“Don’t Know” Responses, Personality and the Measurement of Political Knowledge

Stephen A. Jessee

The University of Texas at Austin

Abstract

A prominent worry in the measurement of political knowledge is that respondents who say they don’t know the answer to a survey question may have partial knowledge about the topic—more than respondents who answer incorrectly, but less than those who answer correctly. It has also been asserted that differentials in respondents’ willingness to guess, driven strongly by personality, can bias traditional knowledge measures. Using a multinomial probit item response model, I show that, contrary to previous claims that “don’t know” responses to political knowledge questions conceal a good deal of “hidden knowledge,” these responses are actually reflective of less knowledge, not only than correct responses, but also than incorrect answers. Furthermore, arguments that the meaning of “don’t know” responses varies strongly by respondent personality type are incorrect. In fact, these results hold for high- and low-trait respondents on each of the five most commonly used core personality measures.

The author would like to thank John Bullock for extensive input and assistance throughout the course of this project and Rob McCulloch for sharing his code and providing extremely helpful advice. This project also benefited greatly from the input of Doug Rivers, Adam Bonica, James Scott, Scott Moser, Tse-min Lin, Chris Wlezian, Seth Hill, Alex Tahk, Christopher Fariss and participants at the UCSD Methods Workshop and University of Texas American Politics Workshop.

1 Introduction

Knowledge of politics is a prominent variable in the study of political behavior, but the field is increasingly beset by concerns that ordinary measures of knowledge are in fact mismeasures of knowledge [Mondak, 1999, Gibson and Caldeira, 2009, Prior and Lupia, 2008, Krosnick et al., 2008, Boudreau and Lupia, 2011]. Conventional knowledge measurement starts with questions that are thought to tap subjects' levels of knowledge. These questions may be about institutions and processes ("If the president vetoes a bill, can Congress override his veto?"), public figures ("What job or office is currently held by Joe Biden?"), foreign affairs ("Is the United States a member of the United Nations?"), or other aspects of politics. Common practice is to define one answer to each question as correct. Subjects who give that answer are given a "score" of 1. Other subjects are given a score of 0 whether they answer incorrectly or say "don't know." Summing a respondent's scores across all questions gives his "knowledge score": a nonnegative integer indicating his level of knowledge.[1]

Several specific concerns have been raised about this method of measurement. One is that it rests heavily on "trivia questions" about matters that are irrelevant to political choices [Lupia, 2006]. Another is that political knowledge is multidimensional—for example, one may know much about domestic politics but little of foreign affairs—and that it is therefore misleading to assign a single "knowledge score" to any survey respondent (Iyengar 1986; but see Delli Carpini and Keeter 1996). Another objection, and the one that is focused on here, is that basing political knowledge measures solely on the number of correct answers, thus combining "don't know" and incorrect answers together, will produce undesirable or inappropriate results, either because these two response categories suggest different levels of political information or because the meaning of "don't know" responses differs systematically across respondents based on individual-level characteristics such as

[1]Following Luskin [2002], I use "knowledge," "sophistication," and "information" interchangeably.

personality (e.g. Mondak 1999). A central contention of Mondak's work on this subject is that "don't know" responses should be taken to indicate more knowledge than incorrect answers. Others, such as Sturgis et al. [2008] and Luskin and Bullock [2011], however, have suggested that this argument may be misguided.

While much of the extant literature related to these topics focuses on whether survey writers should encourage or discourage "don't know" responses when writing new surveys, this paper is centrally concerned with assessing the level of political knowledge suggested by correct, incorrect and "don't know" responses in existing survey data and whether this varies by respondent personality as has been suggested in the literature. For example, how should researchers think about "don't know" responses in existing data such as those from the American National Election Studies (ANES) in which they are relatively common? Simply discarding any respondent who gave a "don't know" answer to one or more questions seems undesirable in the abstract and in practice would often result in the loss of a huge percentage of observations. For example, in the 1988 ANES, nearly 90% of respondents said "don't know" to at least one of the 13 survey's political knowledge questions. But if these "don't know" responses are to be used in the construction of measures of political knowledge, researchers must know how much knowledge they indicate relative to correct and incorrect answers. To this end, I present a multinomial item response model of answers to political knowledge questions that allows for the estimation of respondents' underlying levels of political knowledge as well as the direct comparison of the levels of knowledge implied by correct, incorrect and "don't know" responses. The model is estimated for several waves of the ANES, strongly indicating that "don't know" responses actually reflect less political knowledge—not more—than incorrect responses.

A more specific concern about "don't know" responses is that their meaning may vary across respondents. For example, one type of respondent may offer a "don't know" response even if she is fairly confident, but not completely certain of her knowledge about a certain question, while a different respondent may only say "don't know" if he is completely clueless. The most prominent of these claims in the literature has been that respondent personality plays a large role in determining the level of knowledge that should be inferred from a "don't know" response [Mondak and Halperin, 2008, Mondak, 2010]. Clearly, personality factors, such as how extroverted or how agreeable a respondent is seem like plausible candidates for mediating the relationship between knowledge levels and the likelihood of saying "don't know."

The model used here provides a direct way to investigate such hypotheses. Specifically, I estimate the model using a newer national study that includes both factual questions about politics and a rich battery of personality items. This battery measures the "Big Five" personality traits, which are the most widely studied personality traits in psychology [Funder, 2001] and are increasingly used in political science [e.g., Gerber et al., 2010, Mondak, 2010]. By separately estimating the model for high- and low-trait respondents on each of these five major personality traits, I find that the conclusions drawn from looking at respondents of all personality types together—that "don't know" answers are indicative of less knowledge than incorrect answers—is highly robust to variations in personality: it holds for extroverts and introverts, the agreeable and the less agreeable, the highly conscientious and the less conscientious, and for other personality factors. Contrary to previous arguments in the literature, the meaning of these "don't know" answers does not differ meaningfully by respondent personality.

The paper is organized as follows. I begin by describing conventional knowledge measurement and arguing that one of the main methods used to advance arguments about the meaning of "don't know" responses is inappropriate. I proceed by developing an explicit model of responses to knowledge questions. I then estimate this model using data from several waves of the ANES and discuss the findings. Next, I estimate the same model on a large new dataset, separating respondents by their level of each of the Big Five personality traits. I conclude by considering the implications of these findings and avenues for future research.

2 Concerns about Conventional Knowledge Measurement

Previous work has cast doubt on the wisdom of combining incorrect and "don't know" responses when analyzing political knowledge items. Those who make such arguments have claimed that scoring incorrect and "don't know" responses in the same way is misleading if the two types of response are related to non-knowledge variables in different ways (e.g., Mondak 1999). For example, the argument goes, it is misleading to assign the same score to incorrect and "don't know" responses if education is positively correlated with incorrect answers but negatively correlated with "don't know" responses. This sort of objection figures in several broadsides against conventional knowledge measurement, and Mondak [1999, 66] calls the comparison of such correlations "the critical diagnostic test."

This type of test, however, can easily lead to improper conclusions. To see why, consider the task of estimating knowledge levels in a population that has three types of people: the *uninformed*, who know nothing; the *informed*, who know the answer to any question that might be put to them; and the *partly informed*, whose level of knowledge lies in between. Let the probabilities of knowing the correct answer to a given question be 0, 1, and .5 for the uninformed, informed, and partly informed, respectively. Assume that the uninformed always answer "don't know," the informed always answer correctly, and the partly informed answer correctly half the time and incorrectly half the time.[2] Assume as well that the relationship between education and knowledge is as shown in the left-hand panel of Figure 1. At the lowest level of education, three-fourths of the population is uninformed, one-fourth is partly informed, and no one is informed. At the highest level of education, no one is uninformed, half the population is partly informed, and half the pop-

[2]The behavior of the partly informed could described those who, for example, know that the correct answer to a four option multiple choice question is one of two choices, but do not know which specific answer out of these two is correct. Thus, these people have partial knowledge—more than someone who, for example, cannot eliminate any of the four choices as being incorrect, but less than someone who knows the specific correct answer.

