worchel, Cooper, & Goethals, 1991), in particular among theorists in social cognition and attitude change:

When people with strong prior beliefs encounter mixed or inconclusive evidence, they may reinterpret the evidence as if it were firm support for their schema, causing their beliefs to persevere or become even more extreme. (Fiske & Taylor, 1991, p. 150)

Moreover, the ultimate effect of being exposed to the contradictory evidence was that subjects believed even *more* strongly in their original position! (Higgins & Bargh, 1987, p. 383, italics in original)

Lord et al. (1979) showed that subjects who encountered ambiguous attitude-relevant information interpreted that information in support of their prior attitude positions and became even more convinced of the appropriateness of their views. (Sherman, Judd, & Park, 1989, p. 304)

Thus, both proponents and opponents of capital punishment who were asked to read a mixed package of evidence about the de-

Arthur G. Miller, John W. McHoskey, Cynthia M. Bane, and Timothy G. Dowd, Department of Psychology, Miami University.

We thank Douglas Fawley and Elizabeth McPartland for their help in conducting the experiments reported in this article.

Correspondence concerning this article should be addressed to Arthur G. Miller, Department of Psychology, Miami University, Oxford, Ohio 45056.

terrent value of the death penalty came away with their views strengthened. (Ross & Nisbett, 1991, p. 73)

Subjects who disagreed about capital punishment *became even further separated* after reading the mixed evidence (Lord et al., 1979). Those in favor became more pro; those opposed, more con. (Zimbardo & Leippe, 1991, p. 163, italics in original)

Biased assimilation and resulting attitude polarization are typically discussed in the context of a number of related, fundamental processes in social cognition, including belief perseverance (e.g., Anderson & Sechler, 1986; Jelalian & Miller, 1984), the effect of stored judgments on the encoding of information (e.g., Bodenhausen, 1988; Darley & Gross, 1983), processes of hypothesis confirmation (e.g., Snyder, 1984), and motivated reasoning and subjective construal (e.g., Griffin & Ross, 1991; Kunda, 1990). The Lord et al. (1979) study has also been featured in a recent debate on dissonance theory, specifically concerning the alleged (Berkowitz & Devine, 1989) but disputed (Lord, 1989) failure of Lord et al. (1979) to interpret the phenomenon of biased assimilation in terms of cognitive dissonance reduction.

In addition to its rich conceptual base, the Lord et al. (1979) experiment has powerful and sobering implications in terms of clarifying the basis of attitudinal conflict and the nature of seemingly intractable sociopolitical conflicts. It is difficult to think about their study without immediately bringing to mind a host of contemporary issues that receive endless, often vitriolic debate without the slightest hint of compromise or moderation of views (e.g., abortion rights, affirmative action, and gun control). The concepts of biased assimilation and attitude polarization have, moreover, instructive implications for flawed judgment in a wide array of contexts, including the testing of scientific theories (e.g., Mahoney, 1977).

There are, however, several issues that, in our view, make the extraordinary reception of the Lord et al. (1979) study premature. The distinction between self-reported attitude change and directly assessed attitude change is a particularly critical issue. Lord et al. (1979) dealt specifically with "self-reported attitude change." They justified this measure by noting that their use of subjects with extreme attitudes created scale attenuation effects, thereby preventing the use of difference scores comparing initial attitudes and attitudes at the experiment's conclusion (p. 2101). In their Discussion section, Lord et al. (1979) acknowledged that "it could be argued that subjects were not reporting real shifts in attitudes but instead were merely reporting what they believed to be a rational or appropriate response to . . . the available evidence" (p. 2108).

Two studies speak to the measurement issue raised by the Lord et al. (1979) study. Plous (1991), using a different attitudinal issue (nuclear energy) and the same self-report measure of attitude change used by Lord et al. (1979), observed significant polarization. In his conclusions, however, he noted that "there is always the chance that *perceived* changes in attitudes do not reflect *actual* changes in attitudes . . . thus, all that can be concluded from the present studies is that subjects perceived changes in attitude in the direction of previously held views" (p. 1078, italics in original). Pyszczynski, Greenberg, and Holt (1985) experimentally induced subjects to succeed or fail on a bogus test of social sensitivity. Noting that the procedure used

by Lord et al. might be "subject to a demand interpretation" (p. 182), Pyszczynski et al. asked subjects instead to assess the validity of the social sensitivity test as well as their own level of sensitivity *both before and after* exposure to the mixed research evidence. Subjects in the success condition rated their own sensitivity higher after exposure to the research evidence than before exposure. This polarization effect was not, however, observed in the failure condition nor with respect to ratings of the test's validity, prompting the investigators to conclude "the finding that ratings of the validity of the test were not affected by exposure to the evidence suggests caution in generalizing the polarization effect obtained in the present study on social sensitivity ratings and by Lord et al. (1979) on attitude ratings to other measures and other contexts" (p. 188).

Given the meticulous conceptual and prodigious methodological analyses that social psychologists have given to the attitude construct (e.g., Ajzen, 1988; McGuire, 1985; Petty & Cacioppo, 1981, 1986; Zimbardo & Leippe, 1991), it is surprising that the distinction between self-reported and directly assessed attitude change has been essentially ignored in response to the Lord et al. (1979) experiment. In this context, one might point to the extensive research literature on attitude polarization stemming from Tesser's influential (1978) conceptualization of self-generated thought (e.g., Chaiken & Yates, 1985; Millar & Tesser, 1986). This research has uniformly used direct attitude assessment (i.e., pretest vs. posttest designs) as the dependent variable. In addition, there has been intense debate and considerable skepticism concerning the degree to which individuals have access to the cognitive processes involved in their attitude change (e.g., Goethals & Reckman, 1973; Nisbett & Ross, 1980; Nisbett & Wilson, 1977; Uleman & Bargh, 1989).

Thus, the precise meaning that should be attached to the subject's testimony regarding the degree to which his or her attitude has changed in response to information provided by an experimenter is unclear. This is not to suggest that the perception or self-presentation of attitude change is necessarily artifactual or less significant than "real" change. Nevertheless, self-reported and direct attitude assessment are distinctly different procedures and may not be measures of the same construct. Experiment 1 was primarily concerned with a comparison of these two methods of assessing attitude polarization.

A second major objective of this research was to examine a relevant behavioral consequence of attitude polarization. To this point, reviewers of the Lord et al. experiment appear to accept the attitude-polarization result—that is, the verbal report of attitude change—as having self-evident importance or validity. To our knowledge, however, there is no empirical evidence that links reported polarization, its absence, or attitude depolarization to a relevant behavior.

Experiments 2 and 3 were designed to examine a possible behavioral consequence of polarization. Subjects were first administered variations of the paradigm used by Lord et al. (1979) and were then asked to write essays indicating their final attitude toward capital punishment (Experiment 2) or affirmative action (Experiment 3). These essays were then rated by observers (Experiment 4). The major interest was whether the essays written by subjects who had indicated attitude polarization would be viewed as more extreme than essays written by subjects indicating either no change or a depolarization in their

beliefs. In addition to asking subjects to write essays, Experiments 2 and 3 considered a number of specific issues that will be noted as each study is described below.

### Experiment 1

The major purpose of Experiment 1 was to conduct an experiment in which subjects were randomly assigned to either a *self-reported attitude change* or to a pretest-versus-posttest *direct attitude assessment* condition. Two groups of subjects not considered in the Lord et al. (1979) study, namely those with relatively moderate attitudes toward capital punishment, were included, in addition to those with extreme attitudes. This would not only avoid the potential methodological problems in restricting the analysis to those with extreme attitudes, but would speak to an important question in its own right. Namely, do the phenomena of biased assimilation and polarization occur in individuals with less extreme beliefs? There is a limited amount of evidence suggesting that these phenomena would be attenuated for subjects with moderate attitudes (Houston & Fazio, 1989; Plous, 1991, Experiment 3; Tesser & Leone, 1977). The basic assumption is that for individuals with relatively weak attitudes, there is a reduced motivation to endorse consistent and to reject inconsistent cognitions, or a reduced tendency to engage in schema-driven information processing.

To assess further the generality of the Lord et al. (1979) study, instead of being shown counterbalanced sets of methods and results pertaining to hypothetical research findings (as used in Lord et al., 1979), subjects were presented with two opposing essays on capital punishment that had originally appeared in the *New York Times*. Given the powerful generalizations that have been made to the Lord et al. study, it was of interest to examine whether their basic findings would occur in a context similar to that of observing a debate on a controversial issue, a context with as much or more mundane realism as that of reading social-scientific research on the deterrent effects of capital punishment. In addition, subjects in the self-reported attitude condition responded to only one measure. Lord et al. used several repeated probes ("running record") to assess self-reported attitude change.

