

Estimating Individuals' Political Perceptions While Adjusting for Differential Item Functioning

Stephen A. Jessee¹

The University of Texas at Austin, Government, 1 University Station A1800, Austin, TX 78712-1139, USA.
Email: sjessee@utexas.edu

Abstract

Questions about people's perceptions of politicians or other political actors are of central interest in a wide variety of research areas. But measuring these perceptions is difficult in part because respondents may use survey response scales in different ways. In a classic article, Aldrich and McKelvey (1977) introduce a method adjusting for such differential item functioning by assuming that all respondents perceive political stimuli identically. I propose a modeling approach built on the Aldrich and McKelvey framework but incorporating anchoring vignettes. This approach allows for scale use adjustments without assuming that all respondents perceive a given politician identically. I apply this model to data on Americans' perceptions of parties, elected officials, and other political actors, showing that, contrary to previous arguments, most variation in ideology ratings is due not to differing scale use, but to differences in underlying perceptions. Specifically, while perceptions of Republican politicians and the Republican party show no significant differences by respondent partisanship, Democratic and Republican respondents differ strongly in their perceptions of the ideology of Democratic political actors as well as the Supreme Court.

Keywords: ideology scores, Bayesian analysis, latent variables

1 Introduction

Understanding how people perceive politicians and other political actors is of central importance to many research areas in political science. But measuring these perceptions is far from straightforward. Of central concern is whether survey respondents may interpret the meaning of response scales in different ways. The use of these response scales is common in many areas of survey research. Perhaps most notably, in American politics, the seven-point ideological rating scale has been analyzed in countless studies. But it has long been known that different people may interpret the meaning of this and other scales in different ways (e.g. Brady 1985). For example, if one respondent rates the ideology of Hillary Clinton as a 3 (somewhat liberal) and another respondent rates her as a 1 (very liberal), this difference could be caused by two different things. First, the two respondents might perceive these two candidates identically but use the rating scale differently to express their perceptions. This is the classic version of differential item functioning (DIF). In this case, it could be that the first respondent thinks about the terms "somewhat liberal" and "very liberal" differently than does the second respondent. But this difference in responses could also be due to differing perceptions of Clinton rather than scale use differences. In other words, even if these two respondents actually interpreted and used the seven-point ideology scale in the same way, it could be that the first respondent actually thinks that Clinton has a different ideology than the second respondent does perhaps because they hold different understandings of the policy positions Clinton supports or otherwise differ in their perceptions of her.

Aldrich and McKelvey (1977), in a classic article, introduce a method for estimating and adjusting for differences in scale use. Later, Hare *et al.* (2015) propose a Bayesian version of a similar model, which retains the benefits of the original, while better handling missing data and more appropriately quantifying uncertainty in the estimates. This Aldrich–McKelvey (hereafter A–M) approach assumes that all respondents perceive the same ideology for political

Political Analysis (2020)

DOI: 10.1017/pan.2019.47

Corresponding author
Stephen A. Jessee

Edited by
Jeff Gill

© The Author(s) 2020. Published by Cambridge University Press on behalf of the Society for Political Methodology.

stimuli (e.g. candidates, elected officials, or political parties), but that their ratings may not agree due to differences in the way each person interprets the response scale.¹ In other words, the works of Aldrich and McKelvey (1977) and Hare *et al.* (2015) are focused on estimating and correcting for the nature of respondent-specific DIF and do this by assuming that all respondents perceive the stimuli mentioned in surveys (typically politicians or political parties) identically.

The political actors thus serve as “anchors” that allow for the adjustment of individuals’ responses onto an interpersonally comparable scale. Any difference in respondents’ ratings of a politician’s ideology is assumed to be driven not by one respondent actually thinking of the politician as more liberal or conservative than another respondent does, but instead by respondents thinking of the meaning of the response scale differently. This necessarily prevents the estimation of varying perceptions of a given stimulus across respondents. It also prevents the investigation of how much of respondents’ response differences are the result of DIF and how much they are caused by differences in actual perceptions. If, for example, Republican respondents tend to rate Clinton as more liberal on the seven-point scale than Democrats do, the A–M approach cannot determine whether this is due to differences in scale use, differences in underlying perceptions, or both. This is because the A–M framework is built on the assumption that these differences are driven by DIF rather than true perceptual differences.

