

# Subject bias and the retrospective pretest in retrospect

MIRJAM SPRANGERS

University of Amsterdam, Amsterdam, The Netherlands

(Pieter Koele, Sponsor)

Studies employing designs including a thentest were reviewed to examine to what extent they demonstrated a response shift, while controlling for subject bias by including a placebo control condition. A computer search was executed to detect articles that included the Howard, Ralph, et al. (1979) article in their reference list. Of 24 studies, 16 did not warrant any statement about the occurrence of a response-shift bias, relative to a subject bias, because they lacked a placebo control condition. Support for the subject bias hypothesis, relative to the response-shift hypothesis, was found in one of eight studies that did include a placebo control condition. With respect to these results, three supplementary considerations were taken into account. It was concluded that subject bias has not proved to be an unqualified alternative explanation of the response-shift interpretation in the reported studies.

The research design most frequently employed to evaluate training effectiveness is the two-group pretest-posttest design (Campbell & Stanley, 1966, Design 4). Cronbach and Furby (1970) noted that for pretest and posttest scores to be comparable, a common metric must exist for the two sets of scores. In using self-report instruments, Howard, Ralph, et al. (1979) found that a *response shift*—a change in internal standard from pretest to posttest—renders pretest and posttest scores incompatible. To obtain self-report measures of pretest and posttest levels of functioning with respect to the same internal standard, Howard, Ralph, et al. (1979) administered a retrospective pretest—a *thentest*—at the time of posttesting. Subjects were asked to report how they then perceived themselves to have been prior to the intervention. A significant mean difference between pre- and thenscores in the experimental condition would indicate a response shift. Since a response shift is dependent on the treatment, it should occur in the experimental and not in the control condition. If a response shift is demonstrated, a valid estimate of a treatment effect is given by the mean difference between post- and thenscores (Howard, Ralph, et al., 1979). However, designs including a thentest may be more susceptible to subject bias resulting from subjects' perception of the demand characteristics of the experimental situation (Orne, 1969) than Design 4 of Campbell and Stanley (1966). Since the thentest is administered at the same time as the posttest, subjects may, in an attempt to indicate favorable results, adjust their thenscores in a downward direction. Consequently, a mean pre/then difference score may not reflect a shift in internal standard, but may represent subject bias instead.

Traditionally, one of the best solutions to the problems raised by demand characteristics is to employ the medical placebo model, in which manipulations should appear identical to the subjects in all conditions (Carlsmith, Ellsworth, & Aronson, 1976; Rosenthal & Rosnow, 1969). An adequate design, therefore, should include a placebo control condition. Since placebo subjects devote the same amount of time and effort to the placebo treatment as do experimental subjects to the experimental intervention, a significant mean pre/then difference score in the placebo condition can be attributed to subject bias, thus invalidating the response-shift interpretation. To control for rival hypotheses, such as testing effects, maturation, and history, the design should also include a no-treatment control condition. To be able to exclude subject bias as a viable alternative explanation of the response-shift interpretation, the mean pre/then difference score should not be significant in either placebo or no-treatment control conditions.

The purpose of the present study was to explore the status of subject bias as a viable alternative explanation of the response-shift interpretation in the reported studies. We therefore examined to what extent these studies claimed to have demonstrated a response shift, and to what extent these studies controlled for subject bias by including a placebo control condition.

## REVIEW OF RESEARCH

With respect to the inclusion of control groups in designs including a thentest, three relevant categories can be distinguished: (1) designs that include experimental conditions only, (2) designs that include a no-treatment control condition in addition to the experimental condition(s), and (3) designs that include a placebo control condition and the experimental condition(s). These three categories are used in this paper to organize the review of empirical

---

This study was supported by grants from the Foundation for Educational Research (SVO), Project 1107. Requests for reprints should be sent to Mirjam Sprangers, Department of Psychology, University of Amsterdam, Weesperplein 8, 1018 XA Amsterdam, the Netherlands.

research pertinent to subject bias. A computer search on the *Social Sciences Citation Index* was executed to locate journal articles that included the Howard, Ralph, et al. (1979) article in their reference lists.

### 1. Absence of a Control Condition

Six studies did not include any control group at all (Howard & Dailey, 1979; Howard, Dailey, & Gulanick, 1979, Study 2; Howard, Ralph, et al., 1979, Studies 2 and 5; Howard, Schmeck, & Bray, 1979; Pohl, 1982). In all six studies a significant mean pre/then difference score was found, which led the researchers to conclude that a response shift had taken place. It must be clear, however, that designs not including a control condition do not warrant any statement about the cause of the findings, whether due to response shift or subject bias.

