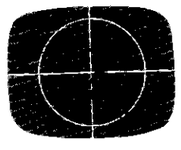


Pathways to Knowledge: Experimentation and the Analysis of Television's Power

The word "experiment" means many things, not only to the proverbial person in the street, but to social scientists as well. Consequently, it is important for us to be clear at the outset about what we mean—and do not mean—by experiment. For us, the essence of the true experiment is control. Experiments of the sort we have undertaken here are distinguished from other systematic empirical methods in the special measure of control they give to the investigator. In the first place, the experimenter *creates* the conditions under investigation, rather than waiting for them to occur naturally. In the second place, the experimenter *randomly assigns* individuals to those conditions, thereby superseding natural processes of selection. By creating the conditions of interest, the experimenter holds extraneous factors constant and ensures that individuals will encounter conditions that differ only in theoretically decisive ways. By assigning individuals to conditions randomly, the experimenter can be confident that any resulting differences between individuals assigned to varying conditions must be caused by differences in the conditions themselves.

Although these features may seem innocuous, they are not: experiments possess genuine advantages over alternative approaches, particularly in the insight they provide into causation. Because this claim may appear pretentious, we will try to illustrate it with a simple example.

Suppose (as is true) that we are interested in the influence television news might exercise over the American public. Like many researchers before us, we therefore decide to interview a sample of Americans, carefully selected to be representative of the nation as a whole. We partition our sample into two groups, those who tell us that they rely primarily upon television news for their information about politics and those who say they rely on other sources. We then compare the political views of the two groups and discover that the television-reliant group regards unemployment as the country's most serious problem while the other group names inflation. Simultaneously, we undertake a content analysis of television news coverage, finding that during the

unemployment. We conclude that television news does indeed shape its viewers' conceptions of political reality.

There is nothing wildly foolish about this hypothetical study. It does have serious limitations, however, just those that a true experiment counteracts. First and foremost, our hypothetical study cannot establish causal relationships. Observing that television news coverage and viewers' beliefs correspond is not the same as establishing that television coverage influences viewers' beliefs. No doubt the television-reliant group differs in many ways from those who obtain their information elsewhere, and it may be these differences that are responsible for generating different outlooks on national problems. If the television group is disproportionately working-class, for example, their special concern about unemployment might be due not to television coverage but to their own experiences in the labor force, or the experiences of their friends and coworkers. Of course, we might be able to test this particular explanation by partitioning the sample into different occupational groups and then examining whether the same relationship between television viewing and political outlook was maintained within each occupational group. But we could never know for certain whether *all* such plausible rival explanations had been ruled out.

This limitation of nonexperimental research is debilitating for causal inference, and it is exactly the problem overcome in experimentation through random assignment. By randomly assigning some people to television and others to newspapers (ignoring for the moment any ethical and practical difficulties that might stand in the way), the experimenter can be certain that whatever differences are detected between the two groups can be traced to differences in the treatments. Alternative explanations due to preexposure differences—associated with class, unemployment experiences, *or anything else*—are made untenable by this simple procedure.

Suppose we could, through some miracle of omniscience, rule out all plausible rival interpretations for our survey result. We would then know that television's agenda influenced the public—but little more than that. We could only speculate about which characteristics of television news are responsible for shaping public opinion. This limitation highlights another advantage of experimentation. By creating treatment and control conditions, the experimenter can isolate one causal variable at a time, thereby opening the way for a more refined understanding of how television influences its audience. Unlike the survey analyst, the experimenter need not wait for history to provide crucial tests.

Despite these advantages, true experiments are few and far be-

tween in the study of politics. There are notable exceptions: Gosnell's (1927) studies of voter turnout in Chicago more than a half-century ago (also see Eldersveld 1956); Campbell's (1969b) advocacy of an experimental approach to evaluating social reforms; and recent experiments on political bargaining and committee decision-making (e.g., Fiorina and Plott 1978). But by and large, students of politics have been reluctant to intervene experimentally in natural political processes. We suffer no such hesitation; we move ahead, confident that experiments bestow advantages unattainable by other means.

The Experimental Designs[1]

overview

All our experiments were advertised as studies of citizens' reactions to television news programs and followed one of two basic designs. In *sequential* experiments, participants were exposed to a sequence of unobtrusively altered network newscasts. Participants viewed one thirty-minute newscast every day over the course of one week. On the experiment's first day, participants came to a building on the campus of Yale University and were instructed concerning the objectives and procedures of the study. They were told that it was necessary for them to watch the news at the University in order to avoid distractions present at home and to ensure that everyone watched the same broadcasts under identical conditions, and that during the length of the experiment they should *not* watch the evening national news at home. They then completed a questionnaire that covered a variety of political topics. Most germane to our interest in agenda-setting and priming, participants were asked to name the country's most important problems and to evaluate the president's performance in office. After completing the questionnaire, participants were shown an unedited videotape recording of the previous evening's national newscast drawn from one of the three major networks. Over the next four days, participants continued to view what they believed to be simply a recording of the previous evening's network newscast. In fact, unknown to participants, we had altered sections of the newscast ahead of time in order to achieve systematic experimental variations in the amount and nature of coverage given national problems. In experiment 2, for example, participants randomly assigned to one condition received a steady dose of news alleging inadequacies in United States defense capability; a second group watched newscasts that paid special attention to pollution; a third group saw newscasts that emphasized economic problems. Each condition was characterized not only by sustained coverage of the target problem (e.g., defense), but also by systematic exclusion of news stories dealing with the other two (pollution and the economy). Thus each condition served as a control group for the others. On the final (sixth) day of the experiment, participants completed a second questionnaire that again inquired into their beliefs about the country's problems and the president's performance.

In *assemblage* experiments, in contrast, participants viewed a collection of news stories taken from the three networks at a single sitting—in different studies, as few as eight stories to as many as thirteen. The presentations were described as a cross-section of typical news stories broadcast by the major networks during the past year. As in sequential designs, the presentations in assemblage experimentations were put together in order to test propositions about agenda-setting and priming. Immediately after the presentations—and *only* afterwards, in contrast to sequential experiments—participants were questioned about their political views.