DIF is most commonly defined in the context of standard item response or similar models (e.g. Rasch 1966). In the A–M framework, the model is somewhat different. In particular, the slope and intercept parameters vary by individual respondent rather than by item, while what is typically called the “ability” parameter in the context of item response models (here the true ideological location of a political stimulus) varies by item. Therefore, the exact interpretation of DIF is different and, one could argue, somewhat less clear in the A–M framework. Here I think of DIF as occurring when two respondents who hold identical perceptions of a political actor’s ideological position would provide different answers in expectation when rating that actor on the same response scale (e.g. the seven-point ideological rating scale). DIF is thus defined here based on the relationship for individual respondents between their perceptions and their responses. Under this definition, examining the degree of DIF in Americans’ ratings of political actors requires being able to separately estimate individual respondents’ perceptions of political actors and also estimating the characteristics of the relationship between each individual’s perceptions and her responses (i.e. her seven-point ideology ratings of each stimulus).

In their discussion of the A–M framework, Hare *et al.* (2015) consider this identical perceptions assumption, arguing, for example, that “ideological centrists and extremists view the political world differently” (p. 761). They further suggest that people may rate politicians they generally agree with as more moderate, while pushing ones they disagree with to the extremes of the ideological scale. But it is not clear whether these sorts of dynamics represent pure DIF or differences in underlying perceptions. In other words, do extremists (or other types of respondents) simply interpret the values of the rating scale differently or do they hold different perceptions of the stimuli being rated? Under the standard A–M setup, it is not possible to differentiate between these two things without making very strong assumptions. For example, one could obtain estimates of individual DIF characteristics if one assumes that all respondents perceive political stimuli in exactly the same way. To the extent that Aldrich and McKelvey (1977) and Hare *et al.* (2015) maintain the assumption of identical perceptions, it appears to be due more to the limitations of the models used rather than a true belief that individuals hold identical perceptions in reality.

¹ I use A–M to refer to the general framework for adjusting for DIF, including the work of Aldrich and McKelvey (1977) and Hare *et al.* (2015).

One strategy that has been used to estimate and adjust for differing interpretations of various response scales, for example, political efficacy, is the inclusion of so-called anchoring vignettes. These vignettes are survey items in which hypothetical stimuli are described and respondents are asked to rate them on a given responses scale (King *et al.* 2004; King and Wand 2007). By giving respondents the same vignettes, we can learn about their response scale use from the variation in their ratings of the vignettes. Then, their ratings of other stimuli can be adjusted based on the nature of the DIF revealed by the vignette responses. This approach has been used in a variety of settings, including the estimation of ideological positions of European political parties (e.g. Bakker *et al.* 2014; Struthers, Hare, and Bakker 2019). A related approach was taken by Simas (2017), who analyzes ideological ratings data with vignettes using a compound hierarchical ordered probit and finds that Republican respondents are more likely to use lower response options to rate the same vignettes that Democratic respondents would rate higher.

In this paper, I propose an approach built on the A–M framework, which incorporates anchoring vignettes to allow for the simultaneous estimation of each respondent’s particular scale use characteristics as well as their perceptions of each political stimulus. After introducing the modeling approach and describing its advantages over the standard A–M framework, I illustrate its potential using data on citizens’ perceptions of American political parties and politicians from the 2016 Cooperative Congressional Election Study (CCES). When compared to the standard A–M setup, the results of this new model indicate that the in ideology ratings between Democratic and Republican respondents are mostly due not to DIF, as the A–M approach would assume, but rather to differing perceptions of the actual ideology of politicians, political parties, or other things. While Democratic and Republican respondents do have some differences in the way they use ideological rating scales, these differences are dwarfed by the large inter-party differences in perceptions, particularly of liberal stimuli such as Hillary Clinton or Barack Obama. Analyzing additional data from the American National Election Studies (ANES), I also provide evidence that these basic patterns of partisan perception have held for several decades. Finally, I discuss how the results of this application cast doubt on the assumptions underpinning the A–M framework, suggesting that the use of vignettes can help overcome these problems.

2 Model: Estimating DIF-Adjusted Individual Perceptions

The model here begins with the basic setup of Aldrich and McKelvey (1977) and the Bayesian version later proposed by Hare *et al.* (2015), with some minor changes. This model is related to a standard latent factor model but with the slope and intercept terms indexed by individual and the latent trait value indexed by item. The central benefit of this setup is that respondents’ ratings of a stimulus (a politician, political party, or, here, a hypothetical individual described in a short vignette) are assumed to depend on the stimulus’s true position on an underlying ideological scale as well as the individual respondents’ scale use (or DIF) parameters.

Formally, the rating for respondent i of stimulus j , where a stimulus can be either an actual political actor (candidate, party, etc.) or a vignette describing a hypothetical person, is denoted as z_{ij} and assumed to be generated according to the model

$$z_{ij} = \alpha_i + \beta_i \zeta_j + \varepsilon_{ij}, \quad (1)$$

where $\varepsilon_{ij} \sim N(0, \tau_j^2)$ with these error terms assumed to be independent across both individuals and stimuli. The actual (underlying) ideology of stimulus j is ζ_j , while α_i and β_i are called the shift and stretch terms, respectively. Intuitively, we can think of a respondent’s shift term α_i as the expected rating she would give to a stimulus whose underlying ideology is zero. A higher (lower) value of α_i indicates a respondent that is more prone to use the higher (lower) values on the seven-point ideology scale, i.e. someone who tends to rate stimuli more conservatively (liberally).

