THINKING ABOUT POLITICAL PSYCHOLOGY

EDITED BY
JAMES H. KUKLINSKI
University of Illinois at Urbana-Champaign

CAMBRIDGE
UNIVERSITY PRESS

2

Who Can Persuade Whom?: Implications from the Nexus of Psychology and Rational Choice Theory

ARTHUR LUPIA

Political psychologists and rational choice theorists do not interact very much. This silence is particularly ironic when it comes to explaining political behavior, as such explanations are a core concern of both groups.

Consider, for example, the topic of persuasion, and in particular how people in political settings choose whom to believe. Voters, legislators, and jurors are but a few of the many political decision makers who have opportunities to base what they do on the written or oral statements of others. To explain the decisions that people who can learn from others make, we should understand what makes some statements more persuasive than others. An irony of extant persuasion research is that while it represents an active field of study for both political psychologists and rational choice theorists, most treatments of the topic cite the contributions of no more than one of the two scholarly traditions. It is as if political psychologists and rational choice theorists have nothing to teach each other about persuasion.

Do intellectual differences between these scholarly traditions negate the possibility of constructive dialogue? As a formal theorist who grapples with political psychology's substantive challenges, I face this question often. I have come to believe that a constructive dialogue is not only possible but also worthwhile.

I thank James Druckman, Mathew McCubbins, Sam Popkin, and Paul Sniderman for guidance offered during all stages of this project. I thank Scott Basinger, Kathleen Bawn, Henry Brady, Andrea Campbell, Christopher Den Hartog, Robert Erikson, Elisabeth Gerber, Jennifer Kuhn, Jim Kuklinski, Susanne Lohmann, Ted Miller, Jim Stimson, and Michael Thies for insightful comments. I thank Karen Garrett and the staff of the Survey Research Center at the University of California at Berkeley for administering the survey. I also acknowledge the financial support of the National Science Foundation, Grants 9309946 and 9422831.

In what follows, I first argue and then demonstrate that more frequent and serious interactions between political psychologists and rational choice theorists can generate substantial gains from trade. That is, I first offer a brief argument about what political psychologists and rational choice theorists have to offer each other. I then use a psychologically informed rational choice model and a rational choice–informed psychology experiment to construct an explanation of persuasion that is more powerful and general than other explanations that ignore at least one of the two scholarly traditions. I conclude that there are gains from trade to be had from sustained and serious interaction between political psychology and formal theory.

GAINS FROM TRADE?

The characteristic that most differentiates political psychologists and rational choice theorists is the method of inference. Political psychologists, like their methodological counterparts in social and cognitive psychology, tend to rely on experiments.[1] When these experiments show a relationship between stimulus and response, they provide evidence for psychologists' behavioral claims. Rational choice theorists, like microeconomists, rely on formal models. These models allow theorists to show how certain behavioral conclusions follow from precise assumptions.

Like all methods of inference, even the most carefully designed experiments and models have substantive limitations. These limitations can be used for different purposes. They can, for example, give members of one scholarly tradition an excuse to discount or ignore the arguments of others. This practice, while easy to justify when in a room with like-minded individuals, is seldom constructive. Alternatively, if the strengths of the political psychology approach are the foil of the rational choice approach or vice versa, then these limitations can be the source of intellectual gains from trade. To see if such gains exist, let's examine the limitations in question.

A typical psychology experiment consists of a control condition and some treatment conditions in an otherwise constant laboratory setting. Experimenters draw behavioral inferences by comparing the relations

1 This representation is admittedly stereotypical. Political psychologists also use quasi-experiments and nonexperimental surveys. For the argument I want to make here, however, an argument assessing the contrasts and complementarity of psychological and rational choice approaches, the focus on experiments gives sufficient representation to the main methodological difference between the two traditions.

between controlled stimuli and subjects' responses across these conditions. Since performing an experiment usually requires substantial time and money, it is usually not feasible for an experimenter to run a large number of treatment conditions – certainly less than the number it would take to identify general properties of most behaviors observed in experiments. Therefore, most experiments, *taken literally*, provide but a few isolated *examples* of reactions to selected stimuli.

Experiments, however, are not designed to be viewed literally. An experiment is supposed to be a simple, controllable analogy to a larger, uncontrollable set of circumstances.[2] Therefore, an experiment should be viewed in terms of whether scholars can use it to achieve a constructive scientific purpose.

An experiment's purpose is to demonstrate that a specific variation in a stimulus corresponds to a specific pattern of behavior. For an experiment to achieve its purpose, the audience must believe that the experimental stimuli, subjects, and environment are reasonable analogies to the stimuli, decision makers, and environments in the larger set of circumstances. In cases where the audience does not debate the analogy's quality, the experiment can achieve its purpose. However, some questions are so complex that their comprehension requires a considerable difference between what happens in and out of the laboratory. When such differences stir debate about the quality of the analogy, overcoming the literal limit of experiments requires a stronger remedy. I will describe such a remedy after briefly discussing the very similar literal limits of formal models.

A typical formal model is a set of mathematical statements. Each statement represents an assumption about human desires, opportunities, or knowledge.[3] Theorists use deductive logic to show that certain conclusions follow from these assumptions. An advantage of a formal model

2 Some people prefer to call these general sets of phenomena "reality." I resist this practice because the counterfactuals required to verify the validity of reality are beyond human comprehension (e.g., external validity is often created in the eye of the beholder).

3 *Webster's New Collegiate Dictionary* defines the term "rational" as "(a) having reason or understanding, (b) relating to, based on, or agreeable to reason." The same source defines "reason," the synonym of "intelligence," as "the proper exercise of the mind." I therefore define "rational choice theory" as an attempt to understand and explain behavior using the assumption that people do the best they can with the knowledge and skills they have. I mention this because it is common for both supporters and critics of rational choice approaches to confound rationality and omniscience. When this mistake is made, rationality ceases to be a useful concept, as the people whose behavior we seek to understand are far less than omniscient.

is generality. A simple spatial model, for example, can be used to generate predictions about behavior in an infinite number of cases. But what is the value of these predictions?

A truth of formal modeling is that analytical tractability often requires simple assumptions. As a result, a model's predictions about human behavior can also be simple. Many formal models, *taken literally*, describe *simple, logical correspondences* between assumptions and conclusions about human behavior. Since people need not be simple, a model's predictions may provide insufficient descriptions of human behavior.

Formal models, however, are not designed to be viewed literally. Like an experiment, a formal model is supposed to be a simple, controllable analogy to a larger, uncontrollable set of circumstances. Therefore, a formal model should be viewed in terms of whether scholars can use it to achieve a constructive, scientific purpose.

A formal model's purpose is to clarify correspondences that are difficult to see in the usual cacophony of social interaction. For a formal model to achieve its purpose, the audience must believe that the model's assumptions and conclusions are reasonable analogies to the stimuli, decision makers, and environments in the larger set of circumstances. In cases where the audience does not debate the analogy's quality, the model can achieve its purpose. However, some questions are so complex that relationships between assumptions, predictions, and observed behaviors can stir debate about a model's analogy. In these cases, overcoming the literal limit of formal modeling requires a stronger remedy.

The gains from trade available to political psychologists and rational choice theorists arise from the complementarity of each tradition's strengths and limitations. These gains are available when one tradition's strengths remedy the other tradition's limitations. For example, a formal model's logically valid correspondences between clearly stated premises and conclusions could make it an effective tool for demonstrating why an experimental behavior should be observed in a particular range of nonexperimental circumstances. Of course, the model must strengthen the believability of the initial experimental analogy. That is, the audience who is questioning the generality of the experiment must perceive the model to be a sufficient analogy to both the experiment and the general set of circumstances in question. Put another way, a formal model is no panacea for a badly designed experiment. However, if you

- present an audience with behavioral premises that they have difficulty refuting

- and prove the logical validity of the relationship between these assumptions and conclusions that pertain to the experiment's generality,
- and if greater interaction between political psychologists and rational choice theorists makes such models easier to design,

then greater interaction provides a remedy for the literal limit of the experiment, and participating scholars experience gains from trade.

Similarly, psychological insights and methods can clarify the validity of a formal model's analogy. In domains where many variations of a well-designed experiment reveal specific boundaries of human ability, opportunity, and information, psychology can inform rational choice theorists about where to start when making assumptions about political decision makers. Of course, the psychologists' contributions must be compelling to the theorist's audience. That is, the audience who is questioning the model's value must perceive the psychologist's data to be relevant to the model and the empirical circumstances in question – an experiment is no panacea for a badly designed model. However, if a modeler presents an audience with an experiment whose similarity to the set of relevant circumstances is difficult to refute, and if greater interaction between political psychologists and rational choice theorists makes such experiments easier to design, then greater interaction will remedy the literal limits of the formal model and participating scholars experience gains from trade.

It is one thing to claim that an explanation that combines political psychology and rational choice theory trumps explanations that ignore either or both approaches. It is quite another to show this claim to be true. What follows is one attempt to accomplish the latter.

THE PROBLEM OF PERSUASION

We know that people lack information about politics (e.g., Converse 1964; Delli Carpini and Keeter 1996). We also know that people have little incentive to acquire more information when doing so is costly (e.g., Downs 1957). When we put these two facts together, we can conclude that people who want greater knowledge have an incentive to substitute low-cost cues for the detailed information that they lack (e.g., McKelvey and Ordeshook 1985; Popkin 1991).[4]

4 I define "knowledge" as the ability to predict accurately the consequences of actions, "information" as the data from which knowledge may be derived, and an "attitude" as a person's general evaluation of an object (also see Lupia and McCubbins 1998).

In many situations of interest to political scientists (e.g., elections for national or state office), many low-cost cues are available.[5] For example, in the weeks before an election, television, radio, newspapers, and casual conversation contain political advertisements, endorsements, and commentaries for a wide range of candidates and causes. But no person in the midst of such a barrage can use all available cues. Each person must choose which cues to use. If we want to understand how cue usage affects behavior, then we should endeavor to explain how people sort among the many cues that are available to them. We must be precise about what differentiates a persuasive cue, one that changes attitudes, from a cue that does not persuade.

Our discipline is often vague about what makes a cue persuasive. For example, it is widely taken for granted that conservatives tend to find other conservatives' cues more credible, that African Americans tend to find other African American elites more credible, and so on. While certain cue-giver attributes sometimes correlate with cue persuasiveness, the extant persuasion literature does not reveal when these correlations are evidence of causality and when they are spurious. This literature does not provide answers to questions such as "When will a speaker's ideology affect the persuasiveness of his or her cues more than does his or her race, likability, or level of education?" Yet to explain how people behave when confronted with the many cues that political settings often proffer, we must find a way to answer such questions.