Although the assemblage design simulates less faithfully the ordinary American's encounters with television news than does its sequential counterpart, it does have one decisive advantage. Because assemblage treatment conditions can be prepared well in advance of the running of the experiment itself, they can be calibrated precisely. Hence assemblage experiments are particularly useful for testing relatively subtle ideas about how television might work its influence. Experiment 3, for example, examined whether priming is enhanced when television news both pays substantial attention to a problem and implies that the president is responsible for the problem. Participants in this experiment saw either many stories describing America's growing dependence on foreign oil or only a few; moreover, the stories implied either that the president had a great deal to do with the nation's energy problems or that the causes of and solutions for the country's energy predicament lay elsewhere. Crossing the two factors—exposure and responsibility—results in four experimental treatments. (Participants in a fifth, control, condition saw no stories about energy at all.)

In summary, participants in sequential experiments viewed unobtrusively edited newscasts over the course of a week. In assemblage experiments, participants watched a single collection of news stories in one session. The assemblage design is obviously less realistic, but in partial compensation, it enables us to pursue more subtle propositions regarding the power of television news. As a general strategy we de-

ployed the two designs together hoping to take advantage of their complementary strengths.

RECRUITING EXPERIMENTAL PARTICIPANTS

We recruited experimental participants by placing advertisements in local newspapers and by displaying posters in various public locations. The notices promised payment in return for participating in research on television news (typically $20 for sequential experiments and $7 for assemblage experiments). When individuals responded to the advertisements, we obtained information about their demographic characteristics (so as to exclude students, noncitizens, and those under the age of eighteen). Participants selected one of several daily viewing sessions; we then randomly assigned sessions to experimental condition.

MANIPULATING THE NETWORK'S AGENDA

A critical feature of both experimental designs is the creation of realistic newscasts. Because the procedure is more elaborate for sequential experiments, we will describe it in some detail. On the evening prior to each session, we videotaped the national newscasts of two of the major networks. To create slightly different broadcasts for each experimental condition, we then edited the actual newscast. We inserted stories into the newscasts, meanwhile deleting innocuous material of roughly equivalent length. The stories we inserted had been broadcast six to eight months earlier by the same network. We made certain that these stories contained no clues as to their actual date of broadcast by selecting feature stories that were relatively "timeless." (In many cases, these stories were themselves edited in order to remove temporal markers.) To do so we accumulated a large pool of news stories, dealing with ten separate problems, from the Vanderbilt University Television News Archive and from the audiovisual facilities at Yale University. For each network and for each problem, we compiled stories from several different reporters introduced by different anchorpersons. Some were winter stories and others were summer stories; some implicated the president and others did not; some displayed "talking heads" while others featured dramatic action. In short, on any given day, we had a large and diverse pool of potential stories from which to draw.

The treatment story (usually one each day) was inserted during the middle portion of the newscast and usually ran for two to four minutes. In practice the actual newscast was left substantially intact except for the insertion of the treatment story and the deletion of a story or two

treatment consisted of four implanted stories spread across four days, totaling about twelve minutes of news time (treatments ranged from seven to seventeen minutes of total coverage). From the networks' point of view, this level of attention represents a substantial commitment, but not an extraordinary one. During 1985, for example, the federal deficit, various hijackings, the crisis in Lebanon, the Nicaraguan conflict, the summit meeting between President Reagan and Secretary General Gorbachev, all drew at least as much weekly coverage as did our experimental problems. It would have been easy, and not particularly interesting, to demonstrate television effects by overwhelming our participants with an avalanche of stories about one problem or another. Such a demonstration would tell us about the potential power of television under extraordinary circumstances. Our interest rather is in the real power of television under ordinary circumstances, and so we designed our experimental treatments to fall within television news's normal range.

AVOIDING EXPERIMENTAL ARTIFACTS

In any experimental procedure it is important to guard against "demand characteristics"—cues in the setting that suggest to participants what is expected of them (Orne 1962). In order to limit the impact of demand characteristics in our experiments, we undertook several precautions. First, we disguised the purpose of our experiments. We began both sequential and assemblage experiments by presenting to participants an entirely plausible but false account of our purpose. Participants were informed that the study was about how people interpret and understand television news and that we were particularly interested in "what social scientists call selective perception. Do viewers' political opinions color what they see in the news? Do Republicans and Democrats really see the same news story?" Because this description provided a compelling explanation for what the participants were in fact later asked to do—namely, reveal their political opinions and evaluate news stories—we hoped it would discourage them from wondering what we were really doing.

Second, to lessen their prominence, we embedded our key measures of problem importance and presidential performance in a lengthy questionnaire. In addition to these questions of real interest, we also asked participants for their opinions on current issues, their explanations for the nation's problems, their perceptions of an ideal president, their partisan leanings, their recent political activities, and so forth.

We believe our precautions were successful. At the conclusion of each study (after the questionnaires had been collected and the partici-

pants paid), all participants were asked to describe their perceptions of what the experiment was about now that they had completed it, and whether their perceptions were consistent with their initial expectations. Across all our experiments, only a handful of participants expressed any skepticism about what we were up to—and most of these skeptics supposed that we were actually engaged in market research for the networks.

Finally, after learning what we could about our participants' perceptions of our purpose, we then gently revealed our real purposes to them. We told them that we had altered the newscasts and described how and why we had done so. We tried to explain the value of their participation to our understanding of the political effects of television news. Participants who indicated interest in our work were sent copies of our papers. We regard this "debriefing" as an essential part of our experiments. As a general matter, our procedures adhered scrupulously to the American Psychological Association's guidelines governing the protection of human subjects in experimental research.

Methodological Pluralism

Experiments have their limitations, of course, which ours do not avoid. While experiments are uniquely strong on matters of *internal validity*, providing evidence on causal relationships, they are typically weak on matters of *external validity*, providing assurance about the generalizability of results. Naturally, we seek generalizable results. The various experimental findings reported in the chapters to come are of interest only insofar as they bear on the workings of television news and public opinion in their natural settings. But to generalize from our experimental arrangements and populations to the American living room and ordinary Americans is to participate in what Campbell (1969a) has called "the scandal of induction." Generalizations always entail a leap of faith; even if informed, they are inescapably matters of opinion.

