Public Choice **106:** 183–189, 2001.

A response to Austen-Smith

In a previous issue of this journal (October 1999), David Austen-Smith (hereafter, A-S) reviewed *The Democratic Dilemma* (Lupia and McCubbins, 1998). His review focused on the book's contribution to and treatment of formal theory. While A-S is highly regarded on such matters, his review contains errors. We have been invited to offer a response. In it, we focus primarily on error correction.

We begin, however, with a brief description of the book. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* starts by recognizing that many citizens lack information about politics. American voters, for example, are notorious for giving incorrect answers to questions such as "Who is your member of the House of Representatives?" For many observers, such evidence implies that citizens make inferior decisions. For other analysts, limited information does not imply inferior choice.

The ongoing debate about the political implications of limited information motivates *The Democratic Dilemma*. Prior to our collaboration in *The Democratic Dilemma*, each of us was among the growing group of scholars who argued that citizens can adapt to their limited information. Yet, we had questions about the conditions under which citizens could adapt successfully. While reading studies of choice throughout the social sciences, we found many examples of people finding quick and effective cues about the consequences of their actions (e.g., traffic lights, brand names, and reputations). We learned that this short-cut mode of reasoning is not something that people turn on and off – it is a fundamental characteristic of how people reason (e.g., Churchland and Sejnowski, 1992). A consequence of this fact is that a person's inability to answer common survey questions does not imply that they cast inferior votes. Such an inference is accurate only if no available short cuts allow that person to emulate the same decision he or she would make given different information. Of course, short cuts are no panacea. If used incorrectly, reliance on short cuts can lead to grave errors.

The purpose of *The Democratic Dilemma* is to answer questions about the conditions under which people who use short cuts can make the same decisions they would have made if better informed. In it, we base our answer on a formal theory and we use experiments to test the theory's predictions. The experiments, over forty of them involving more than 2200 subjects, reveal that the theory – which makes concrete statements about how institutional and other factors affect what and who people believe – holds up well. They highlight conditions under which a person's prior knowledge interacts with institutional factors to affect their competence. In the end, *The Democratic Dilemma* provides a means for better distinguishing the citizens who use short cuts as effective substitutes for more detailed political information from those who do not.

Now we turn to A-S's review. The review is narrowly focused. Over half of it is devoted to parts of one of the book's ten chapters. The chapter in question, Chapter 3, is theoretical. A-S's review of the chapter does not raise questions such as "Did the theorists model the situation correctly?" and "Did they do their math correctly?" Indeed, A-S finds no logical errors in our proofs and suggests no ways in which we should have modeled things differently. Instead, he questions only our treatment of existing theory and questions whether our book differs from work such as his own. While stating his case, he commits errors and makes misrepresentations.

Consider, for example, A-S's claim that our signaling models are not unique. Working in his favor is the fact that it is easy to prove that a model is not unique – just find the *same* model elsewhere. Finding a *similar* model is not sufficient, and yet this is the mode of A-S's critique.

Of course, our signaling models are similar to many others, as are his own. They cannot help but be, since the narrow range of equilibrium concepts, utility functions, and event sequences used in such models make all of them similar to each other. However, resemblance is insufficient to prove non-uniqueness. To prove that an argument is not unique, it is necessary to produce another that has a logically equivalent set of premises and a logically equivalent set of conclusions. A-S does not do this. In fact, as we will now show, in the one case where he attempts to demonstrate non-uniqueness, he accomplishes just the opposite.

At the end of the review's most sustained demonstration, A-S concludes that "To all intents and purposes, therefore, [the] model is strategically and consequentially identical to Sobel's 1985 stage-game." This conclusion is false, as A-S himself proves.

A quick comparison of our "Basic Model" and the model in Sobel (1985) to which A-S refers reveals several differences. In both models, two possible states of the world determine player payoffs. In Sobel, these two states are equally likely (i.e., $b = .5$). In our model, b can take on any value in $[0, 1]$. In Sobel (1985), the sender knows the true state; in our model, he may not. Moreover, in our model, the receiver can be uncertain about what the sender knows. We define as $k \in [0, 1]$ the receiver's prior belief about the probability that the sender knows the true state, $1 - k$ as the receiver's prior probability the sender has no private information about the true state, and $n_k \in \{0, 1\}$ is the realization of k (i.e., the sender's true knowledge). In Sobel's model, by contrast, the receiver knows that the sender is perfectly informed (i.e., $n_k = k = 1$).

There are other differences. The two models have different sets of premises, distinct statements of equilibria and distinct behavioral implications. In what follows, however, we focus only on the differences listed above, as they are the differences that A-S mentions in his critique.

