

IS IT TIME TO LAY THE SLEEPER EFFECT TO REST? ¹

PAULETTE M. GILLIG AND ANTHONY G. GREENWALD ²

Ohio State University

The "sleeper effect" has been defined as a delayed increase in the persuasive impact of a communication from a source low in credibility. This effect failed to appear in the present series of seven replications (total $N = 656$) of a sleeper effect experiment. However, the effectiveness of communications attributed to a source high in credibility did significantly decrease with passage of time when subjects were unprepared to counterargue with the communication. On the basis of these results and a retrospective review of related studies, it was concluded that the latter effect, by itself, could account satisfactorily for previous findings that source credibility effects dissipate or disappear with passage of time.

As many undergraduate psychology students have by now learned, the "sleeper effect" is an established phenomenon in the attitude change literature. This effect is said to occur when a communication from a source of low credibility has a greater persuasive impact after some time delay than on original exposure. The communication works, that is, while the audience "sleeps" on it. Descriptions of the sleeper effect may be found in a variety of social psychology texts. Examination of the authors' bookshelves indicated that approximately 75% of recent survey texts make explicit reference to the sleeper effect as an established phenomenon.

The negative results of seven replications of a sleeper effect experiment to be reported here prompted the authors to reexamine the empirical literature on this phenomenon. Remarkably, this retrospective review led to the conclusion that the sleeper effect has never had a truly satisfactory empirical basis. The original report of a sleeper effect by Hovland, Lumsdaine, and Sheffield (1949) was based on only a subset of the opinion items they

used; these authors did not report any criteria that could be used to sort items, a priori, into ones that would show a sleeper effect and ones that would not. In subsequent reports of a sleeper effect, a statistically significant increase in opinion for a group receiving a communication from a low-credibility source has *never* been reported!³ Typically, statistical significance has been achieved in sleeper effect studies by contrasting the nonsignificant (or absent) delayed impact of a communication from a low-credibility source with a substantial decrease in impact over time in a group receiving the same communication from a source high in credibility (e.g., Hovland & Weiss, 1951; Kelman & Hovland, 1953; Shulman & Worrall, 1970; Watts & McGuire, 1964). This type of interaction result (i.e., an interaction effect of source credibility and time of posttest on an opinion measure) appears to be reliable—it was also found in the present research—but should not be taken to be an equivalent of the phenomenon that was originally designated by Hovland et al. (1949) as the sleeper effect. That phenomenon was an increase, over time, in the persuasive impact of a communication from an untrustworthy source.

The original aim of the research to be reported here was not to test the sleeper effect. Rather the sleeper effect was intended to provide the basis for a study of the persistence of rejection of persuasion. The hypothe-

¹ This research was supported in part by Grants GS-1601 and GS-3050 from the National Science Foundation and by Grant 1-R01-MH20527-01 from the United States Public Health Service. The authors are grateful to Thomas D. Cook and Walter Weiss for critically reading an earlier draft of this article; to David D. S. Poor for assistance in data analysis; and to James A. Chase, Merry Edison, Shiloh Glass, Peter Jay, Kurt Kahler, and Richard Prée for assistance in conducting the studies.

² Requests for reprints should be sent to Anthony G. Greenwald, Department of Psychology, Ohio State University, 404C West 17th Avenue, Columbus, Ohio 43210.

³ The authors are indebted to an unpublished review of the sleeper effect literature, by Thomas D. Cook, for first drawing attention to this surprising absence of significant effects.

sis for the originally planned study derived from a distinction between two categories of cognitive responses (Greenwald, 1968) that might serve as differentially persistent bases for rejecting persuasion. A counterargument was defined as rejection of a communicated argument occurring as a response to the argument's content, such as through recall of information in logical conflict with argument content; a discounting response was defined as rejection of an argument on the basis of context rather than content, such as when the argument is attributed to a distrusted source. The hypotheses formulated in connection with these response categories could not be tested in the present research because of the failure to obtain any sleeper effect. Also, the quest for the sleeper effect gradually became the primary aim of the research as one replication led to the next.

METHOD

Subjects

The 717 subjects who participated in the present series of experiments were either paid volunteers solicited by advertisement in the Ohio State student newspaper (Replications 1, 6, and 7) or were volunteers recruited from the introductory psychology course (Replications 2, 3, 4, and 5). Those who received the experimental communication for Replication 4 provided control data on the opinion topic of Replication 5 and vice versa; Replications 6 and 7 were similarly paired. Data for a total of 61 subjects were omitted from analyses: 19 who could not be recontacted for delayed posttests, 8 who provided incomplete questionnaire responses, and 34 (almost all of these in Replication 2) whose responses to a source-identification question indicated they were not attending when the source was described.