Concern about the generalizability of experimental results usually takes three principal forms. First, because experimental participants ordinarily know that they are taking part in the study of something (even if they're not sure what), this knowledge alone may induce alterations in their behavior. They might become more attentive; they might become less. They might defer to the experimenter's authority; they might react against it. Each represents a response to the special and in some respects *artificial* nature of the research setting. Second, experiments are often conducted with populations of convenience. Because convenient populations are often special populations, such practice has

naturally led to skepticism regarding whether experimental results can be safely generalized to populations of real interest. For social scientists stationed at universities, no population is of course more convenient than the student body. And the typical college sophomore, as Hovland (1959) warned some years ago, may be very far from typical of the average adult American. Third, experimental results are always subject to the charge that they depend precariously on exactly how the variables under investigation are created or measured. Perhaps an intriguing result would simply disappear in another experiment, with conditions realized or questions posed in slightly different ways. Results that conform neatly to expectations in one experiment may become inexplicable in another.

These threats to the external validity of experimental results—whether they can be generalized across settings, populations, independent and dependent variables—can certainly be raised with respect to our experiments in particular. We tried to anticipate and defend against them in the following ways.

In the first place, to minimize the artificial and perhaps reactive nature of our research setting, we made certain that our experimental manipulations were as unobtrusive as possible. The cutting, splicing, and rearranging of news stories were accomplished with state-of-the-art editing equipment that left behind no telltale traces. Even the most dedicated television news afficionado would have been hard pressed to detect any of the changes we introduced. We also encouraged participants to watch the news in a relaxed manner, and provided them with coffee, newspapers, and magazines. Many participants came with friends, spouses, or other family members. As we had intended, the sessions took on a casual atmosphere, with participants chatting among themselves, glancing around the room, browsing through the newspapers and magazines, and occasionally hooting at the commercials.

Second, because we were interested in assessing the impact of television news on ordinary Americans, not on college sophomores, we generally avoided that most convenient of populations.[2] Instead, by circulating advertisements widely in two quite different communities (New Haven, Connecticut and Ann Arbor, Michigan), we had hoped to lure a wide cross-section of people into our experiments.

In this respect, our experiments were highly successful. Table 2.1 presents a demographic and political profile of sequential experimental participants; table 2.2 does the same for assemblage experiments. As revealed in the two tables, participants in our experiments did indeed come in all varieties: young and old; black and white; men and women; poorly educated and well-educated; blue collar workers and profes-

sionals; Democrats, Independents, and Republicans; devoted to and oblivious of television news. Moreover, in aggregate terms, the profiles drawn in tables 2.1 and 2.2 are quite faithful to national figures. It would of course be absurd to pretend that participants in our experiments constitute a representative sample of the national population—or of any population, for that matter. They do not. Nevertheless, it is fair to say that the experimental results presented in the chapters ahead are based upon the reactions of a diverse group of ordinary Americans.

Third, none of our conclusions or interpretations depends upon a single (and perhaps peculiar) comparison. As a general rule, we have followed a strategy of conceptual replication: repeated tests with conceptually identical but empirically different realizations of the variables under investigation. Thus the same hypothesis was tested with different national problems and with different news stories. Similarly, we assessed participants' political opinions in a variety of ways. Within each experiment, we sometimes invited participants to define problems in their own terms; at other times they were taken through a battery of questions with built-in response alternatives. Where possible, we made use of questions that had undergone extensive testing and development by the Center for Political Studies, as part of the periodic National Election Studies. These procedures provide some assurance that any particular result is not specific to a particular problem, to a particular newscast, or to the particular questions we put to our participants.

TABLE 2.1
Demographic and Political Profile of Sequential Experiment Participants

Age		*Occupation**	
Range	19–68	Blue collar	41%
Average	31	Service/clerical	35
Race		Managerial/professional	24
White	82%	*Party Identification*	
Nonwhite	18	Democrat	39%
Sex		Independent	35
Male	48%	Republican	12
Female	52	Other/no preference	14
Education		*TV News Viewing*	
High school or less	35%	Never/hardly ever	23%
Some college	29	Two-three/week	48
College graduate	36	Every evening	29

Note: Number of participants = 259.
*Among those with full-time employment

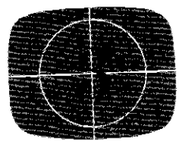
TABLE 2.2
Demographic and Political Profile of Assemblage Experiment Participants

Age		*Occupation**	
Range	18–81	Blue collar	26%
Average	33	Service/clerical	46
Race		Managerial/professional	28
White	82%	*Party Identification*	
Nonwhite	18	Democrat	39%
Sex		Independent	35
Male	51%	Republican	10
Female	49	Other/no preference	15
Education		*TV News Viewing*	
High school or less	34%	Never/hardly ever	23%
Some college	31	Two-three/week	40
College graduate	35	Every evening	37

Note: Number of participants = 772. The table excludes two assemblage experiments (3 and 11), because they were based on college undergraduates.
*Among those with full-time employment

By diminishing the artificiality of the experimental setting, drawing upon heterogenous groups of experimental participants, and following a strategy of conceptual replication, we have reduced the hazards of generalizing from our experimental results. However, the scandal of induction is still with us. The risks of generalizing from experiments can never be entirely eliminated.

Fortunately, the limitations of experimentation correspond to strengths in other methodological approaches. Although the results we report in the chapters ahead come predominantly from experiments, they are supplemented at several critical junctures by results from our own analysis of national surveys. In this we are trying to practice the methodological pluralism the late Carl Hovland (1959) was urging a quarter century ago. By examining television news from several methodological angles, we escape the limitations inherent in any single approach. We emphasize and promote experiments because of their real advantages and because for the most part media specialists have so assiduously ignored them. But we are methodological pluralists at heart.



The Agenda-Setting Effect

In *The Phantom Public*, Walter Lippmann characterized the political sensibilities of the ordinary American this way:

> The private citizen today has come to feel rather like a deaf spectator in the back row, who ought to keep his mind on the mystery off there, but cannot quite manage to keep awake. He knows he is somehow affected by what is going on. Rules and regulations continually, taxes annually and wars occasionally remind him that he is being swept along by great drifts of circumstance.
>
> Yet these public affairs are in no convincing way his affairs. They are for the most part invisible. They are managed, if they are managed at all, at distant centers, from behind the scenes, by unnamed powers. As a private person he does not know for certain what is going on, or who is doing it, or where he is being carried (1925, 13).