The stretch parameter β_i indicates how strongly a respondent's rating of a stimulus is related to these stimuli's underlying ideology. A stretch parameter of zero would mean that the respondent's ratings are unrelated to the ideology of stimuli. A positive stretch parameter indicates that a respondent will tend to rate more conservative stimuli higher on the liberal–conservative scale. Some respondents could even have negative β_i 's which would suggest that as a stimulus becomes more liberal, the respondent is likely to rate them as more conservative, and vice versa.

Because each respondent in applications of the A–M model typically provides a relatively small number of ratings, estimating these respondent-specific parameters can be difficult. Here, I employ a hierarchical structure that allows for partial pooling of these DIF parameters, which should result in more accurate estimates. I adopt a Bayesian approach and use the following hierarchical prior structure for the shift and stretch parameters:

$$\begin{aligned}\alpha_i &\sim N(\mu_\alpha, \sigma_\alpha^2) \\ \beta_i &\sim N(\mu_\beta, \sigma_\beta^2)\end{aligned}$$

with accompanying hyperpriors:

$$\begin{aligned}\tau_j, \sigma_\alpha, \sigma_\beta &\sim U(0, 100) \\ \mu_\alpha, \mu_\beta &\sim N(0, 10^2).\end{aligned}$$

Standard normal priors are used for the latent ideologies of the stimuli ζ_j subject to the restriction that they have mean zero, variance 1 across all stimuli, which identifies the model. This restriction also allows for the interpretation that values of ζ_j near zero are moderate or centrist, while positive (negative) ones are more conservative (liberal).²

As discussed above, in the standard A–M approach, all respondents are assumed to base their ideological ratings on the same positions for a given stimulus as represented by η_j . Any systematic differences between respondents in their ratings are thus attributed to DIF as captured by α_i and β_i .³ In order to relax the assumption of common perceptions across individuals, we require a separate set of “anchors” that can be used to estimate and adjust for each respondent's particular response scale use characteristics. Imagine, then, that in addition to respondents stating their perceptions of the ideology of several different political actors, each respondent is also asked to rate the ideology of several hypothetical individuals, each described in a vignette by their positions on multiple policies. The approach here uses these vignettes, rather than actual politicians, as anchors, allowing for the adjustment of respondent ratings to correct for DIF.

Using anchoring vignettes in this way has two key advantages. First, assuming that all respondents will view a hypothetical individual, described in the same way to each respondent, as having the same ideology may be more plausible than assuming that all respondents would view a real world political actor (e.g. Donald Trump or Nancy Pelosi) in the same way. Second, by adding a second set of stimuli (the vignettes), it is possible to estimate respondent-specific DIF parameters and then to use those estimates to adjust each respondent's ratings of other (nonvignette) stimuli in order to estimate each respondent's potentially different perception of a given politician while also adjusting these perceptions for respondent-specific DIF.

To accomplish this, I estimate the parameters of the model described above, using respondents' ratings of the hypothetical individuals described in the vignettes. We thus obtain

- 2 The directionality restriction is imposed through initial values and at each iteration of the sampler, it is checked whether the scale has flipped. In all model runs presented later, samples for μ_β are positive in each iteration, confirming that the scale does not flip.
- 3 As mentioned above, another interpretation of this (and one that Hare *et al.* 2015 appear sympathetic to) is that the shift and stretch parameters may capture both DIF, as I define it here, and differing perceptions across individuals.

estimates of the ideological position of each vignette and, more importantly for our purposes, estimates of each respondent's DIF parameters. This later set of parameters tells us how each respondent uses the ideology rating response scale. Using these estimates, it is then possible to obtain DIF-adjusted estimates of each respondent's perceptions of each politician based on their estimated individual shift and stretch terms. This is done simply by transforming respondent i 's rating of politician k , denoted as z_{ik} according to

$$x_{ik} = \frac{z_{ik} - \alpha_i}{\beta_i}. \quad (2)$$

In other words, the vignette anchors setup allows us to gain the benefits of the A–M framework—specifically adjusting ideology estimates for each respondent's scale use—while also estimating each respondent's perceptions of the ideological positions of various real world political actors. We simply transform their ratings by “undoing” the particular shift and stretch that each individual utilizes in her responses on the seven-point ideology response scale.⁴ Specifically, for each iteration of the sampling procedure used to estimate the model's parameters, we take the values of α_i and β_i for that iteration and, for each respondent's perception of each politician z_{ik} , use the formula in equation (2) to produce a corresponding draw from the posterior for each of these DIF-adjusted perceptions x_{ik} . From these sets of posterior draws, we can easily obtain estimates (e.g. posterior means) and measures of uncertainty for each respondent's DIF-adjusted perception of each politician.