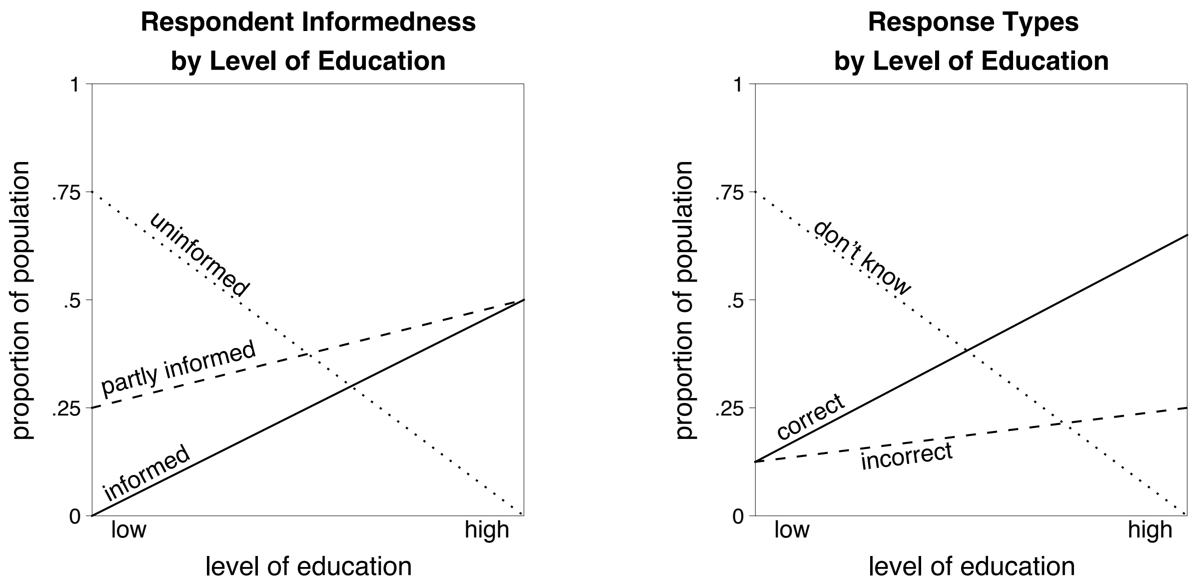

Figure 1: Conventional Knowledge Measures Can Be Unbiased Even if Incorrect and "Don't Know" Responses are Not Equally Correlated with a Third Variable. *The figure depicts relationships between political information, education, and responses to knowledge questions in a hypothetical population. The left-hand panel shows the proportion of respondents who are informed, partly informed or uninformed as a function of education. Under the assumptions described in the text, these relations between education and political information lead to the relations between education and response type—correct, incorrect, or don't know—that are depicted in the right-hand panel. Importantly, the right-hand panel shows that "don't know" responses decline with education while incorrect responses increase with education. When such patterns are found in survey data, they are sometimes taken as evidence that knowledge measures are biased if they score incorrect and "don't know" answers in the same way. As this example demonstrates, this conclusion is too strong.*

ulation is fully informed. These assumptions are consistent with the finding that education and political information are positively correlated [Delli Carpini and Keeter, 1993].

The preceding assumptions produce the patterns that appear in the right-hand panel of Figure 1. As education increases, "don't know" responses become less common while incorrect answers become more common.[3] Some would take this as evidence that "don't know" and incorrect responses must be scored differently, but it is not. To see this, consider conventional knowledge scores for the population described in Figure 1. Uninformed respondents in this population answer "don't know" and therefore receive a score of 0. Informed respondents answer correctly and therefore receive a score of 1. Partly informed

[3]These patterns resemble those observed among ANES subjects [Mondak, 1999].

respondents have an equal chance of answering correctly and receiving a score of 1 or answering incorrectly and receiving a score of 0; the average score of this group is therefore .5. We thus obtain unbiased estimates of respondents' probabilities of knowing the answer to this question, which presumably would serve as a reasonable, if not necessarily ideal, estimate of each respondent's underlying political knowledge level.[4] The estimates are unbiased even though conventional scoring, which treats incorrect and "don't know" answers in the same way, is used. And these estimates are unbiased even though education is positively correlated with incorrect responses but negatively correlated with "don't know" responses. This example thus demonstrates that such patterns do not necessarily imply a flaw in conventional methods of knowledge measurement.

Perhaps more importantly, this example illustrates the difficulty of evaluating the relative levels of knowledge implied by correct, incorrect and "don't know" responses to political knowledge questions by only looking at the relationships between respondent characteristics and their probabilities of answering in different ways. Of course, it remains possible that conventional measurement methods are biased despite the potential for misleading conclusions that exists in these tests. The preceding example, then, should also emphasize the need for investigation through more direct means of the relationship between underlying levels of political knowledge and the probability of giving correct, incorrect or "don't know" responses. The following section addresses this through the development of a multinomial probit (MNP) item response model of political information.

3 Directly Modeling Political Knowledge Item Response

Item response models, also called latent traits or ideal point models, have become increasingly common in the political science literature. They have been used to measure such im-

[4] It is straightforward to extend the logic of this example from one knowledge item to batteries that contain multiple items or to more complicated distributions of respondent knowledge levels.

portant concepts as the ideology of members of Congress [e.g., Poole and Rosenthal, 1985, Heckman and Snyder, 1997, Clinton et al., 2004], Supreme Court justices [Martin and Quinn, 2002], and ordinary citizens [Jessee, 2009, 2010, Bafumi and Shapiro, 2009, Shor, 2009] as well as the level of democracy across countries [Treier and Jackman, 2008], the level of respect for human rights Schnakenberg and Fariss [2014], and the biases of media outlets [Groseclose and Milyo, 2005]. These models provide a valuable tool for estimating the characteristics of variables which are not directly observable, but which affect observable data. Political knowledge is a prime example of this type of variable. There is no way to directly measure a person's level of knowledge, but it would expected that responses to survey questions asking respondents about political facts would be strongly affected by their knowledge levels.

Simplicity is the allure of the conventional approach to knowledge measurement. Few things are simpler than totaling the number of correct answers that a respondent gives to series of factual questions.[5] And while this approach is rarely justified by reference to a model, it can be. In particular, item response models conceptualize respondents' answers to political knowledge survey items as being generated stochastically based on each individual's underlying (latent) level of knowledge. Item response theory models such as the Rasch model [Rasch, 1960a,b] originated in educational measurement research, but in political science they are rarely used to study political knowledge (see Jackman 2000, Lawrence 2009 and Jessee 2009, 2010 for exceptions). Using them in this way, however, is closer to the purpose for which the models were originally designed.

In most political science applications of item response theory the main quantities of interest are respondents' individual levels of the latent attribute. The model presented here is much closer to the spirit of most educational testing analyses because the central

[5] Of course, the simplicity of the task depends on the questions. As Krosnick et al. [2008] point out, judging the "correctness" of responses to some open-ended questions is not always straightforward. But it is straightforward with many, and with closed-ended questions in which respondents choose from a small set of responses, it is typically an easy task.

goal is not to estimate the political knowledge level of each respondent, but instead to determine how much knowledge is indicated by each response type: correct, incorrect and "don't know."

Most latent traits models in political science have utilized binary data. For example, congressional voting ("yea" or "nay") or Supreme Court decisions ("reverse" or "affirm") take only one of two values (but see Noel [2007] for an example of ideal point estimation with abstentions as a third vote category using a multinomial logistic setup). The key issue here, however, is how to understand the level of political knowledge that is suggested by each of three types of responses: correct, incorrect and "don't know". Because there are more than two categories and because the goal is to determine (rather than to assume) the ordering of these categories in terms of how much knowledge each indicates, a multinomial model is used. These sorts of models were first analyzed by Bock [1972] who employed a multinomial logistic setup. Taking a somewhat similar approach, Anderson et al. [2010] use log-multiplicative association models, which are extensions of the multinomial logistic setup, to examine political knowledge levels.

Here, it is assumed that respondents' latent propensities to give each response type are vectors of length equal to the number of response types p, which for us will be 3: correct, "don't know," and incorrect. The propensities, called u_{ij} for each respondent $i = 1, \ldots, n$ on each survey question $j = 1, \ldots, m$ are assumed to have a multivariate normal distribution:

$$u_{ij} = \begin{bmatrix} u_{ij1} \\ u_{ij2} \\ \vdots \\ u_{ijp} \end{bmatrix} \sim N\left(\begin{bmatrix} \gamma_{j1} + \delta_{j1} x_i \\ \gamma_{j2} + \delta_{j2} x_i \\ \vdots \\ \gamma_{jp} + \delta_{jp} x_i \end{bmatrix}, \Omega_j \right), \tag{1}$$

where x_i is respondent i's level of political information, Ω_j is a positive definite covariance matrix. The distribution of the u_{ij}'s are assumed to be independent across questions and respondent given γ, δ, x and Ω.

The only observed data consist of an n by m response matrix y, which contains each respondent's answer type for each question, indexed by an integer from 1 to p. It is assumed that on each question, respondents select the response option for which u_{ij} is largest. Formally, we can write

$$y_{ij} = \sum_{k=1}^{p} k \times I(\max(u_{ij}) = u_{ijk}), \tag{2}$$

where y_{ij} is respondent i's answer to question j.