From this perspective, that ordinary citizens achieve any understanding of public affairs seems rather remarkable. Moreover, the "swarming confusion of problems" that, according to Lippmann, constituted political life more than a half century ago, has grown only more confusing today. Surely the democratic predicament of the ordinary citizen has deepened.

Television news may provide citizens with a convenient escape from this predicament. In this chapter we being to investigate how, if at all, television news influences Americans' conceptions of political reality—their sense of "the mystery off there." Our point of departure is the *agenda-setting hypothesis: those problems that receive prominent attention on the national news become the problems the viewing public regards as the nation's most important.* We pursue this hypothesis with sequential experiments, assemblage experiments, and a longitudinal analysis of national surveys.

Experimental Tests of Agenda-Setting

sequential experiments

Four sequential experiments provide evidence relevant to the agenda-setting hypothesis. Each systematically varied the amount of coverage

that the evening news devoted to various national problems. Experiment 1 is the prototype; we will therefore describe it in detail and then move expeditiously through the rest. (Detailed summaries of all fourteen experiments are presented in Appendix A.)

Experiment 1 was conducted in New Haven, Connecticut during six consecutive days in November of 1980, shortly after the presidential election. Participants watched recordings of the previous evening's network newscasts that had been edited in advance, as described in chapter 2. The thirteen participants randomly assigned to the *treatment condition* in experiment 1 were shown stories that described inadequacies in American defense preparedness. The first edited broadcast included a report on the increase in defense spending to be proposed by the incoming Reagan administration. The next day's newscast featured a "special assignment" report on the declining role of the U.S. as the arsenal of democracy. Spliced into the third day's broadcast was a pessimistic analysis of U.S. military options in the event of Soviet aggression in the Persian Gulf. And the fourth day's broadcast included a story that set out the considerable difficulty the U.S. Army was encountering in finding recruits qualified to operate its increasingly sophisticated equipment. Over the four days, then, participants in the treatment condition saw four stories on defense, totaling seventeen minutes of news coverage. The fifteen participants randomly assigned to the *control condition*, in contrast, watched broadcasts containing no defense-related stories at all.[1]

In experiment 1, as in all those that follow, participants were, of course, *randomly* assigned to conditions. According to information gathered on the first day of experiment 1, this procedure had the intended result: that is, participants assigned to the defense condition did not differ from their counterparts assigned to the control condition. On demographic characteristics, partisanship, and political engagement, the two groups were indistinguishable.[2] This means that whatever postexperimental differences between groups we detect can confidently be attributed to differences in the newscasts they watched.

In order to test the agenda-setting hypothesis, we measured participants' beliefs about the importance of national problems both before and after the experimental sessions. The preexperimental questionnaire was administered immediately before the first newscast and the postexperimental questionnaire was completed one full day after the last newscast. On both occasions, participants judged the importance of each of eight national problems, indicated their personal concern for each, the extent to which each was deserving of additional government action, and the frequency with which they talked about each in every-

day conversation.[3] Because these four ratings were strongly intercorrelated, they were averaged together to form a composite index of problem importance. A score of zero on the index means that the participant thought the problem not important at all; cared not at all about it; felt that people in government should worry about it not at all; and that the problem never served as a topic of conversation. A score of one hundred means, in contrast, that the participant thought the problem extremely important; cared about it very much; felt that the government should worry about it a lot; and talked about it almost incessantly. Of course, virtually all participants rated the problems somewhere between these two extremes.[4]

The critical test of agenda-setting simply entails observation of change over the experiment in the importance participants accord the problems emphasized by the edited newscasts. In experiment 1, participants who viewed newscasts that described glaring inadequacies in U.S. defense capabilities should become more concerned about defense than control condition participants whose newscasts were purged of such stories.

This is exactly what happened. Participants in the defense condition became more concerned about defense over the experiment's six days, while participants in the control condition showed no change in the importance they attached to defense. This difference is significant both statistically and politically.[5] Consider the evidence in detail: on the first day of experiment 1, *before* seeing any newscasts, participants who were randomly assigned to the defense treatment condition ranked defense sixth in relative importance, behind inflation, pollution, unemployment, energy, and civil rights. *After* exposure to the newscasts, the same participants now believed that defense was the country's second most important problem, trailing only inflation. Among viewers in the control condition, meanwhile, the relative position of defense as a national problem did not change.

Such a dramatic shift in priorities, induced by such a modest and unobtrusive alteration in television news coverage, constitutes powerful confirmation of the agenda-setting hypothesis. Moreover, what we found in experiment 1 we found again and again in three additional sequential experiments.

Experiment 2 took place in late February 1981 and focused on three problems. Depending on condition, participants viewed newscasts that emphasized either inadequacies in U.S. defense preparedness, pollution of the environment, or soaring inflation. Experiment 8, administered in July 1982, featured newscasts that concentrated either on unemployment, nuclear arms control, or civil rights. And in experiment 9, which took place in August 1982, participants either viewed a sequence of newscasts that emphasized unemployment or saw no newscasts at all. In other respects, each of these three sequential experiments followed experiment 1's basic design. The recruitment of participants, the splicing in and editing out of stories, the questionnaires administered on the first day and on the sixth: all these procedures were followed as described previously. Counting experiment 1, then, we have eight separate and independent tests of agenda-setting, distributed across four experiments and six different problems.

The results from all four sequential experiments are displayed together in table 3.1. The table shows the average composite importance rating of the target problem before and after exposure to the newscasts. The message could not be clearer. In *every* instance, participants emerged from our experiments believing that the target problem was more important than they did when they began. Seven of the eight changes are statistically significant.

The single exception to this pattern, which occurred in the inflation condition of experiment 2, is no great mystery. In February of 1981, when experiment 2 took place, inflation was running at an annual rate of more than 10 percent. In the preexperimental questionnaire, before they glimpsed a single inflation news story, participants gave inflation an average score of ninety-two on our composite scale of zero to one hundred. Thus we had virtually no opportunity to convince participants of inflation's importance: everyone was already convinced.