These estimated perceptions can be used to answer important questions about whether different types of people tend to perceive political actors differently. Furthermore, these estimates can actually be used to test the assumption, inherent in the A–M approach, that respondents only differ in their scale use, but all respondents will perceive a given politician identically. Perhaps more interestingly, we can test not whether this statement is exactly true, but whether it is a reasonable approximation. For example, estimates from the vignette anchors approach can help determine how much of the variation in reported perceptions across respondents is attributable to DIF and how much is instead due to differences in actual perceptions of the ideology of the political actors being rated. In the following section, I apply this approach to data on Americans' perceptions of parties, elected officials, and other political actors.

3 Application: Partisan Perceptual Differences in the 2016 CCES

Popular perceptions of political actors play a central role in democratic accountability. In particular, citizens must be able to form perceptions of the policy positions and overall ideologies of elected officials and candidates for office in order to vote in a way that encourages representation of their views. The work of Campbell *et al.* (1960) and others, however, emphasizes the role of partisanship as a “perceptual screen” through which citizens view the political world. If citizens' political perceptions differ systematically and, in particular, if these systematic differences are strongly related to voters' partisan predispositions, this could distort the basic mechanics of democratic accountability.

I use the vignette-based A–M model introduced above to investigate a key component of partisan perceptual influence in the United States, namely the degree to which individuals' party identification (PID) relates to their perceptions of the ideology of politicians and political parties. The results show that while partisan differences in ideological evaluations have sometimes

4 Note that this is the same approach used by Aldrich and McKelvey (1977) to obtain DIF-adjusted estimates of respondents' own self-placed ideology under the assumption that all respondents perceive politicians identically. It should be noted, though, that if one believes α_i and β_i capture not only DIF but also differences in true perceptions, it is not clear that these rescaled self-perceptions should be adjusted based on perceptual differences. Of course, without a method for parsing out DIF and differences in true underlying perceptions, one is forced to choose between unadjusted self-placements and ones that are adjusted based on some unknown mix of DIF and perceptual differences.

previously been attributed to DIF, under which survey different respondents perceive response scales in different ways, this cannot account for the partisan discrepancies found in perceptions of politicians' ideologies. While partisanship does appear to be related to response scale use, this relationship is shown to be quite small in magnitude when compared with partisan differences in actual perceptions of political actors.

Specifically, while perceptions of more conservative political actors are similar for respondents of all PID, Democratic and Republican respondents perceive other politicians quite differently. Republican respondents perceive Democratic politicians and the Democratic Party as significantly more liberal than do Democratic respondents. These differences remain strong even after adjusting ratings for DIF using respondents' ratings of anchoring vignettes. In a later section, I also show that these findings appear to hold similarly over the past several decades, although the partisan differences are estimated to be smaller in magnitude in earlier years.

Since the classic work of Converse (1964), several notable accounts have been provided for how partisans may form evaluations of the ideology of political actors. Brody and Page (1972) discuss projection, in which people assume that a politician who shares their party affiliation has the same views as themselves (see also Conover and Feldman 1982). Relatedly, Brady and Sniderman (1985) argue that respondents will tend to assign characteristics similar to their own to politicians they like and do the reverse for those they dislike. Levendusky and Malhotra (2015) show that people view opposing partisans as more ideologically extreme and Westfall *et al.* (2015) argue that perceptions of polarization differ based on respondent characteristics such as political engagement and strength of partisanship. Simas (2018) finds that individuals perceive more ideological heterogeneity within their own political party than in the opposing party. Other work has also focused on asymmetries in the nature of ideology between the two parties. For example, Grossmann and Hopkins (2016) argue that the modern Republican party can most appropriately be thought of as an ideological movement, while Democrats represent more of a coalition of group interests. These arguments might suggest that ideological perceptions may not be symmetric across the two parties.

Here, I focus on perceptions of overall ideology as represented by the so-called ideological scale. The use of ordinal ideology scales has a long history in American survey research, and questions regarding how liberal or conservative a political actor is perceived to be, as well as whether these perceptions differ by respondent characteristics such as partisanship, are of obvious importance. But it has long been known that different people may interpret the meaning of this scale in different ways (e.g. Brady 1985). For example, Democratic respondents could associate a different meaning with a term such as "somewhat liberal" than do Republican respondents. Therefore, looking at the raw values of these placements may produce inappropriate conclusions if the observed differences result from some unknown combination of DIF and actual differences in perceptions.