As in the standard multinomial probit model, however, the latent response propensities u, and hence the other parameters γ, δ, and Ω, are not identified. Adding any arbitrary constant to each element of u_{ij} or multiplying each one by a positive constant will result in exactly the same likelihood for the data given corresponding transformations of γ, δ, and Ω. Because of this, it is common to work with "differenced" propensities (see Rossi et al. 2005 Section 4.2). Accordingly, the model is not written in terms of the response propensities u_{ij}, but instead in terms of the difference between the propensity for a given "baseline" response category and the propensities for each of the other response categories. If we label the categories such that the baseline category is the last one p, we assume that $w_{ij} = u_{ij-p} - u_{ijp}$, which yields

$$w_{ij} = \begin{bmatrix} u_{ij1} - u_{ijp} \\ u_{ij2} - u_{ijp} \\ \vdots \\ u_{ij(p-1)} - u_{ijp} \end{bmatrix} \sim N\left(\begin{bmatrix} \alpha_{j1} + \beta_{j1}x_i \\ \alpha_{j2} + \beta_{j2}x_i \\ \vdots \\ \alpha_{j(p-1)} + \beta_{j(p-1)}x_i \end{bmatrix}, \Sigma_j\right) \tag{3}$$

where $\alpha_{jk} = \gamma_{jk} - \gamma_{jp}$ and $\beta_{jk} = \delta_{jk} - \delta_{kp}$ for each question j and each response type $k = 1, \ldots, p-1$. The observed responses are now determined by these differenced response propensities w according to

$$y_{ij} = \left[\sum_{k=1}^{p-1} k \times I(\max(w_{ij}) = w_{ijk} \text{ and } w_{ijk} > 0)\right] + p \times I(w_{ijk} < 0 \text{ for } k = 1, \ldots, p-1). \tag{4}$$

In other words, if all of the elements of w_{ij} are less than zero, indicating that the response propensity for the baseline category p is larger than that for all other categories, then the baseline option is chosen. If any of the elements of w_{ij} is positive, the alternative corresponding to the largest value of w_{ij} is chosen.

In the standard multinomial probit model with observed (fixed) predictors, a restriction on the variance-covariance matrix Σ is typically imposed such that the first diagonal element is one. Absent this restriction, elements of w_{ij} can simply be multiplied by any positive constant and the corresponding elements of Σ_j adjusted accordingly to give the same likelihood for the data. In our situation, the predictor x is a latent quantity that is itself estimated along with the other parameters of the model. Therefore further restriction is necessary in order to identify the model.

To see this, consider the simplified example of

$$w = x_1\beta + x_2\lambda + \varepsilon, \tag{5}$$

where w is a $n \times 2$ matrix, x_1 and x_2 are $n \times 1$ vectors that are independently drawn from a $N(\mathbf{0}, \mathbf{I})$ distribution, β and λ are 1×2 coefficient vectors, and ε is a $n \times 2$ matrix drawn from a $N(\mathbf{0}, \mathbf{I})$ distribution independent of x_1 and x_2. This setup parallels the one used in Equation 3 above, but omits the intercept terms for convenience.[6] Given this model, we

[6]This simplification does not alter the basic properties of the example. To see this, we can simply consider w here as the latent response propensity differences subtracting off the intercept terms.

could define $\xi = x_2\lambda + \varepsilon$ and rewrite Equation 5 as

$$w = x_1\beta + \xi, \tag{6}$$

where

$$\xi \sim N\left(\mathbf{0}, \begin{bmatrix} \lambda_1^2 + 1 & \lambda_1\lambda_2 \\ \lambda_1\lambda_2 & \lambda_2^2 + 1 \end{bmatrix}\right). \tag{7}$$

But this exercise could also be performed by letting $\xi = x_1\beta + \varepsilon$, demonstrating the lack of identification of this model. Allowing the covariance matrices Σ_j to be non-diagonal is analogous to allowing for a second latent dimension of political information. Moreover, because these covariance matrices are allowed to differ across questions, a potentially different second dimension would be estimated for each question j. It should be noted that the above example involves a situation in which the latent response propensity differences w are directly observed. In the present model, we only observe which element of w is largest for each respondent on a given question—an even less informative situation.

Allowing Σ_j to be unrestricted, then, will result in an unidentified model. A natural identifying restriction is to fix Σ_j to an identity matrix for each question j. In addition to identifying the model, it provides a useful partition of the variation in the latent response propensity differences w. Because $\Sigma_j = \mathbf{I}$, all of the systematic variation in w is assumed to be accounted for my the latent knowledge levels x_i. The only other component of these response propensity differences is independent random noise. In the standard MNP model, the error variance is identified (albeit often tenuously, see Keane 1992) because the systematic portion of the latent propensities is restricted to be some linear combination of the observed predictors included in the model. Because the predictor used here is unobserved (estimated), the variation in the w due to knowledge levels and that due to the random errors are not separately identified.

Formally, fixing $\Sigma_j = \mathbf{I}$ for a single question would define the latent political knowledge dimension x and allow Σ_j for other questions to be estimated. This, however, would again be equivalent to estimating a two-dimensional model in the case of three response categories. Both because a single-dimensional conceptualization of political knowledge is employed here and also because a model with two or more dimensions is only very weakly identified (recall that we do not actually observe the latent response propensity differences w, but only an indicator of which element is the largest for each respondent on each question), this approach is not used here. As discussed below, examination of the correlation between estimated response propensity errors can be used as a sort of check on this unidimensionality assumption. I therefore fix Σ_j as a $p-1 \times p-1$ identity matrix for all questions j. To the extent that the latent errors appear to be uncorrelated, this suggests that the systematic variation in response propensities is well accounted for by the single dimension of estimated for political knowledge, leaving only idiosyncratic independent noise in the observed responses.

The random errors in w, which are assumed to come from a bivariate normal distribution with mean zero and variance I_2 under this restriction, were also examined for model runs in order to assess the reasonability of this assumption. These errors were quite consistent with the assumptions, showing little evidence of correlation between the correct and "don't know" response options. This may be thought to suggest that the single political knowledge dimension estimated by the model accounts well for the variation in responses, leaving little other than random idiosyncratic (uncorrelated) errors.

A final consideration is the fact that the political knowledge levels x are not themselves identified. I follow the common strategy of estimating the model in this unidentified state and post-processing the results to impose the identification restriction across each iteration that the knowledge levels have a mean of 0 and a variance of 1 for the full sample and that higher values of x indicate higher levels of political knowledge.[7]

[7]The last of these restrictions is implemented by choosing the orientation of x that results in the discrimination parameter for the correct response option for the first ques-

I employ independent conditionally conjugate priors for the model's unknown parameters:

$$\begin{bmatrix} \alpha_{jk} \\ \beta_{jk} \end{bmatrix} \sim N(\mathbf{b}_0, \mathbf{B}_0) \tag{8}$$

$$x \sim N\left(\mu_0, \sigma_0^2 \mathbf{I}\right) \tag{9}$$

with the priors for α and β assumed to be independent across both questions $j = 1, \ldots m$, and response options $k = 1, \ldots, p-1$. For all estimations below, I set $\boldsymbol{b}_0 = \mathbf{0}, \boldsymbol{B}_0 = 100\boldsymbol{I}, \mu_0 = 0$ and $\sigma_0^2 = 1$.

Although the setup described above can be motivated as a utility model, it can also be thought of, as with the standard MNP, as simply a probability model for observational data. Just as a standard probit model can be used to analyze shipwreck deaths (even though passengers' "utilities" for survival would almost certainly always be greater than those for drowning) or a standard multinomial probit can be used to analyze a baseball batter's chances of various outcomes in an at-bat (even though it seems unlikely that their "utility" for striking out or being thrown out at first would ever be higher than that for hitting a home run), the present model can be thought of one that simply analyzes the relationship between one's political knowledge level and one's probability of giving each type of response. The key difference here is that knowledge is estimated, rather than observed. In this way, the approach used here is closer to that of educational testing models where, although their decisions can be described in utilitarian terms, the actors being modeled would be expected to prefer to answer correctly rather than incorrectly. The model, then, is thought of as a probability model for the relationship between intelligence or knowledge and the likelihood of answering questions correctly or incorrectly (or in addition here, saying "don't know").

tion being positive. In practice, this also yields positive estimates for the correct response discrimination parameters for all other questions.

This setup models the relationship between respondents' levels of political knowledge and their likelihood of giving each type of answer. These relationships, rather than those between other variables and response probabilities, speak directly to the level of knowledge suggested by correct, incorrect and "don't know" responses. Accordingly, the discrimination parameters in the model merit special attention. These parameters permit different items to tap political knowledge to different degrees. And because they vary across response type (correct, incorrect, or "don't know"), they permit different responses to the same item to indicate different levels of knowledge.

While the general model has been described above for the case of an arbitrary number (p) of alternatives, all of the analyses presented here involve three response categories representing correct, incorrect, and "don't know" responses. Therefore we have p equaling three and thus two categories for which we estimate discrimination and difficulty parameters. Because the baseline (which is the p^{th} in the notation of equations 1 through 4) response category is set to be the incorrect response for each question, the discrimination parameters for all other response categories will be estimated relative to this one. Therefore, positive values for discrimination parameters will imply that a given response category indicates higher levels of political knowledge than does an incorrect response for that question. Conversely, a negative discrimination parameter would imply that the response type implies less information than does an incorrect answer. Because the central question here is how much knowledge "don't know" responses indicate, these discrimination parameter estimates will be the main focus.

The model is estimated using a Gibbs sampler, which produces draws from the posterior distribution over all the model's unknown parameters.[8] The details of this procedure can be found in Appendix A.

[8]The sampler is implemented in R [R Core Team, 2013] and relies in part on code from the rmnpGibbs function in the bayesm library [Rossi, 2012].