Putting this exception aside, the four sequential experiments yield striking evidence of agenda-setting. As in experiment 1, the changes observed in experiments 2, 8, and 9 correspond to substantial shifts in

TABLE 3.1
Change in Problem Importance

Experiment	Problem	Importance Rating of Problem: Before the Experiment	After the Experiment	Change: Pre- to Post-
1	Defense	47	67	20*
2	Defense	48	58	10*
	Inflation	92	93	01
	Pollution	63	76	13*
8	Arms control	76	82	06*
	Civil rights	64	69	05*
	Unemployment	75	82	07*
9	Unemployment	78	83	05*

*$p < .05$

problem hierarchies. In experiment 2, for example, pollution as a national problem moved up from fifth to second most important among participants shown news about pollution, while defense rose from sixth to fourth among participants who watched newscasts that emphasized U.S. defense weaknesses. Similar shifts occurred in the relative importance of arms control, civil rights, and unemployment in experiments 8 and 9.

The agenda-setting hypothesis can be tested in a second way. In experiments 2, 8, and 9 (though regrettably not in 1), the questionnaire asked participants to name "the three most important problems facing the nation." As a second test, therefore, we can compare the proportion that mentioned the target problem in the preexperimental questionnaire with the proportion naming it in the postexperimental questionnaire, following exposure to the altered newscasts. These comparisons are shown in table 3.2.

As indicated there, the evidence in support of agenda-setting is even more striking for this measure than it was for composite ratings. Except, once again, for the inflation condition in experiment 2—where *every* participant named inflation as one of the country's most important problems, both before and after the experiment—references to the target problem were more numerous after the newscasts than before. Some of these increases are massive. In experiment 8, for example, after exposure to coverage of the perils of the arms race, the percentage of participants naming arms control as one of the country's three most important problems rose from 35 percent to 65 percent; in experiment 9, the corresponding percentage, this time for unemployment, increased from 50 percent to 86 percent. Over the seven independent tests, an average of 37 percent nominated the target problem as one of the nation's most important in the preexperimental questionnaire; 57 percent did so in the postexperimental questionnaire.[6]

TABLE 3.2
Change in Problem Importance

Experiment	Problem	Percentage Naming Problem as One of Country's Most Serious		
		Before the Experiment	After the Experiment	Change: Pre- to Post-
2	Defense	33	53	20*
	Inflation	100	100	00
	Pollution	0	14	14*
8	Arms control	35	65	30*
	Civil rights	0	10	10*
	Unemployment	43	71	28*
9	Unemployment	50	86	36*

*$p < .05$

We also assessed the *specificity* of these effects. In general, we looked for "spillover" in agenda-setting. We supposed that drawing viewers' attention to a particular problem might enhance not only the importance they ascribe to that problem but to related problems as well. For example, stories emphasizing dependence on foreign sources of oil might reasonably be expected to raise concern about rising prices, since the public seems to regard the two problems as causally linked (Hendricks and Denney 1979). Reasonable or not, we encountered such spillover effects in only two instances. In experiment 8, participants exposed to news about the arms race became more concerned not only with arms control but also with the conflict in the Middle East. In experiment 9, participants furnished with coverage of unemployment became more concerned about economic problems in general. As a rule, however, the agenda-setting effects we uncovered are notable for their specificity. News about energy influenced viewers' beliefs about the importance of energy and energy alone; news about defense influenced viewers' beliefs about defense and defense alone; and so on.

In sum, the evidence from the four sequential experiments strongly supports the agenda-setting hypothesis. With a single and understandable exception, problems given steady news coverage grow more important, at least in the minds of the viewers. The evening news would seem to possess a powerful capacity to shape the public's national priorities.

ASSEMBLAGE EXPERIMENTS

In sequential experiments viewers are exposed either to a sustained dose of news about a particular problem or to no news at all. One virtue of assemblage experiments is that they permit a more precise calibration of treatment conditions. Here we examine six such experiments in an effort to learn more about the functional relationship between the amount of news coverage and the size of the agenda-setting effect.

Experiment 3 was conducted in New Haven during April and May of 1981 with Yale University undergraduates. Students viewed a forty-minute collection of "typical" news stories that paid either no attention to the nation's energy problems (zero stories), some attention (three stories) or considerable attention (six stories).

Experiment 4 was run in New Haven during late September to early

October 1981. This time participants were recruited from the general community and randomly assigned to one of six experimental treatments. Participants watched a collection of fifteen news stories that gave either moderate attention (three stories) or extensive attention (six stories) to one of three national problems: defense, energy, or inflation. Participants assigned to either the moderate or extensive treatment conditions for any one problem (say defense) saw *no* stories about the other two (energy, inflation). This design enables us to assess the agenda-setting effect induced by some exposure to a problem versus none, as well as the impact induced by incremental increases in coverage.

Experiment 5, which took place in New Haven during August–September of 1981, followed this same design, with two amendments. First, in place of stories about defense, energy, and inflation, we substituted stories about unemployment, civil rights, and social security. Second, we reduced the number of stories bearing on the target problem in the moderate and extensive coverage conditions to two and four, respectively.

Experiments 6, 13, and 14 represent the natural culmination of this trend of diminishing experimental interventions. In experiment 6, conducted in New Haven in May and June of 1981, participants watched a collection of news stories that included either just a *single* story about the target problem—this time either pollution or unemployment—or no stories at all. Likewise, in experiment 13, run in Ann Arbor in June 1983, and in experiment 14, conducted in New Haven in August of 1983: in the former instance, participants watched a collection that included one story either about unemployment or energy; in the latter, participants were exposed to a collection that featured a single story either about government efforts to halt drug smuggling or about the difficulties facing public schools.[7]

We measured problem importance in these six assemblage experiments just as we did in the sequential experiments: i.e., by composite ratings and spontaneous mentions.[8] The test of agenda-setting is different here, though, because assemblage designs forego the preexperimental questionnaire that is a standard fixture of the sequential design. Participants in assemblage experiments complete only one questionnaire, immediately following exposure to the news presentations. Therefore, the appropriate test of agenda-setting here is to compare the importance participants attach to a target problem across different experimental conditions representing different levels of coverage.

The results for the composite importance ratings are shown in table

3.3. The rows reflect different problems across the six experiments; the columns reflect intensity of coverage, from no news stories at all on the left to a maximum of six stories on the right. If the agenda-setting hypothesis holds, the importance ratings of the various target problems should increase from left to right as coverage intensifies—and they generally do. In fact, twelve of the thirteen ratings increase, ten to a statistically significantly degree. As was true in sequential experiments, agenda-setting proved elusive only for those problems that were regarded as highly important at the outset. In the case of inflation in experiment 4, for example, a virtual bombardment of coverage—six stories in a collection of fifteen—was required to boost ratings still higher than those offered by viewers who saw no stories about inflation.