I offer an answer that comes from an integrated foundation of psychological and rational choice insights. My answer is that all cue-giver attributes – such as race, gender, ideology, partisanship, reputation, or likability – affect a cue's persuasiveness only if they are necessary to inform a cue-seeker's perceptions of a cue-giver's knowledge or interests (it need not inform the cue seeker of the cue-giver's actual knowledge or interests). If an attribute is not necessary for this effect, then any correlation between it and a cue's persuasiveness is indirect or spurious.

I support this claim in two ways. First, I derive it from a formal model of political communication. Then I use a survey experiment on over 1,400 randomly selected Americans to show that respondents' percep-

5 By the term "low-cost cue," I mean a statement of the form "Mr. M says 'Vote for candidate C,'" where Mr. M may himself be Candidate C. This is opposed to a higher-cost cue such as a long argument or report by Mr. M about Candidate C. In the language of social psychology, I focus on the peripheral (or heuristic) route to persuasion rather than the central (or systematic) route (Eagly and Chaiken 1993; Petty and Cacioppo 1986). As a result, I focus on the persuasive effect of cue-giver attributes (i.e., source effects) rather than on the persuasive effect of the cue-giver's argument style (i.e., message effects).

tions of a speaker's interests and knowledge explain much more about which cues they use than do cue-giver attributes such as a speaker's likability, ideology, or party.

Together, the theory and the data provide an insight that people skeptical of rational choice theory's contribution to persuasion research may regard as ironic. The insight is that, contrary to common practice, persuasion scholars should adhere to a strict interpretation of seminal psychological claims about persuasion. In particular, the claim that a cue seeker's perception of a cue-giver's knowledge and interests affects a statement's persuasiveness has famous historical antecedents, most notably Aristole (1954 translation) and Hovland, Janis, and Kelley (1953). While these antecedents are widely recognized, their substantive impact on political science is stalled by loose interpretations. As I will show, the aggregation of persuasion research that claims Hovland et al. in its lineage reveals that an incredibly large number of cue-giver attributes cause cue persuasivesness. The net effect of this research has been to bury seminal insights about persuasion so deeply that our discipline finds it hard to answer questions such as "When will a speaker's ideology affect the persuasiveness of his or her cues more than does his or her race, likability, or level of education?" Theory clarifies how the seminal insights should be interpreted, and experiments reveal that if we want to predict human behavior accurately, then careful use of the seminal insights is essential.

Next, I review seminal persuasion insights. Then I introduce a formal model of political communication and use it to clarify the determinants of persuasion. Next, I use a survey experiment on elite endorsements to test the theory's predictions. Then I offer a brief conclusion. Appendixes A and B contain more technical material pertaining to the model and the experiment.

SEMINAL FOUNDATIONS IN PERSUASION RESEARCH

I begin in the modern era with the insights offered by the Yale Communication and Attitude Change Program (e.g., Hovland et al. 1953). One of the program's most important insights is that a speaker's *credibility* determines his or her persuasive power. To evaluate this finding empirically, Hovland and his colleagues reduced credibility to a set of speaker attributes. "The Yale group stated that credibility is mainly based on two factors: expertise, which is the amount of knowledge that a communicator is assumed to possess, and trustworthiness, which is the perceived intention of the communicator to deceive" (Franzoi 1996:214).

While the claim that certain speaker attributes affect persuasion is sensible, it has a mixed legacy. On the positive side, there is a consensus

that attributes such as expertise and trustworthiness can affect whose cues people use. On the negative side, the studies are therefore insufficient to answer questions about what specific role these attributes play in determining cue persuasion. One manifestation of this mixed legacy is the long list of speaker attributes that scholars have correlated with cue persuasiveness. Consider, for example, Klapper's (1960:99) summary:

> In general, sources which the audience holds in high esteem appear to facilitate persuasion, while sources which the audience holds in low esteem appear to constitute at least a temporary handicap. The possible bases of such esteem are perhaps infinitely variable. Audiences have been shown, for example, to respond particularly well to specific sources because they considered them of high prestige, highly credible, expert, trustworthy, close to themselves, or just plain likable.

In another review of the persuasion literature twenty-five years later, McGuire (1985:263) could be no more precise. Not only did attractiveness join expertise and trustworthiness as causal factors, but each factor was represented empirically by its own extensive list of speaker attributes! The list of speaker attributes that scholars had correlated with cue persuasiveness now included social status, professional attainment, tallness, and erect posture.

The 1980s provided new innovations in persuasion research. These innovations, however, did not reveal what made some cues more persuasive than others. For example, Petty and Cacioppo's Elaboration Likelihood Model (ELM) defined two routes to persuasion. As the authors explain (1986:3), "The first type of persuasion was that which likely occurred as a result of a person's careful and thoughtful consideration of the true merits of the information presented in support of advocacy (central route). The other type of persuasion, however, was that which more likely occurred as a result of some simple cue in the persuasion context (e.g., an attractive source) that induced change without necessitating scrutiny of the central merits of the issue-relevant information presented (peripheral route)."

The ELM's distinction proved useful for many scholars, including political scientists. For the ELM's peripheral route resembled the cue-taking behavior that political scientists had been describing for decades (e.g., Berelson, Lazarsfeld, and McPhee 1954; Calvert 1985; Downs 1957; McKelvey and Ordeshook 1985; Popkin, Gorman, Phillips, and Smith 1976). While the ELM helped refocus political scientists' attention on persuasion, its authors recognized (1986:32) that ELM research "postponed the question of what specific qualities make arguments per-

suasive." Eagly and Chaiken (1993:323) gave a more direct assessment: "Because the elaboration likelihood model specifies when peripheral route persuasion should and has occurred, but not which peripheral mechanism has operated (or why), it leaves numerous mediational issues unaddressed."

In the 1990s, political psychologists became increasingly visible players in the study of cue-taking. Some of them showed that people can use cues as effective substitutes for the political information they lack.[6] However, a generalizable explanation of why only certain cues work has been absent. To be sure, there are careful arguments that suggest a strong relationship between certain attributes, such as a speaker's likability (e.g., Brady and Sniderman 1985) or ideology (e.g., Tetlock 1993), and persuasion. Still, our discipline cannot recite the conditions under which some other attribute – such as a cue-giver's race, personality, level of education, or economic interests – will overwhelm ideology or likability as determinants of persuasion. To this end, I offer the following formal model.

Formal Model

The model is that of Lupia and McCubbins (1998). Its logical lineage is most directly traceable to economic models of incomplete information (e.g., Harsanyi 1967, 1968a, 1968b), signaling models in economics and political science (e.g., Spence 1973), and economic cheap-talk models (Crawford and Sobel 1982; Farrell and Gibbons 1989). The main difference between the new framework and prior modeling efforts is in the assumptions about what cue givers and cue seekers know and in the conclusions about who can persuade whom. Most extant formal models of communication focus on the case where communicators know each other well. The new framework, by contrast, also explains cue persuasiveness when communicators know much less.

I focus on the case where cue persuasiveness is a function of *motivated choice*. I say that cue persuasiveness is a function of *choice* when cue seekers have the option of believing, ignoring, or rejecting a cue. I say that these choices are *motivated* when they are consequential from the cue-seeker's perspective (i.e., a voter may be motivated by the belief that his vote can be decisive, the belief that the mere act of participation is valuable, or something else). Put another way, I assume that cue seekers are either "motivated to hold correct attitudes" (Petty

6 See, for example, Kuklinski, Metlay, and May (1982), Kuklinski and Hurley (1994), Popkin (1991), Sniderman, Brody, and Tetlock (1991), and Lupia 1994.

and Cacioppo 1986:5) or "'economy minded souls' who wish to satisfy their goal-related needs in the most efficient ways possible" (Eagly and Chaiken 1993:330). In what follows, I describe the theory's premises and predictions. In Appendix A, I define the theory in mathematical terms.

Premises

I model communication as a game between two players, a *speaker* and a *listener.* The game begins when the speaker provides one of two cues: "x is better than y" or "x is worse than y." The listener then chooses x or y. The listener's choice of x or y ends the game and determines a payoff for each player. The model is meant to be analogous to the wide range of political situations in which one person has an opportunity to base his or her actions or opinions on the statements offered by another (e.g., the speaker runs a political campaign and the listener, who can follow the speaker's advice or ignore it, is a person whose support the speaker desires).

The speaker's attempt to persuade the listener occurs in the midst of up to four types of uncertainty. Each type of uncertainty is common to political contexts.[7] First, the listener is uncertain about the consequences of her actions. That is, the listener may not know whether choosing x or y makes her better off. Second, the listener may be uncertain about the speaker's knowledge of x and y (e.g., the listener may believe that the speaker is ignorant about which candidate is better for her). Third, the speaker may not in fact know which candidate is better for the listener.

Fourth, and finally, the listener may be uncertain about the speaker's motives. I say that the speaker and listener have *common interests* when the speaker benefits from the listener's choice only if the listener chooses what is better for her. So, if x is better than y for the listener, and if the listener and speaker have common interests, then the speaker benefits from the listener's decision only if x is chosen. Otherwise, I say that the speaker and listener have conflicting interests. Stated another way, the fourth type of uncertainty pertains to the listener's belief about whether she and the speaker have common or conflicting interests.

7 I treat each type of uncertainty as a probability distribution. As a result, the model allows me to describe persuasion in a broad range of contexts, including contexts in which there is no uncertainty. Note that the inclusion of these four types of uncertainty is what differentiates this chapter's model from most economic models of communication. The seminal model by Crawford and Sobel (1982), for example, includes only the first type of uncertainty.

Predictions

I first characterize persuasion in the model with a theorem. Then I describe the theorem's main implication for the question "Who can persuade whom?" Persuasion occurs only if the speaker's statement causes the listener to change her beliefs about which alternative is better for her.

> **Theorem 1:** Perceived common interests and perceived speaker knowledge are each necessary for persuasion.[8]

Note that perceived common interests are the listener's prior belief about the probability that she and the speaker have common interests. If the listener believes that this probability is low (less than .5), then the speaker will not persuade the listener – regardless of the speaker's actual knowledge. Similarly, perceived speaker knowledge is the listener's prior belief about the probability that the speaker, in fact, knows whether x or y is better for the listener. If this probability is 0, then persuasion does not occur.[9]

> **Implication:** If a speaker attribute is not necessary to change the listener's perception of the speaker's knowledge or interests, then it is irrelevant to cue persuasiveness. Thus, perceived common interests and perceived speaker knowledge explain when all other speaker attributes affect cue persuasiveness, while the converse of this statement is not true.

The logic behind this implication is as follows. A speaker's knowledge and interests are often impossible to observe directly. When this is true, a listener's perception of a speaker's interests and knowledge will be

8 The sufficient condition is the satisfaction of the necessary conditions plus the listener being so uncertain about x and y that the speaker's statement, if true, will cause her to hold a correct belief about which alternative is better for her.