4 Political Knowledge and the ANES

I begin by estimating the model with data from the 1988 ANES. I focus on the ANES because most studies of political knowledge have relied on it, and on the 1988 ANES in particular because it contains an unusually large set of knowledge items with a very high percent of respondents having non-missing values for all questions. Four of these items are closed-ended questions about recent trends (in defense spending, unemployment, inflation, and the deficit), two are closed-ended questions about party control of Congress, and seven are open-ended items that asked subjects to identify the offices held by prominent politicians (Yasser Arafat, Mikhail Gorbachev, Ted Kennedy, William Rehnquist, George Shultz, Margaret Thatcher, and Jim Wright). A small fraction of subjects were not asked some questions, typically because they were not available for the post-election interviews in which most questions were asked. Because these people represent a small proportion of the respondents, and because the vast majority of them had missing values for all or most questions, only respondents who have non-missing responses for all questions are analyzed here.[9] The text of all 13 questions is listed in Appendix section B.1.

The baseline response category for each question is chosen to be incorrect, which simplifies interpretation of the resulting estimates. This restriction implies that the discrimination parameter for correct responses for each question j, which for clarity can be called β_j^C, indicates the increase in the propensity to give a correct answer relative to an incorrect answer that results from moving one unit (one sample standard deviation) higher on the political knowledge scale. Similarly, the discrimination parameters for "don't know" responses, now called β_j^{DK}, indicates the increase in the propensity to give an answer of "don't know" relative to an incorrect answer that results from a one unit increase in x.

[9] 1,750 of the 2,040 total respondents were used. Of those who were dropped for having one or more missing values, over ninety percent (265) of them were missing ten or more of the 13 questions analyzed and so were unlikely to contribute a meaningful amount of information. Furthermore, because missing values could mean multiple things, it is not obvious how they should be treated. Although this could represent an interesting line of inquiry, it is beyond the scope of this article.

The intercepts α, often called "difficulty parameters," indicate the relative propensities of giving correct or "don't know" responses for people with political knowledge levels of zero (the sample average).[10]

The main focus of attention will be on the "don't know" response discrimination parameters. Positive values of β_j^{DK} imply that "don't know" responses for the corresponding question are indicative of more knowledge than are incorrect answers. Negative values of β_j^{DK} would mean that "don't know" responses suggest lower levels of knowledge than incorrect answers. Although not of central interest here, estimates of the difficulty parameters α_j^{I} and α_j^{DK} for each question, which are related to the relative baseline likelihoods of giving each response, are also presented.

Most obviously, we would expect that the response type that indicates the most knowledge is correct, which would imply that β_j^C should be greater than both β_j^I and β_j^{DK}. The common strategy of measuring knowledge as the percent of questions answered correctly, thus lumping incorrect and "don't know" responses together, implicitly assumes that these two response types suggest the same level of knowledge. If this were the case, their discrimination parameters β_j^I and β_j^{DK} should be equal. By contrast, if "don't know" responses reflected "partial knowledge" as is sometimes suggested, then we should see $\beta_j^C > \beta_j^{DK} > \beta_j^I$. Finally, if we observe $\beta_j^C > \beta_j^I > \beta_j^{DK}$, this would suggest that don't know responses actually reveal less political knowledge not only than do correct responses, but also incorrect ones. The most important aspect of these results, then, should be the relative size of the discrimination parameters for incorrect and don't know responses.

Figure 2 displays the discrimination and difficulty parameter estimates for the 13 knowledge questions in the 1988 ANES.[11] As expected, the discrimination parameters for

[10] In many item response models, particularly those in educational testing, these intercepts α are assumed to be subtracted from, rather than added to, the utility difference equations (those for w_{ij}). In this case, higher values of α generally imply more difficult questions (hence the term "difficulty parameter"). Although the reverse is true here, I still refer to them with this more traditional terminology.

[11] These estimates are based on a run of 200,000 iterations of the Gibbs sampler, dropping the first 50,000 iterations as burn-in and then storing every 50th iteration thereafter,

correct responses $\left(\beta_j^C\right)$ are estimated to be positive in every case, indicating that correct answers are associated with higher levels of knowledge than are incorrect answers. The estimates for "don't know" responses $\left(\beta_j^{DK}\right)$, by contrast, are negative in every case, suggesting that "don't know" responses are associated with lower levels of political knowledge not only than correct answers but also incorrect ones. For some questions, such as those about Gorbachev and Thatcher, there is a strong negative relationship between political knowledge and the probability of saying "don't know." For other items, such as those about inflation and the deficit, the relationship is only weakly negative. But the most important variation in Figure 2 is not between items—it is between response types. Across all questions, the discrimination parameters for correct responses are estimated to be positive with exceedingly high probability, while those for "don't know" responses are estimated to be negative with near certainty.[12]

I also repeat this analysis using the political knowledge items from the 1992 and 1996 ANES studies.[13] The text of these survey questions are given in Appendix sections B.2 and B.3. These surveys are used because they have a relatively large number of political knowledge items (though not as many as the 1988 ANES) and also because a very high percentage of respondents provided answers (correct, incorrect, or "don't know") to all of these questions. Figure 3 reports the results, which are quite similar to those from the 1988 study. For both the 1992 and 1996 ANES data, discrimination parameters for correct responses are consistently estimated to be positive, while those for "don't know" responses are negative. This again implies that respondents who say they "don't know" an answer, rather than those who offer an incorrect response, are typically the least informed.

Across all three of the ANES studies analyzed here, which together include a total of

yielding a total of 3,000 draws from the posterior for analysis.

[12]A similar multinomial logistic (MNL) item response model was also estimated for this dataset and also for all others analyzed below. In all cases, the overall substantive results were the same for the MNL and MNP specifications.

[13]These estimates are each based on a run of 110,000 iterations of the Gibbs sampler, dropping the first 25,000 iterations as burn-in and then storing every 25th iteration thereafter, yielding a total of 3,400 draws from the posterior for analysis.

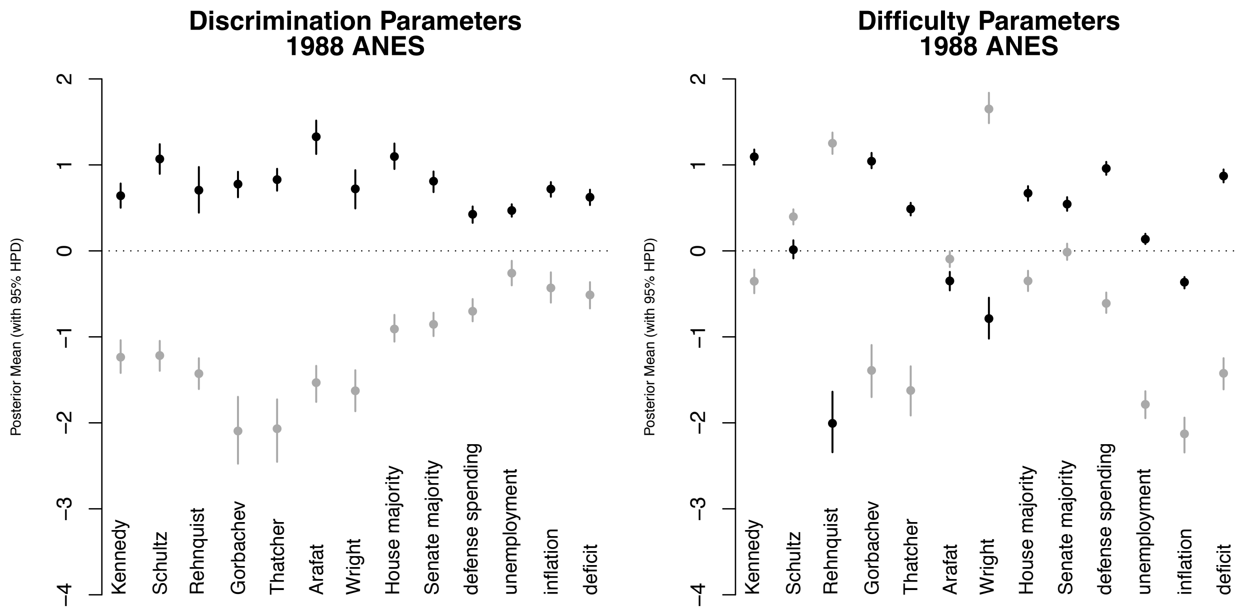

Figure 2: Question Parameters for 1988 ANES Knowledge Items. *Black dots represent discrimination parameters* β_j^C *(left pane) and difficulty parameters* α_j^C *(right pane) for correct response options. Grey dots represent discrimination parameters* β_j^{DK} *(left pane) and difficulty parameters* α_j^{DK} *(right pane) for "don't know" response options. Vertical line segments indicate 95% highest posterior density regions (HPDs). A positive discrimination parameter indicates that a given type of response to a question is associated with more political knowledge than an incorrect response, while negative discrimination parameters indicate that a response type reflects less knowledge than an incorrect answer. Discrimination parameters for correct responses are all estimated to be positive, indicating that these responses imply more knowledge than would an incorrect response. Discrimination parameters for "don't know" responses are all negative, implying that these responses are indicative of less knowledge than both correct and incorrect answers. Difficulty parameter values are not of direct interest for the question asked here but are presented in the interest of completeness.*