In order to assess how results from the standard A–M setup, which uses politicians, parties, or other political actors to anchor the ideological scale, change when using descriptions of hypothetical individuals (so-called vignettes), I analyze data from the University of Texas module of the 2016 CCES. This module was administered to 1000 respondents and included several sets of ideological placement items, each asking respondents to indicate a rating on a seven-point scale ranging from "very liberal" to "very conservative," following the standard format.⁵ In addition to rating their own ideology on this scale, respondents were asked to rate six specific politicians: Barack Obama, Hillary Clinton, Donald Trump, the Democratic Party, the Republican Party,

⁵ Respondents who did not rate at least half of the political stimuli and at least half of the hypothetical vignettes were dropped, leaving a total of 840 respondents.

[NAMERAND3] has the following policy positions:

- **opposes** a ban on assault rifles
- **opposes** raising the federal minimum wage to \$12 an hour by 2020
- **supports** allowing gays and lesbians to marry legally
- **supports** eliminating mandatory minimum sentences for non-violent drug offenders
- **supports** giving Environmental Protection Agency power to regulate Carbon Dioxide emissions

How would you rate this individual?

Figure 1. Example of anchoring vignette. Respondents were given six of these hypothetical individuals and asked to rate each one on a 1–7 ideology scale. The full list of vignette wordings can be found in the appendix.

and the Supreme Court.⁶ Beyond these standard ideological rating questions, I also showed each respondent six vignettes, each one describing a hypothetical individual through their positions (support or oppose) on five different specific policies such as banning assault weapons, raising the minimum wage, or allowing same-sex couples to marry. The hypothetical vignettes, which were presented in random order to respondents, ranged from consistently liberal (five out of five policy positions) to consistently conservative (five out of five conservative positions). While some of the policies listed overlapped across vignettes, they were not identical. An example of these items, each of which described a person whose name was randomized from a list of six possible names, is shown in Figure 1.⁷ The full list of questions, as well as a discussion of the motivation behind the design of these vignettes, is available in the online appendix.

Figure 2 plots the average of raw (unadjusted) seven-point scale ratings of the six politicians (left pane) and the six vignettes about hypothetical individuals (right pane) separated by respondent PID.⁸ The ratings of Obama, the Democratic Party, Clinton, and the Supreme Court show quite strong associations with PID. It is clear that Republican identifiers rate these stimuli as much more liberal on average than other respondents do. These partisan differences are much less prominent for ratings of Trump and the Republican Party.

Ratings of hypothetical individuals, as described in the vignettes, show much smaller differences by respondent partisanship. This suggests that respondents' scale use may not be the only, or even the primary, driver of these partisan differences in respondents' ratings of politicians. If the differences, for example, the tendency of Republicans to rate more liberal stimuli quite differently from the way Democrats do, were due to differences in how Democratic and Republican respondents interpret the meaning of the rating scale's values, we would expect to see similar partisan differences in their ratings of the vignettes.

The findings presented in Figure 2 are suggestive, but in order to better understand these patterns, we need a more rigorous analysis. The goal of this analysis is to assess whether the differences in scale use previously found between partisans in the United States (e.g. Aldrich

6 Although the Supreme Court is not a political institution (at least not an elected or partisan one), I include this item because it was asked of all respondents and also because it is informative to see whether perceptions of it follow similar patterns to perceptions of elected officials or parties. An item asking respondents to rate the ideology of Merrick Garland was also included in the CCES common content, but for some reason, the vast majority of respondents are listed as not having been shown this question. Therefore, it is not included in these analyses.

7 These names were included to prompt respondents to think about a (hypothetical) person who held the policy views listed. Respondent ratings of the vignettes' ideologies did not appear to be influenced by which name was randomly assigned to each one. F-statistics from a pooled regression of respondents' ratings of all six vignettes on a set of dummy variables for which name was shown did not come close to statistical significance. Moreover, in separate regressions predicting deviations from the average rating for each of the six vignettes with no intercept and dummy variables for the randomly assigned names, none of the 36 total coefficients estimated were statistically significant ($p < 0.05$) and only two of them were significant at the 0.1 level.

8 I use the term "politicians" for simplicity to refer to the actual political actors—candidates, elected officials, political parties, and the Supreme Court. "Vignettes" refers to the hypothetical individuals described to respondents through their issue positions.