A final concern of this research was an inquiry regarding the basis of attitude polarization. Lord et al. (1979) interpreted attitude polarization as a consequence of the biased assimilation.<sup>1</sup> We also examined this relationship. However, we assigned half of the subjects to a no-evaluation condition, in which they simply read the essays but did not actively record their evaluations. Although this procedure could not rule out a covert biased processing of the essays, it would indicate whether committing to a written evaluation of the mixed information was essential to the subsequent appearance of polarization.

### Method

#### Overview

Subjects were identified as having extreme or moderate attitudes in the pro or anti direction on the capital punishment issue and were asked to read two opposing essays on this issue published in the *New York Times*. They were assigned to an evaluation or no-evaluation condition, with those in the former asked to evaluate the persuasiveness of

several points in each essay. Subjects assigned to the self-reported attitude change condition then responded to the measure used by Lord et al. (1979). Those in the direct assessment condition responded to the same measure that had been used initially to assess their attitude toward capital punishment.

#### Subjects

The subjects were 337 male and female undergraduate students participating as part of a research experience requirement in introductory psychology.

#### Procedure

Subjects were seen in groups of approximately 12. They were informed that the study concerned their opinions about controversial social issues, that there would be several parts to the study, that all of their responses would be confidential, and that they were free to ask questions or to withdraw from participation at any time. Subjects received a brief oral and more extended written debriefing at the conclusion of the experiment and were told where they might receive information about the results.

#### Initial Attitude: Direction and Magnitude

Subjects responded to a 16-item attitude survey dealing with controversial social and political issues (e.g., gun control legislation, affirmative action, and abortion). Each item was on an *extreme disagreement* (–50) to *extreme agreement* (50) scale, with the 0 point labeled *neutral* or *undecided*. The capital punishment item was the 8th in the 16-item survey. Attitudes between 36 and 50 were defined as *extreme pro*, between 1 and 35 as *moderate pro*, between –1 and –35 as *moderate anti*, and between –36 and –50 as *extreme anti*. Subjects indicating a neutral attitude were infrequent ( $N = 13$ ) and were not included in the analyses. The majority of subjects were in favor of capital punishment: extreme pro = 129, moderate pro = 122, moderate anti = 44, and extreme anti = 42.

#### Essays

The essays were obtained from the *New York Times*.<sup>2</sup> The pro-capital-punishment essay, "For the Death Penalty," was written by Ernest van den Haag; the anti-capital-punishment essay, "The Intolerable Horror of Legalized Killing," was written by Irwin Stark. The essays were approximately 1,000 words long and constituted thorough, articulate statements zealously defending one or another side of the issue. Both writers frequently noted but then refuted counterarguments to their preferred orientation. Subjects were randomly assigned to read the pro essay either first or second.

#### Essay Evaluation

Half of the subjects were randomly assigned to an *essay evaluation condition*. After reading each essay, they responded to a 6-item survey asking for an evaluation of the "persuasiveness and convincingness" of six arguments contained in the essay. These items were quoted verbatim or closely paraphrased from each essay. For the anti-capital-pun-

<sup>1</sup> The existing evidence for this relationship is strictly correlational, given that biased assimilation was assessed in terms of individual differences in the biased evaluation of the mixed research evidence.

<sup>2</sup> The essays appeared in the early 1980s and are available from us on request.

ishment essay, an illustrative item was the following: "Murder and capital punishment are not opposites that cancel one another out, but similars that breed their kind;" illustrative of the pro-capital-punishment essay was the following: "The death penalty may have flaws, such as being used more against criminals from minority groups, but this does not mean that the death penalty idea is bad, but rather that it needs to be improved in its implementation."

Each rating was on a -8 (*extremely unpersuasive and unconvincing*) to 8 (*extremely persuasive and convincing*) scale; the 0 point was labeled *average*. After each set of ratings, subjects were asked to write a short paragraph briefly summarizing their views and explaining their persuasiveness ratings. This procedure was similar to that of Lord et al. (1979).

Subjects assigned to the *no-evaluation essay condition* read the essays but did not record their evaluations. Instead, they responded to a filler task dealing with a long-standing controversial campus rule prohibiting undergraduate students from operating cars on campus. Subjects were asked to write a convincing essay, expressing their opinion as if they were in a debate on the no-car rule. The no-car rule was one of the items on the initial attitude survey, giving this exercise credibility as a formal part of the study.

After the initial attitude survey, booklets containing the essays and their associated evaluation pages, or the filler task, were randomly administered to subjects, permitting the different experimental conditions to be run in the same session. The experimenter was not aware of the subject's condition (evaluation or no evaluation) or attitude.

### Primary Dependent Measure

Booklets containing the essays and evaluation or filler-task pages were collected, and a third booklet was distributed. The first item constituted the primary dependent variable of this study and consisted, by means of random assignment, of either (a) a self-report measure of attitude *change* or (b) a direct assessment of the subject's *current attitude*.

**Reported attitude change.** For subjects assigned to the self-reported attitude change condition, the measure was identical to that used by Lord et al. (1979). They were asked the following: "How would you compare your current attitude toward capital punishment with the attitude you had at the very start of this experiment?" (-8 = *much more against capital punishment*, 8 = *much more in favor of capital punishment*, 0 = *no change in attitude*).

**Directly assessed attitude change.** Subjects assigned to this condition were asked to respond to the same item, on a -50 to 50 scale, that had appeared on the initial attitude survey: "What is your current opinion of capital punishment, the use of the death penalty (electric chair or lethal injection) for convicted murderers?"

## Results

### Essay Evaluations

Subjects assigned to the essay evaluation condition rated the persuasiveness of six arguments in each essay. These ratings were highly correlated and were therefore averaged to create a general persuasiveness index for each essay. These indices were analyzed in a 2 (Direction of Subject's Attitude)  $\times$  2 (Magnitude of Attitude)  $\times$  2 (Essay Order)  $\times$  2 (Essay) analysis of variance (ANOVA); the latter variable was treated as a repeated measure. Cell means for this analysis are shown in Figure 1.

The Attitude  $\times$  Essay interaction was significant,  $F(1, 142) = 146.54$ ,  $p < .01$ . Subjects in favor of capital punishment rated the pro essay as more persuasive than the anti essay (pro = 3.77

and anti = 0.94). Subjects opposed to capital punishment rated the anti essay as more persuasive (anti = 4.63 and pro = -0.13). These results document what Lord et al. (1979) defined as biased assimilation.

The degree of biased assimilation was modified, however, by the extremity of the subject's attitude. The Attitude  $\times$  Magnitude  $\times$  Essay interaction was significant,  $F(1, 142) = 15.37$ ,  $p < .01$ . The differential evaluation of the pro and anti essays was significantly greater for subjects holding extreme as opposed to moderate attitudes (extreme pro = 4.44, moderate pro = 1.31, moderate anti = 3.95, and extreme anti = 5.61). Each of these difference scores was significantly different from 0 ( $p < .01$ ). Thus, biased assimilation was significantly observed, although to an attenuated degree, in subjects with moderate attitudes.

### Reported Attitude Change

Reported attitude change was analyzed in an Attitude  $\times$  Magnitude  $\times$  Evaluation or No Evaluation  $\times$  Essay Order ANOVA. Cell means are shown in Table 1 and Figure 2.

Subjects in favor of capital punishment reported that their attitudes had become more favorable toward capital punishment, whereas subjects opposed to capital punishment reported an increasingly negative attitude (pro = 1.15 and anti = -1.02),  $F(1, 149) = 20.09$ ,  $p < .01$ . These results clearly document the phenomenon of attitude polarization and are in close agreement with the magnitude of this effect reported by Lord et al. (1979, Figures 1 and 2, pp. 2106-2107).

Although the Attitude  $\times$  Magnitude interaction was not significant,  $F(1, 149) = 2.34$ ,  $p < .13$ , the results suggest that the degree of attitude polarization was attenuated for subjects with moderate attitudes: extreme pro,  $t(68) = 5.25$ ,  $p < .01$ ; moderate pro,  $t(53) = 1.22$ ,  $p < .22$ ; moderate anti,  $t(22) = -1.59$ ,  $p < .13$ ; extreme anti,  $t(18) = -2.55$ ,  $p < .02$ . Moreover, attitude polarization was statistically significant only in subjects with extreme attitudes. The essay evaluation versus no evaluation manipulation had no influence on self-reported attitude change.