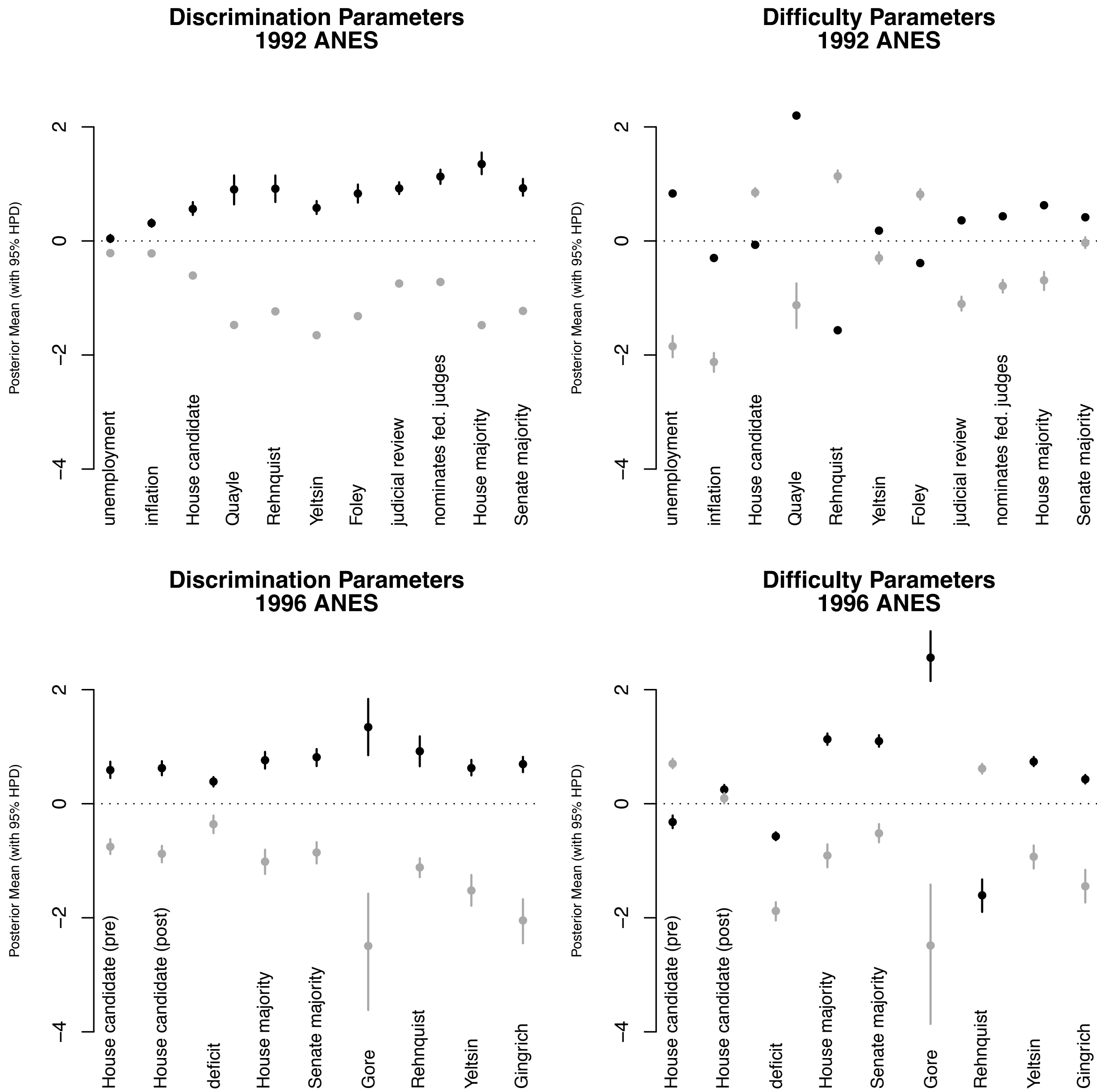

Figure 3: Discrimination Parameters for Knowledge Items in the 1992 and 1996 ANES. *Black dots represent discrimination parameters β_j^C (left pane) and difficulty parameters α_j^C (right pane) for correct response options. Grey dots represent discrimination parameters β_j^{DK} (left pane) and difficulty parameters α_j^{DK} (right pane) for "don't know" response options. Vertical line segments indicate 95% highest posterior density regions (HPDs). A positive discrimination parameter indicates that a given type of response to a question is associated with more political knowledge than an incorrect response, while negative discrimination parameters indicate that a response type reflects less knowledge than an incorrect answer. Discrimination parameters for correct responses are all estimated to be positive, indicating that these responses imply more knowledge than would an incorrect response. Discrimination parameters for "don't know" responses are all negative, implying that these responses are indicative of less knowledge than both correct and incorrect answers. Difficulty parameter values are not of direct interest for the question asked here but are presented in the interest of completeness.*

33 political knowledge items, all of the discrimination parameters for correct items were estimated to be positive and all of the discrimination parameters for "don't know" responses were estimated to be negative. In fact, in only one case—the unemployment question on the 1992 ANES which nearly 80 percent of respondents answered correctly—did the 95 percent highest posterior density region for any of these discrimination parameters overlap zero. Although the finding that correct responses suggest higher levels of political knowledge than incorrect ones is hardly surprising, the finding that "don't know" responses are associated with lower levels of political knowledge than both correct and incorrect answers is not only at odds with conventional practice, which awards the same score to incorrect and "don't know" responses, but is even more strongly at odds with claims that people who say "don't know" in response to knowledge questions are typically more knowledgeable than those who give the wrong answers.[14]

5 Does the Meaning of "Don't Know" Vary by Respondent Personality?

As discussed above, a central contention of some previous works on the measurement of political knowledge has been that "don't know' responses may indicate different levels of knowledge for different types of people. In particular, it has been claimed that personality plays a large role in this process (see e.g., Mondak and Halperin, 2008, Mondak, 2010). For example, these arguments state, respondents who are more extroverted may provide responses to a question (guess) even when they have little or no idea of the correct answer, while introverts may say "don't know" unless they are virtually certain of their answer. It is argued that this process can bias the measurement of political knowledge.

To explore the role of personality in knowledge measurement, we focus on the Big

[14]Note, however, that this finding is in keeping with a few other works [e.g., Nadeau and Niemi, 1995].

Five, the model of personality that has commanded most attention from psychologists in the last three decades [Funder, 2001]. The claim underpinning Big Five research is that personality can be partitioned into five distinct traits: openness to experience, conscientiousness, extroversion, agreeableness and neuroticism. The Big Five is now the most widely used model of personality in the social sciences [Funder, 2001, Gosling et al., 2003, 506], and other traits can be explained in significant part as combinations of Big Five traits [Ozer and Reise, 1994].

Two of the Big Five traits, extroversion and agreeableness, seem most likely to distinguish people for whom "don't know" responses suggest an unusually high degree of knowledge. Extroversion may make people talkative and thus likely to offer substantive responses to questions even when they have limited knowledge. By contrast, introverts—those who score low on measures of extroversion—may be likely to say "don't know" under the same conditions (see Furnham and Medhurst 1995, Lovallo and Pishkin 1980, Swann and Rentfrow 2001). Introverts may therefore be a distinct group of people for whom "don't know" responses reflect an unusually high degree of knowledge. Agreeableness might have the same effect: agreeable subjects who intuit that the interviewer wants them to give a substantive answer might do so even when they have limited knowledge, while less agreeable respondents might say "don't know" under the same conditions. In this case, less agreeable subjects would also be a distinct group of people for whom "don't know" responses reflect an unusually high degree of knowledge.

To test these hypotheses, we need a dataset that contains measures of both political knowledge and Big Five personality traits. The 2008 Cooperative Campaign Analysis Project (CCAP) meets this requirement. The CCAP is an Internet survey of 18,250 registered voters; collectively, these voters closely match the population of registered voters in the United States in 2008 with respect to age, race, gender, education, metropolitan area, party ID, and self-reported ideology (see Ansolabehere and Rivers, 2013 for more information). The full text of all questions along with further information about its design and sampling procedure is in Appendix section B.4.

The CCAP contains 14 factual questions about politics. All are closed-ended. For 11 of the questions, a prominent figure is named and subjects are asked whether he is a Senator, a "House Member," or neither. (The figures are John Boehner, Susan Collins, John Dingell, Bill Gates, Ted Kennedy, Dennis Kucinich, Jon Kyl, Patrick Leahy, Nancy Pelosi, Condolezza Rice, and Henry Waxman.) Subjects are also asked to identify the correct description of Guantanamo Bay from a set of choices (e.g., "Fidel Castro is in a hospital there and is quite ill"), whether America imports more manufactured goods than it did a couple of years ago, and whether the dollar had become stronger or weaker in the past year.