9 The probability of perceived common interests and perceived speaker knowledge required for persuasion varies with the listener's uncertainty. If, for example, the listener is very uncertain about which alternative is better for her, then the levels of perceived speaker knowledge and common interests necessary for persuasion approximate those described in the text. If, however, the listener is more certain, then persuasion requires stronger perceptions. So, for example, if the listener is *certain* that the speaker lacks the knowledge she desires, then persuasion is impossible. This is true even if the listener is certain that she and the speaker share common interests. By contrast, if the listener believes that the speaker *might* possess the knowledge she requires, then persuasion is *possible*.

affected by other factors. However, the fact that speaker knowledge and interests are unobservable directly does not negate the fact that each factor is necessary for persuasion. Therefore, any speaker attribute can have a nonspurious correlation with cue persuasiveness *only if* the attribute determines the listener's perception of a speaker's knowledge or interests.

Put another way, the listener's perception of the speaker's knowledge and motives is the fundamental source effect in the context of cue persuasiveness.[10] All other speaker attributes (such as a speaker's attractiveness, party, race, likability, ideology, or reputation) affect cue selection when they do because they affect the listener's perception of the speaker's knowledge or motives. If an attribute does not have this effect, then any correlation between it and cue persuasiveness is indirect or spurious.

For example, you may be Extremely Liberal, and another person who you know to be Extremely Liberal in the same way is attempting to convince you that welfare reform is a great idea. If, however, you believe that the other person knows only a subset of what you know about welfare reform, then even your shared ideology gives you an insufficient basis for following that person's cue. Put another way, a speaker's ideology makes a cue persuasive *only if* it influences the listener's assessment of the speaker's interests or knowledge. By contrast, the speaker's interests and knowledge are always relevant to a cue's persuasiveness regardless of whether or not they convey any information about a speaker's ideology. At best, a speaker attribute such as likability, ideology, race, or gender can be said to *cause* cue persuasiveness *only if* the attribute is the listener's sole means of assessing the speaker's interests or knowledge.

CUE PERSUASION IN MASS COMMUNICATION CONTEXTS

In some political contexts, one person attempts to convince another about what he or she should do. More commonly, however, attempts at political persuasion often involve more people. In particular, when we

10 Neither *actual* speaker knowledge nor *actual* common interests is necessary or sufficient for the cue to be persuasive. This implies that a speaker can persuade a listener even if the speaker actually knows nothing. This occurs when the listener mistakenly perceives the speaker to be knowledgeable. More generally, a speaker's actual knowledge may have *absolutely nothing* to do with his ability to persuade. As a result, and as Hovland et al. (1953:21) recognized, knowledge is not necessarily (persuasive) power. For similar reasons, *a knowledgeable speaker who shares common interests with a listener will fail to persuade* if the listener does not perceive these interests accurately.

think of the persuasive attempts of candidates, campaigns, or advertisers, it is a case of one or more people attempting to persuade a large audience. A simple extension of the model just described reveals how people choose whom to believe in mass communication settings.

When a persuasive attempt involves only one speaker and one listener, as described earlier, persuasion is impossible when the speaker and listener have conflicting interests.[11] For when the listener perceives a speaker to have conflicting interests, her best response is to ignore the cue.[12] However, this claim contradicts a common experience – sometimes a cue is persuasive because a speaker has conflicting interests. For example, a person may oppose environmental regulation but might have her attitude changed by the claims of a pro-environment group; she may learn from their claims which way not to vote. I call this outcome "negative third-party persuasion."

To derive the conditions for third-party persuasion, it is sufficient to add third parties, called "observers," to the original model. The only difference between observers and the (original) listener is that the observers cannot directly affect the speaker's utility. Examples of observers include individuals at a mass rally or people watching a nationally televised political speech.

Amending the theory in this way allows me to revise Theorem 1 in the following way:

> **Theorem 1′:** If the observer believes that her interests conflict with both the listener's and the speaker's and that the speaker has an incentive to make truthful, knowledgeable statements to the listener, then the speaker can persuade.

11 Lupia and McCubbins (1998) extend the model described here in a different way to identify conditions under which certain external forces – inherent in common political institutions – substitute for common interests as a determinant of persuasion.

12 Formal models of communication, such as the one presented here, reveal why this is so. For example, a listener should not follow the advice of a speaker who she believes wants to deceive her. If she did, the speaker's best response would be to make a false statement. However, it is reasonable to ask why the listener might not then do the opposite of what the speaker recommends. The answer to this question is that the speaker in the model can anticipate such behavior and would adapt by making a true statement (which, if the listener did the opposite of what this statement recommends, would make her worse off). In equilibrium, the listener's best response is to make her choice of x or y independently of what the speaker says. In the case where the listener perceives the speaker to have conflicting interests, this strategy is the only one that prevents the speaker from deceiving her.

So, when an observer comes across a speaker in such a situation, she has an incentive to take the speaker's advice and do the opposite. For example, suppose that you are a Democrat who observes a prominent politician addressing an important group of Republican supporters. If you believe that the politician is knowledgeable, that he and the group perceive themselves to have common interests, and that your interests conflict with both, then you ought to do the opposite of what he recommends. *Positive third-party persuasion* occurs when an observer perceives that she shares common interests with both the listener and the speaker and that the speaker has an incentive to make truthful statements to the listener. In this case, the observer should follow the speaker's advice.

SURVEY EXPERIMENT

Of course, Theorem 1 and its implication are, at the moment, mere logical implications from a theory about how listeners who are either "motivated to hold correct attitudes" or "economy-minded souls" should process cues. To see whether people actually process cues in this way, I designed a simple experiment.

My experiment was part of the Multi-Investigator Study on Political Persuasion and Attitude Change. The study contained twelve separate and independently conceived experiments along with a set of core questions. The University of California's Survey Research Center conducted the study. The survey population consisted of all English-speaking adults eighteen years of age or older, residing in households with telephones, within the forty-eight contiguous U.S. states. Professional interviewers conducted all interviews between June 15 and November 4, 1994. The interviewers randomly contacted 2,234 households using computer-assisted telephone interviewing technology. In these households, 686 persons refused to participate, 68 were never at home, and 16 were unable to participate. The remaining 1,464 households constitute the sample.

The experiment consisted of three questions. It began with respondents hearing one version of the Attitude Question.

> **Attitude Question with No Cue:** *"Now I am going to ask you a couple of questions about a new issue in American politics – spending money to build prisons. What do you think? Is spending money to build prisons a good idea or a bad idea?"*
>
> **Attitude Question with Cue:** *"Now I am going to ask you a couple of questions about a new issue in American*

> *politics – spending money to build prisons. It's been reported that talk show host [SPEAKER] [POSITION] spending money to build more prisons. What do you think? Is spending money to build prisons a good idea or a bad idea?"*

The Attitude Question elicited data on the dependent variable of interest: respondent attitudes. It also contained the experimental variation, which had two components. The variation's first component, [SPEAKER], determined the speaker's identity. Respondents in the treatment group heard a version of the Attitude Question with a cue by Rush Limbaugh or Phil Donahue. Respondents in the control group heard a version of the Attitude Question with no cue. The variation's second component, [POSITION], determined the cue's content. Some respondents in the treatment group heard a cue (by either Limbaugh or Donahue) that supported spending on prison construction; others heard a cue that opposed it. Random assignment determined which version of the question each respondent heard.[13]

Although the Attitude Question contained the sole experimental intervention, the purpose of the experiment was to evaluate the relationship between cue persuasiveness and respondent perceptions. I designed the Interest Question and Knowledge Question to elicit measures of perceived speaker interests and knowledge, respectively.

> **Interest Question:** *"Now I am going to ask you a couple of questions about [SPEAKER]. On most political issues would you say that you and [SPEAKER] agree all of the time, most of the time, only some of the time, or never?"*

13 Each respondent had a 10% chance of hearing the Attitude Question without a cue and a 22.5% chance of hearing the Attitude Question with each of the four possible cues: "Rush Limbaugh supports," "Rush Limbaugh opposes," "Phil Donahue supports," and "Phil Donahue opposes." Note that this experiment employs a "posttest only" design. I can draw inferences about persuasion in such an experiment when the random assignment of respondents across experimental conditions results in each condition-specific subsample having roughly comparable initial attitudes. Note also that respondents had the option of responding "a very good idea" or "a very bad idea." To simplify the presentation, we collapsed such responses into the categories "a good idea," and "a bad idea" respectively. Only eleven respondents replied "don't know" to the Attitude Question. Each of them was then asked a similarly worded question in an attempt to elicit an opinion. Only three respondents subsequently replied "don't know." I dropped their interviews from the analysis.

Knowledge Question: *"How much would you say that [SPEAKER] knows about what will happen if this country spends money to build more prisons – a lot, some, a little, or nothing?"*

Since the Interest and Knowledge Questions offer only crude measures (as opposed to precise point estimates) of respondent perceptions, Theorem 1 implies that respondent attitudes are most likely to match the speaker's cue when respondents perceive the speaker to agree with them "all of the time" and to know "a lot." It also implies that the incidence of such matches should decrease as *either* of these perceptions becomes less favorable.[14]

Motivation for Choice of Issue and Speakers

I chose prison spending as the experiment's issue for three reasons. First, I expected it to be salient for many respondents; indeed, many public opinion polls showed crime to be a primary concern for many Americans at the time of the study. Second, I expected subjects to be unclear about what side to take on this issue. While building prisons is consistent with the law-and-order caricature of contemporary conservatism, spending money to solve problems fits better with the contemporary liberal stereotype of supporting government intervention. In other words, I chose the issue so that we could represent each of our speakers as either a supporter or an opponent. Third, I expected that many respondents would be uncertain about the effect of such a policy change. By contrast, had we chosen an issue whose consequences were transparent, we would

14 Some critics have questioned the experimental design on the grounds that respondents' answers to the Interest and Knowledge Questions are affected by the experimental manipulation in the Attitude Question rather than mediating the effects, as I argue. I concur that prior survey questions can affect answers to later ones, so this critique affects any ordering of the questions – not just this one. I chose this ordering, as it was *the least likely* to contaminate the experiment. That is, asking the Interest and Knowledge Questions first suggests that something approximating the Attitude Question is forthcoming, while the converse of this statement is not true. Moreover, if the critics are correct, then why did so few people who agreed with the speaker give him the highest ranking on the *Agrees* or *Knows* scales? Table 2.2 indicates that nothing of the sort occurred. And when we asked the respondent feeling thermometer questions about the same speaker, why did so few who agreed (disagreed) with the speaker offer a score of 100 (0)? Table 2.6 indicates that nothing of the sort occurred. The critics' complaint is, at best, a curious speculation and, at worst, unsupported by the data and theoretically baseless.

Table 2.1. *Responses to Opinion Question*

Category	% Replying "Good Idea"	*N*
All respondents	60	1427
Heard "supports" from either source	60	639
Heard "opposes" from either source	59	666
Heard no endorsement	61	122
Heard Phil Donahue supports	59	339
Heard Phil Donahue opposes	61	344
Heard Rush Limbaugh supports	61	300
Heard Rush Limbaugh opposes	57	322

expect, and Theorem 1 predicts, no persuasion – as it is difficult to persuade people of something they think they know.