### Frequency of Reported Attitude Polarization

Of the 165 subjects assigned to the self-reported attitude change condition, 41% indicated that their attitude had polarized, that is, had changed by 1 or -1, or more in the direction of their initial attitude; 46% indicated that their opinion had not changed, and 13% reported a depolarization of their attitude.

The extremity of a subject's attitude was also related to reported attitude polarization,  $\chi^2(2, N = 165) = 8.96$ ,  $p < .01$ . Subjects with more extreme attitudes indicated a greater incidence of attitude polarization (extreme pro = 52%, moderate pro = 26%, moderate anti = 39%, and extreme anti = 47%). The frequency of the no-change response was in the 40% to 60% range in all attitude conditions, thus constituting, overall, the modal response. The depolarization of attitudes was relatively infrequent in all conditions (extreme pro = 7%, moderate pro = 18%, moderate anti = 17%, and extreme anti = 11%).<sup>3</sup>

<sup>3</sup> In the Lord, Ross, and Lepper (1979) study, 75% of the subjects reported attitude polarization at the end of the experiment, 19% indicated no change, and 6% reported depolarization.

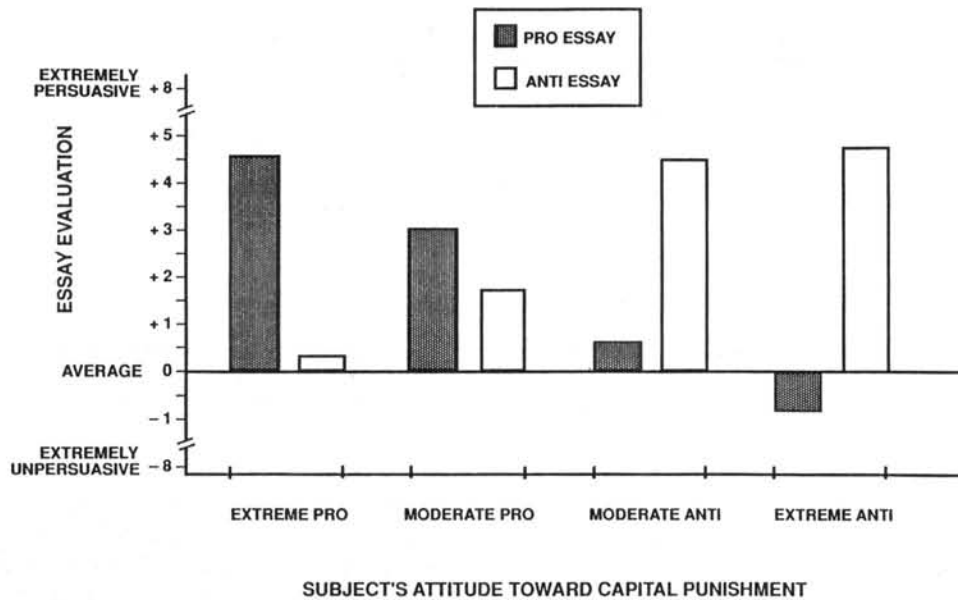


Figure 1. Mean essay evaluations by subjects with extreme or moderate attitudes toward capital punishment.

attitudes toward capital punishment.

### Directly Assessed Attitude Change

As in Tesser's (1978) approach to assessing polarization in the "self-generated or thought-induced" attitude-change paradigm, polarization scores were based on the "trinary index" (e.g., Chaiken & Yates, 1985; Liberman & Chaiken, 1991; Millar & Tesser, 1986). Comparing each subject's pre- and postessay attitude responses, an increase in extremity on the same side of the subject's initial position on the issue (polarization) was coded as 1, a change in a direction of less extremity (depolarization) was coded as -1, and no change was coded as 0.

Trinary scores were analyzed by ANOVA; cell means are shown in Table 1. No significant effects were associated with the experimental conditions. Of primary interest was the absence of a significant attitude polarization effect in this data set (pro = 0.04, anti = -0.05), main effect for attitude,  $F(1, 156) = 0.18$ , *ns*. Further analysis indicated that net polarization, as assessed by a mean greater than 0, was found only for subjects who were moderately pro capital punishment,  $t(67) = 2.93$ ,  $p < .01$ . A significant depolarization was observed for subjects in the extreme pro-capital-punishment condition,  $t(59) = -2.91$ ,  $p < .01$ . Means in the extreme anti and moderate anti conditions were not significantly different from 0.<sup>4</sup>

### Relationship of Biased Assimilation to Attitude Change

Correlations were computed between subjects' differential evaluations of the essays (biased assimilation) and both reported attitude change and polarization with respect to directly assessed attitudes. Higher scores on biased assimilation were associated with relatively greater self-reported attitude polarization ( $r = .36$ ,  $N = 74$ ,  $p < .01$ ), confirming the results reported by Lord et al. (1979). There was also a significant correlation be-

tween biased assimilation and directly assessed attitude change ( $r = .29$ ,  $N = 76$ ,  $p < .01$ ).

### Estimates of Essay Writing Behavior

Subjects were asked to imagine being assigned to write an essay defending as well as opposing capital punishment and to estimate how difficult or easy it would be to prepare convincing statements on each position (1 = *extremely difficult* and 10 = *extremely easy*). Subjects were categorized on the basis of their reported attitude change response as having indicated no change, polarization, or depolarization. Estimates of the ease of essay-writing behavior were analyzed using ANOVA, including self-reported change as a factor and essay direction as a repeated measure in the design.

A significant Attitude  $\times$  Assignment  $\times$  Change interaction was observed,  $F(2, 142) = 13.72$ ,  $p < .01$ . Those in favor of capital punishment indicated, in general, that it would be easier to write an essay in favor of capital punishment than against it, but this was true only for those who had indicated a polarization (mean difference = 3.8) or no change (mean difference = 2.3) in their attitudes. Those indicating depolarization estimated that it would be somewhat easier to write on the opposing side (mean difference = -0.53) of the issue. Similarly, those opposed to capital punishment indicated that it would be easier to write against rather than in favor of capital punishment, but

<sup>4</sup> The algebraic differences between pre and post attitude responses were also analyzed. No polarization effects were observed—extreme pro: pre = 45.3, post = 41.7,  $t(59) = -2.65$ ,  $p < .01$ ; moderate pro: pre = 22.2, post = 24.55,  $t(67) = 1.17$ , *ns*; moderate anti: pre = -22.6, post = -16.6,  $t(20) = -1.02$ , *ns*; extreme anti: pre = -45.0, post = -39.3,  $t(22) = 1.97$ ,  $p < .06$ . Differences in the extreme groups were in the direction of depolarization.

Table 1  
*Attitude Change Assessed by Self-Report and Direct (Pre and Post) Measurement: Experiment 1*

Subject attitude	Self-reported attitude change <sup>a</sup>		Directly assessed attitude change (trinary index) <sup>b</sup>	
	<i>M</i>	<i>n</i>	<i>M</i>	<i>n</i>
Extreme pro	1.76**	69	-0.27**	60
Moderate pro	0.37	54	0.31**	68
Moderate anti	-0.91	23	0.00	22
Extreme anti	-1.16*	19	-0.09	24

<sup>a</sup> Higher positive numbers indicate reported change in the direction of *more in favor of capital punishment*; higher negative numbers indicate *more against capital punishment* (-8 to 8 scale). <sup>b</sup> Higher positive numbers indicate greater polarization; higher negative numbers indicate greater depolarization.

\*  $p < .05$ . \*\*  $p < .01$ .

this was true only for those indicating a polarization (mean difference = 3.55) or no change (mean difference = 2.05) in their attitudes. Those indicating depolarization estimated that it would be easier to write on the pro side (mean difference = -1.67) of the issue. Thus, there was, with respect to an attitude-relevant behavior, a meaningful relationship between self-reported attitude change and estimates of the ease of engaging in that behavior.

### Discussion

The biased assimilation and attitude polarization phenomena, reported by Lord et al. (1979; Lord, Lepper, & Preston, 1984), were clearly in evidence in the present experiment. At the individual-subject level, the frequency of attitude polariza-

tion was substantial, although somewhat less than that reported by Lord et al. (1979). Correlational evidence, although not as strong as reported by Lord et al., supported the hypothesis that attitude polarization is, at least in part, mediated by biased assimilation. That these results were obtained in a study in which the stimulus materials and procedural details differed in numerous respects from those of Lord et al. (1979) substantiates the generality of their basic findings. Evidence indicating reported attitude polarization in the absence (as well as presence) of recorded essay evaluations also provides generality for the Lord et al. (1979) study. Whereas it is likely that subjects evaluated the essays implicitly in the course of reading them or in recalling their features, the present data indicate that an overt or public commitment to such evaluations is not a necessary precondition for reported polarization.