General Procedure

The experimental procedures were administered almost entirely by means of videotaped presentations of instructions, communications, and dependent measures. The only exceptions to this were the use, for some replications, of telephoned interviews for opinion posttests. For those telephoned posttests, opinion measures were administered by audiotape after the experimenter had ascertained that the subject was on the line. The separation between the initial session, during which the subject watched and listened to a communication and provided immediate posttest data, and the collection of delayed posttest data was 10-14 days. Procedures were always administered individually to subjects, with up to four subjects simultaneously viewing the same videotape in separate rooms during laboratory sessions.

Opinion Topics

Three different communications were used. Two of the communications attacked cultural "truisms" and had previously been used by McGuire (1964) in studies of resistance to persuasion. One of these opposed the widespread use of penicillin (used in Replications 1, 3, and 6), and the other argued against the practice of routine annual medical check-ups (used in Replications 2, 5, and 7). Replication 4 employed a communication arguing in favor of massive daily doses of Vitamin C, adapted from Pauling (1970). The criteria for selecting opinion topics were that (a) subjects would not normally have sufficient informational resources to counterargue, that is, to reject the communicated arguments on the basis of their content, and (b) adequate counterarguments which could be administered in pretreatments were available. The counterarguments for the two truism topics were available from McGuire's (1964) "refutational-same" defenses for those topics; for the Vitamin C topic, counterarguments were available from a *Consumer Reports* article (Anonymous, 1971) written in response to Pauling's book.

Independent Variables

Source credibility. Control subjects excluded, in Replications 1-3 half of the subjects received a communication from a source of high credibility and half from a source of low credibility. In Replications 4-7, only the low-source-credibility condition was used. The reason for dropping high-credibility sources from the latter replications was that the research had become focused on the attempt to produce the sleeper effect of delayed impact of communications from sources low in credibility. For this purpose, high-credibility conditions were unnecessary. In Replications 1 and 2, the low-credibility source was described as a "nature therapist," this being elaborated as a person who was generally antagonistic to therapeutic use of drugs and to other modern medical techniques. The high-credibility source was described, in contrast, as a "leading medical researcher," whose research was on the topic of the communication. For the subsequent replications the source-credibility manipulation was strengthened, although its external validity was perhaps reduced, by simply informing subjects in the low-credibility condition that the content of the communication was not necessarily drawn from factual material. For the high-credibility condition of Replication 3 subjects were informed that the communication consisted strictly of factually accurate material.

Defense pretreatment. In all seven replications, for half of the subjects the persuasive communication was preceded by a counterargument defense. This consisted of an integrated series of arguments that could be used, on a point-by-point basis, to refute every major argument in the communication that would subsequently be received. The counterargument defense preceded the persuasive communication by just a few minutes and was not attributed to any source; rather it was presented simply as a

TABLE 1
COGNITIVE RESPONSES AS A FUNCTION OF SOURCE CREDIBILITY AND DEFENSE PRETREATMENT

Condition	N	Response category			
		Agreement	Discounting	Counterarguing	Neutral
High credibility source					
No defense	49	5.6	2.5	1.7	2.2
Counterargument defense	41	5.8	2.2	1.8	2.2
Low credibility source					
No defense	44	3.2	3.9	2.3	2.6
Counterargument defense	52	2.5	3.3	3.8	2.4
Combined means					
High credibility	90	5.7	2.4	1.7	2.2
Low credibility	96	2.9	3.6	3.0	2.5
No defense	93	4.4	3.2	2.0	2.4
Counterargument defense	93	4.2	2.8	2.8	2.3

Note. Data for Replications 1-3 only.

"communication" that was being used in the research program in which the subjects were participating. The remaining half of the subjects received an irrelevant communication on genetic determinants of behavior in place of the counterargument defense. The purpose of the counterargument defense pretreatment was to enable half of the subjects to counterargue during the subsequent communication; those not receiving the prior defense were expected to lack the informational resources necessary to engage in such counterarguing.

