"Don't Know" Means "Don't Know": DK Responses and the Public's Level of Political Knowledge

Robert C. Luskin University of Texas at Austin
John G. Bullock Yale University

Does the public know much more about politics than conventionally thought? A number of studies have recently argued, on various grounds, that the "don't know" (DK) and incorrect responses to traditionally designed and scored survey knowledge items conceal a good deal of knowledge. This paper examines these claims, focusing on the prominent and influential argument that discouraging DKs would reveal a substantially more knowledgeable public. Using two experimental surveys with national random samples, we show that discouraging DKs does little to affect our picture of how much the public knows about politics. For closed-ended items, the increase in correct responses is large but mainly illusory. For open-ended items, it is genuine but minor. We close by examining the other recent evidence for a substantially more knowledgeable public, showing that it too holds little water.

Contemporary politics owes much of its character to the severely right-skewed distribution of political knowledge. The policy deficiency of many campaigns and most campaign coverage, the prevalence and effectiveness of sound bites, the formidable impact of candidate personalities and simple retrospections on votes, the fierce but transient rallying around the president in foreign crises, and, in admittedly debatable ways, the outcomes of many elections—all this and more would be different if the public had the same values and interests it does but the political knowledge of the average legislative aide, political commentator, or even mere reader of the political content of a serious daily newspaper. Or so many of us who study mass politics believe.

Many, but not all. Both the fact of widespread political ignorance and its consequences have been disputed—the latter, in recent times, much more than the former. A few years ago, Luskin (2002) described the "Panglossian side" in these debates as having retreated from "denial" to "extenuation"—from the position that most people are not really ignorant (Nie, Verba, and Petrocik 1979) to the position that they are but that their ignorance makes little difference to their policy and electoral preferences (Lupia and McCubbins 1998; Page and Shapiro 1992; Popkin 1991). It turns out he may have spoken too soon. Efforts at "extenuation" continue, but some research seems to be slipping back into "denial."

A number of recent studies have used untraditionally designed or scored knowledge items to claim that the public knows much more about politics than conventionally thought. The untraditional features include forestalling "don't know" (DK) responses (Mondak 1999, 2001; Mondak and Davis 2001); randomly rescoring DKs as correct (Mondak 1999); giving partially correct answers part credit (Mondak and Davis 2001 and, implicitly, Krosnick et al. 2008); abandoning open-ended items (Gibson and Caldeira 2009; Mondak 2001; Mondak and Anderson 2004); and offering monetary incentives for correct answers (and the opportunity to research them) (Prior and Lupia 2008).

We return to this longer story after homing in on one of its most prominent chapters: the contention that a great deal of knowledge lies hidden in DK responses. At least until recently, the conventional practice has been to permit or even encourage respondents to say DK and to take those who do so at their word, pooling DK responses with wrong answers. But perhaps many of those saying they don't know actually do know. Perhaps they are just too unconfident or impatient to give the right answer, even in response to closed-ended questions. Perhaps they are momentarily unable to recall the right answer in response to open-ended questions. The

The Journal of Politics, Vol. 73, No. 2, April 2011, Pp. 547–557 doi:10.1017/S0022381611000132

ISSN 0022-3816

result, on this view, is a chronic undercount of the percentage knowing the answer.

Thus Mondak (1999, 2001; Mondak and Anderson 2004; Mondak and Davis 2001) advocates changes in the traditional design and scoring of knowledge items. On the first count, he recommends discouraging DKs—urging respondents who think they don't know to guess, and then, if they still say DK, asking them again. On the second, he recommends randomly rescoring some DKs (to closed-ended items) as correct as a way of simulating the heightened guessing from discouraging DKs.

We have addressed the scoring recommendation elsewhere (Luskin and Bullock 2005); here we focus on design. The issue is particularly important because it is not something the secondary analyst has any choice about. The scoring can be fiddled with, but the questions are what they are. The American National Election Studies (ANES) adopted a version of Mondak's DK-discouraging format for a random third of its sample for four of its open-ended knowledge items in 2000 and has continued using it for some of its open-ended knowledge items ever since.

Our evidence comes from two survey experiments, both on national random samples. One, conducted through Time-Sharing Experiments in the Social Sciences (TESS), included DK-discouraging, -neutral, and -encouraging versions of open-ended knowledge items; the other, embedded in the 2000 ANES, included DK-discouraging and -neutral versions of open-ended ones. In both cases, discouraging DKs produces more correct answers, but, as we expect and show, the magnitude and legitimacy of the increase (the extent to which it actually reflects previously hidden knowledge) varies by format. For closed-ended items, the increase is large but consists mainly of lucky guesses; for open-ended items, it consists mainly of hidden knowledge but is small. In neither case, therefore, is our picture of how much the public knows about politics much affected.

Conceptual Preliminaries: Knowing, Not Knowing, and Partially Knowing

Most work on political knowledge implicitly takes "knowledge" as fundamentally binary: one either knows something, or one doesn't. Closed-ended knowledge items present little occasion to reconsider this assumption. The responses, assuming well drawn categories, are either correct or incorrect. But open-ended items, as we have recently been reminded (Gibson and Caldeira 2009; Krosnick et al. 2008), may elicit some nontrivial number of partially correct responses. Asked, in the vein of the ANES's open-ended items, to identify, say, Mitch McConnell, some people may describe him simply as a congressional leader or a U.S. Senator without indicating that he is the Minority Leader. The knowledge underlying such responses, if not merely partially retrieved, is partial. The respondent knows some but not all of the relevant information.