A-S refers to the differences he lists as "substantively irrelevant," "purely for expository convenience," and as "only modifying the boundary conditions defining the equilibrium set of the game." If he is correct, then two things about his argument are ironic, at best. First, it takes A-S two full paragraphs to describe the differences between the two models. In these paragraphs, he argues that relaxing Sobel's assumptions about k and b causes changes in the equilibrium boundary conditions and alters the curvature of the set of informative equilibria. The irony here is that A-S shows that *relaxing Sobel's assumptions leads to a logically distinct conclusion* about sender and receiver behavior; for if the models were "strategically and consequentially identical," then the boundary and curvature changes that A-S describes could not occur.[1]

Second, to make his argument A-S must choose very precise ways of relaxing k and b. The irony here is that the relaxations he chooses are *precisely the premises that distinguish our model from Sobel's* – premises that he earlier refers to as "purely for expository convenience" and "irrelevant." Unstated is the fact that other means of relaxing Sobel's model do not substantiate A-S's claim. A-S seems unaware of this inconsistency; for at the end of his argument, he concludes that the two models are identical. However, he has proven the opposite, showing that the models' conclusions are logically distinct and that imposing our premises is the reason for the difference.

Other inaccuracies are present in A-S's claims about our treatment of the literature. In his opening paragraph, for example, A-S claims that our depiction of "the canonic cheap talk signaling model and the existing literature are demonstrably false." However, A-S never offers the demonstrations. Instead, he critiques our literature treatments only after substituting his own questionable definitions of key terms for the ones that we use in the book. Such moves are standard rhetorical sleights-of-hand, but they do not substantiate his assertions.

For example, A-S asserts that "*It is also odd to read LM's claim that in the standard cheap talk model, people cannot deceive each other in equilibrium*." However, our claim is true. We follow social science citation patterns and define Crawford and Sobel (1982) as the standard cheap talk model (Lupia and McCubbins, 1998: 46). We are also consistent throughout the book: we refer to Crawford and Sobel as the seminal cheap talk model on pages 45 and 50, we never apply the "standard" or "seminal" cheap talk model label to any other source, and nowhere in the text do we use the term standard cheap talk *models* (so as to imply that there is more than one seminal model). A-S, by contrast, makes his argument by substituting his own definition for ours without alerting the reader.[2] Of course, A-S is correct to say that deception is possible in other models (nowhere do we say otherwise). However, he is incorrect to say the same of Crawford and Sobel (1982). There, the sender and receiver know each other's preferences and informative communication occurs in the equilibrium they report only if the sender and receiver have spatially proximate ideal points (see their Corollary 1). Indeed, that paper contains no equilibrium in which the sender makes a false statement that the receiver treats as true.

In the next paragraph, A-S asserts that "*Given the large literature on cheap talk and costly signaling games, one wonders why LM continue to insist that* lying *in equilibrium is impossible in* these games (emphasis added)." But the last phrase of this statement is erroneous – we make no such claim about the equilibrium status of lying. Instead, we describe the equilibrium status of deception (e.g., see our quote in the previous item). Throughout the book, we define deception as, "the process by which the testimony we hear reduces our ability to predict accurately the consequences of our actions. For example, we are deceived when someone lies to us and we believe that individual" (p. 8, emphasis in original). So, while lying can be part of deception, lying is not sufficient for deception. This distinction is common in dictionaries – lying is an act by a speaker, deception is one possible result of an interaction between a speaker and receiver. We are clear about the difference throughout the book. This clarity is particularly manifest in the book's experimental chapters (pp. 97–201), where, for example, the tables in Chapters 7 and 8 provide separate measures for the frequencies of deception and lying.

In other cases, A-S simply substitutes conjecture for fact. For example, A-S uses a long passage to criticize us for misusing a concept *that never appears in the book*. A-S argues that we "describe Farrell's concept of a neologism incorrectly" and that we apply Farrell's neologism proof equilibrium refinement concept incorrectly. But *Webster's* defines neologism as "A new word, phrase, or expression or a new meaning for an existing word." We use the term in the last sense, as does Farrell (1993), which we point out. Farrell then went on to develop an equilibrium refinement concept around the idea of a neologism, a concept that we neither adopt in the book nor even mention. A-S, however, argues that when we used the word neologism, we must have meant to use Farrell's concept (to rule out pooling equilibria) and applied it incorrectly. A-S's claim is entirely speculative – and false.

A-S makes several other assertions for which he offers no evidence. For example, A-S makes the following claim about our extensions of the Basic Model:

> The main results here are that typically more information can be credibly revealed and the probability of successful dissembling by the sender diminishes; *neither result is novel by any stretch of the imagination (see any recent text book on game theory*).

Of course, it would be difficult for any formal theorist to reject such a broad and imprecise characterization of their model's dynamics. For if we allow critics to ignore the mathematics and related details, then such claims are surely correct. Move to a finer level of detail, however, and the outcome is different. In particular, and over a trial of many years, A-S has yet to produce the "recent text book on game theory" whose findings he claims we replicate. Is it Kreps? Binmore? Rasmussen? Myerson? Fudenberg and Tirole? Certainly, each of these important texts cover games of incomplete information, but our results – while sharing broad themes with previous work – are distinct. Indeed, we have searched and have found no textbook on game theory that states as we do the conditions under which the threat of verification, the introduction of penalties for lying, and costly effort affect the conditions under which a speaker speaks truthfully and the conditions under which receivers find speakers to be credible. To the extent that many of our results, taken one at a time, are similar to others in the literature, these "others" are dispersed over a wide set of papers and models, many of which we cite. By contrast, we produce many results about cheap talk models, signaling games, and how institutions affect communication from a single framework (a framework developed for the purpose of clarifying the political consequences of limited information). And while Chapter 3 presents these results in an accessible manner, their descriptions in the appendix are precise and unique. Our many experiments also attest to our results' distinct properties. In these experiments, we do more than just test vague statements about the effect of penalties for lying, verification, and costly effort. Instead, we generate and test precise explanations of the tradeoffs that these forces engender – explanations that explain our experiments' very unlikely strings of observed behaviors in ways that other models do not.