Support for agenda-setting is generally more striking when importance is measured by the spontaneous nomination of national problems, shown in table 3.4. In every instance but one, participants shown some stories about a particular problem—as many as six stories or as

TABLE 3.3
Problem Importance as a Function of Intensity of Coverage

		Composite Ratings							
		Number of Stories							Difference: Maximum Coverage Minus No Coverage
Experiment	Problem	0	1	2	3	4	5	6	
3	Energy	64			66			74	10***
4	Defense	58			63			70	12***
	Energy	72			67			72	00
	Inflation	81			81			90	09***
5	Civil rights	69		71		86			17***
	Social security	77		84		88			11***
	Unemployment	78		87		84			06*
6	Pollution	77	81						04***
	Unemployment	88	89						01
13	Unemployment	90	95						05***
	Energy	75	68						−07
14	Drugs	43	53						10***
	Education	70	74						04*

*$p < .20$
**$p < .10$
***$p < .05$

TABLE 3.4
Problem Importance as a Function of Intensity of Coverage

		Percentage Naming Problem as One of Country's Most Serious							
		Number of Stories							Difference: Maximum Coverage Minus No Coverage
Experiment	Problem	0	1	2	3	4	5	6	
3	Energy	24			50			65	41***
4	Defense	33			57			64	31***
	Energy	21			46			46	25***
	Inflation	45			50			79	34***
5	Civil rights	15		29		33			18**
	Social security	10		41		44			34***
	Unemployment	30		30		67			37***
6	Pollution	10	27						17***
	Unemployment	53	73						20***
13	Unemployment	50	68						18**
	Energy	0	23						23
14	Drugs	0	11						11***
	Education	14	11						−03

*$p < .20$
**$p < .10$
***$p < .05$

few as one—were more likely to name that problem as one of the country's most important than were those whose attention was directed elsewhere. All but one of these differences surpass statistical significance and some of them are extraordinary.[9] In experiment 5, for example, whereas less than one-third of the participants exposed to two stories on unemployment named it as one of the country's most serious problems, fully two-thirds of those exposed to four stories on unemployment did so. Perhaps the most arresting result of all is that agenda-setting can be triggered by such ostensibly innocuous provocations. In experiments 6, 13, and 14, viewers' priorities were significantly affected by a *single* news story.

Persistence of Agenda-Setting

Measurable immediate influence is not the same as influence that lasts, of course. We assessed the influence of television news in assemblage

experiments immediately following the broadcasts. Sequential experiments are somewhat more informative, but they tell us only that television news's influence is detectable twenty-four hours after the experimental intervention is completed. That the effects survive this long is certainly important. Television dispenses news periodically, typically on cycles of twenty-four hours or less. The regularity and frequency of broadcasts means that for many viewers, agenda-setting is a continuous process. When the networks develop priorities, viewers' beliefs are affected—and affected again as new priorities arise. Having said that, however, we are still left with the question of how long our experimentally-induced effects last.

We designed experiments 13 and 14 partly to investigate the persistence of agenda-setting effects. In the former, we reinterviewed as many participants as possible over the telephone one week after they had been exposed to our news broadcast. Participants were told that we were conducting an opinion poll of the Ann Arbor community. Virtually everyone we were able to reach agreed to participate (75 percent of the original group). In the followup to experiment 14, we mailed to each participant a second questionnaire one week after their experimental session. Eighty-three of the original 121 participants (69 percent) completed and returned the questionnaire. In both followup sessions, among many other questions, participants were asked to name the country's most serious problems. Experiments 13 and 14 thereby afford a test of the persistence of agenda-setting effects—and a stringent one at that. The two experiments are not only assemblage designs, which produce less powerful effects than do sequential designs, but they represent the weakest of the assemblage designs, involving as they do only a single story.[10]

Nevertheless, both experiments reveal evidence of persistence. These results are displayed in table 3.5. As indicated there, participants in experiment 13 who had been exposed to a single story about unemployment continued, one week later, to nominate unemployment more frequently as one of the country's most important problems than did those who saw no news about unemployment. This difference was virtually as great at one week's remove from the experimental intervention as it was immediately afterwards. More generally, the table shows that the agenda-setting effect was maintained over the one week period in two instances, diminished in one, and actually strengthened in another. Keeping in mind that alterations in viewers' political priorities were prompted originally by a single story, the degree of persistence revealed here is remarkable.

TABLE 3.5
Immediate and Delayed Effects of Coverage on Problem Importance

		Percentage Naming Problem as One of Country's Most Serious					
		Immediate:			One Week Later:		
		Number of Stories			Number of Stories		
Experiment	Problem	0	1	Difference	0	1	Difference
13	Unemployment	46	72	+26**	54	73	+19**
	Energy	0	15	+15	4	8	+ 4
14	Drugs	0	14	+14***	0	14	+14***
	Education	14	13	− 1	8	26	+18***

Note: Table includes only those participants interviewed immediately after the experiment *and* one week later.

*$p < .20$
**$p < .10$
***$p < .05$

Time Series Tests of Agenda-Setting[11]

Our experimental results suggest that television newscasts shape and intensify viewers' sense of which national problems are important and which are not. But do our experimental results generalize to the natural setting that is our real interest? We think they do—partly because of the convergence of findings across experiments, problems, and populations; partly because of the steps we took to diminish the artificiality of our experiments—but we cannot be completely confident.

To bolster our confidence and complement our experimental results, we undertook a nonexperimental test of agenda-setting. We examined trends in television news coverage over time, and compared them with changes over comparable periods in public opinion. Prior efforts of this sort suggest that there should be a correspondence between the two, and a strong one. Thus Funkhouser (1973) discovered striking concurrence between the amount and timing of attention paid to various problems in the national press between 1960 and 1970 and the importance accorded those problems by the American public. Across the decade, public opinion seemed to follow, not lead, the press's agenda, results that were substantially fortified by the more sophisticated analyses that followed (MacKuen 1981, 1984).