So what does it mean to know, not know, or partially know something? If knowledge is something like the accurate long-term memory representation of real-world objects and their attributes, any single object-attribute relation (like Mitch McConnell's being a U.S. Senator) may be regarded as either known or not known (accurately represented in long-term memory or not). But since given objects typically have multiple attributes, knowledge of multi-attribute objects like Mitch McConnell may be regarded as the proportion of his attributes one knows—and thus a matter of degree. In the present context, where the relevant knowledge is of McConnell as political leader, his attributes may need to be weighted by their centrality to his political role, with U.S. Senate Minority Leader, for example, counting for more than U.S. Senator. This, we suggest, is the light in which partially correct responses to open-ended knowledge items should be viewed. What makes them partially correct is that they refer to one or more relevant attributes; what makes them only partially correct is that they miss one or more of the most important.[1]

DK Responses to Open- versus Closed-Ended Items

The issue of partial knowledge will recur, but let us return to DKs. Here too the cases of open- and closed-ended items are quite different. In both cases, at least some DKs must conceal some knowledge, while others must be candid admissions of ignorance, and in both cases discouraging DKs may therefore coax some of the timid or impatient but knowledgeable into venturing the correct response and, at the same time, some of the genuinely ignorant into guessing. One difference lies in the ease of retrieval. The right answer may be known but temporarily inaccessible.[2] Another,

[1]This necessarily light sketch presumes a standard associative network model of semantic memory (as in Collins and Loftus 1975 or Anderson 1983).

[2]The "tip of the tongue" phenomenon (Brown and McNeill 1966).

we suspect bigger difference lies in the ease of guessing—and, given that one guesses, of guessing the right answer (''lucky guessing'').

On closed-ended items, guessing is a snap, lucky guessing a reasonable chance, and retrieval not a factor. Here the question is of balance—of the extent to which discouraging DKs exposes hidden knowledge more or less than it stimulates lucky guessing. Our money is on the lucky guessing, and by a wide margin. The task is simply to choose from a menu, and most respondents want to oblige the interviewer and/or avoid looking ignorant. Hence almost everyone who knows the answer can be expected to give it, which is to say that the vast majority of those saying ''don't know'' really don't know.[3]

On open-ended items, guessing is much harder and rarer, lucky guessing still harder and rarer, and retrieval uncertain. In this case, the great bulk of any additional correct answers from discouraging DKs figure to reflect genuine knowledge. The question here is of magnitude. Discouraging DKs must uncover some hidden knowledge, but how much? Past results suggest that it may not be much. The DK responses (to open- as to closed-ended items) come from the least knowledgeable respondents, as gauged by other, independent knowledge items (Luskin and Bullock 2005).[4] Here, too, therefore, ''don't know'' may preponderantly mean ''don't know.'' But let us see.

The TESS and ANES Experiments

Our data on closed-ended items come from a Time-Sharing Experiments in the Social Sciences (TESS) survey administered by telephone from October, 2004 through January, 2005. Interviewers from the Indiana University Center for Survey Research collected data from a national household sample of 1,507 respondents, with a response rate of 30.8%.

The experimental manipulation, applying to three closed-ended knowledge items, lay in randomly assigning respondents to one of three conditions, differing only in the encouragement versus discouragement of DK responses. The items ask whether the federal budget deficit has ''got smaller,'' ''stayed about the same,'' or ''got larger'' over the past four years (roughly George W. Bush's first term); whether ''the term of a United States Senator'' is two, four, six, or eight years; and whether the responsibility for ''nominat[ing] judges to the federal courts'' is ''the President's, the House of Representatives', the Senate's, [or] the Supreme Court's.'' We refer to these three items, for which the DK treatment is experimentally varied, as ''dependent.''

In the first, *DK-discouraging* condition, 893 respondents were told, ''If you aren't sure of the answer, we'd be grateful if you could just give your best guess,'' were given no explicit DK option, and were probed once for a substantive response if they nonetheless said DK. In the second, *DK-encouraging* condition, 321 respondents were told ''Many people have trouble answering questions like these. So if you can't think of the answer, don't worry about it. Just tell me, and we'll move on to the next one.''[5] They were then offered an explicit DK option for each question. In the third, *DK-neutral* condition, 293 respondents were simply asked the questions, with no preamble encouraging or discouraging DKs and neither any explicit DK response option nor any probing of initial DKs. The ANES used this last, DK-neutral approach for factual knowledge items until 2000. As we shall see, this condition turns out to be important for interpreting the effects of the other two.

Another seven knowledge items are ''independent,'' in the sense of being outside the experimental manipulation—asked of all respondents in exactly the same, DK-neutral form. These ask ''what job or political office is currently held by Dick Cheney,'' whether it is ''the President's, the Attorney General's, the Senate's, or the Supreme Court's'' responsibility ''to determine if a law is constitutional or not,'' ''what fraction or percentage of the U.S. Senate and House is required to override a presidential veto,'' whether the respondent's state could prohibit an adult woman from having an abortion during the first three months of her pregnancy, whether Bush was closer to the position that ''the government

[3]Rampant guessing on knowledge items is of a piece with cognitively empty responses to closed-ended attitude items. Many people claim familiarity with fictional historical figures, books, scientific terms, and all manner of other objects (Paulhus et al. 2003). Many claim to favor or oppose fictional legislation (Bishop 2005), and providing such pseudo-opinions is negatively correlated with actual knowledge of real objects in the same domain (Sturgis and Smith 2009).

[4]Respondents giving DK responses to the DK-neutral open-ended items in the 1988 ANES average only 26.5% correct on an independent closed-ended knowledge measure, compared to 41.2% of those answering incorrectly and 50.4% of those answering correctly. Note that even the DK-responders' achieve a distinctly nonzero mean on the independent knowledge index because even those who know absolutely nothing about the item to which they are saying DK may nonetheless know the answers to—or guess luckily on—other items. We present the corresponding figures for the open-ended knowledge items in the 2000 ANES in n. 17 below.