Dependent Variables

Opinion. Four self-rating items, each with a 15-point response scale ranging from definitely true to definitely false, were used to measure opinions on the communication topic in each experiment. For the topics used previously by McGuire (1964), these items were identical to his. Comparable items were devised for the Vitamin C topic. In all replications, subjects who were tested immediately following the communication also received delayed retests with the same items 10-14 days later. In Replications 6 and 7, additional groups of subjects received an opinion posttest only at the end of the delay interval. In Replications 1 and 2, the delayed posttest was administered by telephone; in Replications 3-5, both immediate and delayed posttests took place in the laboratory; in Replications 6 and 7, both posttests were conducted by telephone, the immediate posttest being within 24 hours of the laboratory session in which the persuasive communication was received.

Cognitive responses. Half of the subjects within each condition in Replications 1-5 received the opinion posttests as described above, while the remainder provided information used to infer the nature of cognitive responding that had occurred during receipt of the communication. This assessment used an "instant replay" technique. Immediately after the first exposure to the communication, it was replayed in 12 short paragraph segments. The subject was asked to attempt to recall and report in writing, after

each segment, the thoughts that had occurred to him upon the original presentation of that segment of the communication. Subsequently these written reactions were sorted, using a four-category system, by judges who were kept ignorant of the subjects' condition assignments. The four categories were (a) agreement, (b) discounting response, (c) counterargument, and (d) neutral. The discounting response and counterargument categories were described to the judges in a manner consistent with the definitions of these categories presented earlier. Interjudge agreement on the application of the category system to a randomly selected subset of 20 subjects was satisfactorily high, with agreement on 81% of judgments (chance agreement base line = 38%).

Summary of design. For Replications 1-3, the design was a 2³ factorial, with two levels each of source credibility (high or low), pretreatment (counterargument defense or no defense), and type of dependent measure (opinion or cognitive response). For subjects receiving the opinion measure, time or measurement was a fourth (repeated-measure) factor added to the 2³ design. For Replications 4-7, this design was reduced by the omission of the high-credibility conditions. For Replications 6 and 7, groups receiving the cognitive response dependent measure were replaced by groups receiving the delayed posttest opinion measurement only. Beyond these basic designs, Replication 3 included two control groups, providing opinion scores (a) in the absence of a communication or (b) after receiving the communication with no source mentioned. Additionally, as mentioned earlier, Replications 4 and 6 were paired with 5 and 7, respectively, in that the subjects in each provided control (no-communication) opinion data for the replication with which they were paired.

RESULTS

In order to simplify the presentation of results and also to reduce the noise contained

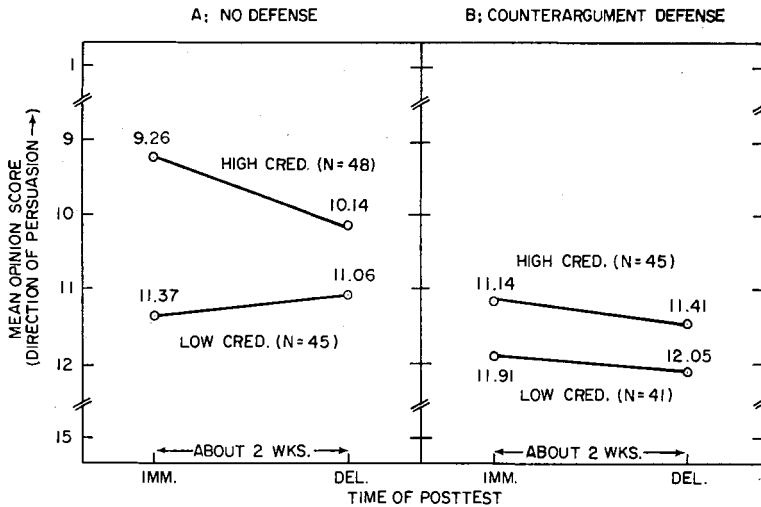


FIGURE 1. Posttest opinion as a function of source credibility and time of posttest for no defense (A) and counterargument defense (B) pretreatments. (Data combined over Replications 1-3.)

in the variations of results from replication to replication, analyses to be reported were combined across all replications that shared a given design.⁴

Effects on Cognitive Responses

Table 1 presents a summary of data for Replications 1-3 illustrating the effects of source credibility and defense pretreatment on the content of cognitive responses to the communication. Recall that these data were obtained very shortly following exposure to the persuasive communication and were obtained by the instant replay method from groups of subjects who did not provide opinion posttest data.