Funkhouser and MacKuen presumed, as we do, that agenda-setting effects should be observed and estimated over time, as problems ap-

pear and disappear, and as network news coverage shifts accordingly. What we have attempted to do in our experiments is convert the variation in coverage that occurs naturally over time to contemporaneous variation across experimental conditions. We create and then offer to our viewers alternative portrayals of political reality. As an important check on the experimental results, here we will determine through time-series analysis the extent to which the preoccupations of network news become the political preoccupations of the American public.

For this purpose, we compiled results from national surveys between 1974 to 1980 pertaining to three prominent national problems: energy, inflation, and unemployment. By ransacking Gallup, Yankelovich, and Center for Political Studies surveys, we were able to obtain a measure of the importance attached by the public to each of the three problems for every two-month period between January 1974 and December 1980.[12] Our specific measure of problem importance stems from "the most important problem facing the nation." Unfortunately, the exact wording, format, and coding of the question varies across survey organizations. Gallup and Yankelovich accept multiple answers while CPS does not; and Yankelovich interviewers consistently "pull" more answers from survey respondents than do Gallup's. To ensure comparability in results across the three survey organizations, we took as our dependent variable the percentage of *responses* to the question rather than the percentage of *respondents*. (For a detailed explanation of this procedure, see Appendix B).

We measured television news coverage of the three problems by recording the number of pertinent news stories appearing in the weekday CBS Evening News.[13] Using the Vanderbilt Television Archive's *Abstracts* of daily newscasts as our source, we classified news stories on the basis of their major focus (news stories that lasted less than thirty seconds were excluded). The number of news stories for each problem was totaled for every month and then averaged for each bimonthly observation.

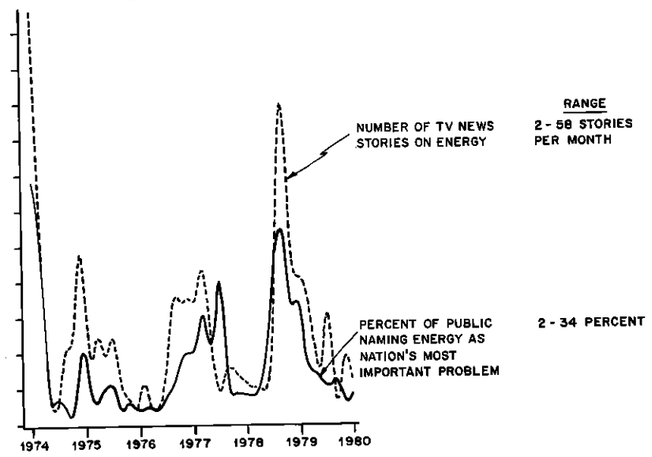
Measured in this way, the attention provided these three problems by television news underwent dramatic changes between 1974 and 1980. Figure 3.1 displays the 1974–80 time graphs for energy; figure 3.2 does so for inflation; figure 3.3 provides the same information for unemployment. Between 1974 and 1980, CBS's coverage of energy ranged from two stories per month to fifty-eight stories per month. Inflation received as few as six stories per month to as many as thirty-seven. Monthly coverage of unemployment ranged between no stories at all to a modest peak of seven.

Over the same period there were also striking changes in the impor-

tance the American public ascribed to the three problems. The proportion of the public naming energy as one of the country's most important problems fluctuated from a low of 2 percent to a high of 34 percent; inflation, from 19 to 72 percent; and unemployment, from 2 to 32 percent (see figures 3.1, 3.2, and 3.3). And to the naked eye, at least, these fluctuations in public concern seem to move roughly in tandem with fluctuations in television news coverage.

That the trends move together does not, of course, tell us anything about the *causal* impact of television news coverage on problem importance. The parallel trends might mean that news coverage influences public opinion, but it could mean just the reverse: that news organizations respond to the public's priorities. In order to attract the largest audience, the networks might feature stories about inflation when the public seems concerned about inflation and stories about unemployment when the public seems preoccupied by unemployment. Or the correspondence in the over-time trends might reflect that the networks and the public are responding in concert to real changes in the world. Soaring prices are noticed in New York as in Peoria, with implications that are easy to imagine for both the networks and the public. Our task

FIGURE 3.1
TV News Coverage and Public Opinion toward Energy, 1974–1980

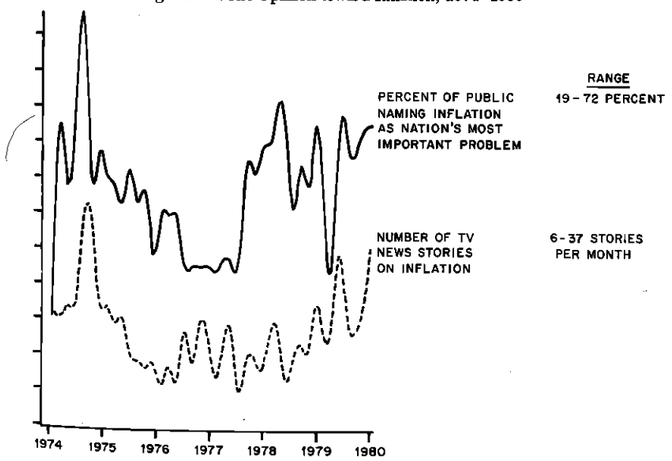


Source: AIPO, Yankelovich, National Election Studies Surveys; Vanderbilt TV News Abstracts

must be to determine not just the association between television news and public opinion, but rather the precise causal impact (if any) of television news on public opinion.