[5]Similar to the preface recommended by Delli Carpini and Keeter (1996).

should pay for all medical and hospital expenses for everyone'' or the position that ''individuals or private insurance companies should pay for all medical and hospital expenses,'' which of the same two positions Kerry was closer to,[6] and whether the ''federal government currently spend[s] the least'' on ''Medicare, foreign aid, national defense, or education.''[7]

Table 1 shows the percentages answering correctly, incorrectly, and DK. With the exceptions of the deficit-trend dependent item and the Cheney independent item, which were answered correctly by more than 80% and more than 90%, the percentages answering correctly are in the range familiar from other studies (Delli Carpini and Keeter 1996). But the most important point here is the experimental manipulation's effect. The DK-encouraging condition produced the most DKs (averaging 21.8%), the DK-discouraging condition the fewest (averaging 1.4%). The DK-neutral condition was in-between but far closer to the DK-discouraging one (averaging 4.2%). All these differences are pairwise significant at $p < .05$ by a one-tailed test.

The asymmetry, however, is important. The DK-encouraging condition increased DKs far more than the DK-discouraging condition decreased them.[8] Urged neither to admit ignorance nor to venture an answer despite feeling uncertain, almost everyone chooses a substantive answer. These are multiple-choice questions, after all, so why not? In practice, therefore, DK-discouragement is not very different from DK-neutrality. Mondak (2001) and Sturgis et al. (2008), contrasting DK-discouraging and DK-encouraging conditions, find a big difference between them—as do we—but the addition here of the DK-neutral condition shows that it is the encouraging, not the discouraging, that is responsible.

Our data on open-ended items come from a national random sample experiment within the 2000 ANES (also analyzed by Bennett 2001 and Mondak and Davis 2001). Here there were two conditions, DK-discouraging and DK-neutral. The dependent knowledge items asked respondents to identify the ''job or political office[s]'' held by Tony Blair, Trent Lott, William Rehnquist, and Janet Reno. The DK-discouragement consisted only of a probe of initial DK responses; there was no DK-discouraging preface. For additional design details, we refer readers to Mondak and Davis (2001) and the ANES codebook.

TABLE 1 Responses to Closed-Ended Knowledge Items

Dependent Items	Correct	Incorrect	DK
Senate Term			
DK-encouraging	31.2%	47.0%	21.8%
DK-neutral	43.0	53.6	3.4
DK-discouraging	41.9	57.6	0.6
Nominating Judges			
DK-encouraging	44.9	20.2	34.9
DK-neutral	54.8	38.0	7.2
DK-discouraging	54.0	43.0	3.0
Deficit Trend			
DK-encouraging	81.6	9.7	8.7
DK-neutral	88.1	9.9	2.0
DK-discouraging	86.1	13.3	0.6
Mean			
DK-encouraging	52.5	25.7	21.8
DK-neutral	62.0	33.8	4.2
DK-discouraging	60.7	38.0	1.4
Independent Items			
Cheney name recognition	94.5	5.5	0.0
Determining constitutionality	77.3	19.7	3.0
Overriding veto	37.8	62.2	0.0
Abortion legal, first three months	70.5	22.8	6.7
Bush health care	50.8	45.0	4.2
Kerry health care	41.3	52.3	6.5
Budget allocations	9.8	87.0	3.2
Mean	54.0	42.6	3.4

As has recently emerged, the Rehnquist item elicited a sizable number of partially correct responses, identifying him, for example, as ''a judge'' or a ''member of the Supreme Court'' (Gibson and Caldeira 2009), and that discovery has also served as a reminder that open-ended items generally elicit at least some partially correct responses, most of which the ANES has traditionally coded as incorrect (Krosnick et al. 2008). There may therefore be some knowledge hidden in the ''incorrect'' responses, in addition to whatever is hidden in the DKs—and the effect of discouraging DKs may seem smaller than it is, to the extent that some of the additional substantive responses it produces are also partially correct but coded as incorrect.

The ANES is now embarked on a project to recode the responses to past open-ended knowledge questions more finely. In the meantime, we take two tacks. The first is to take the ANES's existing codings at face value, on the grounds that if we make only the

[6]Being ''right in the middle'' was also a response option for the Bush and Kerry health care placements and was coded as wrong (in keeping with the results in Luskin and Bullock 2005).

[7]These all followed, and thus could not affect, the dependent knowledge items.

[8]The DK-neutral *independent* knowledge items average only 3.4% DKs, still slightly closer to the 1.4% in the DK-discouraging condition.

binary correct-incorrect distinction, the vast majority of the partially correct responses are just that, not quite correct. Of course a better solution, once all the original transcripts have been suitably recoded, would be to give these responses the part credit they deserve. Pending that, our second tack is to use side evidence and deliberately too-generous assumptions to get a sense of the largest difference an appropriate part-credit scoring could possibly make.

Descriptive Validity

The DK treatment may affect both *description* (of the public's level of knowledge) and *correlation* (between measured and actual knowledge). Here we focus on description. The consequences for correlation, described in the online Appendix A, turn out to be similar.[9]

Closed-Ended Items

Discouraging DKs to closed-ended items does produce a higher percentage of correct answers, at least in comparison to encouraging them. A glance back at Table 1 shows that the mean percentage of dependent knowledge items answered correctly is 8.2% higher in the DK-discouraging than in the DK-encouraging condition. The DK-neutral condition, surprisingly, averages still slightly higher, but the reason seems to be simply that the gods of random assignment have given us a DK-neutral sample that is slightly more knowledgeable than the DK-discouraging one. The DK-neutral sample's mean percentage correct is 1.3% higher than the DK-discouraging one's on the dependent knowledge index but also 1.8% higher on the independent knowledge index, consisting of items not subject to the experimental manipulation. So DK-discouragement does not actually reveal *less* knowledge than DK-neutrality—but neither does it reveal detectibly more. Again the treatment that makes a difference is DK-encouragement.