Elsewhere, A-S asserts that we are incorrect to claim that "many of [economists' and political scientists'] definitions confound rationality and omniscience, assuming that rational actors must be limitless calculators." While A-S offers no evidence for his assertion, there is plenty of evidence supporting our claim. Consider, for example, Simon's (1995) definition, "A decision is only rational if it is supported by the best reasons and achieves the best possible outcome in terms of all the goals." Or Kreps' (1990) definition of irrationality, "Such an unaware monopolist is not fully rational, at least in the sense that she doesn't quite understand the full implications of her actions.... [We] call irrational any behavior in the game that flies in the face of the player's own 'best interests' as determined by the player's payoffs." While not everyone ties rationality to omniscience, it is easy to find numerous examples of scholars who do – many times to great effect.

A-S then claims that: "The *entirely standard* use of rationality in contemporary economics and political science requires individuals to have consistent (i.e., at least acyclic) preferences over consequences and to choose more preferred over less preferred feasible alternatives; *in particular, rationality imposes no constraints on how individuals respond to or evaluate uncertainty per se*" (emphasis added). We have no quarrel with the first part of his assertion. But to say that the entirely standard definition of rationality "imposes no constraints on how individuals respond to or evaluate uncertainty" is tone deaf to one of the most critical debates in contemporary economic theory and political science. Certainly, Simon, Kreps, Aumann, and Rubenstein, among others, would disagree.

Looking beyond all the rhetoric, A-S focuses his critique on our Chapter 3 signaling model and the associated literature review. His argument boils down to a personal judgment about where to draw the line between new insights and rehash. As is not uncommon in scientific debate, A-S wants us to draw the line after his research, but before that of those who challenge him. Is he correct?

As A-S himself demonstrates, the distinct aspects of our model affect the conditions under which speakers tell truths and the conditions under which receivers believe what they hear. In the book's lengthy experimental section, we use these differences as the basis of many experiments. In so doing, we provide evidence that these differences are far from trivial in substance – our model predicts the very unlikely behavior strings observed in Chapters 7 and 8 better than do rival explanations. These strings involve precise patterns of behavior regarding the tradeoffs that people make when choosing to lie, tell the truth, believe others, and ignore others. Our predictive success testifies to the unique substantive implications of our modeling choices.

Therefore, whether or not our book meets A-S's self-selected criterion for theoretical uniqueness is a question whose answer has little apparent scientific value. The point of *The Democratic Dilemma* remains to use formal theories, experiments, and lessons from several academic disciplines to help others understand the conditions under which voters, legislators, and jurors who appear uninformed can learn what they need to know.

Notes

1. A-S treats his argument as if it derives from comparative statics on the Sobel model. However, Sobel explicitly defines both k and b as constants. Since varying parameters are required for a comparative statics analysis, since A-S's analysis depends entirely on variations in k and b, and since it is impossible to vary a constant, it is incorrect to interpret A-S's effort as a comparative statics analysis on the Sobel model. His analysis, however, can be interpreted as comparative statics on our model.
2. In his description of "the standard model," A-S substitutes Sobel (1985) and Austen-Smith (1992) for Crawford and Sobel (1982). However, recent citation patterns support our definition. Over the last four complete years (1994–1997), the years in which electronic versions of the index were available to us, the *Social Science Citation Index* reports 54 citations of Crawford and Sobel (1982), 11 citations of Sobel (1985), and 1 citation of Austen-Smith (1992).

References

Austen-Smith, D. (1992). Strategic models of talk in political decision making. *International Political Science Review* 13: 45–58.

Churchland, P.S. and Sejnowski, T.J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

Crawford, V.P. and Sobel, J. (1982). Strategic information transmission. *Econometrica* 50: 1431–1451.

Farrell, J. (1993). Meaning and credibility in cheap-talk games. *Games and Economic Behavior* 5: 514–531.

Kreps, D.M. (1990). *A course in microeconomic theory*. Princeton, NJ: Princeton University Press.

Lupia, A. and McCubbins, M.D. (1998). *The democratic dilemma: Can citizens learn what they need to know?* New York: Cambridge University Press.

Simon, H.A. (1995). Rationality in political behavior. *Political Psychology* 16: 45–61.

Sobel, J. (1985). A theory of credibility. *Review of Economic Studies* 52: 557–573.

ARTHUR LUPIA and MATHEW D. MCCUBBINS, Political Science, University of California, San Diego, La Jolla, CA 92093, USA