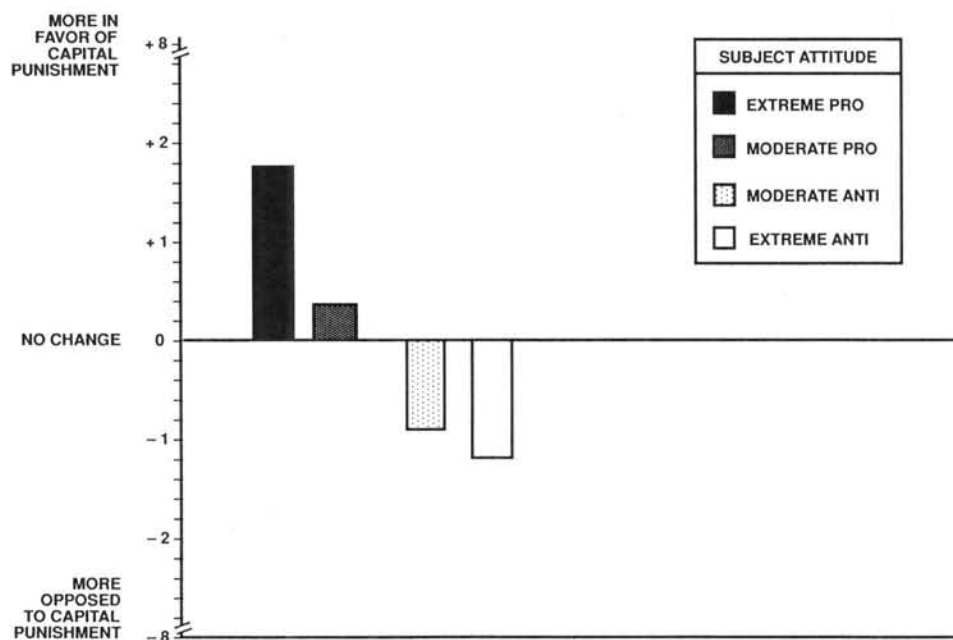


Figure 2. Mean self-reported attitude change relative to the start of the experiment by subjects with extreme or moderate attitudes toward capital punishment.



An important feature of this study was the inclusion of subjects with moderate attitudes toward capital punishment. At least with respect to this issue and subject sample, the moderate attitude category included a substantial number of subjects (49% of the sample). These individuals displayed significantly less biased assimilation, as well as significantly less attitude polarization. The latter findings are consistent with the assumption that subjects with moderate attitudes have relatively weak cognitive schemata or less accessible attitudes (e.g., Houston & Fazio, 1989; Plous, 1991, Experiment 3). The failure of these subjects to report significant attitude polarization is also consistent with research on thought-induced attitude change (e.g., Tesser, 1978; Tesser & Leone, 1977).

The strongest departure from the observations of Lord et al. (1979) pertains to the direct assessment of subjects' attitudes. Considering the overall results for subjects assigned to this condition, attitude polarization was not in evidence. Although Lord et al. (1979) clearly acknowledged the difference between reported and directly assessed attitude polarization (1979, p. 2107), this distinction has essentially been ignored by reviewers of their study. On the basis of the present results, one cannot safely generalize from self-reported attitude change to directly assessed attitudes.

This is not to suggest that self-reported attitude polarization is uninteresting or inconsequential. As noted, estimates of the ease of writing consistent or counterattitudinal essays suggested a relationship to self-reported attitude change, specifically with respect to subjects reporting attitude depolarization. From a self-perception perspective, one can hardly dismiss the potential importance of circumstances that would induce individuals to tell others, or themselves, that their attitude on an issue has strengthened. Genuine attitude change and a variety of behavioral effects might well result from such self-reports. This issue will be reconsidered in a discussion of Experiment 4.

The correlations between biased assimilation and self-reported, as well as directly assessed, attitude change support the hypothesis of Lord et al. that attitude polarization is driven by subjects' biased assimilation of mixed evidence. However, the relatively modest size of the correlations and the minimal evidence for attitude polarization in the directly assessed attitude condition suggest that other influences may also be operative. For some subjects, reporting a strengthening of their attitude might, in their eyes, lead others to view them as intelligent or as having strong convictions.<sup>5</sup> In some instances, subjects might presume that they are expected to change their opinions in response to the information given to them—a point raised by Pyszczynski et al. (1985). Reporting a strengthening of their attitude might also, for some individuals, reduce cognitive dissonance experienced in being exposed to (and evaluating) strongly discrepant views on an issue of strong emotional or moral significance (Berkowitz & Devine, 1989; Sherman & Gorkin, 1980).

## Experiment 2

To assess the reliability of the findings of Experiment 1, particularly with respect to the direct assessment of attitude change, we conducted a second experiment. A key element in this study was the assessment of attitude polarization by means

of reported as well as directly measured (pre vs. post) attitude change in the same subjects. Additional features of this study, including those associated with obtaining an attitude-relevant behavior, will be noted in a description of this experiment.

## Method

### Overview

Subjects expressed their attitude toward capital punishment, evaluated two essays endorsing and opposing capital punishment, and responded to the measure of attitude polarization used by Lord et al. (1979). Subjects then wrote essays expressing their opinion on capital punishment. The study concluded by asking subjects to evaluate their essays and to indicate their current attitude toward capital punishment on the same measure administered at the start of the experiment. (As noted earlier, a major purpose of this study, as well as Experiment 3, was to assess whether the essays written by subjects who reported attitude polarization would be rated by observers as stronger than essays written by subjects not indicating polarization. This analysis, involving the essays written in both Experiment 2 and Experiment 3, will be discussed in Experiment 4).

### Subjects

Male and female introductory psychology students ( $N = 224$ ) participated as part of a research experience requirement.

### Procedure

The initial phase was identical to that in Experiment 1. Four groups were identified: extreme pro,  $n = 89$ ; moderate pro,  $n = 86$ ; moderate anti,  $n = 29$ ; and extreme anti,  $n = 20$ .

Subjects evaluated the same essays used in Experiment 1 and responded to the measure of attitude polarization used by Lord et al. (1979). Subjects were then given 20 min to write an essay expressing their current opinion about capital punishment. On completing their essays, subjects rated the position of their essay on an *extremely against* (−5) to *extremely in favor* (5) scale, in addition to several other items. Finally, subjects were asked to express their current opinion on capital punishment on the same measure (−50 to 50) administered at the start of the experiment.

## Results

### Biased Assimilation

Biased assimilation was observed: Essay  $\times$  Attitude interaction,  $F(1, 215) = 100.55$ ,  $p < .01$ . Proponents of capital punishment rated the pro essay as more persuasive than the anti essay (pro = 3.25 and anti = 1.20); the converse was true for opponents (pro = −0.05 and anti = 3.93). As in Experiment 1, the degree of biased assimilation was modified by the extremity of the subject's attitude, Attitude  $\times$  Magnitude  $\times$  Essay interaction,  $F(1, 215) = 17.58$ ,  $p < .01$ . The differential evaluation of the pro and anti essays was greater for subjects holding extreme (pro  $M = 3.63$ , anti  $M = -5.22$ ) as opposed to moderate (pro  $M = 0.42$ , anti  $M = -3.13$ ) attitudes.

<sup>5</sup> We have collected unpublished data addressing the consequence of reported attitude change on impression formation, which are consistent with this interpretation.

### Reported Attitude Change

A significant attitude polarization effect was observed,  $F(1, 223) = 7.74, p < .01$ . As in Experiment 1, proponents of capital punishment reported an increase in the favorability of their attitudes, whereas opponents reported an increasingly negative attitude (pro = 0.67 and anti = -0.63). As shown in Table 2, for proponents, the polarization effect was only significant for those with extreme attitudes: Extreme,  $t(88) = 5.43, p < .01$ ; moderate, *ns*. However, for opponents, the polarization effect was significant only for those with moderate attitudes, moderate,  $t(28) = -2.75, p < .01$ ; extreme, *ns*, reversing the pattern displayed in Experiment 1. The correlation between biased assimilation and polarization was significant ( $r = .24, p < .01$ ), replicating the results of Experiment 1.