But does encouraging DKs or discouraging them (or, almost equivalently, doing neither) make for more accurate description? The answer depends on the extent to which discouraging DKs actually reveals hidden knowledge, versus merely giving that impression, on account of lucky guessing. Let the percentages responding correctly, incorrectly, and DK in the three conditions be C_e, C_n, and C_d; I_e, I_n, and I_d; and D_e, D_n, and D_d. If all but D_d of the DK-responders in the DK-encouraging and DK-neutral conditions were instead to guess blindly, thus equalizing the percentages of DKs at D_d, the expected increases in the percentages answering correctly by guessing luckily would be $G_n \equiv (D_n - D_d)/J$ and $G_e \equiv (D_e - D_d)/J$, where J is the number of response categories. The expected percentages correct under this scenario, $C'_n = C_n + G_n$ and $C'_e = C_e + G_e$, can then be compared to the percentage actually answering correctly in the DK-discouraging condition, C_d. The differences $C_d - C'_n$ and $C_d - C'_e$ estimate the extent to which C_d exceeds C_n and C_e because of hidden knowledge, as opposed to lucky guessing. Since $C_d - C_n = (C_d - C'_n) + G_n$ and $C_d - C_e = (C_d - C'_e) + G_e$, the differences between the percentages answering correctly in the DK-discouraging and the other two conditions can also be partitioned into the gains from hidden knowledge ($C_d - C'_n$and $C_d - C'_e$) versus lucky guessing (G_n and G_e).

The results, in Table 2, show blind guessing to be a pretty good approximation of what DK-responders do when compelled to give substantive answers. The last four columns give the punch lines. Discouraging the DKs in the DK-neutral condition would apparently uncover no appreciable hidden knowledge. As previously noted, there is no increase here to partition. Discouraging the DKs in the DK-encouraging condition would reveal some hidden knowledge (5.4%) on the Senate item but very little (1.1% and 1.6%) on the other two items. The average is only 2.4%.[10]

But this is still only half the story. Discouraging DKs not only reveals hidden knowledge but increases lucky guessing. Thus the most important lesson from Table 2 lies in the comparison of the $C_d - C'_e$and G_e columns, which show that the increase in lucky guessing is roughly as great as the uncovering of hidden knowledge for the Senate item and far greater for the other two items. On average, discouraging DKs does nearly twice as much harm (5.7%) as good (2.8%). On closed-ended items, we do best to take those saying DK at their word.

[9]They need not have been. To illustrate, let two indicators of the conceptual variable x^* be $x_{i1} = a_1 + x_i^* + u_{i1}$ and $x_{i2} = a_2 + x_i^* + u_{i2,}$ where the error terms u_1 and u_2 both have zero means and are independent of x^*. If a_1 exceeds a_2 in absolute value but the variance of u_2 exceeds that of u_1, the expected value of x_1 will be further from that of x^*, yet the correlation between x_1 and x^* will be higher than that between x_2 and x^*.

[10]For similar results on a quota sample of the United Kingdom, see Sturgis, Allum, and Smith (2008, 96–97). Sanchez and Morchio (1992) also report similar results for the effectively closed-ended items about party control of the House and Senate in the 1984 ANES.

TABLE 2 **Hidden Knowledge vs. Lucky Guessing on Closed-Ended Items**

	Percentage Correct			With Blind Guessing		Improvement over Blind Guessing		Increase in Lucky Guessing	
	DK-encouraging (C_e)	DK-neutral (C_n)	DK-discouraging (C_d)	From DK-encouraging (C'_e)	From DK-neutral (C'_n)	From DK-encouraging (C_d - C'_e)	From DK-neutral (C_d - C'_n)	From DK-encouraging (G_e)	From DK-neutral (G_n)
Senate term	31.2	43.0	41.9	36.5	43.7	5.4	−1.8	5.3	0.7
Nominating judges	44.9	54.8	54.0	52.8	55.8	1.1	−1.9	8.0	1.0
Deficit trend	81.6	88.1	86.1	84.3	88.6	1.8	−2.4	2.7	0.5
Mean	52.5	62.0	60.7	57.9	62.7	2.8	−2.0	5.3	0.8

Note: For none of the three items is the improvement over blind guessing is significantly different from 0 (at the .05 level., two-tailed). Only the Senate item comes close (p = .09), For two of the three, by contrast, the increase in lucky guessing is significantly different from 0 (again at the .05 level., two-tailed), and the third (the deficit trend item) also comes close (p also = .09).

Open-Ended Items

In the open-ended case, where the question is simply how much hidden knowledge is revealed, Table 3 shows that probing open-ended DKs converts very few to substantive responses, and still fewer to *correct* substantive responses (as Bennett 2001 also observes). The percentage of DKs remaining DKs exceeds 80% for every political figure but Reno and is 62% even for Reno. Of those venturing substantive answers, moreover, the great majority—66.7% for Blair, 95.6% for Lott, 94.9% for Rehnquist, and 61.2% for Reno—answer incorrectly.[11] Overall, then, the percentage of the probed DKs giving a substantive response *and* getting it right was only 14.6% for Reno, 6.0% for Blair, and under 1.0% for both Lott and Rehnquist. Across the four items, only 36, or 4.4%, of the 809 probed DKs became correct answers. Even these small increments, moreover, are percentages only of the initial DKs. The impact on the percentage of the whole sample answering correctly is still smaller. The largest increase is only 2.9% (for Rehnquist), the average only 1.7%.