### 2. No-Treatment Control Condition

Ten studies included either a waiting-list or a no-treatment control condition. The results are summarized

in Table 1.<sup>1</sup> Nicholson, Belcastro, and Gold (1985) and Sprangers and Hoogstraten (in press, Study 1) did not find evidence for a response shift. Wexley and Baldwin (1986) showed that a comparison of post- and then-scores led to significant differences across experimental and control conditions. Since a pretest was not administered, there is no clue as to whether a response shift did occur. Perry and Apostol (1986) found that differences between the experimental and control groups reached significance only when the then-test was the covariate in an analysis of covariance on posttest scores. They did not directly test the mean pre/then difference score, but discussed the possibility of a response shift as an explanation for the results. A significant mean pre/then difference score was found only in the experimental conditions of the Porras and Singh (1986) study, and in the Sprangers and Hoogstraten (in press, Study 2) experiment. In both studies, the researchers concluded that a response shift had taken place. In the remaining four studies (Bray & Howard, 1980; Howard, Millham, Slaten, & O'Donnell, 1981;

**Table 1**  
Studies Utilizing Posttest-Then Test Designs, Including No-Treatment or Waiting-List Control Conditions

Research	Experimental Treatment	Control Treatment	Design	<i>n</i>	Pre/Then Difference	Response Shift
Howard, Ralph, et al. (1979), Study 3	Assertiveness training for women	Waiting list	Pr O X1 Po T O	17	?	Yes
			Pr O X2 Po T O	18	?	Yes
			Pr O -- Po T O	17	?	No
	Study 4	Assertiveness training for women	No-treatment	Pr O X Po T O	12	?
Pr O - Po T O	13			?	No	
Bray & Howard (1980)	Teacher training	No-treatment	O Pr X1 O Po T	<i>N</i> =38	?	Yes
			O Pr X2 O Po T		?	Yes
			O Pr X3 O Po T		?	Yes
			O Pr -- O Po T		?	?
Howard, Millham, Slaten, & O'Donnell (1981)	Assertiveness training	Waiting list	Pr X1 X2 Po T O	9	?	Yes
			Pr -- X2 Po T O	9	?	Yes
			Pr X1 -- Po T O	9	?	No
			Pr -- -- Po T O	9	?	No
Nicholson, Belcastro, & Gold (1985)	Stress counseling	No-treatment	Pr X Po T	<i>N</i> =45	?	No
			Pr - Po T		?	No
Perry & Apostol (1986)	Disability awareness seminar	Waiting list	Pr X Po T	32	?	Yes
			Pr - Po T	13	?	No
Porras & Singh (1986)	Communication skills training	No-treatment	Pr X Po T	38	Yes	Yes
			Pr - Po T	21	No	No
Wexley & Baldwin (1986)	Time-management skills	No-treatment	- X X1 Po T O	60	—	—
			- X X2 Po T O	65	—	—
			- X X3 Po T O	63	—	—
			- X -- Po T O	68	—	—
Sprangers & Hoogstraten (in press) Study 1	Communication skills training	Waiting list	Pr X Po T O	9	No	No
			-- X Po T O	9	—	—
			Pr - Po T O	9	No	No
			-- - Po T O	10	—	—
Study 2	Communication skills training	Waiting list	O Pr X O Po T	20	No	No
			- Pr X O Po T	9	Yes	Yes
			O Pr - O Po T	19	No	No
			- Pr - O Po T	10	No	No

Note—Pr = self-report pretest; Po = self-report posttest; T = then-test; O = objective measure; X = experimental treatment; X1, X2, X3 = alternative experimental treatments; - in design = not administered or no treatment; — in last two columns = not relevant; ? = not (directly) tested.