The cognitive response data showed that the manipulations had their intended effects. High-source credibility produced about twice as many agreement responses as did low-source credibility ($F = 58.3$, $df = 1/174$, $p < .001$) and correspondingly fewer disagreeing responses (discounting responses plus counterarguments). The counterargument defense pretreatment enhanced counterarguments relative to the no-defense condition ($F = 7.0$, $df = 1/174$, $p < .01$). This latter effect was

⁴ A replication-by-replication summary of the data may be obtained by writing to the second author (see Footnote 2).

strong, however, only within the low-source-credibility condition, an observation probably associated with the general suppression of both categories of disagreeing responses in the high-credibility condition.⁵

An alternative way of examining the cognitive response data is to observe that (a) within both high-credibility conditions, the most frequent response by far was the agreement response, while (b) within the low-credibility-counterargument defense condition the counterarguing response category was the most frequent, and (c) in the low-credibility-no-defense condition, discounting was the most frequent response. These observations indicated that both the credibility and defense manipulations had impact as intended on cognitive responses to the persuasive communication.

Effects on Opinion Posttests

Figure 1 presents the immediate and delayed opinion posttest data for Replications

⁵ The cognitive response data for Replications 4 and 5 have not been included in the statistical tests reported here or in Table 1 because these replications did not include high-credibility conditions. The effect of the defense pretreatment on counterarguments was, in fact, essentially the same for Replications 4 and 5 as that shown in Table 1 for low-source-credibility conditions of Replications 1-3.

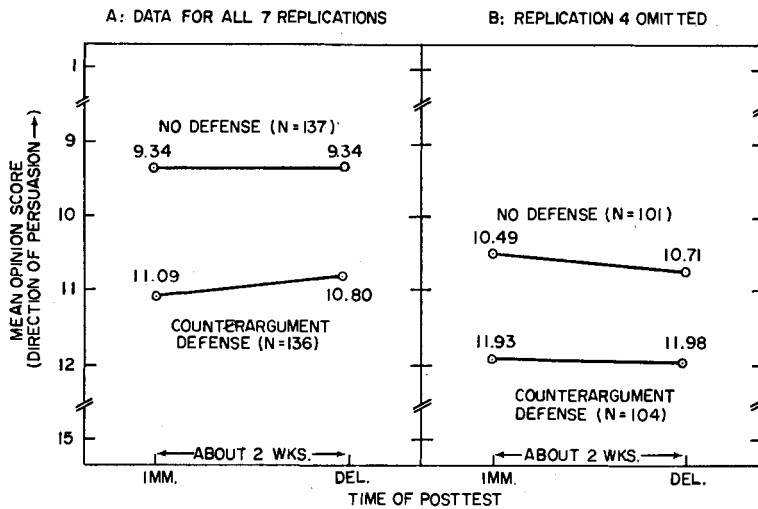


FIGURE 2. Posttest opinion as a function of time of posttest and defense pretreatment. (Low-source-credibility conditions only combined over Replications 1-7, A, and with Replication 4, in which a boomerang effect occurred, omitted, B.)

1-3 in which source credibility was a design factor. Figure 1A gives the results for the no-defense conditions only. The pattern shown here is a very familiar one in sleeper effect studies. There was (a) a slight (nonsignificant) increase in agreement with the communication between immediate and delayed posttest for the low-credibility condition ($F < 1$), (b) a more substantial decrease in agreement for the high-credibility condition ($F = 7.6$, $df = 1/167$, $p < .01$), and (c) a significant Source Credibility \times Time of Posttest interaction ($F = 7.7$, $df = 1/167$, $p < .01$). In contrast, none of the components of this pattern was apparent in the data for the counterargument defense conditions as shown in Figure 1B. When all of these data were combined in a single analysis, the Credibility \times Defense \times Time of Opinion Measurement interaction approached significance ($F = 2.8$, $df = 1/167$, $p < .10$). This suggests that the interaction pattern of Figure 1A, a pattern typical of prior sleeper effect studies, may depend on the use of topics for which audiences are unprepared to counterargue.

Figure 2A presents a test of the sleeper effect, including data for all conditions in which that effect might have been expected

to occur.⁶ These include the defense and no-defense variations of the low-credibility conditions in each of Replications 1 through 7. As may be seen readily, the obtained functions relating opinion to time since posttest were essentially flat. The standard deviation for the mean difference between immediate and delayed opinion posttest was .14 on the 15-point opinion scale for both defense conditions combined and .19 for each defense condition separately. Accordingly, the series of replications had experimental power sufficient to detect a true sleeper effect as small as .50 scale points overall (or .71 scale points within either defense condition) with 95% probability, using a two-tailed .05 alpha criterion.