In a couple of ways, to be sure, these numbers are a shade too low. The first lies in the interviewing. The ANES interviewers did not actually probe quite all the DKs in the DK-discouraging condition, and some of the un-probed DKs might have become correct answers. The second lies in the coding. Some of the second-try substantive responses coded as incorrect may be partially correct.

But let us bound the effects of these distortions by making some deliberately generous assumptions. In the first case, we assume that the un-probed DKs would have become correct answers just as often as the probed (whereas in fact the interviewers often neglected to probe because they could see it would be futile).[12] In the second case, we assume all partially correct responses to be worth the .725 estimated by Mondak and Davis (2001) (whereas many—Rehnquist as merely "a judge," e.g.—are arguably worth less); that *every* response coded as incorrect is partially correct (whereas many, presumably most, are not); and that no partially correct response is oppositely miscoded as *correct* (whereas some, as we know from Krosnick et al. 2008, clearly are).[13] Combined, the resulting adjustments increase the

[11] Computed from Table 3.

[12] Personal communication from Patricia Luevano, Senior Systems Analyst at the ANES.

[13] Examples include the "English pres or whatever" and the "Head of England" for Tony Blair—the first always coded as correct, the second sometimes so. See Krosnick et al. (2008).

TABLE 3 DK-Discouraging versus DK-Neutral Responses to Open-Ended Items

	DK-Neutral (n = 1025)			DK-Discouraging (n = 530)			Probed DK-Discouraged DKs		
	Correct	Incorrect	DK	Correct	Incorrect	DK	Correct	Incorrect	DK
Tony Blair	33.9 (347)	7.1 (73)	59.0 (605)	35.7 (189)	13.0 (69)	51.3 (272)	6.0 (13)	11.9 (26)	82.1 (179)
Trent Lott	8.3 (85)	25.4 (260)	66.3 (680)	9.4 (50)	31.5 (167)	59.0 (313)	0.8 (2)	17.8 (43)	81.3 (196)
William Rehnquist	9.6 (98)	29.0 (297)	61.5 (630)	12.5 (66)	35.7 (189)	51.9 (275)	0.9 (2)	16.8 (37)	82.3 (181)
Janet Reno	54.7 (561)	16.1 (165)	29.2 (299)	55.8 (296)	20.4 (108)	23.8 (126)	14.6 (19)	23.1 (30)	62.3 (81)
Mean	26.7	19.4	54.0	28.3	25.1	46.5	4.4	16.8	78.7

Note: Entries are percentages of the relevant sample, followed by the raw numbers of respondents in parentheses.

estimated effect of discouraging DKs to an average of 5.9%. The online Appendices B and C trace the calculations in greater detail. But again the generous assumptions make this a ceiling—well above the actual effect, which might more plausibly be something like half-way back toward the raw mean difference of 1.7%, say around 3.8% (strikingly close to the 3.7% estimate derivable from Mondak and Davis's 2001 Tallahassee data, which do distinguish partially correct responses).[14]

Distributional Implications: How Knowledgeable a Public?

Discouraging DKs does paint a more comforting picture of the public's knowledge of politics—but, as the foregoing shows, only slightly so in the open-ended case and spuriously so in the closed-ended one. Anyone searching for large caches of hidden knowledge, it appears, should look elsewhere.

Consider first our closed-ended items. With DKs discouraged, the mean percentage answering correctly is 60.7%, but that includes a lot of lucky guessing. With DKs encouraged, the mean drops to 52.5%, although that misses a smaller amount (2.8%) of hidden knowledge. Adding that last back in edges the mean back up to 55.3%. But that is still too high. There is a lot of guessing, some of it lucky, even when DKs are encouraged. A global correction for guessing—applied to all substantive responses, not just the increment from discouraging DKs—drops the mean DK-discouraging and -encouraging percentages correct to 47.1% and 40.7%, a reasonable estimated range for the mean percentage actually knowing the answers.[15]

The percentages answering our open-ended items correctly run still lower, as the percentages answering open-ended items correctly characteristically do (Delli Carpini and Keeter 1996, Appendix 2; Mondak 2001). Even with DKs discouraged, the mean is only 28.4%. And even our deliberately overshot adjustments for inconsistent probing and partially correct second-try responses increase it to only 32.6%. More realistic adjustments would knock it back again, by at least a percent or two. The surviving DKs, moreover, are unlikely to be hiding much additional knowledge. In both conditions, those initially responding DK score decidedly lower on the independent knowledge index described in the online Appendix A than those responding incorrectly, which suggests that that there was not much knowledge in the initial DKs to begin with.[16] And, as we have just seen, probing those initial DKs only rarely produces additional correct answers, which suggests that there is not much more to extract.

We may also briefly note the existence of parallel consequences for the *conditional* distributions of knowledge with respect to gender (men versus women), education (college or more versus no college), and race (white versus nonwhite). For

[14]Based on their Table 2, the mean percentage of DK-discouraged respondents giving partially correct responses to open-ended ANES-type items about Al Gore, Newt Gingrich, Trent Lott, and William Rehnquist is 5.1%. Valued at .725 (their estimate), the partially correct responses increase the mean percentage responding correctly by only 3.7%.

[15]And broadly consistent with Delli Carpini and Keeter (1996), whose round-up of 612 (231 distinct) knowledge items averages 45.5% correct. Some of these items are open-ended, bringing the average down, but the closed-ended items were uncorrected for guessing, bringing it up.