**Table 2**  
**Studies Utilizing Posttest–Thentest Designs, Including Placebo Control Conditions**

Research	Experimental Treatment	Control Treatment	Design	<i>n</i>	Pre/Then Difference	Response Shift
Terborg & Davis (1982)	Judging selection interviews	Part of experimental training	Pr X1 X2 Po T	12	No	No
			-- X1 X2 Po T	8	—	—
			Pr X1 X2 T Po	9	No	No
			Pr X1 X2 Po T	9	No	No
			Pr X1 X1 Po T	10	?	?
Hoogstraten (1982)	Problem solving strategies	Ethical permissibility research	Pr O X O Po T	25	Yes	Yes
			-- O X O Po T	23	—	—
			Pr O P O Po T	21	No	No
Hoogstraten (1985), Study 1	Film on parent-child interaction	Film on communication skills	Pr X1 X2 Po T O	22	No	No
			Pr -- X2 Po T O	20	Yes	Yes
			Pr X1 P Po T O	25	No	No
Study 2	Program on counseling	Program on free-attitude interview	Pr O X Po T O	28	Yes	Yes
			Pr - X Po T O	28	Yes	Yes
			Pr O P Po T O	29	No	No
Sprangers & Hoogstraten (1987)	Problem solving strategies	Ethical permissibility research	X1 Pr O X2 Po T O	24	No	No
			-- Pr O X2 Po T O	25	Yes	Yes
			X1 Pr O P Po T O	24	No	No
Sprangers & Hoogstraten (1988a)	Film on reanimation	Film on air pollution	X1 Pr O X2 Po T O	18	No	No
			-- Pr O X2 Po T O	18	No	No
			X1 Pr O P Po T O	17	No	No
Sprangers & Hoogstraten (1988b)	Text study training	Study planning training + no treatment	Pr O X O Po T	10	No	No
			Pr O P O Po T	10	No	No
			Pr O - O Po T	14	No	No
Sprangers & Hoogstraten (1988c)	Film on parent-child interaction	Film on communication skills + no treatment	Pr O X Po T O	17	Yes	Yes
			Pr O P Po T O	14	Yes	No
			Pr O - Po T O	15	No	No

Note—Pr = self-report pretest; Po = self-report posttest; T = thentest; O = objective measure; X = experimental treatment; X1, X2 = alternative experimental treatments; P = placebo treatment; - in design = not administered; — in last two columns = not relevant; ? = not (directly) tested.

Howard, Ralph, et al., 1979, Studies 3 and 4), the mean pre/then difference score was not directly tested. Because the post/then difference scores correlated more highly with objective measures of change than did the pre/post difference scores, the researchers tended to conclude that a response shift had occurred. Since these 10 designs did not control for subject bias, this alternative explanation can still account for these results. However, with respect to the last four studies, the high correlation of post/then difference scores with objective indices of change, is indeed an indication of the thentest's validity.

### 3. Placebo Control Condition

Of the eight relevant studies, summarized in Table 2 (see Note 1), Terborg and Davis (1982) offered their control subjects part of the experimental training, and the remaining studies included a real placebo condition. In two studies (Sprangers & Hoogstraten, 1988a; Terborg & Davis, 1982) significant mean pre/then difference scores were not found in the experimental conditions, although the treatments were effective. In two Hoogstraten experiments (1982; 1985, Study 2), a significant mean pre/then difference score was found in the experimental and not in the placebo control conditions. Therefore, it was concluded that a response shift had occurred. In addition, in the unmanipulated experimental conditions (conditions 2)

of the Hoogstraten (1985, Study 1) and the Sprangers and Hoogstraten (1987) experiments, a significant mean pre/then difference score was found. These differences were interpreted as response shifts. Taken together, these findings do not support the alternative hypothesis of subject bias, since a significant mean pre/then difference score was not found in the placebo control conditions, whereas in four of these six studies, it was found in the experimental conditions. Two studies included a placebo condition in addition to a no-treatment and experimental condition (Sprangers & Hoogstraten, 1988b, 1988c). Sprangers and Hoogstraten (1988b) found no significant mean pre/then difference in either condition. In Sprangers and Hoogstraten's (1988c) study, a significant mean pre/then difference score was found in both the experimental and placebo control conditions and not in the no-treatment control condition. Thus, in this latter experiment, the hypothesis of subject bias was supported.

## DISCUSSION

To summarize this overview, 16 of the 24 studies do not warrant any statement about the occurrence of a response-shift bias, relative to subject bias, because they lacked a placebo control condition. In only one of eight studies that did include a placebo control condition were

significant mean pre/then difference scores found in both experimental and placebo control conditions, thus supporting the hypothesis of subject bias.

Three supplementary considerations should be taken into account. First, in the Sprangers and Hoogstraten (1988b) study, significant mean post/then difference scores were found in both the experimental and placebo control conditions, and not in the no-treatment control condition. It was concluded that mean post/then difference scores were not free from subject bias.

Second, because the thentest is administered at the time of the posttesting, subject bias may be an artifact of the administration procedure. However, in three studies the order of the thentest and posttest was reversed (Sprangers & Hoogstraten, 1987; in press, Study 2; Terborg & Davis, 1982).<sup>2</sup> Results suggest that neither post- nor then-scores were affected. It was concluded that the thentest seems rather robust against order manipulations.

Third, Howard et al. (1981) investigated the operation of subject's response-style effects on then-scores using an alternative procedure. A prevention or reduction of response-style effects was planned, utilizing a bogus-pipeline technique at the posttesting. The results did not support the hypothesis of greater bias in then-scores. Sprangers and Hoogstraten (1987, 1988a), who induced the bogus-pipeline procedure at the pretesting, also found no support for subject's response-style effects in then-scores.