Control opinion data. Because opinion data were obtained for no-communication control groups only for Replications 3-7, control data were not included in Figures 1 and 2. The chief importance of the control groups was to assure that the low-credibility communications were not producing boomerang effects,

⁶ The results for subjects who, in Replications 6 and 7, provided opinion data only at the delayed posttest are not included in Figure 2. However, these data agreed closely with those for subjects observed at both posttests and would not at all alter the pattern shown.

that is, less agreement with the viewpoint of the communication than would have been shown by subjects not receiving it. If such boomerang effects were occurring, then an increase in agreement over time could be attributed to the dissipation of a boomerang effect rather than to a delayed persuasive impact. The control data indicated that, as desired, the communications for the medical checkup and penicillin topics were not producing any boomerang effects. However, the Vitamin C communication used in Replication 4 did produce a boomerang effect; that is, mean immediate posttest opinion scores for groups receiving this communication showed significantly less agreement with the communication than was manifested by the no-communication control group. Further, the Replication 4 data did show an increase of agreement on the delayed posttest, but this was not a sleeper effect since the delayed posttest opinion means were still below the mean for the no-communication control group. Accordingly, inclusion of Replication 4 together with the other six replications in Figure 2A has introduced some artifactual bias in favor of finding a pattern of results that would look like a sleeper effect. The fact that the sleeper effect was nonetheless quite apparently absent strengthens the basis for interpreting the present results as questioning the existence of that phenomenon. With the Replication 4 data removed (see Figure 2B), the curves for both defense conditions exhibit slight downward agreement trends over time, in clear opposition to the expected upward trend that would be characteristic of a sleeper effect.

DISCUSSION

A few positive findings of this study should not be totally overshadowed by the major negative finding, that is, the failure to obtain a sleeper effect. Before discussing this major finding the positive findings are reviewed briefly.

Use of the Cognitive Response Measure

A novel methodological aspect of this study was the use of a set of groups that received the measure of cognitive responding to persuasion run in parallel with the groups tested for opinion effects of the communication

under varying conditions of source credibility and prior defense. This procedure was hoped to provide as direct an assessment as possible of the cognitive response processes that mediate the effects of communication independent variables on the dependent variable of opinion. The results of this series of experiments indicated that this goal was successfully accomplished. The credibility and defense pretreatment manipulations produced easily detectable effects on cognitive responding as determined by both multivariate (not reported here) and univariate F tests of the several cognitive response categories.

Effects of Source Credibility and Defense Pretreatment

The main effects on opinion of source credibility and defense pretreatment were quite in line with previous research on the effects of these variables. Briefly, persuasion was reduced by either inducing a belief in the untrustworthiness of the message or by providing the audience with informational resources adequate to counterargue. The results for the cognitive response measurement groups that received these inductions indicated that the reduction in persuasion associated with the low-credibility source was accompanied by a reduction in agreeing responses to the communication and a corresponding increase in discounting responses and counterarguments. The effect on cognitive responses of the defense pretreatment was particularly apparent only within the low-credibility conditions; in these the defense, as expected, increased the frequency of counterarguments. The fact that the defense pretreatment did not produce a noticeable increase in counterarguing for the high-source-credibility conditions (see Table 1) suggests that the effectiveness of a communication from a high-credibility source may in part be due to the suppression of counterarguments of which the audience is aware.

Distinction between Discounting Responses and Counterarguments

The theoretical analysis of differences between discounting and counterargument responses as bases for persuasion rejection has not been developed in this report, primarily

because the series of experiments turned out not to create the situations necessary for testing this analysis. Nonetheless, two aspects of the present data suggest that this may be a distinction worth making. First, the defense pretreatment affected measures of discounting and counterargument responses differently; within the low-credibility conditions, the defense increased counterarguing while slightly decreasing discounting responses. Second, the interaction pattern of Figure 1A was not replicated in the counterargument defense conditions (Figure 1B), suggesting that the capacity to counterargue might prevent the occurrence of this interaction effect.

Epitaph for the Sleeper Effect?

With the reader's indulgence, the authors would prefer to relax into first person in order to offer some commentary on the present findings and on the sleeper effect more generally.

The fact that we were unable to obtain a sleeper effect—that is, a greater delayed than immediate impact of a communication from a low-credibility source—can be interpreted in two ways: (a) the phenomenon may not exist or (b) our experimental tests may have been inadequate to detect it. Because we have done our best to create the conditions which, according to past reports, should have been sufficient to obtain a sleeper effect, we have difficulty in accepting the second interpretation. Additionally, we have employed large numbers of subjects, many more than in the prior laboratory investigations of this phenomenon, so that a true sleeper effect as small as .5 point on the 15-point opinion scale might have been detected with 95% probability (power) using a two-tailed .05 alpha criterion. While it is true that a restricted range of issues was used in our research—we used three health-related topics—there was certainly no reason to suspect, a priori, that the sleeper effect should not have applied to such issues.