[16]Those responding DK, incorrectly, and correctly to the dependent knowledge items average 37.9%, 48.2%, and 60.5% correct and 40.5%, 45.6%, and 59.2% correct on the independent knowledge index in the DK-neutral and -discouraging conditions, respectively. The percentages for the DK-responders are nonzero for the reasons given in n. 4.

open-ended items, discouraging DKs again makes little difference, perceptibly increasing the percentage answering correctly only among men and the less well educated. Again we take these small increments as descriptive improvements. For closed-ended items, on the other hand, discouraging DKs again worsens the measurement, creating more lucky guessing than it uncovers hidden knowledge, and does so for everyone, albeit more for some groups than others. (See Luskin and Bullock 2010).[17]

Discussion

Ironically, discouraging DKs turns out to improve the measurement only for the open-ended knowledge items we have been urged to shun (Gibson and Caldeira 2009; Mondak 2001; Mondak and Anderson 2004). In the open-ended case, discouraging DKs unearths some hidden knowledge without reburying it in lucky guessing. It also slightly increases the correlations with criterion variables (as shown in the online Appendix A). In the closed-ended case, however, it worsens the measurement, especially in comparison with encouraging DKs. Here the uncovering of hidden knowledge is overwhelmed by lucky guessing. Correspondingly, the correlations with criterion variables are highest for the DK-encouraging treatment and lowest for the DK-discouraging one (again see the online Appendix A).

At a technical level, the moral for designing closed-ended items is clear. DKs should *not* be discouraged. The necessary preambles and probes consume interview time without producing appreciably fewer DKs.[18] The only question is whether to *en*courage them. That also consumes interview time but reduces the measurement error introduced by guessing and increases the correlations with criterion variables (again see the online Appendix A). For open-ended items, the moral is less clear. Discouraging DKs improves the measurement, but the improvement is small and counterbalanced by practical considerations. The prefaces and probes, which still take time, may mean not getting to ask other items.

For already-collected DK-discouraged data, the morals for both formats are clear. For open-ended items, it makes sense to use the ''final answers.'' They are a small, nonreturnable gift that may or may not have been worth the price but may as well be accepted. For closed-ended items, on the other hand, the ''final answers'' should be ignored. They are overwhelmingly guesses, best left as DKs (and scored as incorrect). Unfortunately, this is only a partial remedy, since, for any given item, there is no way of recovering the would-have-been DKs emboldened to guess, even before any probing, by the DK-discouraging preamble (if there is one) or by the probing of DKs to preceding items. Indeed it would be better to have *en*couraged DKs. But this is the best the secondary analyst can do—this, plus, for description, a correction for guessing.

Substantively, these results are important for what they refute. Mondak's claim, recently echoed by others (Gibson and Caldeira 2009; Prior and Lupia 2008) on other grounds, is that the public's level of political knowledge is substantially higher than most of us have been thinking. If true, this would rotate our understanding of much of the rest of politics. Not 180°—knowledge levels, like all the voter, campaign, and media behaviors they are commonly thought to condition, are continua, to which mere adjectives do imperfect justice—but to some appreciable degree.[19] The more people actually know, the less phenomena like the policy-innocence of most ''news'' coverage can be understood in terms of public ignorance.

As we have begun to show here, however, no such reorientation seems necessary. Discouraging DKs reveals precious little hidden knowledge. It increases the chronically low percentages answering open-ended items correctly, but not very much. It increases the somewhat higher percentages answering closed-ended items correctly more sharply (in comparison with DK-encouragement), but mostly on account of lucky guessing. Indeed, what this suggests is that the

[17]To some extent, these effects vary by group, thus affecting some of the demographic ''gaps'' in knowledge. For open-ended items, according to which these gaps are generally smaller, discouraging DKs actually widens the gender gap, producing a distinctly larger increase in correct answers from men than from women. For closed-ended items, according to which the gaps are generally larger, discouraging DKs narrows both the gender and education gaps—partly because it uncovers more hidden knowledge among women and the less well educated, partly because it worsens the measurement less (but still worsens it) for those groups. Again see Luskin and Bullock (2010) for further detail.

[18]Mondak (2001) presents evidence that discouraging DKs did not significantly increase interview length in his Tallahassee study or the 1998 ANES Pilot. But his comparison is to a DK-*encouraging* treatment, with its own preamble and an explicit DK response option, the reading of which counterbalances the probing in the DK-discouraging treatment. A DK-discouraging treatment *must* take more time than a DK-*neutral* one. The clock does not pause while interviewer reads the preamble or probes DKs.

[19]It also undermines the ''extenuationist'' literature To the extent that people do know a lot about politics, what is to be made of accounts of how they circumvent their ignorance?

standard, bleakish picture of the public's level of political knowledge, to the extent that it is based on closed-ended items uncorrected for guessing, may not be quite bleak enough.

But what of the other recent efforts at finding hidden knowledge? These too do little to alter the standard picture. Randomly recoding some DKs as correct injects a dose of simulated guessing, biasing the measurement upward and enfeebling the correlations with criterion variables (see the online Appendix A and Luskin and Bullock 2005). Shifting entirely to closed-ended items would trade the problem of temporarily inaccessible knowledge for that of lucky guessing—on the evidence here a bad trade, which may be why the correlations with criterion variables run noticeably higher with open-ended items (again see the online Appendix A and Luskin and Bullock 2005).

Nor does there appear to be a great deal of knowledge hidden in partially correct responses. Precise estimates await a universal recoding of the ANES's open-ended items, but Gibson and Caldeira (2009) have at least recoded the 2000 Rehnquist item, from which upper bounds can be can extrapolated. To do so, we assume, again generously, that every open-ended item elicits as many partially correct responses as the 2000 Rehnquist item.[20] We also continue all the deliberately too-generous assumptions above. In the 2000 ANES, these calculations yield a maximum adjustment for partially correct responses of 10.9%, which leaves the maximum adjusted percentage answering correctly at 38.3% (see the online Appendix C for more detail). While 10.9% may seem impressive, note again that it is only a very high ceiling. A more realistic adjustment would be much lower. And note too that 38.3% is still low—lower, in fact, than the already-low guessing-corrected percentages answering closed-ended items correctly.