The CCAP also includes the Ten-Item Personality Inventory (TIPI), a battery of items that tap Big Five personality traits. The items ask subjects how strongly they agree or disagree that different pairs of traits apply to them:

1. Extroverted, enthusiastic
2. Critical, quarrelsome
3. Dependable, self-disciplined
4. Anxious, easily upset
5. Open to new experiences, complex
6. Reserved, quiet
7. Sympathetic, warm
8. Disorganized, careless
9. Calm, emotionally stable
10. Conventional, uncreative

For each trait pair, subjects could choose from seven response options ranging from "disagree strongly" to "agree strongly." Responses were coded and summed to create indicators of the extent to which subjects exhibited each Big Five trait.[15] Measures of personality that are based on the TIPI correspond closely to measures based on much larger batteries [Gosling et al., 2003].

To examine the possibility that "don't know" responses mean different things for respondents of different personality types, I estimate the MNP model of political knowledge separately for ten subgroups: the groups of subjects whose TIPI answers place them in the upper or lower eights of the distributions of each of the Big Five including ties.[16] Looking only at the most extreme 25% of respondents on each of the five personality traits allows for the isolation of high- and low-trait groups on each factor, while, because of the survey's large sample size, also keeping enough respondents per subgroup to obtain relatively precise estimates of the item parameters. The ten subgroups each contain between 1,738 and 3,638 respondents.[17] This strategy was chosen instead of others—for example, modeling the effects of personality with interaction terms—for two reasons. First, it is not clear how personality might alter the relationship between knowledge and respondent answers. The choice of a specific functional form for modeling these effects would therefore be somewhat arbitrary. By contrast, separately estimating the MNP political knowledge model for

[15]Items 1 and 6 were used to measure extroversion; 2 and 7, agreeableness; 3 and 8, conscientiousness; 4 and 9, neuroticism; 5 and 10, openness to new experience. For items 1, 3, 4, 5, and 7, "disagree strongly" was coded as 1, "disagree moderately" was coded as 2, and so on. The other items were reverse-coded: "disagree strongly" was coded as 7, "disagree moderately" was coded as 6, and so on. In fact, some work in psychology has used item response modeling or similar approaches to analyze personality data (see e.g., Goldberg, 1990). Because the goal here is to separately analyze high- and low-trait groups, however, a simpler and more standard approach is used instead.

[16]This is done after dropping respondents who skipped or were not administered one or more of the political knowledge questions, which results in the retention of over 90% of observations).

[17]Because the five TIPI scales are discrete, we construct high- and low-trait groups by taking the lowest (or highest) set of values that include at least one eighth of respondents. Each subgroup therefore contains a different number of respondents.

"high" and "low" levels of each personality trait does not require assumptions about the specific nature of possible personality effects. Moreover, comparisons between "high" and "low" personality subgroups are a straightforward way of detecting such effects. If we detect them, we can proceed to further investigations of their nature.[18]

The black and grey dots in each panel of Figure 4 show estimated discrimination parameters for correct and "don't know" responses, respectively.[19] Unsurprisingly, the estimates for correct responses are overwhelmingly positive for every personality subgroup, indicating that correct answers imply a significantly higher amount of political knowledge than do incorrect answers. The estimates for "don't know" responses are overwhelmingly negative for each subgroup of these Big Five traits. This suggests that the main result of Figures 2 and 3—"don't know" responses reflect even less knowledge than incorrect answers—holds not just on average across all types of respondents but for a wide range of personality types. In particular, it holds strongly for introverts and for those who score low on agreeableness—exactly the people for whom it seems most likely that the main result would not hold.

This result is weakest for the questions about Condolezza Rice, Guantanamo Bay, imports of manufactured goods, and the strength of the U.S. dollar. These questions differed in several ways from others in the battery: they asked about different topics, had different response options, and appeared later in the survey. Any of these factors may account for the questions' relatively high "don't know" discrimination parameters. But even for these items, "don't know" discrimination parameter estimates are negative in all but three of the 140 cases (roughly 98%), with all of the HPDs for the positive estimates overlapping zero. In fact, all but nine of the "don't know" discrimination parameters (roughly

[18]Direct investigation of the connection between personality and knowledge measurement is impossible with the ANES because the ANES does not contain relevant personality measures. Even if it did, its small sample size would make investigation difficult.

[19]These estimates are based on a run of 100,000 iterations of the Gibbs sampler, dropping the first 25,000 iterations as burn-in and then storing every 25th iteration thereafter, yielding a total of 3,000 draws from the posterior for analysis.

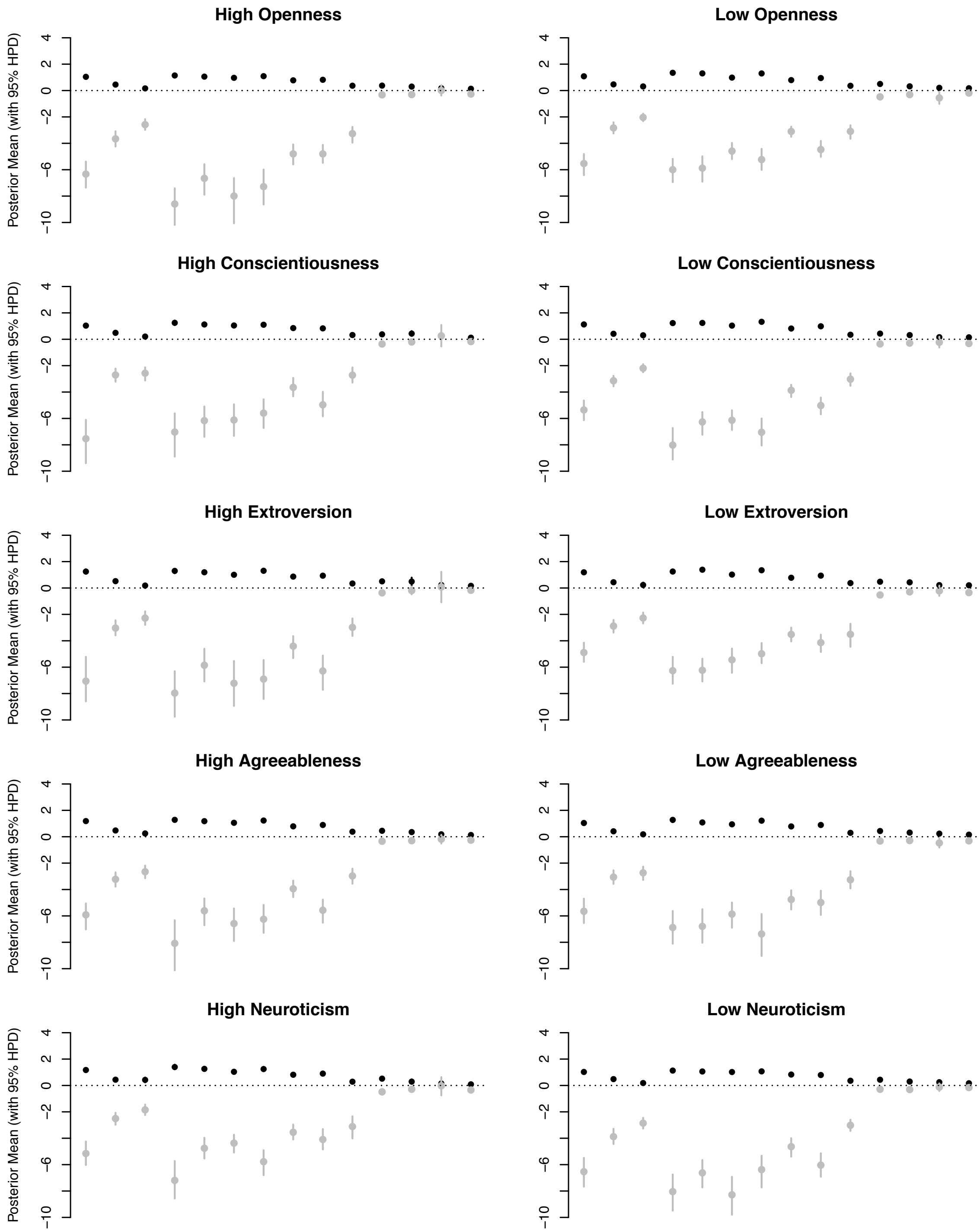

Figure 4: Discrimination Parameters and Personality Traits in the 2008 CCAP. *Each panel plots discrimination parameters for subjects scoring in the highest or lowest eighth (including ties) of a given Big Five trait. Black dots are discrimination parameter estimates for correct responses* (β_j^C)*; grey dots are discrimination parameter estimates for "don't know" responses* (β_j^{DK})*. Vertical line segments indicate 95% highest posterior density regions (HPDs). HPDs for some estimates are hard to discern because they are small.*

94%) have 95% HPDs that are entirely below zero. Furthermore, all of the discrimination parameters for correct responses are estimated to be positive and only one has a 95% HPD that overlaps zero.

Figure 5 presents the estimated difficulty parameters for correct and "don't know" responses by personality type. Although these parameters are not of central interest in determining the level of knowledge implied by each response type, it is clear that they are estimated to have very similar values across personality types for both correct and "don't know" responses. The difficulty parameter estimates for the ANES data (presented in Figures 2 and 3) appear somewhat different than those for the CCAP. The likely explanation is that while the CCAP contained only closed-ended knowledge questions, the ANES datasets used here included both closed- and open-ended items. The discrimination parameter estimates vary strongly across the open-ended ANES questions, but show a relatively strong pattern—with the parameter for correct responses being larger than the corresponding "don't know" parameter—across both the ANES and CCAP datasets.