To do so, we relied on a procedure that computes a consistent estimate of the impact of television news coverage on public opinion purged of the reverse effect, if any, of public opinion on news coverage (see Appendix B for the technical details). This procedure also estimates the impact of television coverage over and above the effects due to real world conditions. Because energy shortages, price increases, and job loss can all be experienced personally, they may influence public opinion directly. To take such effects into account, we coded various measures of real world conditions and incorporated them into our analysis. They included the cost and availability of energy, American dependence on foreign sources of energy, meetings of OPEC oil ministers in the energy analysis; various aggregate indicators of prices and interest rates in the inflation analysis; and aggregate measures of the extent of unemployment and change in unemployment in the unemployment analysis. Finally, we also included a measure of major presidential speeches devoted primarily to energy, inflation, or unemployment (see

FIGURE 3.2
TV News Coverage and Public Opinion toward Inflation, 1974–1980

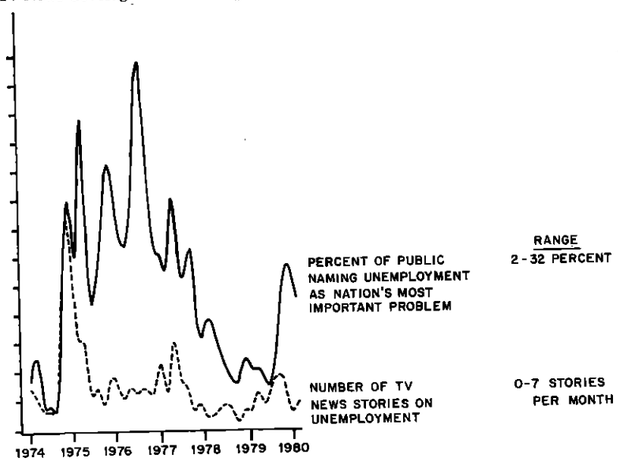


Source: AIPO, Yankelovich, National Election Studies Surveys; Vanderbilt TV News Abstracts

Appendix B for details on all these measures). In short, our analysis attempts to reveal the degree to which television news influences public opinion, independent of the effects due to actual conditions and to presidential efforts to mobilize public opinion.

The results for energy are shown in table 3.6.[14] As indicated there, television news coverage does indeed influence the importance the American public attaches to energy. For every seven stories broadcast, public responses citing energy as one of the country's most important problems increased by about 1 percent. Notice that this is a *contemporaneous* effect: television coverage in the current period influences public opinion in the current period. We also tested for but could not find lagged effects: i.e., the amount of coverage devoted to energy during any two-month period apparently had no effect on opinion toward energy expressed during the next two-month period. Perhaps surprisingly, public opinion on energy was unaffected by real world conditions. Energy costs, fuel oil costs in particular, dependence on OPEC imports, OPEC oil minister meetings: none of these boosted public concern with energy independently of television news coverage.[15] The public's concern for energy *was* shaped independently by the presi-

FIGURE 3.3
TV News Coverage and Public Opinion toward Unemployment, 1974–1980



Source: AIPO, Yankelovich, National Election Studies Surveys; Vanderbilt TV News Abstracts

TABLE 3.6
Predictors of Public Opinion toward Energy, January 1974–December 1980 (Two-stage, maximum likelihood estimates)

Predictors	Coefficient
Number of stories on energy	.13*
Presidential speeches on energy	4.44*
Constant	12.52*
Adjusted R^2 = .55	
Standard error of regression = 3.20	
Durbin-Watson statistic = 2.06	

Number of observations = 42.
*$p < .05$

dent, however. When the president chose to address the nation on the subject of energy, he succeeded in raising the level of public concern by over 4 percent.

The American public's preoccupation with inflation between 1974 and 1980 was determined by a similar combination of news coverage and presidential rhetoric (see table 3.7). The number of television news stories about inflation significantly increased the percent of responses naming inflation as the nation's most important problem. On average, five stories per month on inflation elevated public concern by 1 percent (again, an entirely contemporaneous effect), whereas a presidential address to the nation on the economy increased the degree of public concern about inflation by over 8 percent. Again, actual conditions had no direct impact on public opinion: changes in the consumer price index, the consumer price index for food, and interest rates were

TABLE 3.7
Predictors of Public Opinion toward Inflation, January 1974–December 1980 (Ordinary least square estimates)

Predictors	Coefficient
Number of news stories on inflation	.21*
Presidential speeches on inflation	8.26*
Constant	41.92*
Adjusted R^2 = .49	
Standard error of regression = 7.38	
Durbin-Watson statistic = 1.54	

Number of observations = 42.
*$p < .05$

all unrelated to the importance the American public attached to inflation, once the influence of television news coverage was taken into account.[16]

This brings us finally to the case of unemployment, where television news effects appear to be weaker. As shown in table 3.8, eleven stories per month were required to boost public concern about unemployment by a single percentage point. Moreover, on statistical grounds, we cannot be certain that television news coverage had *any* effect at all. And unlike public opinion on energy and inflation, the American public's concern about unemployment was unaffected by presidential addresses. Instead, the importance of unemployment to the American public was determined by actual conditions. As unemployment spread and deepened, more and more Americans considered it to be among the country's most pressing problems, largely independently of trends in television news coverage.

The comparatively frail results uncovered for television's impact on public concern about unemployment may be due to the chronically low level of news coverage. Over the seven-year period under examination here (which preceded the dramatic increases in unemployment that occurred in 1981 and 1982), CBS broadcast an average of just four stories on unemployment every two months. This represents less than one-third of the coverage CBS gave to energy and less than one-fourth of the attention the network devoted to inflation. If the *networks* regard unemployment as less newsworthy than rising prices or energy shortages, then so, too, may the public.

This point aside, we should not be deflected from the central message carried by the time-series results. Here we find strong convergent

TABLE 3.8
Predictors of Public Opinion toward Unemployment, January 1974–December 1980 (Maximum likelihood estimates)

Predictors	Coefficient
Number of news stories on unemployment	.09
Unemployment rate	3.18*
Average duration of unemployment (weeks)	1.41*
Constant	−23.34*
Adjusted R^2 = .50	
Standard error of regression = 3.98	
Durbin-Watson statistic = 1.89	

Number of observations = 42.
*$p < .05$

support for television agenda-setting. Between 1974 and 1980, the American public's political preoccupations underwent sharp changes, changes that we have traced in part to changing patterns of television news coverage.[17]

Conclusion

Taken all together, our evidence decisively sustains the agenda-setting hypothesis. The verdict is clear and unequivocal: It issues from sequential experiments that last a week, from assemblage experiments that last an hour, and from time-series data that span seven years; it holds across different measures of importance; and it is confirmed for a variety of problems, from national defense to social security. By attending to some problems and ignoring others, television news shapes the American public's political priorities. These effects appear to be neither momentary, as our experimental results indicate, nor permanent, as our time-series results reveal.

All told our evidence implies an American public with a limited memory for last month's news and a recurrent vulnerability to today's. When television news focuses on a problem, the public's priorities are altered, and altered again as television news moves on to something new.