Even providing monetary incentives for correct answers and much more time to summon them up does surprisingly little. Prior and Lupia (2008) report that paying a dollar for each correct answer increases the percentage answering correctly by 11%, that allowing a full day to respond increases it by 18%, and that doing both increases it by 24%, but these figures are *percentage increases*, rather than *increases in the percentages* (the conventionally defined *treatment effects*). These latter, which can be calculated from Prior and Lupia's Tables 1 and 3, are much smaller. The monetary incentive increases the percentage answering correctly by only 3.6% (from 32.1% to 35.7%), the 24-hour response interval increases it by only 6.5% (from 32.1% to 38.6%), and the two together increase it by only 7.9% (from 32.1% to 40.0%). It is hard to regard these effects, whatever they may be as percentage increases, as anything but small.[21]

As estimates of second-effort recall, moreover, they are overestimates. Being able to infer or research the right answer (never mind blindly guessing it) is not the same as knowing it. But the monetary incentive can be expected to induce more frequent, sometimes lucky guessing (10 of Prior and Lupia's 14 items are closed-ended) and more determined and hence sometimes shrewder inference, while the 24-hour response interval can be expected to induce easy research (asking others, looking online—this was an online survey, after all), as well as still more guessing and inference. So not all—indeed, it is fair to suspect, not much—of even these already modest increments represents actual knowledge.[22]

This may be why Prior and Lupia wind up retreating to describing their results as showing ''political learning skills''—in effect conceding that people do not actually appear to know much more than we thought but arguing that at least they can learn. No doubt they can. But *do* they? The potential is only important insofar as it leads, at appropriate junctures, to the reality. As we have seen, Prior and Lupia's results do not actually show much learning. Greater incentives or longer response intervals might produce more, but the real question is what can be expected from the incentives and learning periods voters realistically face, and the evidence on this score is not encouraging. The surveys showing how little most people know are mostly conducted just before or after national elections. If people do not exercise their ''political learning skills'' enough to find the

[20]The Rehnquist item elicited more than twice as many partially correct responses as the average of the other three open-ended items in Mondak and Davis's (2001) Tallahassee data. See their Table 2.

[21]There may be cases where a percentage increase is worth citing for elaboration or rhetorical flourish, but since percentage increases are always larger, sometimes vastly larger, than the underlying treatment effects, focusing on the percentage increases can make small effects look much larger than they are. A 1-to-2% increase in the percentage answering correctly is a 100% percentage increase, even though the percentage answering correctly has only increased by 1%. Still worse, smaller treatment effects can look larger than larger ones. As a percentage increase, a 1-to-2% increase in the percentage answering correctly would also be larger (100%) than a 20-to-35% increase (75%), even though the treatment effect is only 1% in the first case, as compared to 15% in the second.

[22]A correction for guessing would help, but not enough, since it would not correct for inference or research.

right answers then, when do they?[23] Indeed, the timing of these surveys suggests that their results may be something of a high-water mark. Between elections, people presumably know somewhat *less* than the usual survey results suggest.

Let us also not forget that all knowledge measures are subject to two further discounts. First, the people we actually get to interview, a distinctly nonrandom subsample (and, at most, these days, a slender majority) of the designated interviewees, are more knowledgeable about politics than the population as a whole (Brehm 1993, 61–63). Second, the knowledge questions we ask, a distinctly nonrandom sample of the universe of potential knowledge questions, are on the grand scale of things extremely easy. The universe includes many items only policy specialists, if anyone, would know. Thus all survey questionnaires, including ours and the ANES's, involve a severe item sampling bias, toward easy knowledge items (Converse 2000; Luskin, Sood, and Helfer 2010). One needn't be an expert—needn't even be paying more than slight and occasional attention—to know who the incumbent Vice-President is, what party holds the majority in the House of Representatives, or which party is more in favor of redistributive policies. Optimists may see the glass as 35–40% full rather than as 60–65% empty, but it is a thimble glass, not a *ballon*. For both these reasons, all knowledge measures, even corrected for guessing, run high.

Sadly, many even of the more than averagely knowledgeable respondents we actually get to interview still miss or confess they don't know the answer to many of even the relatively easy knowledge questions we ask. By and large, as we argue here, those saying they don't know really don't know. Why, when more than 20% of the population, in both the United States (Miller 2004) and Europe (Pardo and Calvo 2004), do not know that the earth revolves around the sun rather than vice versa,[24] should we find it so hard to believe that substantially larger percentages don't know salient facts about politics?

[23]Prior and Lupia, evidently recognizing this problem, argue that much learning is motivated by the need to decide how to vote and that many interviews occur too early to catch all the predecision learning. But this argument rests on the implicit assumptions (a) that the preelection interview *precedes* the decision; (b) that there *is* such a decision, rather than just an obviating nondecision to abstain; and (c) that many respondents learn a lot between the preelection interview and deciding how to vote, yet then forget it all by their postelection interview. The first two assumptions are known to be wrong for many voters, and the third is simply implausible.

[24]Incorporating, in the European case, a correction for guessing. Miller (2004) reports only the percentage answering incorrectly. If the ratio of incorrect responses to DKs is roughly the same as in Europe (5: 3), the guessing-corrected U.S. percentage not knowing would approach thirty. Of course all these percentages are still too low, on account of the same unit sampling bias described above.