### Frequency of Reported Polarization

Of the 224 subjects in this study, 36% indicated that their attitude had polarized in the direction of their initial attitude, 45% indicated that their attitude had not changed, and 19% reported a depolarization of their view. These figures are similar to those observed in Experiment 1 (polarization = 41%, no change = 46%, and depolarization = 13%). The lack of a significant reported polarization effect for the extreme anti group reflects that of these subjects, an equal number (5) polarized and depolarized; for the moderate anti group, however, more subjects reported polarization (12) than depolarization (2).

### Directly Assessed Attitude Change

As in Experiment 1, the trinary index was used as a measure of directly assessed attitude change. A key finding was the absence of a significant attitude polarization effect (pro = -0.02, anti = 0.12, main effect for attitude,  $F(1, 216) = 0.54, ns$ ). As shown in Table 2, further analysis indicated depolarization in the extreme pro condition,  $t(88) = -3.39, p < .01$ , polarization in the moderate pro condition,  $t(85) = 2.40, p < .01$ , and nonsignificant effects in the moderate anti condition,  $t(28) = 1.68$ , and extreme anti condition,  $t(19) = -0.70$ .<sup>6</sup>

Thus, in both Experiment 1 and Experiment 2, attitude polarization occurred in moderate pro subjects, whereas depolariza-

tion occurred in extreme pro subjects. These results are consistent with a regression toward the mean interpretation. Also, in both experiments, subjects opposed to capital punishment displayed nonsignificant trinary scores. From this perspective, attitude polarization as assessed by the trinary index clearly did not constitute the modal response of subjects in either experiment. The correlation between biased assimilation and the trinary index of attitude change was not significant ( $r = .12, N = 224, p < .07$ ).

### Relationship Between the Two Measures of Attitude Change

The correlation between the two measures of attitude change was significant ( $r = .24, N = 224, p < .01$ ). To clarify this relationship, subjects were first categorized, on the basis of the self-report measure, as having polarized, depolarized, or not changed their attitude. The trinary index of attitude change for each of these three groups was then analyzed using ANOVA. A significant effect for self-reported change was observed,  $F(2, 201) = 12.29, p < .01$ . The mean trinary index for subjects who reported no attitude change was 0.13 ( $t = 1.53, ns$ ); a similar result, 0.14 ( $t = 1.44, ns$ ), was observed for subjects who reported polarization. However, for subjects who reported a depolarization of their beliefs, the trinary index was significantly different from 0 and in the logical direction ( $-0.49, t = -4.01, p < .01$ ).

### Self-Ratings of Essays

Subjects indicated that their essays reflected their initially stated beliefs, Attitude  $\times$  Magnitude interaction,  $F(1, 201) = 6.32, p < .01$  (extreme pro  $M = 3.52$ , moderate pro  $M = 1.50$ , moderate anti  $M = -2.79$ , extreme anti  $M = -4.05$ ). Of primary interest was the question of whether a subject's reported attitude change (i.e., depolarization, no change, or polarization) would be reflected in his or her essay ratings. Although the interaction for this analysis was not significant, Attitude  $\times$  Reported Attitude Change interaction,  $F(2, 201) = 2.09, p < .12$ , there was a trend suggesting that subjects who reported a depolarization of their attitude rated their essays as less extreme than subjects in the other two categories (pro: depolarization = 0.13, no change = 3.39, polarization = 2.86; anti: depolarization = -2.00, no change = -3.32, polarization = -3.82). There were no effects of interest on other essay ratings (e.g., strength, persuasiveness, or ease of writing).

### Discussion

Regarding the phenomena of biased assimilation and attitude polarization (assessed in terms of self-reported change),

Table 2  
Attitude Change Assessed by Self-Report and Direct (Pre and Post) Measurement: Experiment 2

Subject attitude	Self-reported attitude change <sup>a</sup>		Directly assessed attitude change $M^b$
	<i>M</i>	<i>n</i>	
Extreme pro	1.44*	89	-.27*
Moderate pro	-0.13	86	.24*
Moderate anti	-1.14*	29	.28
Extreme anti	-0.05	20	-.10

<sup>a</sup> Positive numbers indicate more in favor of capital punishment; negative numbers indicate more against capital punishment (-8 to 8 scale). <sup>b</sup> Positive numbers indicate polarization; negative numbers indicate depolarization (trinary index).

\*  $p < .01$ .

<sup>6</sup> The algebraic differences between pre and post attitude responses were analyzed. No polarization effects were observed—extreme pro: pre = 45.9, post = 41.8,  $t(88) = -3.06, p < .01$ ; moderate pro: pre = 20.9, post = 15.5,  $t(85) = -1.76, p < .08$ ; moderate anti: pre = -21.0, post = -21.6,  $t(28) = 0.11, ns$ ; extreme anti: pre = -44.5, post = -31.4,  $t(19) = -1.92, p < .07$ . Differences in the extreme groups were in the direction of depolarization, as in Experiment 1.



the present findings in general replicated those of Experiment 1. Subjects processed the essays in a manner that confirmed their initial attitudes. The polarization effect was also observed, although the influence of the magnitude of the subject's attitude, shown in Experiment 1, was not replicated in the case of subjects opposed to capital punishment. A comparison of the two measures of attitude change—obtained here from the same subjects—indicated a relatively weak relationship ( $r = .24$ ). A significant polarization effect occurred for reported attitude change but not for directly assessed change.

There was, however, evidence suggesting that a subject's report of attitude depolarization, in contrast to reports of either no change or polarization, was meaningfully related to other responses. This was shown in terms of the relationship of reported depolarization to attitude depolarization assessed by the trinary index, as well as to the subjects' ratings of the positions of their essays. The latter finding is consistent with the result of Experiment 1, in which subjects who reported a depolarization of their attitudes estimated that they would have relatively more difficulty writing an essay consistent with their initial attitude than subjects reporting no change or polarization.

### Experiment 3

The major objectives of this experiment were to examine the phenomena of biased assimilation and attitude polarization using a different attitude issue and to obtain essays from subjects to be subsequently evaluated (with those from Experiment 2) in an effort to assess the behavioral consequences of reported attitude polarization.

### Method

#### Subjects

Female and male introductory psychology students ( $N = 109$ ) participated as part of a research experience requirement.

#### Procedure

The general procedure followed closely that of Experiment 1 with the exception of the attitude issue, in this instance that of affirmative action. Only one measure of attitude polarization was obtained: the self-report measure used by Lord et al. (1979).

**Initial attitude assessment.** Four groups were identified: extreme pro,  $n = 16$ ; moderate pro,  $n = 26$ ; moderate anti,  $n = 43$ ; and extreme anti,  $n = 24$ . (Twelve subjects expressed a neutral attitude on affirmative action).

**Essay evaluations.** The pro essay, "The Great White Myth," was written by Anna Quindlen (1992). The anti essay, "A Trio of Sensible Thinkers on the Subject of Affirmative Action," was written by Paul Greenberg (1990). The essays were similar in length and constituted articulate statements advocating one or the other side of this issue. Subjects were asked to read each essay (counterbalanced across subjects) and to evaluate the persuasiveness of five major arguments mentioned in the essay (on  $-8$  to  $8$  scales). An illustrative argument in Quindlen's (pro) essay was "The new myth is that the world is full of Black Americans prospering unfairly at White expense, and anecdotal evidence abounds. The stories about incompetent Black co-workers always leave out two things: the incompetent White co-workers and the talented Black ones." Illustrative of Greenberg's (anti) essay was the following: "Affirmative action started out as a good idea: Take extra

pains to recruit members of minorities for jobs and schools lest they be overlooked or discriminated against. Now it has become a rank form of discrimination itself."

**Self-reported attitude change.** After evaluating the essays, subjects responded to the self-reported attitude change measure.

**Essays.** The experiment concluded by asking subjects to write an essay expressing their current position on affirmative action and to respond to several items on a final rating sheet.

### Results

#### Biased Assimilation

Biased assimilation of the essays was observed, Attitude  $\times$  Essay interaction,  $F(1, 100) = 13.88$ ,  $p < .01$ . Proponents of affirmative action rated the pro essay as more persuasive than the anti essay (pro  $M = 4.37$ , and anti  $M = 2.98$ ); the converse was true for opponents (pro  $M = 2.40$ , and anti  $M = 4.16$ ). Although the Essay  $\times$  Attitude  $\times$  Magnitude interaction was not significant,  $F(1, 100) = 2.08$ ,  $p < .15$ , there was a clear trend for the differential evaluation of the pro and anti essays to be greater for those holding extreme attitudes (pro  $M = 2.55$ , and anti  $M = -2.21$ ) than for those holding moderate attitudes (pro  $M = 0.68$ , and anti  $M = -1.40$ ). These results replicate the pattern shown in Experiments 1 and 2.