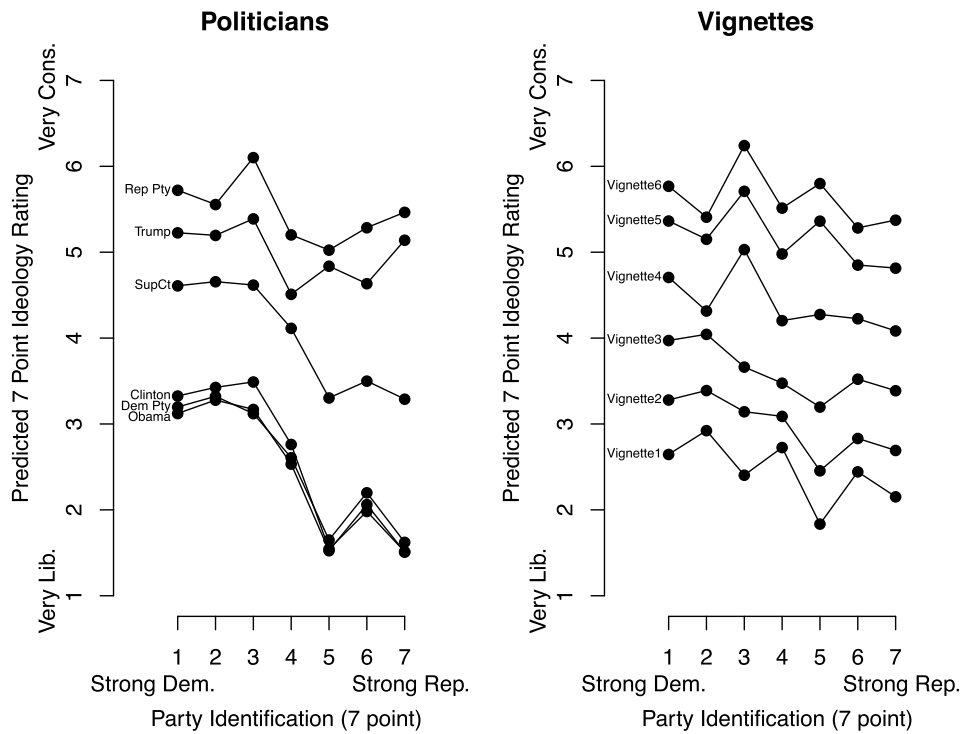


Figure 2. Seven-point scale ratings of politicians and vignettes by PID. *Left (right) panel plots averages of raw, i.e. unadjusted, seven-point ideology ratings for each politician (vignette) calculated separately for each respondent PID category. The appendix shows full question wordings for the six vignettes.*

and McKelvey 1977, Hare *et al.* 2015) are actually due to DIF or whether they result from differences in actual perceptions. To this end, I estimate the model described in the previous section in two ways. First, I estimate the model in “politician anchors” setup that is standard in the A–M framework, where all respondents are assumed to view the six political actors identically, but respondents are allowed to differ in their use of the seven-point scale. In the notation of equation (1), the stimuli terms ζ_j in this politician anchors setup represent the true underlying ideological positions (assumed to be invariant across respondents) of Obama, Clinton, and others, while α_i and β_i are the respondent-specific “shift” and “stretch” terms, respectively, describing how each respondent translates a stimulus’s underlying ideology into a seven-point scale rating. In this “politician anchors” setup, the data used (z in equation (1)) is a matrix of ideological ratings, where each row represents a respondent and each column represents ratings of a given politician.

Second, I estimate the same model, but instead of using actual political actors to anchor the scale, I use the six vignettes describing the policy views of hypothetical individuals as the anchors. In this “vignette anchors” setup, η_j from equation (1) represents the true position of a hypothetical individual described in a vignette, while z_{ij} is respondent i ’s rating of the ideology of that hypothetical individual. The assumption in this setup is that respondents perceive these vignettes in the same way. This might be seen as more plausible given that they all receive the same information about each vignette. By contrast, respondents may differ in the information they have about actual politicians or parties and respondents may harbor biases, partisan or otherwise, that color their perceptions of real world political actors. In addition to resting on this more plausible anchoring assumption, the vignette anchors model allows for the estimation of DIF parameters α_i and β_i for each respondent while also providing estimates of each respondent’s perception of the ideology of politicians. This later quantity can be estimated by using the same approach typically used in the standard A–M setup to estimate a respondent’s ideology based on her ideological self-placement, given her estimated DIF parameters (see equation (2) above).