I chose Rush Limbaugh and Phil Donahue as speakers for similar reasons. First, I wanted speakers with whom most respondents were likely to be familiar. In fact, only three respondents volunteered that they had never heard of the speaker whose cue we presented to them. Second, I wanted speakers for whom it was reasonable to expect variance in responses to the Interest and Knowledge Questions. Instead of choosing speakers who were likely to be universally trusted or reviled, I chose talk show hosts who were not widely recognized experts on crime-related issues.[15]

Analysis

In this section, I analyze the data. First, I present summary statistics about the responses of the 1,464 participants (Tables 2.1 and 2.2). Then I describe the relationship between perceived speaker knowledge, perceived speaker interests, and persuasion (Tables 2.3 and 2.4). Next, I examine the cue–persuasion relationship in the context of two alternative explanations – ideological similarity and the likability heuristic (Tables 2.5 and 2.6). Then I present a regression analysis that allows me to evaluate a broader range of explanations (Table 2.7).

Table 2.1 contains summary statistics of attitudes on prison spending. Overall, just under 60% of the respondents responded "good idea" to

15 Of course, I would have preferred to run many variations of this experiment combining many more issues and speakers. However, in a collaborative effort, such as the Multi-Investigator Study, the time available for questions is both valuable and scarce. I am happy to have run even this variation.

Table 2.2. *Responses to Agrees and Knows Questions*

Rush Limbaugh				
	Knows = a Lot	Knows = Some	Knows = a Little	Knows = Nothing
Agrees = all	8	1	0	0
Agrees = most	48	43	5	0
Agrees = some	30	155	90	26
Agrees = never	4	17	45	52
Phil Donahue				
	Knows = a Lot	Knows = Some	Knows = a Little	Knows = Nothing
Agrees = all	3	0	0	0
Agrees = most	19	28	4	0
Agrees = some	56	191	121	27
Agrees = never	4	16	34	31

the Attitude Question. To see that the experimental intervention alone did not create this division, notice that the level of support for prison spending in the control group (61%) was about the same as the level of support in the entire sample (60%).

These summary statistics also suggest that neither speaker was persuasive. For example, respondents who heard "Limbaugh supports" were only a little more likely (3.3 percentage points) to respond "good idea" than were those who heard "Limbaugh opposes." Respondents who heard "Donahue supports" were slightly less likely to support the issue than were those who heard "Donahue opposes."

Hidden beneath these statistics, however, is substantial variance in respondents' perceptions of speaker interests and knowledge. Theorem 1 implies that this variance is the key to identifying when the cues are most (and least) likely to persuade. Table 2.2 shows the distribution of respondents' perceptions.[16] For simplicity, I henceforth refer to responses to the Interest Question (perceived speaker interests) and responses to the Knowledge Question (perceived speaker knowledge) as the variables *Agrees* and *Knows*, respectively.

16 Tables 2.2 through 2.6 refer exclusively to respondents in the treatment group who answered the Attitude, Interest, and Knowledge questions. Comparisons with the control group are straightforward – 61.4% of the respondents in the control group responded "good idea."

Two elements of Table 2.2 merit attention. First, almost no respondents (12/1,058) reported agreeing with the speaker "all of the time." So, to clarify the tables that follow, I collapse the top two categories of *Agrees* into a single category labeled "*Agrees* = all or some." Second, respondents varied in their perceptions of the speaker. While the modal pair of responses was "*Agrees* = some" and "*Knows* = some," fewer than 35% (346/1,058) of the sample gave this response. This variance is important, as it gives us the opportunity to identify the correspondence between respondent perceptions and cue persuasiveness.

In Tables 2.3 and 2.4, I examine two predictions from Theorem 1. The predictions are "an increase in *Agrees* increases the incidence of persuasion when it has an effect" and "an increase in *Knows* increases the incidence of persuasion when it has an effect."[17]

I evaluate these predictions by looking at the correspondence between *Agrees* or *Knows* and respondent attitudes. Table 2.3 separates respondents by their choice of *Agrees* or *Knows*. For example, in the parts of Table 2.3 that pertain to the variable *Agrees*, I separate respondents who reported that they agree with Limbaugh "all or most" of the time from respondents who said they agree with him "some" of the time.

I further separate respondents who heard "supports" from those who heard "opposes." Then I report the percentage of these respondents who replied "good idea." For example, the top of Table 2.3 shows that, of the respondents who agreed with Limbaugh "all or most" of the time and heard that Limbaugh supports prison spending, 76% responded "good idea." By contrast, only 54% of the respondents who agreed with Limbaugh "all or most" of the time and heard that Limbaugh opposes prison spending responded "good idea."

The column labeled *Effect of Treatment* contains my measure of persuasion. *Effect of Treatment* measures how often the response to the Attitude Question ("good idea" or "bad idea") matches the speaker's cue. I computed *Effect of Treatment* by subtracting the percentage of respondents who heard "opposes" and replied "good idea" from the percentage of respondents who heard "supports" and replied "good idea." As the share of respondents whose attitudes match the speaker statements increases, so does *Effect of Treatment*. So, if all respondent attitudes match speaker statements, then *Effect of Treatment* = 100; if all respondent attitudes are contrary to speaker statements, then *Effect of*

17 *Agrees* and *Knows* are too blunt to allow reasonable interpretation of their lowest values as a 0% chance of common interests or knowledge, respectively. Therefore, the prediction from Theorem 1's implication is that for every category of *Agrees* (*Knows*), an increase in *Knows* (*Agrees*) will either increase or not change the incidence of persuasion.

Table 2.3. *Responses to Opinion Question, Separated by Agrees or Knows*

Rush Limbaugh				
Response to "Agrees" or "Knows" Question	Of Those Who Heard "Supports," % Replying "Good Idea"	Of Those Who Heard "Opposes," % Replying "Good Idea"	Effect of Treatment	*N* (Total, Heard "Supports," Heard "Opposes")
Agrees = all or most	76	54	22	(105, 51, 54)
Agrees = some	66	60	6	(306, 150, 156)
Agrees = never	33	55	–22	(123, 60, 63)
Knows = a lot	72	45	27	(90, 50, 40)
Knows = some	76	65	11	(222, 106, 116)
Knows = a little	46	60	–14	(142, 70, 72)
Knows = nothing	27	48	–21	(81, 37, 44)
Phil Donahue				
Response to "Agrees" or "Knows" Question	Of Those Who Heard "Supports," % Replying "Good Idea"	Of Those Who Heard "Opposes," % Replying "Good Idea"	Effect of Treatment	*N* (Total, Heard "Supports," Heard "Opposes")
Agrees = all or most	76	43	33	(54, 33, 21)
Agrees = some	59	59	0	(401, 191, 210)
Agrees = never	53	77	–24	(85, 38, 47)
Knows = a lot	74	51	23	(89, 46, 43)
Knows = some	62	53	9	(243, 125, 118)
Knows = a little	49	74	–25	(179, 83, 96)
Knows = nothing	48	65	–17	(64, 27, 37)

Treatment = −100, and if the number of respondent attitudes that match the speaker statements equals the number of respondent attitudes that are contrary to the speaker statements, then *Effect of Treatment* = 0.[18]

Theorem 1 predicts that *Effect of Treatment* should increase as we move from a lower category of *Agrees* or *Knows* to a higher category. Table 2.3 reveals that higher values of *Agrees* or *Knows* were associated with increases in *Effect of Treatment*. For example, respondents who agreed with Rush Limbaugh "all or most of the time" and who were

18 The final column gives the total number of respondents for each category, as well as the total number of respondents who heard "supports," and "opposes," respectively.

randomly assigned to hearing "Rush Limbaugh supports" were 22% more likely to respond "good idea" than were subjects with the same attribute who were randomly assigned to hearing "Rush Limbaugh opposes." As the model predicts, our experimental intervention ("supports" or "opposes") had a sizable persuasive effect on this subset of the sample. By contrast, respondents who agreed with Limbaugh "some of the time" were less affected by the cue (*Effect of Treatment* = 6). Moreover, respondents who never agreed with Limbaugh showed a tendency to take his advice and do the opposite (*Effect of Treatment* = −22); these reactions are evidence of *negative third-party persuasion.* With one exception, the relationship between *Agrees*, *Knows*, and *Effect of Treatment* in the Phil Donahue treatment conditions have the same direction and magnitude.

Next, I focus on the combined impact of *Agrees* and *Knows* on *Effect of Treatment*. It is this combined effect about which Theorem 1 is most explicit. It predicts the highest incidence of persuasion when both *Agrees* and *Knows* take on high values. It also predicts a decrease in the incidence of persuasion when either *Agrees* or *Knows* falls.

In Table 2.4, I separate respondents by *Agrees* and *Knows*. To simplify the presentation, I henceforth collapse *Knows* into a binary variable of "a lot or some" or "a little or nothing." I do this because the within-group variance of these combined categories was low relative to the between-group variance, as can be seen in Table 2.3.[19]

As predicted, at every level of *Agrees*, *Effect of Treatment* increases as *Knows* increases. Similarly, at every level of *Knows*, *Effect of Treatment* increases as *Agrees* increases. Moreover, for *Effect of Treatment* to take on positive values (i.e., for the speaker's cue to elicit matching respondent attitudes), both *Agrees* and *Knows* require high values. This is also consistent with the model's predictions, which is quite remarkable given the crudeness of the *Agrees* and *Knows* measures.