Taking these supplementary considerations into account, results suggest that subject bias has not proved to be an unqualified alternative explanation of the response-shift interpretation in the reported studies. However, since subject bias has been demonstrated, researchers should be alert to this potential bias in future research.

#### REFERENCES

- BRAY, J. H., & HOWARD, G. S. (1980). Methodological considerations in the evaluation of a teacher-training program. *Journal of Educational Psychology*, *72*, 62-70.
- CAMPBELL, D. T., & STANLEY, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- CARLSMITH, J. M., ELLSWORTH, P. C., & ARONSON, E. (1976). *Methods of research in social psychology*. California: Addison-Wesley.
- CRONBACH, L. J., & FURBY, L. (1970). How we should measure "change"—or should we? *Psychological Bulletin*, *74*, 68-80.
- HOOGSTRATEN, J. (1982). The retrospective pretest in an educational training context. *Journal of Experimental Education*, *50*, 200-204.
- HOOGSTRATEN, J. (1985). Influence of objective measures on self-reports in a retrospective pretest-posttest design. *Journal of Experimental Education*, *53*, 207-210.
- HOWARD, G. S., & DAILEY, P. R. (1979). Response-shift bias: A source of contamination of self-report measures. *Journal of Applied Psychology*, *64*, 144-150.
- HOWARD, G. S., DAILEY, P. R., & GULANICK, N. A. (1979). The feasi-

bility of informed pretests in attenuating response-shift bias. *Applied Psychological Measurement*, *3*, 481-494.

HOWARD, G. S., MILLHAM, J., SLATEN, S., & O'DONNELL, L. (1981). Influence of subject response style effects on retrospective measures. *Applied Psychological Measurement*, *5*, 89-100.

HOWARD, G. S., RALPH, K. M., GULANICK, N. A., MAXWELL, S. E., NANCE, D. W., & GERBER, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychological Measurement*, *3*, 1-23.

HOWARD, G. S., SCHMECK, R. R., & BRAY, J. H. (1979). Internal invalidity in studies employing self-report instruments: A suggested remedy. *Journal of Educational Measurement*, *16*, 129-135.

NICHOLSON, T., BELCASTRO, P. A., & GOLD, R. S. (1985). Retrospective pretest-posttest analysis versus traditional pretest-posttest analysis. *Psychological Reports*, *57*, 525-526.

ORNE, M. T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143-179). New York: Academic Press.

PERRY, D. C., & APOSTAL, R. A. (1986). Modifying attitudes of business leaders toward disabled persons. *Journal of Rehabilitation*, *52*, 35-38.

POHL, N. F. (1982). Using retrospective pre-ratings to counteract response-shift confounding. *Journal of Experimental Education*, *50*, 211-214.

PORRAS, J. I., & SINGH, J. V. (1986). Alpha, beta, and gamma change in modelling-based organization development. *Journal of Occupational Behaviour*, *7*, 9-24.

ROSENTHAL, R., & ROSNOW, R. L. (Eds.). (1969). *Artifact in behavioral research*. New York: Academic Press.

SPRANGERS, M., & HOOGSTRATEN, J. (1987). Response-style effects, response-shift bias and a bogus pipeline. *Psychological Reports*, *61*, 579-585.

SPRANGERS, M., & HOOGSTRATEN, J. (1988a). Response-style effects, response-shift bias and a bogus pipeline: A replication. *Psychological Reports*, *62*, 11-16.

SPRANGERS, M., & HOOGSTRATEN, J. (1988b). *Subject bias in three self-report measures of change*. Manuscript submitted for publication.

SPRANGERS, M., & HOOGSTRATEN, J. (1988c). On delay and reassessment of retrospective ratings. *Journal of Experimental Education*, *56*, 148-153.

SPRANGERS, M., & HOOGSTRATEN, J. (in press). Pretesting effects in retrospective pretest-posttest designs. *Journal of Applied Psychology*.

TERBORG, J. R., & DAVIS, G. A. (1982). Evaluation of a new method for assessing change to planned job redesign as applied to Hackman and Oldham's job characteristic model. *Organizational Behavior & Human Performance*, *29*, 112-128.

WEXLEY, K. N., & BALDWIN, T. I. (1986). Posttraining strategies for facilitating positive transfer: An empirical exploration. *Academy of Management Journal*, *29*, 503-520.

#### NOTES

1. The designs summarized in Tables 1 and 2 are confined to only one self-report and one performance measure, if they were administered, and to the (combination of) conditions in which the relevant analyses were performed.

2. With respect to the Sprangers and Hoogstraten studies (1987, in press), originally there were more conditions. In these additional conditions the order of the posttest and thentest was reversed. Since data indicated that this manipulation did not affect the scores, conditions were combined, resulting in the designs presented in Tables 1 and 2.

(Manuscript received February 20, 1988.)