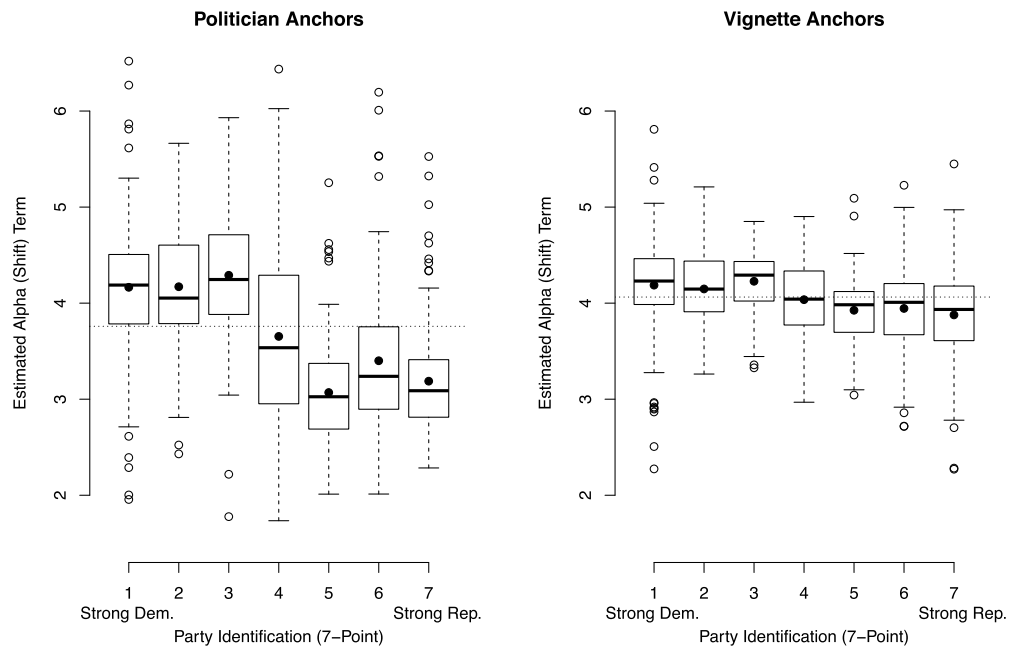


Figure 3. Boxplots of alpha (shift) term by partisanship for politician anchors and vignette anchors estimates. Horizontal tick marks indicate PID category medians and solid dots indicate PID category means. Horizontal dashed lines indicate overall means for all respondents.

Therefore, using the vignette anchors approach, it is possible to separate ordinary DIF from differences in the perceptions of politicians' ideologies. In this “vignette anchors” setup, the data used (z in equation (1)) is a matrix of ideological ratings, where each row represents a respondent and each column represents ratings of a given hypothetical individual as described in a vignette.

Both the standard “politician anchors” model and the “vignette anchors” model are estimated using JAGS (Plummer *et al.* 2003). Each estimation begins with 50,000 burn-in iterations which are discarded, then 250,000 iterations are run, storing every 50th iteration for a total of 5000 iterations saved. Multiple convergence and sampling diagnostics suggest that the sampler converged properly and that the number of iterations stored was sufficient. Replication data and code is available on the *Political Analysis* Dataverse.

Figure 3 plots estimates (posterior means) for respondent-specific α_i (shift) parameters from the model estimated under the politician anchors setup and, separately, those estimated under the vignette anchors setup against respondents' PID.⁹ As the left panel shows, the standard politician anchors approach produces shift parameter estimates that differ strongly by PID. Democratic respondents have much higher α_i values than Republicans do, with independents falling in between. This finding is broadly consistent with Hare *et al.* (2015) and has been interpreted as showing that Democrats and Republicans interpret the seven-point scale and its terminology such as “very liberal” or “somewhat conservative” differently.¹⁰ But as discussed above, these partisan differences could be due to ordinary DIF or due to Democrats and Republicans perceiving politicians as having different ideologies. The standard A–M setup cannot arbitrate between these two possibilities because it estimates the DIF parameters by *assuming* that all respondents perceive a given politician's ideology as being the same on the underlying ideology scale.

⁹ Note that because each respondent provides only six ratings (six politicians or six vignettes) under a given model, the shift and stretch parameters of each respondent are estimated with a good deal of uncertainty for any given respondent. The averages by partisanship, however, which are of central interest here, are estimated with much more precision.

¹⁰ See Hare *et al.* (2015) Figure 1, but note that they transform respondents' ideology ratings by subtracting 4 before analysis, unlike here where the ratings are used in their raw form.

Table 1. Relationship between shift or stretch parameters and party identification for politician anchors and vignette anchors models. Table shows estimates from linear regression models predicting either shift (α_j) or stretch (β_j) terms from politician anchors and vignette anchors models.

DV:	Politician anchors		Vignette anchors	
	α (Shift)	β (Stretch)	α (Shift)	β (Stretch)
Intercept	4.47 (0.05)	0.98 (0.07)	4.27 (0.03)	1.16 (0.05)
Party ID	-0.19 (0.01)	0.10 (0.02)	-0.06 (0.01)	0.01 (0.01)
<i>n</i>	828	828	828	828
Residual SE	0.76	1.01	0.43	0.78
R^2	0.22	0.04	0.07	0.001

The right panel of Figure 3 shows estimates of the same shift parameters by PID, this time from the vignette anchors setup. Although there remain partisan differences, they are much smaller in magnitude than in the politician anchors model. This suggests that while there are some differences in the ways that Democratic, Republican, and independent respondents utilize the ideological rating scale, much of the partisan difference estimated in the candidate anchors setup was not true DIF but rather differing perceptions of the candidates.