19 I exclude from Table 2.4 the category *Agrees* = "all or most" and *Knows* = "a little or nothing" because almost no respondents were in this category. Their behaviors were as follows: of the three respondents who heard Rush Limbaugh supports, none responded "good idea"; of the three respondents who heard that Rush Limbaugh opposes, both responded "good idea" (therefore *Effect of Treatment* for this group = −100); the one respondent who heard that Phil Donahue supports did not respond "good idea"; and of the three respondents who heard that Phil Donahue opposes, two responded "good idea" (therefore *Effect of Treatment* for this group = −67). Note also that in the top and bottom halves of Table 2.4 I switch the order of the category *Agrees* = "some" and *Knows* = "a little or nothing" and the category *Agrees* = "never" and *Knows* = "a lot or some." The theorem implies no prediction about which of these two categories should have a higher *Effect of Treatment.*

Table 2.4. *Responses to Opinion Question, Separated by Agrees and Knows*

Rush Limbaugh					
Response to "Agrees" Question	Response to "Knows" Question	Of Those Who Heard "Supports," % Replying "Good Idea"	Of Those Who Heard "Opposes," % Replying "Good Idea"	Effect of Treatment	*N* (Total, Heard "Supports," Heard "Opposes")
All or most	A lot or some	81	52	**29**	(100, 48, 52)
Some	A lot or some	76	63	**13**	(185, 93, 92)
Some	A little or nothing	49	56	**–7**	(116, 55, 61)
Never	A lot or some	46	62	**–16**	(21, 13, 8)
Never	A little or nothing	30	54	**–24**	(97, 47, 50)
Phil Donahue					
Response to "Agrees" Question	Response to "Knows" Question	Of Those Who Heard "Supports," % Replying "Good Idea"	Of Those Who Heard "Opposes," % Replying "Good Idea"	Effect of Treatment	*N* (Total, Heard "Supports," Heard "Opposes")
All or most	A lot or some	78	39	**39**	(50, 32, 18)
Some	A lot or some	63	52	**11**	(247, 119, 128)
Never	A lot or some	67	75	**–8**	(20, 12, 8)
Some	A little or nothing	52	68	**–16**	(148, 71, 77)
Never	A little or nothing	46	77	**–31**	(65, 26, 39)

Moreover, recall that Theorem 1 implies that actual speaker interests, knowledge, or attributes, such as ideology, are neither necessary nor sufficient for persuasion. Now note that neither Rush Limbaugh's nor Phil Donahue's *actual* knowledge, interests, ideology, personality, or reputation varied within the experiment – they were exogenous constants. As a result, these attributes cannot possibly be the source of the systematic variation in the data.

Table 2.5. *Responses to Opinion Question, Separated by Ideological Similarity, Agrees, and Knows*

Speaker-Respondent Ideology	Agrees Knows Category	Of Those Who Heard "Supports," % Replying "Good Idea"	Of Those Who Heard "Opposes," % Replying "Good Idea"	Effect of Treatment	*N* (Total, Heard "Supports," Heard "Opposes")
Same		61	59	2	(399, 188, 211)
Different		55	64	–9	(423, 211, 212)
Same	High	79	54	25	(95, 47, 48)
	Medium	61	57	4	(147, 64, 83)
	Low	52	68	–16	(101, 44, 57)
Different	High	73	25	48	(19, 15, 4)
	Medium	62	60	9	(123, 61, 62)
	Low	45	68	–23	(207, 100, 107)

Alternative Explanations

In Tables 2.5 through 2.7, I continue to examine the relationship between *Agrees*, *Knows*, and *Effect of Treatment*. However, I now do so in the context of other well-known cue persuasiveness explanations. I will demonstrate that Theorem 1 adds to and clarifies, as opposed to merely restates, these explanations.

To simplify the presentation in these analyses, I group respondents into three self-selected categories: *AK High*, *AK Medium*, and *AK Low*. Respondents classified themselves as *AK High* if *Agrees* = "all or most" and *Knows* = "a lot or some." Respondents classified themselves as *AK Medium* if *Agrees* = "some" and *Knows* = "a lot or some." Otherwise, respondents who answered the *Agrees* and *Knowledge Questions* classified themselves as *AK Low*. Tables 2.3 and 2.4 suggest that these groupings are relatively homogeneous for the purpose of our analysis (i.e., with respect to observed behavior, there is low within-group variance and high between-group variance).

In Table 2.5, I first consider ideological similarity as an alternative explanation of cue persuasiveness. This explanation is manifest in beliefs such as "people who share my ideology usually have correct attitudes."[20] Using ideological similarity to explain cue persuasiveness implies that,

20 In this analysis, I classify Phil Donahue as a liberal and Rush Limbaugh as a conservative and use responses to the standard ideology survey question to measure respondent ideologies. The ideology question is "Generally speaking, do you usually think of yourself as a liberal, a conservative, or what?"

regardless of the value of *Agrees* and *Knows*, *Effect of Treatment* should increase as we move from the case where the speaker and respondent have different ideologies to the case where they share ideology. This prediction finds little support in Table 2.5. In the top part of the table, *Effect of Treatment* is slightly higher for respondents who shared the speaker's ideology (+2) than for those who did not (−9). However, the bottom part of the table reveals that this correspondence is heavily conditioned by *Agrees* and *Knows*. In fact, *Effect of Treatment* was greatest (+48) for the subset of the sample who heard cues by speaker *with different ideologies* for whom *Agrees* and *Knows* were both high.

By contrast, Theorem 1 predicts that, regardless of ideological similarity, *Effect of Treatment* should increase as we move from *AK Low* to *AK Medium* to *AK High*. The data strongly support this prediction. When *Agrees* and *Knows* were high, so was *Effect of Treatment*, regardless of ideological similarity. As *Agrees* and *Knows* decreased, so did *Effect of Treatment*, regardless of ideological similarity.

Table 2.5 shows that claims such as "*Agrees* and *Knows* are mere restatements of the effect of ideology" are plainly false. As Theorem 1 predicts, when *Agrees* and *Knows* were high, respondents were persuaded by speakers with different ideologies, and when *Agrees* and *Knows* were low, ideological similarity was not sufficient for persuasion. This evidence makes clear the primacy of perceived common interests and perceived speaker knowledge as determinants of cue persuasiveness. Put another way, our respondents were far more likely to be persuaded by someone they regarded as knowledgeable and as having common interests than by someone they perceived to be a common ideologue.

I next consider a common measure of affect as an alternative explanation of cue persuasiveness. The "likability heuristic" is expressed by beliefs such as "People should agree with people they like" or "People I like usually have correct opinions." In his review of the relevant social psychology literature, O'Keefe (1990:107) states, "Where this heuristic is invoked, liked sources should prove more persuasive than disliked sources."

Feeling thermometers are the conventional measure of "liking." If feeling thermometers are a good measure of liking, and if the likability heuristic is a good predictor of persuasion, then an increase in the speaker's thermometer should correspond to an increase in *Effect of Treatment*.[21] In Table 2.6, I use feeling thermometers to reevaluate the

21 The Multi-Investigator Study's feeling thermometer question was: "I'll read a name and ask you to rate the person on a thermometer that runs from zero to one hundred. The higher the number, the warmer or more favorable you feel toward that person. The lower the number, the colder or less favorable you feel. If you feel neither warm nor cold toward them, rate that person a fifty."

Table 2.6. *Responses to Opinion Question, Separated by Feeling Thermometer, Agrees, and Knows*

Therm	Agrees Knows Category	Of Those Who Heard "Supports," % Replying "Good Idea"	Of Those Who Heard "Opposes," % Replying "Good Idea"	Effect of Treatment	*N* (Total, Heard "Supports," Heard "Opposes")
0 to 10		45	63	**–18**	(234, 114, 120)
11 to 20		57	50	**7**	(62, 28, 34)
21 to 30		49	52	**–3**	(113, 51, 62)
31 to 40		63	77	**–14**	(98, 54, 44)
41 to 49		56	89	**–33**	(18, 9, 9)
50		66	59	**7**	(355, 163, 192)
51 to 60		56	62	**–6**	(111, 61, 50)
61 to 70		66	53	**13**	(87, 47, 40)
71 to 80		71	51	**20**	(86, 45, 41)
81 to 90		53	41	**12**	(32, 15, 17)
91 to 100		100	67	**33**	(16, 7, 9)
Under 50	High	100	67	**33**	(15, 7, 6)
	Medium	65	59	**6**	(131, 63, 68)
	Low	45	63	**–18**	(302, 146, 156)
Over 50	High	78	47	**31**	(116, 63, 53)
	Medium	64	52	**8**	(144, 77, 67)
	Low	31	70	**–39**	(56, 26, 30)

experimental data. The likability heuristic implies that, regardless of the value of *Agrees* and *Knows*, *Effect of Treatment* should increase as we move from low thermometer scores to high ones. The top and bottom parts of Table 2.6 reveal this prediction's limited success.

By contrast, Theorem 1 predicts that, regardless of the value of the feeling thermometer score, *Effect of Treatment* should increase as we move from *AK Low* to *AK Medium* to *AK High*. The bottom part of Table 2.6 shows that when *Agrees* and *Knows* are high, people can indeed be persuaded by speakers they do not like. Similarly, when *Agrees* and *Knows* are low enough, even liked speakers cannot persuade.

Table 2.6 supports my general conclusion about attribute-based explanations of cue-taking behavior. When people like others because of their knowledge or interests, then liking may well be correlated with cue persuasiveness. However, if a listener likes a speaker but regards the speaker as either lacking knowledge or having conflicting interests, then likability will not affect persuasiveness. Consequently, my explanation for the

failure of this affect-driven variable is that feeling thermometers are a terrible measure of the factors that cause persuasion; they do not allow for a speaker's perceived knowledge or interests to mediate the manner in which a respondent's feelings affect attitudes.[22]

Of course, it is possible that the preceding analysis obscures the fact that some people use the likability heuristic, others condition their willingness to follow a cue on ideological similarity, and still others use other well-known explanations of cue persuasiveness as ways to choose whom to believe. I explore this possibility, and conclude my analysis, by conducting logit analyses that simultaneously incorporate a broad range of alternative explanations of cue persuasiveness. In each logit analysis, the dependent variable is the response to the Attitude Question. The dependent variable equals 1 if the respondent's answer to the Attitude Question was "good idea" and 0 if the response was "bad idea." To simplify the interpretation of the logit coefficients, I scaled *all* independent variables to the range [0,1].

The most important explanatory variables, from the perspective of Theorem 1, are those representing the interaction between *Agrees*, *Knows*, and cue's content. So, for example, the variable *Supports AK High* equals 1 if and only if *Agrees* = "all or most," *Knows* = "a lot or some," and the respondent heard a cue supporting prison spending. Similarly, *Opposes AK Medium* equals 1 if and only if *Agrees* = "some," *Knows* = "a lot or some," and the respondent heard a statement opposing prison spending. The other AK variables have equivalent definitions.

Theorem 1 predicts that *Supports AK High* will have a large positive coefficient, that *Supports AK Medium* will have a smaller positive coefficient, that *Opposes AK High* will have a large negative coefficient, and

22 I have also considered process-based explanations of cue persuasiveness. For example, Petty and Cacioppo (1986) argue that a person's "need for cognition" affects whether he or she will pursue the central or peripheral route to persuasion and, therefore, partially explains cue usage. As O'Keefe (1990:101) summarizes: "it appears that persons low in need for cognition are relatively more influenced by peripheral persuasion cues [e.g., speaker attributes] than are those in high need for cognition; and, correspondingly, those in high need for cognition appear to be more influenced by the quality of the message's arguments than are those low in need for cognition." My analysis of the effect of need for cognition parallels that of the analysis of feeling thermometers in Table 2.6 – both in construct and in consequence. As was true with feeling thermometers, what I draw from the relative failure of the need for cognition variable is that need for cognition explanations of cue persuasiveness should be founded on the premise that people who use the peripheral route to persuasion are not passive recipients of all cues – even people on the peripheral route make systematic choices about whose cues to use.

that *Opposes AK Medium* will have a smaller negative coefficient. If we allow for the possibility of negative third-party persuasion, then Theorem 1′ implies that *Supports AK Low* will have a negative coefficient and that *Opposes AK Low* will have a positive coefficient. That is, Theorem 1 implies that if *Agrees* or *Knows* increases, then so does the likelihood that the respondent's attitude matches the speaker's cue.