The general trend of these personality-based analyses, therefore, remains clear: across items and personality types, knowing more about politics makes subjects more likely to answer incorrectly, less likely to answer "don't know." This again suggests that "don't know" responses are indicative of the lowest amounts of political information possible—lower even than incorrect answers. Moreover, it demonstrates that personality, at least in terms of differences between respondents with high and low levels of each of the heavily-studied Big Five personality types, does not meaningfully affect the amount of political information indicated by correct, incorrect, and "don't know" responses.

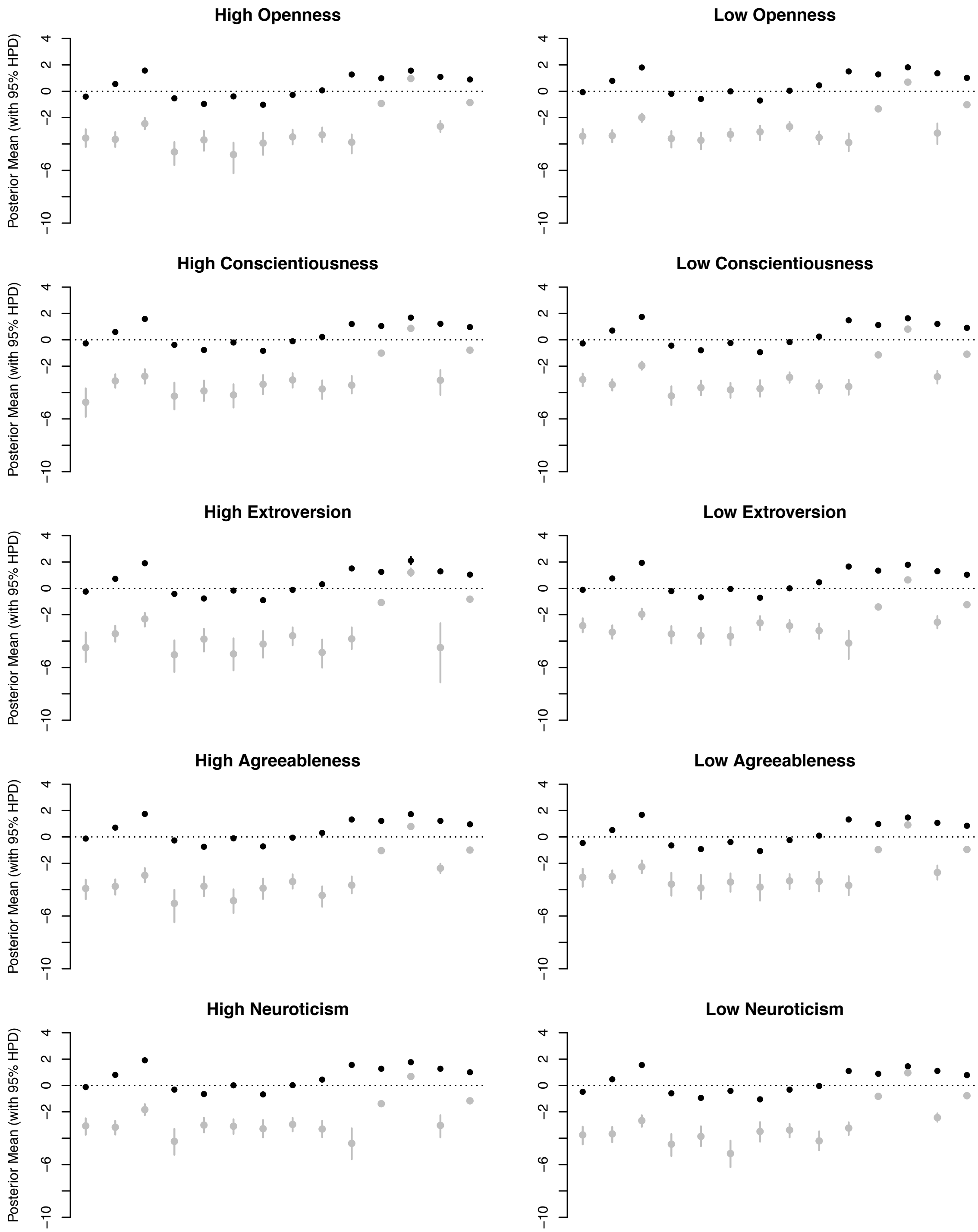

Figure 5: Difficulty Parameters and Personality Traits in the 2008 CCAP.*Each panel plots discrimination parameters for subjects scoring in the highest or lowest eighth (including ties) of a given Big Five trait. Black dots are difficulty parameter estimates for correct responses (α_j^C); grey dots are difficulty parameter estimates for "don't know" responses (α_j^{DK}). Vertical line segments indicate 95% highest posterior density regions (HPDs). HPDs for some estimates are hard to discern because they are small.*

6 How Different are Estimates of Political Knowledge Under Different Approaches?

The central question addressed here has been how much political knowledge is implied by correct, incorrect and "don't know" responses, in particular whether or not "don't know" responses should be thought to indicate partial knowledge—something between what is implied by incorrect and correct answers. The results presented above point to a clear answer to this question: "don't know" responses do not indicate more knowledge than incorrect responses, in fact they indicate less. Whereas standard binary item response models typically either assume that "don't know" responses are equivalent to incorrect ones or drop "don't know" responses, treating them as missing, the MNP item response model used here allowed for the estimation of know "knowledgeable" each of these responses types is.

A subsequent question, then, is how similar political knowledge estimates would look under different approaches. To answer this question, I compare the political knowledge estimates (posterior means for x_i's) for each respondent in the 1988 ANES using the MNP model as well as using a binary probit-link item response model and a simple average (the proportion of questions answered correctly by each respondent). For each of the later two approaches, I produce two sets of estimates: one treating "don't know" responses as incorrect and another treating them as missing. Table 1 shows the correlations between these different estimates.

Estimates of political knowledge from the MNP model are most similar to the binary IRT model in which "don't know" responses are treated as incorrect, with these two measures having a correlation of .97. This indicates that although "don't know" responses indicate less political knowledge than do incorrect responses, a binary item response model looking at whether respondents provide correct answers or not (thus combining incorrect and "don't know" responses) results in very similar estimates to the MNP model which

Table 1: Correlations Between Knowledge Estimates for 1988 ANES

	MNP IRT	Binary IRT (DK=0)	Binary IRT (DK=NA)	Simple Average (DK=0)	Simple Average (DK=NA)
MNP IRT	1	.97	.64	.95	.71
Binary IRT (DK=0)	.97	1	.72	.99	.83
Binary IRT (DK=NA)	.64	.72	1	.78	.84
Simple Average (DK=0)	.95	.99	.78	1	.85
Simple Average (DK=NA)	.71	.83	.84	.85	1

Note: Entries are correlations between estimates of political knowledge estimates. Binary IRT models are estimated using the ideal function in the pscl R library [Jackman, 2009], while simple average estimates are calculated by coding correct responses as 1, incorrect responses as 0, and "don't know" responses as specified (either 0 or missing depending on the process used), and then averaging up each respondent's score across all questions, dropping missing values.

considers each of the three response types separately. Furthermore, we see that taking a simple average (calculating the proportion of correct answers for each respondent) while treating treating DKs as incorrect produces nearly as high a correlation (.95) as estimating a binary IRT model with the same categorizations. This suggests that although there exists clear variation in the difficulty and discrimination across different questions (as shown in Figure 2), taking this into account does not have a strong effect on the estimated knowledge levels.

Political knowledge estimates treating "don't know" responses as missing values show much lower correlations with the MNP estimates, whether based on IRT or simple averaging. These lower correlations should not be surprising given that the missing value approach discards information from all "don't know" responses. Treating "don't know" responses as "partially informed," as has been advocated by some, would entail either assigning some number between zero and one to these answers before calculating simple averages or perhaps positing a full statistical model in which "don't know" responses are more informed than incorrect responses. In either case, the resulting estimates would have lower correlations with the MNP estimates produced here than those treating "don't know" responses as incorrect. In fact, the correlations between the MNP and simple average estimates could be maximized by assigning a number less than zero to "don't know" responses in the averaging procedure (although the largest possible increase is quite small—about .01).

The fact that simpler methods can produce estimates similar to those of the more complex MNP model should not imply that there are no clear benefits to using this approach. For example, model-based estimates allow for the quantification of uncertainty, which may be important if knowledge estimates are to be used either as independent or dependent variables in subsequent analyses. Using a full statistical model also accommodates straightforward hypothesis testing both dealing with item characteristics and individual knowledge levels. Most importantly for the present study, item response modeling, specifically the a multinomial setup, was used to learn about the meaning of the various response types, something that would have to be assumed, rather than tested, when using simple averages.