#### Reported Attitude Change

Subjects' reports of attitude change failed to indicate a significant polarization effect,  $F(1, 101) = 0.74$ , *ns*. The mean reported attitude change for proponents of affirmative action was  $0.12$  ( $n = 42$ ); the mean reported change for opponents was  $-0.12$  ( $n = 67$ ). The Attitude  $\times$  Magnitude interaction was also not significant,  $F(1, 101) = 0.24$ , *ns*. None of the mean reported change scores in each of the subject attitude categories differed significantly from 0 (see Table 3): extreme Pro  $t(15) = 1.54$ , *ns*; moderate pro  $t(25) = -0.60$ , *ns*; moderate anti  $t(42) = -0.78$ , *ns*; extreme anti:  $t(23) = 0.28$ , *ns*. The correlation between the degree of biased assimilation and reported polarization was significant ( $r = 0.22$ ,  $n = 108$ ,  $p < .05$ ), consistent with the findings of Experiments 1 and 2.

Table 3

*Affirmative Action Attitude Change Assessed by Self-Report: Experiment 3*

Subject attitude	Self-reported attitude change	
	<i>M</i>	<i>n</i>
Extreme pro	0.88	16
Moderate pro	-0.35	26
Moderate anti	-0.26	43
Extreme anti	0.13	24

*Note.* Positive numbers indicate more in favor of affirmative action; negative numbers indicate more against affirmative action ( $-8$  to  $8$  scale).

### Frequency of Reported Attitude Polarization

Of the 109 subjects in this study, 31% reported polarization, 43% indicated no change, and 26% indicated depolarization in their attitudes. Relative to Experiments 1 and 2, these figures reflect a somewhat reduced level of polarization and greater incidence of depolarization. These data are consistent with the lack of an overall significant polarization effect.

### Discussion

Although biased assimilation was demonstrated in their essay evaluations, subjects did not report a significant polarization of their attitudes. Biased assimilation thus does not inevitably lead to reported attitude polarization. One explanation for the lack of an overall polarization effect is readily apparent: An approximately equal frequency of subjects reported a depolarization as well as a polarization of their attitudes toward affirmative action. In essence, these groups canceled one another. As in Experiments 1 and 2, approximately half of this sample reported no change in their opinions. Whether an overall polarization effect occurred in these studies thus appears to reflect the relative proportion of subjects indicating a change in the polarized direction. More generally, this finding emphasizes the importance of considering individual differences in the direction of reported attitude change.

Why did relatively more subjects in this study report a depolarization of their attitudes? We have no convincing answer. Subjects may have been less familiar with detailed arguments about affirmative action relative to the capital punishment issue used in Experiments 1 and 2. A larger number of subjects were perhaps more informed by the essays in this study, and, as a result, indicated a reversal of their position.

Consistent with this line of reasoning, although there was a highly significant biased assimilation effect, this effect was smaller relative to that observed in Experiments 1 and 2. For Experiments 1 and 2, the mean biased assimilation indices (i.e., the absolute difference in evaluation of the two essays by opponents and proponents) were 3.80 and 3.01, respectively; for Experiment 3, the index was 1.55. This could, at least in part, account for the absence of an overall polarization effect, although this interpretation remains speculative.

### Experiment 4

Subjects in Experiments 2 and 3 wrote essays expressing their current position on capital punishment or affirmative action. The objective of Experiment 4 was to note whether subjects who had reported attitude polarization in Experiments 2 and 3 wrote essays that would be judged by observers as more extreme than essays written by subjects indicating either no change or attitude depolarization. Should this result occur, it would provide, in our view, a significant validation of the polarization construct as conceptualized by Lord et al. (1979).

### Method

#### Subjects

The subjects were 70 undergraduates (19 men and 51 women) who participated in fulfilling a research experience requirement in introductory psychology.

### Procedure

The 333 essays written by subjects in Experiments 2 and 3 were divided into 14 sets. Within each set, there were 16 essays on capital punishment and 8 essays on affirmative action, written by proponents as well as opponents of the issue, and (as the numbers in each category permitted) by writers who had reported not changing, polarizing, or depolarizing their attitudes before writing their essay. (To create equal numbers of essays in the 14 sets, 3 extra essays written by neutral subjects were included). The 24 essay folders in each set were initially placed in a randomized order. After each set was evaluated by a rater, the set was rearranged by counterbalancing the order of the folders. Each set was evaluated by five subjects who were asked to read each essay and rate its position on the issue ( $-5 = \text{extremely against}$ ,  $5 = \text{extremely in favor}$ ), in addition to its strength and persuasiveness.

After rating the essays, subjects filled out a brief attitude questionnaire. On capital punishment, 49 subjects indicated a pro attitude, 6 indicated a neutral attitude, and 15 expressed an anti attitude; on affirmative action, 25 indicated a pro attitude, 8 indicated a neutral attitude; and 37 expressed an anti attitude. For affirmative action, 13 of the 14 sets of essays were rated by at least one rater on both the pro and anti side of the issue; for capital punishment, 10 of the 14 sets were rated by at least one rater on both the pro and anti sides of the issue.

### Results

On the basis of the five observations of each essay, mean ratings of the essay's position were calculated and analyzed in a Writer's Attitude (pro vs. anti)  $\times$  Magnitude of Writer's Attitude (extreme vs. moderate)  $\times$  Issue (affirmative action or capital punishment)  $\times$  Writer's Category (depolarize, no change, or polarize) ANOVA. Means for this analysis are shown in Figure 3.

The writer's attitude was reflected in the rated position of the essay (Pro  $M = 2.40$ ; Anti  $M = -2.58$ ),  $F(1, 308) = 236.42$ ,  $p < .01$ , as was the magnitude of the writer's attitude (Pro: extreme = 3.54 and moderate = 1.35; Anti: extreme =  $-3.04$  and moderate =  $-2.30$ ); Attitude  $\times$  Magnitude interaction,  $F(1, 308) = 9.93$ ,  $p < .01$ . Of critical importance for this study was the effect of the writer's category, Attitude  $\times$  Category interaction,  $F(2, 308) = 18.54$ ,  $p < .01$ . This interaction was significant for capital punishment,  $F(2, 308) = 5.42$ ,  $p < .01$ , and for affirmative action  $F(2, 308) = 13.53$ ,  $p < .01$ . As shown in Figure 3, the ratings of essays by writers who reported a depolarization of their attitudes were substantially less extreme relative to those reporting either no change or a polarization of their attitudes.<sup>7,8</sup>

<sup>7</sup> There were no significant differences in the initial attitudes of essay writers reporting polarization, no change, or depolarization—extreme pro: polarization = 46.8, no change = 41.7, depolarization = 42.5; moderate pro: polarization = 20.8, no change = 22.1, depolarization = 20.4; moderate anti: polarization =  $-19.8$ , no change =  $-21.2$ , depolarization =  $-18.4$ ; extreme anti: polarization =  $-46.1$ , no change =  $-45.2$ , depolarization =  $-43.5$ .

<sup>8</sup> Simple effects of reported attitude change were performed for the four groups shown in Figure 3. Tukey's range test indicated that for extreme pro, the difference between no change and depolarization was significant ( $p < .05$ ); for extreme anti, the differences between no change and depolarization and between polarization and depolarization were significant; for moderate pro, the differences between no change and depolarization and between polarization and depolarization were significant; for moderate anti, the difference between polarization and depolarization was significant. The difference between position ratings of essays in the no change and polarization conditions was not significant for any of the four groups of writers.