The remaining independent variables represent either alternative explanations of cue-taking or demographic variables that could affect the preferences on prison spending. Since most of these variables are secondary to my argument and have small and insignificant coefficients, I defer their descriptions to Appendix B.

Table 2.7 contains the results. The table describes three logistic regressions. The first logit includes only variables derived from responses to the Interest and Knowledge Questions. The second and third logits add variables representing alternative explanations and demographic variables that, independent of persuasion, could affect attitudes on prison spending.

The signs on the AK coefficients correspond precisely to the model's predictions. Our experimental treatment had its greatest effect on attitudes when *Agrees* and *Knows* were *High*, and its effect decreased as we moved to *AK Medium* and *AK Low*. Moreover, the impact of the AK variables did not vary much across the logits. The consistent impact of the AK variables is evidence that *Agrees* and *Knows* were not merely restatements of the alternative explanations. By contrast, the performance of independent variables derived from alternate explanations of cue persuasiveness was uniformly weak. In sum, the experiment reveals that over 1,000 randomly selected survey respondents made systematic, and seemingly motivated, choices about whose cues to use.

CONCLUSION

Formal models of persuasion demonstrate systematic relationships between simple assumptions about speakers and listeners and clear conclusions about when persuasion occurs. Persuasion experiments generate examples of how certain types of cues affect behavior. Yet many formal models of persuasion are based on premises about human cognition that few psychologists recognize as reasonable, and many persuasion experiments are not attached to clearly stated theory. Fortunately, there are remedies to these ills.

Attention to the empirical foundations of modern psychology can alert modelers to more reasonable assumptions about human decision makers. Attention to the logical requirements of rational choice theory can reveal ways to draw a logically consistent lesson from the wealth of extant

Table 2.7. *Multivariate Analyses of Experimental Data*

Position	Independent Variable	Pred. Sign	Agrees and Knows Only	With *Ideology* Controls	With *Party* Controls
	Constant		0.42[a]	−0.36	−0.38
			(0.10)	(0.55)	(0.56)
Supports	AK High	+, big	**0.97**[a]	**1.26**[a]	**1.13**[a]
			(0.30)	(0.38)	(0.38)
Supports	AK Medium	+	**0.38**[a]	**0.74**[a]	**0.68**[a]
			(0.18)	(0.25)	(0.26)
Supports	AK Low		**−0.59**[a]	−0.05	−0.11
			(0.17)	(0.25)	(0.25)
Opposes	AK High	−, big	**−0.60**[a]	**−0.81**[a]	**−0.73**[a]
			(0.28)	(0.35)	(0.35)
Opposes	AK Medium	−	−0.14	−0.33	−0.19
			(0.17)	(0.24)	(0.24)
Opposes	AK Low		0.13	−0.05	0.06
			(0.17)	(0.24)	(0.24)
Supports	Need for cognition	−		−0.39	−0.59
				(0.32)	(0.32)
Opposes	Need for cognition	+		0.11	0.28
				(0.30)	(0.31)
Supports	Low involvement	+		−0.16	−0.14
				(0.19)	(0.19)
Opposes	Low involvement	−		**−0.53**[a]	**−0.50**[a]
				(0.19)	(0.19)
	Conservative/ Republican			0.29	0.19
				(0.23)	(0.29)
	Liberal/ Democrat			−0.24	−0.08
				(0.24)	(0.28)
	Moderate			**−1.60**[a]	
				(0.63)	
Supports	Same Ideology/Party	+		−0.47	0.23
				(0.31)	(0.36)
Supports	Different Ideology/Party	−		−0.38	0.33
				(0.30)	(0.36)
Opposes	Different Ideology/Party	+		0.03	0.12
				(0.22)	(0.19)
Supports	Thermometer	+		0.67	0.74
				(0.43)	(0.43)
Opposes	Thermometer	−		0.36	0.29
				(0.41)	(0.41)
	African American			**−0.62**[a]	**−0.53**[a]
				(0.21)	(0.22)
	Age			**1.22**[a]	**1.28**[a]
				(0.32)	(0.32)
	Education			0.07	−0.07
				(0.18)	(0.17)
	Observations		1427	1190	1190
	Initial Log Likelihood		−989.12	−824.85	−824.85
	End Log Likelihood		−938.72	−751.33	−759.49

Dependent Variable: = 1; if response to the Opinion question was "good idea."
= 0; if response to the Opinion question was "bad idea."

[a] Indicates that the coefficient is significant at the .05 level.

experiments. The research presented in this chapter is but one modest attempt to attend to both sets of lessons simultaneously. While rational choice theory and political psychology often proceed as though the other tradition has nothing to contribute, for behavioral questions in which each tradition's methods provide insufficient analogies, both traditions are worse off for their lack of interaction. There are important questions about politics that do not allow the luxury of simple argument. For scholars whose goal is the pursuit of clear and reliable explanation, logically organized and empirically grounded analyses are the only path to success. Together, the logic of rational choice theory and the empirical foundations of political psychology provide the foundations for achieving such success.

APPENDIX A: THE FORMAL MODEL

The purpose of this appendix is to supply the notation necessary to present a precise statement of the model's equilibrium. Readers who are interested in knowing more about this theoretical framework should consult Lupia and McCubbins (1998).

The sequence of events begins with three probabilistic choices by nature. These choices are the source of the four types of uncertainty described in the text. I denote these choices $n = \{\beta, A, K\}$. The order of these choices is irrelevant. Unless otherwise stated, all elements of the model are common knowledge.

The choice $\beta \in$ {better, worse} determines whether x is better or worse than y for the listener. Nature chooses the state $\beta =$ better (for the listener) with probability $b \in [0,1]$ and the state $\beta =$ worse with probability $1 - b$. The listener knows b, but not β. If $\beta =$ better and the listener chooses x, then the listener earns utility $U \geq 0$. If $\beta =$ worse and if the listener chooses x, then she earns utility $\underline{U} \leq 0$. If the listener chooses y, then she earns utility 0.

The choice $K \in \{0,1\}$ determines whether or not the speaker knows the true value of β. Nature allows the speaker to know β ($K = 1$) with probability $k \in [0,1]$ and makes no such revelation (chooses $K = 0$), with probability $1 - k$. The speaker knows K, while the listener does not. Both players know k.

The choice $A \in \{0,1\}$ determines whether the speaker and listener have common or conflicting interests. If $A = 1$ (common interests), then the speaker receives utility $Z \geq 0$ when the listener receives utility $U \geq 0$ and receives utility $\underline{Z} \leq 0$ when the listener receives utility $\underline{U} \leq 0$. If $A = 0$ (conflicting interests), then the speaker receives utility $\underline{Z} \leq 0$ when the listener receives utility $U \geq 0$ and receives utility $Z \geq 0$ when the listener receives utility $\underline{U} \leq 0$. If the listener chooses y, then the speaker earns

utility 0. Nature chooses $A = 1$ with probability $a \in [0,1]$ and $A = 0$ with probability $1 - a$. The speaker knows A and knows that the listener does not. Both players know a.

After nature makes its three choices, the speaker sends a signal $s \in \{B,W\}$ to the listener. $s = B$ is the signal "I assert that x is better than y for the listener." $s = W$ is the signal "I assert that x is worse than y for the listener." Next the listener chooses x or y. Then the game ends and both players receive a utility payoff.

I use the perfect Bayesian equilibrium to derive the theoretical results (see Fudenburg and Tirole 1991). Let the vector π be a typical strategy profile, the scalar h a typical information set, and the vector μ a typical system of beliefs. So, $\mu(h)$ is a player's beliefs about which of several unobservable events – the decision nodes within information set h – has led to his present observable situation – the information set h. Formally, μ is a function from $d \in D$, the set of decision nodes, to $[0,1]$, such that for every information set h, $\Sigma_{d \in h}\, \mu(d) = 1$. I make the usual assumption that the game's information sets collectively partition D.

Let π_s denote the speaker's component of strategy profile π. π_s has six scalar elements, one for each speaker information set $h_s \in \{h_1, \ldots, h_6\}$. Let $N = \{\beta,K,A\}$ be the vector of moves by nature. Then at h_1, $N = \{1,1,1\}$, at h_2, $N = \{0,1,1\}$, at h_3, $N = \{1,1,0\}$, at h_4, $N = \{0,1,0\}$, at h_5, $N \in \{\{0,1\}, 0,1\}$, and at h_6, $N \in \{\{0,1\}, 0,0\}$. Speaker information sets are completely determined by nature's choice vector N. Each element, $\pi_s(s;h_j)$, $j = 1, \ldots 6$, is the probability that the speaker signals $s \in \{B,W\}$ if he or she is at information set h_j. These probabilities sum to 1 for each information set.

I use the vector π_r to denote the listener's component of strategy profile π. This vector has two scalar elements, one for each listener information set $h_r = \{h_B,h_W\}$. Note that the listener's information sets are completely determined by the speaker's signal. Each element, $\pi_r(x;s)$, is the probability that the listener chooses x after having heard the signal $s \in \{B,W\}$. $1 - \pi_r(x;s)$ is the probability that the listener chooses y given the same signal. A signal s is "along the path of play" if there exists an information set at which $\pi_s(s;h_s) > 0$.

Definition

A pair of strategy profiles (π_r,π_s) is a perfect Bayesian equilibrium if (1) for each h_s, $\pi_s(s;h_s)$ maximizes the speaker's expected utility given $\pi_r(x;s)$ for all $s \in \{B,W\}$; (2) for each s that is along the path of play, $\pi_r(x;s)$ maximizes the listener's expected utility given $\mu(\text{better}|s)$ and $\mu(\text{worse}|s)$, where μ is computed from π_s by Bayes's rule; and (3) for any s that is

not along the path of play, $\pi_r(x;s)$ maximizes the listener's expected utility given $\mu(\text{better}|s) = b$ and $\mu(\text{worse}|s) = 1 - b$.

I identify the set of nonbabbling perfect Bayesian equilibria. A babbling equilibrium requires either a listener who ignores all signals or a speaker who sends only uninformative signals. In this model, a babbling equilibrium is an equilibrium in which either the speaker does not base his signal on n_b or the listener does not base her response on s. I focus on nonbabbling equilibria because we are interested in determining the conditions under which people *can* persuade each other when they attempt to communicate with each other. I also focus on nonneologistic equilibra. In our model, a neologistic equilibrium *requires* the speaker and listener to agree that the signal B means "worse" and not "better" and that the signal W means "better" and not "worse." Focusing on nonneologistic equilibria is equivalent to assuming that words have focal meanings. Since the speaker can lie, focusing on nonneologistic equilibria is not restrictive. For notational simplicity, let $\pi = (\pi_r,\pi_s)$, $\pi_{14} = (\pi_s(B;h_1), \pi_s(B;h_2), \pi_s(B;h_3), \pi_s(B;h_4))$, $\pi_5 = \pi_s(B;h_5)$, $\pi_6 = \pi_s(B;h_6)$, and $\pi_r = (\pi_r(x;B), \pi_r(x;W))$.