Acknowledgments

Earlier versions of this paper were presented at the 2005 annual meeting of the Midwest Political Science Association and at the 2005 biennial General Conference of the European Consortium for Political Research, Budapest, Hungary. The original data were collected by Time-Sharing Experiments for the Social Sciences, NSF Grant 0094964, Diana C. Mutz and Arthur Lupia, Principal Investigators. The American National Election Studies data were collected by the Center for Political Studies, University of Michigan, and are publicly available at http://www.electionstudies.org/. We have been stimulated by conversations with Larry Bartels, Jon Krosnick, and Simon Jackman and benefited from comments by John Bargh, Samuel Gosling, William Jacoby, Jeffrey Karp, Delroy Paulhus, Markus Prior, Jasjeet Sekhon, and especially James Kuklinski. Patricia Luevano and Matthew DeBell shed valuable light on the ANES data. We alone are responsible for the arguments we make. The online Appendices A, B, and C are available at http://journals.cambridge.org/JOP; the data and supporting materials at http://www.utexas.edu/cola/depts/government/faculty/rluskin?tab=186 or http: //bullock.research.yale.edu/DK/DK.zip.

References

Anderson, John R. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Bennett, Stephen Earl. 2001. "'Reconsidering the Measurement of Political Knowledge' Revisited: A Response to Jeffery Mondak." *American Review of Politics* 22: 327–48.

Bishop, George F. 2005. *The Illusion of Public Opinion: Fact and Artifact in American Public Opinion Polls*. Lanham, MD: Rowan and Littlefield.

Brehm, John. 1993. *The Phantom Respondents: Opinion Surveys and Political Representation*. Ann Arbor: University of Michigan Press.

Brown, R., and McNeill, D. 1966. "The 'Tip of the Tongue' Phenomenon." *Journal of Verbal Learning and Verbal Behavior* 5: 325–37.

Collins, Allan M., and Elizabeth Loftus. 1975. A Spreading-Activation Theory of Semantic Processing. *Psychological Review* 82: 407–28.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: Free Press, 206–61.

Converse, Philip E. 2000. "Assessing the Capacity of Mass Electorates." *Annual Review of Political Science* 3: 331–53.

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

Gibson, James L., and Gregory A. Caldeira. 2009. Knowing the Supreme Court? A Reconsideration of Public Ignorance of the High Court. *Journal of Politics* 71: 429–41.

Krosnick, Jon A., Arthur Lupia, Matthew DeBell, and Darrell Donakowski. 2008. Problems with ANES Questions Measuring Political Knowledge. ANES Report, March 2008.

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

Luskin, Robert C. 2002. ''From Denial to Extenuation (and Finally Beyond): Political Sophistication and Citizen Performance.'' In *Thinking about Political Psychology*, ed. James H. Kuklinski. Cambridge: Cambridge University Press, 281–305.

Luskin, Robert C., Gaurav Sood, and Ariel Helfer. 2010. Measuring Learning in Informative Processes. Presented at the annual meeting of the American Political Science Association, Washington, DC.

Luskin, Robert C., and John G. Bullock. 2005. ''Measuring Political Knowledge.'' Unpublished manuscript, University of Texas.

Luskin, Robert C., and John G. Bullock. 2010. ''Discouraging DKs and the Gender (and Race and Education) Gaps in Political Knowledge.'' University of Texas at Austin.

Miller, Jon D. 2004. ''Public Understanding of, and Attitudes toward, Scientific Research: What We Know and What We Need to Know.'' *Public Understanding of Science* 13: 273–94.

Mondak, Jeffery J. 1999. ''Reconsidering the Measurement of Political Knowledge.'' *Political Analysis* 8: 57–82.

Mondak, Jeffery J. 2001. ''Developing Valid Knowledge Scales.'' *American Journal of Political Science* 45: 224–38.

Mondak, Jeffery J., and Mary R. Anderson. 2004. ''The Knowledge Gap: A Reexamination of Gender-Based Differences in Political Knowledge.'' *Journal of Politics* 66: 492–512.

Mondak, Jeffery J., and Belinda Creel Davis. 2001. ''Asked and Answered: Knowledge Levels When We Will Not Take 'Don't Know' for an Answer.'' *Political Behavior* 23: 199–224.

Nie, Norman, Sidney Verba, and John R. Petrocik. 1979. *The Changing American Voter*. Cambridge, MA: Harvard University Press.

Page, Benjamin I., and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences.* Chicago: University of Chicago Press.

Pardo, Rafael, and Félix, Calvo. 2004. The Cognitive Dimension of Public Perceptions of Science: Methodological Issues. *Public Understanding of Science* 13: 203–27.

Paulhus, Delroy L., P. D. Harms, M. Nadine Bruce, and Daria C. Lysy. 2003. ''The Over-Claiming Technique: Measuring Self-Enhancement Independent of Ability.'' *Journal of Personality and Social Psychology* 84(April): 890–904.

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

Prior, Markus, and Arthur Lupia. 2008. Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills. *American Journal of Political Science* 52 (1): 169–83.

Sanchez, Maria Elena, and Giovanna Morchio. 1992. Probing 'Don't Know Answers': Effects on Survey Estimates and Variable Relationships. *Public Opinion Quarterly* 56: 454–74.

Sturgis, Patrick, Nick Allum, and Patten Smith. 2008. An Experiment on the Measurement of Political Knowledge in Surveys. *Public Opinion Quarterly* 72 (1): 90–102.

Sturgis, Patrick, and Patten Smith. 2010. ''Fictitious Issues Revisited: Political Interest, Knowledge and the Generation of Nonattitudes.'' *Political Studies* 58: 66–84.

John G. Bullock is Assistant Professor at Yale University, New Haven, CT 06511.

Robert C. Luskin is Associate Professor at the University of Texas at Austin, Austin, TX 78712.