7 Discussion

Several recent works have charged that conventional studies understate levels of political knowledge in the mass public [Mondak, 1999, Krosnick et al., 2008, Gibson and Caldeira, 2009]. In particular, these works charge that some people who know the answers to political knowledge questions may say that they don't know, implying that "don't know" responses should be treated as suggesting some level of partial knowledge rather than a lack of knowledge. By treating "don't know" and incorrect responses equivalently, these authors argue, conventional knowledge measurement methods may overlook "hidden knowledge" in the former type of response and thereby understate the public's level of political knowledge. It has also been argued that the meaning of "don't know" responses—in particular how much knowledge they indicate relative to correct or incorrect answers—may differ systematically based on respondents' personality traits [Mondak, 2010].

I bring a new statistical model to bear on these questions. Unlike other—typically implicit—models that have underpinned political knowledge measurement, this model does not rest on any *a priori* assumption about the level of knowledge represented by "don't know" responses. I estimate this model with data from multiple ANES studies as well as the large scale nationally representative 2008 CCAP survey and consistently find that "don't know" responses indicate *less* knowledge, not more, than incorrect answers. This finding is also robust to differences in personality traits, suggesting that personality does not play a meaningful role in altering the meaning of "don't know" responses. The findings presented here are in sharp disagreement with key conclusions reached by Mondak and others, but are largely compatible with those of Sturgis et al. [2008].

In addition to the substantive findings presented here, the general modeling approach suggests a useful framework for the analysis of different response types on common survey items. For questions in which a specific ordering of response options is not obvious, researchers can use the approach outlined here to investigate, rather than assume, which options reflect higher or lower values of some unobserved latent trait. Potential applications could include estimating ideology based on questions with multiple responses or aiding in the construction of response options to include those consistent with a wide range of different latent trait values. This approach may be useful both to scholars seeking to design appropriate response options, perhaps based on pilot tests of different possibilities, or to those who wish to examine existing survey data whose response categories are already determined. These and other techniques could also be applied to investigate potential variation in the meaning of different response categories across various survey situations including factors such as inattention or respondent fatigue or interview modes such as in-person, telephone, or online.

References

Carolyn J. Anderson, Jay Verkuilen, and Buddy L. Peyton. Modeling polytomous item responses using simultaneously estimated multinomial logistic regression models. *Journal of Educational and Behavioral Statistics*, 35(4):422–452, 2010.

Stephen Ansolabehere and Douglas Rivers. Cooperative survey research. *Annual Review of Political Science*, 16:307–329, 2013.

Joeseph Bafumi and Ian Shapiro. A new partisasn voter. *Journal of Politics*, 71(1):1–24, 2009.

Darrell R. Bock. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1):29–51, 1972.

Cheryl Boudreau and Arthur Lupia. *Cambridge Handbook of Experimental Political Science*, chapter Political Knowledge. Cambridge University Press, 2011.

Joshua Clinton, Simon Jackman, and Douglas Rivers. The statistical analysis of roll call data: A unified approach. *American Political Science Review*, 98:355–70, 2004.

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

Michael X. Delli Carpini and Scott Keeter. Measuring political knowledge: Putting first things first. *American Journal ofPolitical Science*, 37(4):1179–1206, 1993.

David C. Funder. Personality. *Annual Review of Psychology*, 52:197–221, 2001.

Adrian Furnham and Sarah Medhurst. Personality correlates of academic seminar behaviour: A study of four instruments. *Personality and Individual Differences*, 19(2): 197–208, August 1995. ISSN 0191-8869. doi: DOI: 10.1016/0191-8869(95)00026-3. URL `http://www.sciencedirect.com/science/article/B6V9F-3YB56P1-2H/2/80c0fee51b8e315e816174`

Alan S. Gerber, Gregory A. Huber, and Ebonya Washington. Party affiliation, parisanship and political beliefs: A field experiment. *American Political Science Review*, 104(4):720–745, 2010.

James L. Gibson and Gregory A. Caldeira. Knowing the supreme court? a reconsideration of public ignorance of the high court. *Journal of Politics*, 71(2):429–41, April 2009.

Lewis R Goldberg. An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, 59(6):1216, 1990.

Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann, Jr. A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37(6):504–528, December 2003.

Tim Groseclose and Jefferey Milyo. A measure of media bias. *Quarterly Journal of Economics*, CXX(4):1191–1237, 2005.

James J. Heckman and James M. Snyder. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *RAND Journal of Economics*, 28:142–89, 1997.

Shanto Iyengar. Whither political information? American National Election Studies: Pilot Study Report 002253., 1986.

Simon Jackman. Estimation and inference are missing data problems: Unifying social science statistics via bayesian simulation. *Political Analysis*, 8:307–32, July 2000.

Simon Jackman. *pscl: Political Science Computational Laboratory (R library)*, 2009.

Stephen A. Jessee. Spatial voting in the 2004 presidential election. *American Political Science Review*, 103(01):59–81, 2009.

Stephen A. Jessee. Partisan bias, political information and spatial voting in the 2008 presidential election. *Journal of Politics*, 72(2):327–340, 2010.

Micheal P. Keane. A note on identification in the multinomial probit model. *Journal of Business & Economic Statistics*, 10(2):193–200, 1992.

Jon A. Krosnick, Arthur Lupia, Matthew DeBell, and Darrell Donakowski. Problems with anes questions measuring political knowledge. ANES Report. `http://electionstudies.org/announce/newsltr/20080324PoliticalKnowledgeMemo.pdf` (May 8, 2010)., 2008.

Christopher N. Lawrence. Should voters be encyclopedias? measuring political knowledge in the united states and the netherlands. Texas A&M International University. Manuscript. `http://www.cnlawrence.com/papers/irt-soph.pdf` (accessed December 27, 2009), 2009.

William R. Lovallo and Vladimir Pishkin. Type a behavior, self-involvement, autonomic activity, and the traits of neuroticism and extraversion. *Psychosomatic Medicine*, 42(3): 329–34, 1980.

Arthur Lupia. Elitism and the study of democratic competence. *Critical Review*, 18(1-3): 1–74, Winter-Summer 2006.

Robert Luskin. *Thinking about Political Psychology*, chapter From Denial to Extenuation (and Finally Beyond): Political Sophistication and Citizen Performance. New York:Cambridge University Press, 2002.

Robert Luskin and John Bullock. 'don't know' means 'don't know': Dk responses and the public's level of political knowledge. *Journal of Politics*, 73(2):547–57, 2011.

Andrew D. Martin and Kevin M. Quinn. Dynamic ideal point estimation via markov chain monte carlo for the us supreme court, 1953-1999. *Political Analysis*, 10(2):134, 2002.

Jeffery J. Mondak and Karen D. Halperin. A framework for the study of personality and political behaviour. *British Journal of Mathematical and Statistical Psychology*, 38(2): 335–362, April 2008. URL `http://dx.doi.org/10.1017/S0007123408000173`.

Jeffrey J. Mondak. *Personality and the Foundations of Political Behavior*. New York: ambridge University Press, 2010.

Jeffrey J. Mondak. Reconsidering the measurement of political knowledge. *Political Analysis*, 8(1):57, 1999.

Richard Nadeau and Richard G. Niemi. Educated guesses: The process of answering factual knowledge questions in surveys. *Public Opinion Quarterly*, 59(3):323–46, Autumn 1995.

Hans Noel. Multinomial ideal point estimation: When the decision to speak is as important as what you say. Poster, Annual Meeting of the Society for Political Methodology, July 2007.

Daniel J. Ozer and Steven P. Reise. Personality assessment. *Annual Review of Psychology*, 45:357–88, 1994.

Keith T. Poole and Howard Rosenthal. A spatial model for legislative roll call analysis. *American Journal of Political Science*, 29(2):357–384, 1985.

Markus Prior and Arthur Lupia. Money, time, and political knowledge: Distinguishing quick recall and political learning skills. *American Journal of Political Science*, 52(1): 169–83, January 2008.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL `http://www.R-project.org/`.

Georg Rasch. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danmarks Paedogogische Institut, Copenhagen, 1960a.

George Rasch. *Probabilistic Model for Some Intelligence and Attainment Tests.* Copenhagen: Danmarks Paedogogische Institut, 1960b.

Peter Rossi. *bayesm: Bayesian Inference for Marketing/Micro-econometrics*, 2012. URL `http://CRAN.R-project.org/package=bayesm`. R package version 2.2-5.

Peter E. Rossi, Greg M. Allenby, and Robert McCulloch. *Bayesian Statistics and Marketing.* John Wiley & Sons, 2005.

Keith E Schnakenberg and Christopher J Fariss. Dynamic patterns of human rights practices. *Political Science Research and Methods*, 2(01):1–31, 2014.

Boris Shor. All together now: Putting congress, state legislatures, and individuals in a common ideological space. Working Paper, 2009.

Patrick Sturgis, Nick Allum, and Patten Smith. An experiment on the measurement of political knowledge in surveys. *Public Opinion Quarterly*, 85(1):90–102, 2008.

William B. Swann, Jr. and Peter J. Rentfrow. Blirtatiousness: Cognitive, behavioral, and physiological consequences of rapid responding. *Journal of Personality and Social Psychology*, 81(6):1160–1175, 2001.

Sean Treier and Simon Jackman. Democracy as a latent variable. *American Journal of Political Science*, 52(1):201–217, 2008.