Perhaps the strongest support for our suggestion that the sleeper effect may not exist comes from the coupling of our present data with a retrospective examination of previous studies (see also Footnote 3). That examination revealed that *no* laboratory investigation

has fulfilled what should be regarded as the sine qua non for concluding that the sleeper effect exists—specifically, a statistically reliable increase in persuasion over time. As we mentioned earlier, the only investigation reporting such a significant increase was the Hovland et al. (1949) study. An important deficiency of that investigation was that the sleeper effect phenomenon was reported for only a subset of a large group of opinion items that had been employed.

It should be stressed that our conclusions apply only to the delayed action effect defined by Hovland et al. (1949) and others in terms of initially ineffective communications from untrustworthy sources. Other delayed action effects are possible (McGuire, 1969, pp. 254–256) but have received much less research attention than the sleeper effect. It would certainly be inappropriate to interpret the present report as casting doubt on the possibility of *any* delayed action persuasion effects.

As an exercise in the history of science, we have asked ourselves why the sleeper effect has achieved such a strong foothold in the traditions and archives of social psychology, given the inadequacy of empirical support for its existence. A few circumstances appear to be responsible for this situation. First, the investigations (Hovland & Weiss, 1952; Kelman & Hovland, 1953) that followed up on the original Hovland et al. (1949) finding altered somewhat the empirical criterion of the sleeper effect from a significant increase in persuasion at a time delay to a significant interaction effect involving source credibility and time of posttest opinion measurement. As a result, those investigations and more recent ones in their wake have reported a "sleeper effect" when the earlier criterion had not been met. Additionally, Kelman and Hovland (1953) offered an explanation for the sleeper effect, which they described as a consequence of a dissociation, over time, of message content from message source. This discovery was perhaps crucial to the sleeper effect's becoming an established phenomenon. In retrospect, it would have been more cautious of textbook writers and literature reviewers to conclude from Kelman and Hovland's important study that (a) yes, source

and content of a message are dissociated over time, and (b) yes, the effect of source credibility is not likely to be apparent a few weeks after exposure to a persuasive communication, and (c) yes, the effect of source credibility can be reinstated by reminding the audience, at the time of a delayed posttest, about the message and its source, but (d) all of this may be perfectly true even if there is no such thing as a sleeper effect! In other words, there is nothing about the dissociation hypothesis that requires an increase in persuasion over time.

A viable alternative to burying the sleeper effect is to speculate that the sleeper effect requires conditions that have not been captured in our own or others' procedures. Although this alternative cannot be excluded, we feel that our results, interpreted along with those of other investigators, (a) indicate surely that the sleeper effect cannot be a robust phenomenon and (b) shed no light on the conditions that might be necessary to produce it. In short, if the sleeper effect is alive, we do not know where it is living.

REFERENCES

- ANONYMOUS. Linus Pauling and the common cold. *Consumer Reports*, 1971, 36, 113-114.
- GREENWALD, A. G. Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes*. New York: Academic Press, 1968.
- HOVLAND, C. I., LUMSDAINE, A. A., & SHEFFIELD, F. D. *Experiments on mass communication*. Princeton, N.J.: Princeton University Press, 1949.
- HOVLAND, C. I., & WEISS, W. The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 1951, 15, 635-650.
- KELMAN, H. C., & HOVLAND, C. I. "Reinstatement" of the communicator in delayed measurement of opinion change. *Journal of Abnormal and Social Psychology*, 1953, 48, 327-335.
- MCGUIRE, W. J. Inducing resistance to persuasion: Some contemporary approaches. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 1. New York: Academic Press, 1964.
- MCGUIRE, W. J. The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology*. (2nd ed.) Vol. 3. Reading, Mass.: Addison-Wesley, 1969.
- PAULING, L. *Vitamin C and the common cold*. San Francisco: Freeman, 1970.
- SHULMAN, G. I., & WORRALL, C. Salience patterns, source credibility, and the sleeper effect. *Public Opinion Quarterly*, 1970, 34, 371-382.
- WATTS, W. A., & MCGUIRE, W. J. Persistency of induced opinion change and retention of the inducing message contents. *Journal of Abnormal and Social Psychology*, 1964, 68, 233-241.

(Received June 12, 1972)