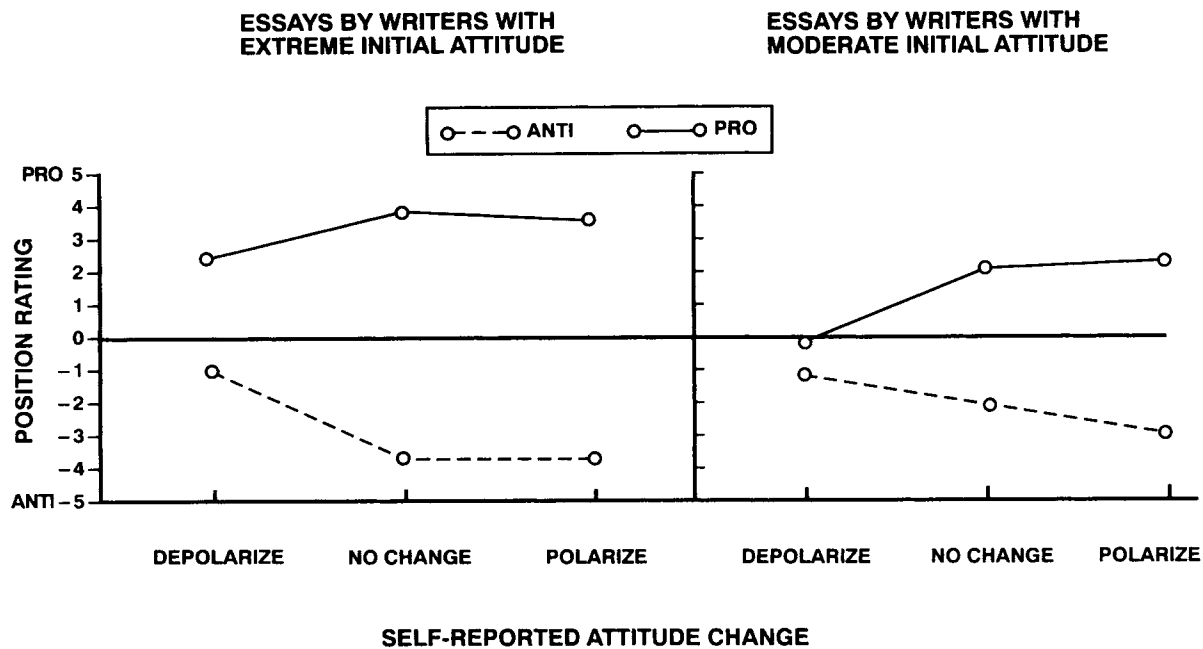


Figure 3. Mean ratings of essay position as a function of the direction and magnitude of writer's attitude and reported attitude change.

For ratings of essays by writers with moderate attitudes, there is the suggestion of somewhat more extreme ratings in the polarization than in the no-change category (Figure 3, right panel). An ANOVA of these ratings failed to show a significant Attitude  $\times$  Category,  $F(1, 308) = 1.25$ , *ns*, or Attitude  $\times$  Category  $\times$  Issue,  $F(1, 308) = 2.14$ , *ns*, interaction. Further analysis indicated that with respect to the essays on capital punishment, the Attitude  $\times$  Category interaction was significant,  $F(1, 308) = 4.38$ ,  $p < .05$ ; this was not true for affirmative action,  $F(1, 308) = 0.04$ . For capital punishment, the means in the polarizing category were more extreme (pro = 2.63, anti = -3.43) than in the no-change category (pro = 2.11, anti = -2.20).

Subjects also rated the strength and persuasiveness of each essay. These ratings ( $r = .91$ ) were summed and analyzed by ANOVA. There were no significant effects associated with the writer's category,  $F(2, 308) = 1.08$ , *ns*. However, the means were in a direction consistent with the ratings for essay position (depolarization = 0.40, no change = 0.98, polarization = 0.93). Further analysis indicated that the rating for essays in the depolarization category was significantly less ( $M = 0.40$ ) than the composite rating of the other groups ( $M = 0.96$ ),  $F(1, 308) = 8.53$ ,  $p < .01$ .

### Discussion

Raters in this study were clearly able to detect the essay writers' initial attitudes as well as the magnitude of those attitudes. The essays were therefore diagnostic of two key attributes of the writers' position, namely, direction and magnitude. From this perspective, the essay task was an appropriate behavioral test of the polarization construct. Had writers reporting polarization in fact written demonstrably stronger essays than

their no-change or depolarizing counterparts, it is likely that this would have been detected by the essay raters. However, the only evidence consistent with a linkage between reported attitude polarization and a meaningful behavioral outcome was observed with respect to position ratings of essays by writers with moderate attitudes toward capital punishment. However, the absence of this effect for essays on affirmative action, in addition to the lack of differences between the polarization and no-change categories with respect to ratings of essay strength or persuasiveness, suggest that in general there were very minimal behavioral consequences to subjects' reports of attitude polarization.

In sharp contrast, however, are the results with respect to essays written by subjects reporting attitude depolarization. These essays were clearly judged to be less extreme (as well as less strong and persuasive) than those written by subjects in the other categories. This finding provides confirming behavioral evidence for the self-report measure of attitude change. Subjects who reported a depolarization of their attitudes subsequently behaved in a manner consistent with these reports.

Why were the behavioral effects of reported depolarization stronger than those of polarization? In the case of essays written by writers with extreme attitudes, one possible reason is a ceiling effect. As shown in Figure 3 (left panel), essays written by writers in the no-change categories were rated as extreme in both the pro and anti directions, leaving virtually no room for an extremity shift for essays in the polarizing condition. This ceiling effect was clearly not present in the case of essays written by subjects with moderate attitudes. There was thus more room for polarization to occur, and as indicated in the right-hand panel of Figure 3, there was a trend for this to occur. It is also conceivable that the report of attitude depolarization in the

current study reflected genuine attitude change, whereas reported polarization did not, at least to the same degree. This account, while highly speculative, is consistent with the result of Experiment 2, in which subjects who reported attitude depolarization also displayed significant depolarization on the trinary index of attitude change. A comparable effect for subjects reporting polarization, however, was not observed.<sup>9</sup>

### General Discussion

The results of the experiments reported here provide equivocal evidence regarding the status of attitude polarization reported by Lord et al. (1979). We consider first the positive evidence. Clearly, the phenomena of biased assimilation and attitude polarization were in evidence, in addition to significant correlations between these two phenomena. That these results were observed in the context of a number of major procedural variations on the Lord et al. (1979) paradigm attests to the robustness of their basic findings.

Evidence supporting the validity of the self-report measure of attitude change used by Lord et al. (1979) may be found in each of the experiments reported in this study. In Experiment 1, subjects' estimates of the relative ease of writing an essay consistent or inconsistent with their attitude was significantly related to self-reported attitude change. In Experiment 2, there was a trend for subjects' ratings of the position of their essays to reflect self-reported attitude change. In addition, there was a significant correlation between the measures of reported and directly assessed attitude change. Finally, ratings of an attitudinally relevant behavior demonstrated a significant effect for self-reported attitude change.

Several features of the results, however, challenge the conceptualization of attitude polarization presented by Lord et al. (1979) and subsequently endorsed by numerous reviewers of that study. First, with respect to the direct assessment of attitude change using a pre-post procedure, the results in general provided no evidence of attitude polarization. The overall trinary index was essentially 0 in both Experiment 1 and Experiment 2. If one agrees that a pre-post measurement of attitude change (i.e., the trinary index) is as legitimate a means for assessing polarization as is the self-report of attitude change, the present findings do not provide evidence for the equivalence of these methods. These results would suggest a note of caution in terms of the strong generalizations concerning attitude change that have been made about the Lord et al. (1979) findings by commentators and reviewers of that study.

A second major result that does not support the validity of reported attitude polarization is the essential absence of any behavioral consequences of this response. Apart from the position ratings of essays written by writers with moderate attitudes toward capital punishment (an effect not replicated with respect to essays on affirmative action), there were no differences between the judgments or behavior of subjects indicating polarization and those reporting no change in their attitudes.

Significant findings relating subjects' self-reported attitude change to other responses and to their essay behavior were specific to reported attitude depolarization. These results are of considerable interest in their own right and point to the importance of examining individual differences in the direction of

self-reported attitude change. In this paradigm, a sizable number of subjects appear to have become less extreme in their beliefs. These findings at least suggest that attitude polarization could, in other contexts, be shown to have meaningful behavioral consequences.<sup>10</sup> However, given that the Lord et al. (1979) study has been widely endorsed as a particularly compelling documentation of attitude polarization, the present findings must be construed as providing essentially no substantial evidence for the generality of this phenomenon. The lack of reported attitude polarization in Experiment 3 is also consistent with this conclusion.

With respect to the direct assessment of attitude change (i.e., trinary index), Tesser's (1978) hypothesis that "mere thought" about an attitude results in attitude polarization appears to receive little confirmation. Attitude polarization was not observed as a significant main effect in either Experiment 1 or Experiment 2. Although research on "mere (self-generated) thought" shows attitude polarization to be a well-documented phenomenon, there are procedural differences between the experiments in that tradition (e.g., Chaiken & Yates, 1985; Millar & Tesser, 1986; Tesser & Leone, 1977) and the paradigm used in Lord et al. (1979) and the current research. For example, the present studies provided, in the form of powerful media essays, external sources for attitudinally discrepant cognitions, a procedure different from that used in thought-induced polarization studies.