Equilibrium

The only nonbabbling, nonneologistic perfect Bayesian equilibrium in the basic model is $\pi_{14} = (1,0,0,1)$; $\pi_5 = 1$; if $bZ + (1 - b)\underline{Z} \geq 0$ and $\pi_5 = 0$ otherwise; $\pi_6 = 1$; if $b\underline{Z} + (1 - b)Z \geq 0$ and $\pi_6 = 0$ otherwise; $\pi_r = (1,0)$. This equilibrium requires Condition A:

$$\frac{[(1-a)k + [(1-k) \times [\pi_s(B;h_5)a + \pi_s(B;h_6)(1-a)]]]}{[ak + [(1-k) \times [\pi_s(B;h_5)a + \pi_s(B;h_6)(1-a)]]]} \leq bU/(b-1)\underline{U}$$

and Condition B:

$$\frac{[ak + [(1-k) \times [(1-\pi_s(B;h_5))a + (1-\pi_s(B;h_6))(1-a)]]]}{[(1-a)k + [(1-k) \times [(1-\pi_s(B;h_5))a + (1-\pi_s(B;h_6))(1-a)]]]} \geq bU/(b-1)\underline{U}$$

where at least one of the inequalities is strict.

Proof

I proceed as follows. First, I define the expected value of every pure strategy at every speaker information set. Second, I identify the boundaries of the set of potential nonbabbling, nonneologistic perfect Bayesian equilibria. Third, I identify the sequentially rational strategy profiles within this set. I find that the named equilibrium is this set's only member.

Finally, I evaluate the consistency of the sequentially rational strategy profiles.

To see the expected value of every pure strategy at every speaker information set, consider the following relationships. At h_1, the expected utility from $\pi_s(B;h_1) = 1$ is $\pi_r(x;B)Z$. The expected utility from $\pi_s(B;h_1) = 0$ is $\pi_r(x;W)Z$. If $\pi_r(x;B) \geq \pi_r(x;W)$, then $\pi_s(B;h_1) = 1$ is the best response. At h_2, the expected utility from $\pi_s(W;h_2) = 1$ is $\pi_r(x;B)\underline{Z}$. The expected utility from $\pi_s(W;h_2) = 0$ is $\pi_r(x;W)\underline{Z}$. If $\pi_r(x;B) \geq \pi_r(x;W)$, then $\pi_s(B;h_2) = 0$ is the best response. At h_3, the expected utility from $\pi_s(B;h_3) = 1$ is $\pi_r(x;B)\underline{Z}$. The expected utility from $\pi_s(B;h_3) = 0$ is $\pi_r(x;W)\underline{Z}$. If $\pi_r(x;B) \geq \pi_r(x;W)$, then $\pi_s(B;h_3) = 0$ is the best response. At h_4, the expected utility from $\pi_s(W;h_4) = 1$ is $\pi_r(x;B)Z$. The expected utility from $\pi_s(W;h_2) = 0$ is $\pi_r(x;W)Z$. If $\pi_r(x;B) \geq \pi_r(x;W) = 0$, then $\pi_s(B;h_4) = 1$ is the best response. At h_5, the expected utility from $\pi_s(B;h_5) = 1$ is $b\pi_r(x;B)Z + (1 - b)\ \pi_r(x;B)\underline{Z}$. The expected utility from $\pi_s(W;h_5) = 0$ is $b\pi_r(x;W)Z + (1 - b)\ \pi_r(x;W)\underline{Z}$. If $\pi_r(x;B) \geq \pi_r(x;W)$ and $bZ + (1 - b)\underline{Z} \leq 0$, then $\pi_s(B;h_5) = 0$ is the best response. At h_6, the expected utility from $\pi_s(B;h_6) = 1$ is $b\pi_r(x;B)\underline{Z} + (1 - b)\pi_r(x;B)Z$. The expected utility from $\pi_s(W;h_6) = 0$ is $b\pi_r(x;W)\underline{Z} + (1 - b)\ \pi_r(x;W)Z$. If $\pi_r(x;B) \geq \pi_r(x;W)$ and $b\underline{Z} + (1 - b)Z \leq 0$, then $\pi_s(B;h_6) = 0$ is the best response.

> **Lemma 1:** All mixed-strategy perfect Bayesian equilibria in the model are babbling equilibria.

Proof of Lemma 1. A mixed-strategy equilibrium requires that each player choose a strategy that makes the other player indifferent between their two pure strategies. A necessary and sufficient condition for rendering the speaker indifferent between his pure strategies at information sets h_1 through h_4 is to set $\pi_r(x;B) = \pi_r(x;W)$. Setting $\pi_r(x;B) = \pi_r(x;W)$ is also necessary and sufficient to make the speaker indifferent between her two strategies at h_5 if $bZ + (1 - b)\underline{Z} \equiv 0$ and at h_6 if $b\underline{Z} + (1 - b)Z \equiv 0$. Setting $\pi_r(x;B) = \pi_r(x;W)$ implies that the listener is not conditioning her strategy on the signal. Anticipating such behavior, the speaker can choose any strategy he likes. These speaker strategies will either make the listener indifferent between her pure strategies, in which case we have a babbling equilibrium, or they will not, in which case we do not have an equilibrium.

If $bZ + (1 - b)\underline{Z} = 0$ or $b\underline{Z} + (1 - b)Z = 0$, then any listener strategy, including $\pi_r(x;B) = \pi_r(x;W)$, makes the speakers at h_5 and h_6 indifferent. Note, however, that the listener has an incentive to choose a mixed strategy other than $0 < \pi_r(x;B) = \pi_r(x;W) < 1$ only if she can induce the speaker at h_5 and h_6 to take distinct and knowledge transferring actions. Since the speaker at h_5 and h_6 has no useful private information at either of

these information sets, by definition, the requirement cannot be met. Therefore, only an equilibrium that could result from such an adaptation is a babbling equilibrium. QED.

From similar logic, it follows that all equilibria for which $\pi_r(x;B) = \pi_r(x;W)$ are babbling equilibria. Moreover, any nonbabbling equilibrium for which $\pi_r(x;B) = 0$ and $\pi_r(x;W) = 1$ requires neologisms (i.e., both players know that B means "worse" and W means "better"). Therefore, nonbabbling, nonneologistic perfect Bayesian equilibria must include $\pi_r = (1,0)$.

Since nonbabbling, nonneologistic perfect Bayesian equilibria must include $\pi_r = (1,0)$, they must also include $\pi_{14} = (1,0,0,1)$. The reason for this is that the expected speaker utility at h_1 through h_4 reveals $\pi_{14} = (1,0,0,1)$ to be the unique profile of best responses when $\pi_r(x;B) > \pi_r(x;W)$. Therefore, the set of nonbabbling, nonneologistic perfect Bayesian equilibria must be contained *within* $\pi = (1,0,0,1, \{0,1\}, \{0,1\}, 1,0)$, where $\{0,1\}$ within strategy profile π is read as "either 0 or 1." First, one must identify the sequentially rational strategy profiles within this set and then evaluate these profiles' consistency.

At h_B, the expected utility from $\pi_r(x;B) = 1$ is

$$[akb\pi_S(B;h_1)U + ak(1-b)\pi_S(B;h_2)U + (1-a)kb\pi_S(B;h_3)U + (1-a)k(1-b)\pi_S(B;h_4)U + a(1-k)b\pi_S(B;h_5)U + a(1-k)(1-b)\pi_S(B;h_5)U + (1-a)(1-k)b\pi_S(B;h_6)U + (1-a)(1-k)(1-b)\pi_S(B;h_6)U]/[akb\pi_S(B;h_1) + ak(1-b)\pi_S(B;h_2) + (1-a)kb\pi_S(B;h_3) + (1-a)k(1-b)\pi_S(B;h_4) + a(1-k)b\pi_S(B;h_5) + a(1-k)(1-b)\pi_S(B;h_5) + (1-a)(1-k)b\pi_S(B;h_6) + (1-a)(1-k)(1-b)\pi_S(B;h_6)]$$

At h_W, the expected utility from $\pi_r(x;W) = 1$ is

$$[akb(1-\pi_S(B;h_1))U + ak(1-b)(1-\pi_S(B;h_2))U + (1-a)kb(1-\pi_S(B;h_3))U + (1-a)k(1-b)(1-\pi_S(B;h_4))U + a(1-k)b(1-\pi_S(B;h_5))U + a(1-k)(1-b)(1-\pi_S(B;h_5))U + (1-a)(1-k)b(1-\pi_S(B;h_6))U + (1-a)(1-k)(1-b)(1-\pi_S(B;h_6))U]/[akb(1-\pi_S(B;h_1)) + ak(1-b)(1-\pi_S(B;h_2)) + (1-a)kb(1-\pi_S(B;h_3)) + (1-a)k(1-b)(1-\pi_S(B;h_4)) + a(1-k)b(1-\pi_S(B;h_5)) + a(1-k)(1-b)(1-\pi_S(B;h_5)) + (1-a)(1-k)b(1-\pi_S(B;h_6)) + (1-a)(1-k)(1-b)(1-\pi_S(B;h_6))]$$

Recall that the listener earns utility zero for choosing y. Therefore, $\pi_r(x;B) = 1$ is the best response only if the expected utility from $\pi_r(x;B) = 1$ is ≥ 0 and $\pi_r(x;W) = 0$ is the best response only if the expected utility from $\pi_r(x;W) = 1$ is ≤ 0. Since a nonbabbling equilibrium requires that the expected utility from $\pi_r(x;B) = 1$ is ≥ 0, that the expected utility from $\pi_r(x;W) = 1$ is ≤ 0, and that one of these inequalities is strict, it requires that one of the inequalities in Conditions A or B be strict.

We can now prove that $\pi = (1,0,0,1,0,0,1,0)$ is a perfect Bayesian equilibrium under the conditions of the equilibrium. The other cases – $\pi = (1,0,0,1,0,1,1,0)$, $\pi = (1,0,0,1,1,0,1,0)$, and $\pi = (1,0,0,1,1,1,1,0)$ – follow equivalent logic. From the expected utility at information sets h_5 and h_6, we know that this equilibrium holds only if $bZ + (1 - b)\underline{Z} \leq 0$ and $b\underline{Z} + (1 - b)Z \leq 0$. This requirement matches the related requirement of the equilibrium. From the expected utility at information sets h_1 through h_4, we know that this equilibrium requires the expected utility of $\pi_r(x;B) = 1 \geq 0 \geq$ the expected utility of $\pi_r(x;W) = 1$. We evaluate the conditions under which this inequality holds subsequently.