Nevertheless, there is a similarity between the conceptualization of thought-induced attitude polarization and the hypothesis of Lord et al. (1979) relating biased assimilation to attitude polarization. Regardless of the source of discrepant cognitions, that is, internal through the subjects' covert reconsideration of their attitude, or external through the subjects' reading of essays, the underlying mechanism would seem to be the same, namely, biased processing.

Biased assimilation, as considered in this research, can, for example, readily be construed in terms of evaluative consistency (e.g., Berkowitz & Devine, 1989; Kunda, 1990) and is clearly similar to the idea of a "reinterpretation of inconsistent cognitions" noted by Chaiken and Yates (1985, p. 1479) as a central mechanism in thought-induced polarization. The correlations between biased assimilation and reported attitude

<sup>9</sup> Another possible explanation for the more sharply defined behavioral effect of reported depolarization is that the magnitude of reported attitude polarization (in terms of the algebraic difference from 0) could have been lower than the magnitude of reported attitude depolarization. The mean reported attitude change scores were therefore computed for the essay writers in both Experiment 2 and Experiment 3. There is no indication of a significant difference in the relative magnitude of reported polarization and depolarization in either study (*Experiment 2*—polarization  $M = 3.17$ , depolarization  $M = -2.53$ ; *Experiment 3*—polarization  $M = 2.74$ , depolarization  $M = -2.76$ ).

<sup>10</sup> The particular behavioral measure used in this study, that is, essay writing, may reflect a number of factors, for example, attitudes, individual differences in the ability to construct a persuasive essay, ego involvement in the task, and so forth. Whether attitude polarization would relate significantly to another conceptually relevant behavioral measure, but differing in form from essay writing, is, of course, a reasonable question, which is not addressed in this research.

polarization, shown in Lord et al. (1979) and replicated in the studies reported here, are consistent with the conceptualization linking self-generated thought to polarization.

The absence of attitude polarization in Experiment 3, on the issue of affirmative action, remains puzzling. A recent study in the "thought-induced" attitude change paradigm by Liberman and Chaiken (1991) may be relevant. They reported an absence of attitude polarization in subjects for whom the particular attitude induced a sense of value conflict. The authors noted that "if an attitude brings two important values into conflict, thinking about it should bring evaluatively inconsistent thoughts to mind, resulting in little attitude polarization" (p. 206).

The issue of affirmative action may have produced such value conflict in a number of the subjects of Experiment 3, given the conflicting values of (a) desiring justice for discriminated minorities on the one hand and (b) not discriminating against deserving members of the majority (e.g., White men) on the other. In addition, the racial context of this issue and what some might view to be its relatively strong "political correctness" implications, could well invoke a measure of conflict in subjects. This account would suggest that subjects considering affirmative action experienced more value conflict than those considering the capital punishment issue. Although we have no direct evidence to support this interpretation, the value-conflict hypothesis of Liberman and Chaiken (1991) remains an intriguing possibility for the results of Experiment 3 as contrasted with Experiments 1 and 2.<sup>11</sup> Of course, this analysis focuses strictly on reported attitude polarization. The lack of a significant trinary index of polarization in Experiments 1 and 2 might, from the perspective of Liberman and Chaiken's study, suggest that the paradigm used in this research tends in general to induce value conflict in a substantial number of subjects.

In closing, we endorse, with others, the provocative and potentially significant implications of the attitude polarization concept as originally reported by Lord et al. (1979). On the basis of our findings, however, a measure of caution regarding this construct is in order. The dire conclusions with respect to the costs of reported attitude polarization, often drawn by reviewers of the Lord et al. (1979) study, seem to be premature.

<sup>11</sup> Although at this point one would be speculating, one might hypothesize that the smaller biased assimilation observed in the essay evaluations of Experiment 3 relative to that of Experiments 1 and 2 is at least consistent with the idea of greater value conflict.

## References

- Ajzen, I. (1988). *Attitudes, personality, and behavior*. Chicago: Dorsey.
- Anderson, C. A., & Sechler, E. S. (1986). Effects of explanation and counterexplanation on the development and use of social theories. *Journal of Personality and Social Psychology*, 50, 24-34.
- Berkowitz, L., & Devine, P. G. (1989). Research traditions, analysis, and synthesis in social psychological theories: The case of dissonance theory. *Personality and Social Psychology Bulletin*, 15, 493-507.
- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models for stereotype use. *Journal of Personality and Social Psychology*, 55, 726-737.
- Chaiken, S., & Yates, S. (1985). Affective-cognitive consistency and thought-induced attitude polarization. *Journal of Personality and Social Psychology*, 49, 1470-1481.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20-33.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Goethals, G. R., & Reckman, R. F. (1973). The perception of consistency in attitudes. *Journal of Experimental Social Psychology*, 9, 491-501.
- Greenberg, P. (1990, September 7). A trio of sensible thinkers on the subject of affirmative action. *Chicago Tribune*, Section 1, p. 21.
- Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 319-359). San Diego, CA: Academic Press.
- Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 38, pp. 369-425). Palo Alto, CA: Annual Reviews.
- Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgments subjectively. *Social Cognition*, 7, 51-66.
- Jelalian, E., & Miller, A. G. (1984). The perseverance of beliefs: Conceptual perspectives and research developments. *Journal of Social and Clinical Psychology*, 2, 25-56.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480-498.
- Liberman, A., & Chaiken, S. (1991). Value conflict and thought-induced attitude change. *Journal of Experimental Social Psychology*, 27, 203-216.
- Lord, C. G. (1989). The "disappearance" of dissonance in an age of relativism. *Personality and Social Psychology Bulletin*, 15, 513-518.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47, 1231-1243.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098-2109.
- Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research*, 1, 161-175.
- McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, 3rd ed., pp. 233-346). New York: Random House.
- Millar, M. G., & Tesser, A. (1986). Thought-induced attitude change: The effects of schema structure and commitment. *Journal of Personality and Social Psychology*, 51, 259-269.
- Myers, D. G. (1990). *Social psychology* (3rd ed.). New York: McGraw-Hill.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal report on mental processes. *Psychological Review*, 84, 231-259.
- Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Brown.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Plous, S. (1991). Biases in the assimilation of technological breakdowns: Do accidents make us safer? *Journal of Applied Social Psychology*, 21, 1058-1082.
- Pyszczynski, T., Greenberg, J., & Holt, K. (1985). Maintaining consistency

- tency between self-serving beliefs and available data: A bias in information evaluation. *Personality and Social Psychology Bulletin*, 11, 179-190.
- Quindlen, A. (1992, January 16). The great white myth. *New York Times*, p. A21.
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York: McGraw-Hill.
- Sherman, S. J., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology*, 16, 388-403.
- Sherman, S. J., Judd, C. M., & Park, B. (1989). Social cognition. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 40, pp. 281-326). Palo Alto, CA: Annual Reviews.
- Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 248-306). San Diego, CA: Academic Press.
- Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 289-338). San Diego, CA: Academic Press.
- Tesser, A., & Leone, C. (1977). Cognitive schemas and thought as determinants of attitude change. *Journal of Experimental Social Psychology*, 13, 262-270.
- Uleman, J. S., & Bargh, J. A. (Eds.). (1989). *Unintended thought*. New York: Guilford Press.
- Worchel, S., Cooper, J., & Goethals, G. R. (1991). *Understanding social psychology*. Pacific Grove, CA: Brooks/Cole.
- Zimbardo, P. G., & Leippe, M. R. (1991). *The psychology of attitude change and social influence*. New York: McGraw-Hill.

Received July 24, 1991

Revision received October 22, 1992

Accepted October 22, 1992 ■

## Low Publication Prices for APA Members and Affiliates

**Keeping You Up-to-Date:** All APA members (Fellows; Members; Associates, and Student Affiliates) receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*.

High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they can subscribe to the *American Psychologist* at a significantly reduced rate.

In addition, all members and affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

**Essential Resources:** APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the APA*, the *Master Lectures*, and *Journals in Psychology: A Resource Listing for Authors*.

**Other Benefits of Membership:** Membership in APA also provides eligibility for low-cost insurance plans covering life, income protection, office overhead, accident protection, health care, hospital indemnity, professional liability, research/academic professional liability, student/school liability, and student health.

**For more information,** write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242, USA