If $\pi_s = (1,0,0,1,0,0)$, then the numerator of the expected utility from $\pi_r(x;W) = 1$ reduces to $ak(1 - b)U + (1 - a)kbU + a(1 - k)bU + a(1 - k)(1 - b)U + (1 - a)(1 - k)bU + (1 - a)(1 - k)(1 - b)U$. It is trivial to show that this quantity is ≤ 0 iff $[1 - k + ak]/[1 - ak] \geq bU/(b - 1)U$, which is true iff Condition B is true. Similarly, if $\pi_s = (1,0,0,1,0,0)$, then the numerator of the expected utility from $\pi_r(x;B) = 1$ reduces to $akbU + (1 - a)k(1 - b)U$. It is trivial to show that this quantity is ≥ 0 iff $bU/(b - 1)\underline{U} \geq [1 - a]/a$, which is true iff Condition A is true. Therefore, $\pi = (1,0,0,1,0,0,1,0)$ is sequentially rational under the conditions of the equilibrium.

If the beliefs required to support this profile are consistent, then the profile and beliefs together constitute a perfect Bayesian equilibrium. Beliefs are consistent iff

$$\mu(better|B) = (b \times probability\ that\ s = B\ if\ B = better)$$

$$[(b \times probability\ that\ s = B\ if\ B = better) + ((1 - b) \times probability\ that\ s = W\ if\ B = better)]$$

In the proposed equilibrium $\mu(\text{better}|B) = 1$, the probability that $s = W$ *if* $B =$ better is zero and the probability that $s = B$ *if* $B =$ better is nonzero. Therefore, beliefs are consistent. Equivalent logic proves consistency for $\mu(\text{better}|W)$, $\mu(\text{worse}|B)$, and $\mu(\text{worse}|W)$. QED.

Lemma: The equilibrium exists only if $a > .5$ and $k > 0$.

Theorem: Perceived common interests and perceived speaker knowledge are necessary, but not sufficient, for persuasion. The sufficient condition is the satisfaction of both necessary conditions plus the listener's uncertainty about x and y leading her to believe that she can benefit from the speaker's knowledge.

APPENDIX B: OTHER INDEPENDENT VARIABLES

Like the *AK* variables, I introduce likability, ideological similarity, and partisan similarity (another popular attribute-based explanation of cue-taking that is analogous to ideological similarity) into the empirical model by interacting them with the content of the cue. The alternative explanations that generate these three sets of variables imply that each set will affect the dependent variable in the same way.

Affect-driven explanations of cue-taking imply that *Supports Thermometer* will have a large positive coefficient and that *Opposes Thermometer* will have a large negative coefficient. For respondents who heard a "supports" cue, I set *Opposes Thermometer* = .5. I used an equivalent transformation for respondents who heard an "opposes" cue. This transformation is consistent with the wording of the thermometer questions; respondents were instructed to rate the person as a 50 if they feel neither warm nor cold toward the person.

The variables *Supports/Same Ideology*, *Supports/Different Ideology*, and *Opposes/Different Ideology* account for the ideological similarity between respondent and speaker. For example, the variable *Supports/Same Ideology* equals 1 if either a conservative respondent hears "Rush Limbaugh supports" or a liberal respondent hears "Phil Donahue supports." It equals 0 otherwise. I use partisan similarity variables in an analogous manner. Ideological similarity-based explanations imply that *Supports/Same Ideology* will have a large positive coefficient, *Opposes/Different Ideology* will have a large positive coefficient, and *Supports/Different Ideology* may have a negative coefficient. Partisan-based similarity explanations of persuasion suggest analogous implications for the variables *Supports/Same Party*, *Opposes/Different Party*, and *Supports/Different Party*, respectively. I ran separate logit analyses for party and ideology because the two factors were too highly correlated to allow their simultaneous inclusion.

The need for cognition variables is analogous to the need for the variables just described. However, this set of variables is hypothesized to have a different coefficient. Petty and Cacioppo (1986), among others, argue that a person's need for cognition affects whether he or she will take the central or peripheral route to persuasion. To evaluate the claim that need for cognition affects cue persuasiveness, I use the Multi-Investigator Study's four "need for cognition" questions. These questions began with the query "How well would you say that the following statement describes you?" where the four statements in question were "I only think as hard as I have to;" "It's enough for me that something gets the job done; I don't care how or why it works"; "I really

enjoy a task that involves coming up with new solutions to problems"; and "I like tasks that require only a little thought once I've learned them." I follow the standard practice in need for cognition studies and use the answers to the four questions to form a need for cognition scale. The scale ranges from 0 to 4, with 4 implying highest in need for cognition and 0 implying lowest in need for cognition. Need for cognition–based explanations imply that as need for cognition increases, the effect of the cue will decrease. That is, high need for cognition will drive the coefficient to 0. Therefore, it implies that *Supports/Need for Cognition* will have a negative sign and *Opposes/Need for Cognition* will have a positive sign.

I also include a variable representing the respondent's prior involvement with prison spending. The "prior involvement" hypothesis that "Source credibility has more impact under low than under high involvement" is common in persuasion research (e.g., Chaiken 1980; Petty, Cacioppo, and Goldman 1981). To measure prior involvement, I used responses to the question "How often would you say that you had thought about this issue before today – often, sometimes, rarely, or never?" In an analysis analogous to that of Table 2.7, the prior involvement prediction fared poorly, while the theorem's prediction performed well regardless of the subject's prior involvement with prison spending. Prior involvement–based explanations imply that the variable *Supports/Low Involvement* will have a positive coefficient and that *Opposes/Low Involvement* will have a high coefficient.

The other independent variables represent additional factors that could affect respondent attitudes on prison spending. Each is drawn from standard survey questions. These factors are the respondent's partisanship, ideology, education, race, and age. *Democrat* = 1 for self-identified Democrats and 0 otherwise. I code *Republican*, *Conservative*, *Liberal*, *Moderate*, and *African American* analogously. *Age* is based on the numerical response to the question "How old were you on your last birthday?" *Education* is based on the response to the question "What is the highest grade or year of school you completed?"

In the analysis, *Opposes/Low Involvement* and *Supports Thermometer* were relatively good performers. Both were large, the former statistically significant, and both had the signs predicted by the alternative explanations. By contrast, the coefficients of *Supports/Low Involvement* and *Opposes Thermometer* had the wrong sign and were not significant. Also of interest is the fact that African Americans were significantly less likely to support prison spending and older people were significantly more likely to support it, all else constant. All other variables performed far worse. While *Age* also had a large coefficient, note that the underlying range of *Age* is eighteen to ninety-six years. Therefore, an age

increase of seventy-eight years was required to get the large effect seen in Table 2.7.

References

Aristotle. 1954. *Rhetoric.* New York: Modern Library.

Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: a Study of Opinion Formation in a Presidential Campaign.* Chicago: University of Chicago Press.

Brady, Henry E., and Paul M. Sniderman. 1985. "Attitude Attribution: A Group Basis for Political Reasoning." *American Political Science Review* 79: 1061–78.

Calvert, Randall L. 1985. "The Value of Biased Information: A Rational Choice Model of Political Advice." *Journal of Politics* 47:530–55.

Chaiken, Shelly. 1980. "Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion." *Journal of Personality and Social Psychology* 39:752–66.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In David E. Apter, ed., *Ideology and Discontent.* New York: Free Press.

Crawford, Vincent, and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50:1431–51.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters.* New Haven, CT: Yale University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper & Row.

Eagly, Alice H., and Shelly Chaiken. 1993. *The Psychology of Attitudes.* Fort Worth, TX: Harcourt Brace Jovanovich.

Farrell, Joseph, and Robert Gibbons. 1989. "Cheap Talk with Two Audiences." *American Economic Review* 79:1214–23.

Franzoi, Stephen L. 1996. *Social Psychology.* Madison, WI: Brown & Benchmark.

Fudenberg, Drew, and Jean Tirole. 1991. "Perfect Bayesian Equilbrium and Sequential Equilibrium." *Journal of Economic Theory* 53:236–60.

Harsanyi, John. 1967. "Games with Incomplete Information Played by 'Bayesian' Players, I: The Basic Model." *Management Science* 14:159–82.

1968a. "Games with Incomplete Information Played by 'Bayesian' Players, II: Bayesian Equilibrium Points." *Management Science* 14:320–34.

1968b. "Games with Incomplete Information Played by 'Bayesian' Players, III: The Basic Probability Distribution of the Game." *Management Science* 14:486–502.

Hovland, Carl I., Irving L. Janis, and Harold H. Kelley. 1953. *Communication and Persuasion: Psychological Studies of Opinion Change.* New Haven, CT: Yale University Press.

Klapper, Joseph T. 1960. *The Effects of Mass Communication.* New York: Free Press.

Kuklinski, James H., and Norman L. Hurley. 1994. "On Hearing and Interpreting Political Messages: A Cautionary Tale of Citizen Cue-Taking." *Journal of Politics* 56:729–51.

Kuklinski, James H., Daniel S. Metlay, and W. D. May. 1982. "Citizen Knowledge and Choice on the Complex Issue of Nuclear Energy." *American Journal of Political Science* 26:615–42.

Lupia, Arthur. 1994. "Shortcuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88:63–76.

Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* New York: Cambridge University Press.

McGuire, William J. 1985. "Attitudes and Attitude Change." In Gardner Lindzey and Elliot Aronson, eds., *Handbook of Social Psychology*, 3rd ed. New York: Random House.

McKelvey, Richard D., and Peter C. Ordeshook. 1985. "Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources." *Journal of Economic Theory* 36:55–85.

O'Keefe, Daniel J. 1990. *Persuasion: Theory and Research*. Newbury Park, CA: Sage Publications.

Petty, Richard E., and John T. Cacioppo. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Petty, Richard E., John T. Cacioppo, and R. Goldman. 1981. "Personal Involvement as a Determinant of Argument-Based Persuasion." *Journal of Personality and Social Psychology* 41:847–55.

Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Popkin, Samuel L., John W. Gorman, Charles Phillips, and Jeffrey A. Smith. 1976. "Comment: What Have You Done for Me Lately? Toward an Investment Theory of Voting." *American Political Science Review* 70:779–805.

Sniderman, Paul M., Richard A. Brody, and Philip E. Tetlock. 1991. *Reasoning and Choice: Explorations in Political Psychology*. Cambridge: Cambridge University Press.

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87:355–74.

Tetlock, Philip E. 1993. "Cognitive Structural Analysis of Political Rhetoric: Methodological and Theoretical Issues." In Shanto Iyengar and William J. McGuire, eds., *Explorations in Political Psychology*. Durham, NC: Duke University Press.