We can also look at these relationships in a more formalized way. Table 1 shows the results of linear regression predicting shift parameters with PID under the candidate anchors and vignette anchors setups. In both cases, there is a negative and highly significant coefficient estimated on PID, but the magnitude is much larger for the politician anchors parameters. Furthermore, the proportion of the shift parameter’s variance across respondents that is explained by PID is roughly three times larger in the politician anchors estimates than in the vignette anchors setup.

Figure 4 plots estimates of the stretch (β_j) parameters against PID from the two setups. For the politician anchors setup, Republicans appear to have a larger stretch term, meaning that increasing the ideology of a politician by a given amount would cause them to increase their seven-point scale rating more than a Democrat would. The vignette anchors estimates, by contrast, do not show a strong relationship between PID and the stretch terms. Table 1 also shows the results of linear regressions predicting the stretch parameters with PID for both politician anchors and vignette anchors estimates. The corresponding vignette anchors estimates of β_j show essentially no relationship with respondents’ partisanship.

Another way to look at these results is to ask what the predicted relationship would be between a politician’s or vignette’s underlying ideological position (ζ_j) and respondents’ seven-point ratings of them for Democrats, independents, and Republicans. Figure 5 plots these predicted relationships, using the fitted values for the shift and stretch terms for strong Democrats, pure independents, and strong Republicans, based on Table 1’s estimates. The left panel of Figure 5 plots relationships based on the politician anchors setup. The most obvious feature of this plot is the high degree of divergence between partisans, particularly when evaluating more liberal politicians. For example, a politician at -1 (roughly where Obama’s ideological position is estimated under the politician anchors setup) would be predicted to be rated nearly two points more liberal on the seven-point scale by a strong Republican than by a strong Democrat—an expected rating of 1.5 compared to 3.2, with independents falling roughly midway between the two partisan groups. This divergence narrows as respondents assess more conservative stimuli such that a Democratic and a Republican respondent presented with a stimulus whose ideology is 1.28 (the estimated position of the Republican party from the politician anchors setup) would differ by 0.4 in their expected seven-point rating. While this difference is still important, the partisan

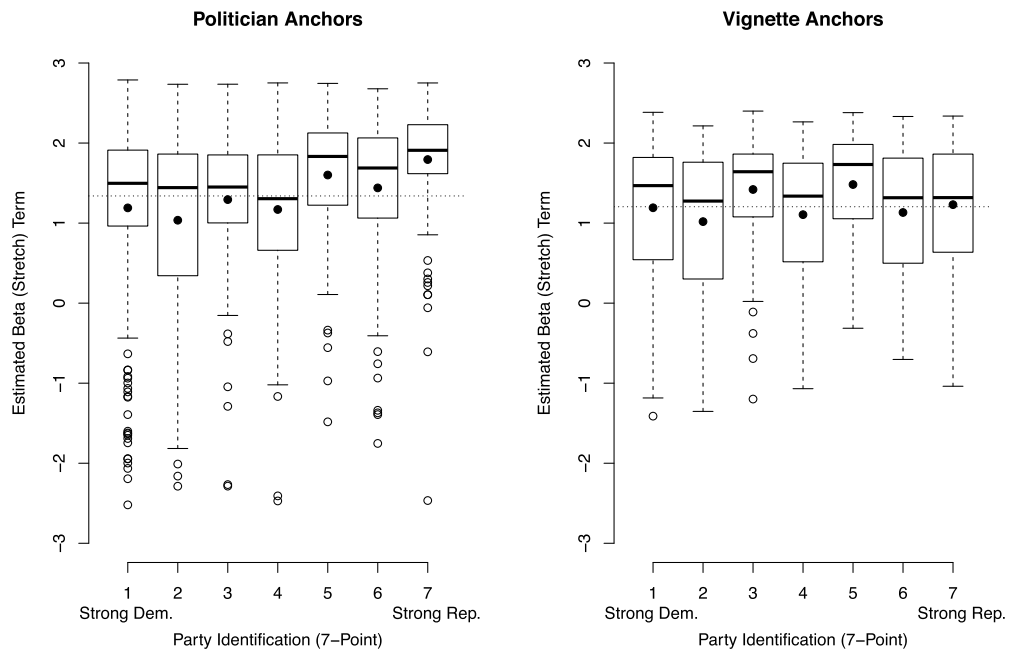


Figure 4. Boxplots of beta (stretch) term by partisanship for politician anchors and vignette anchors estimates. Horizontal tick marks indicate PID category medians and solid dots indicate PID category means. Horizontal dashed lines indicate overall means for all respondents.

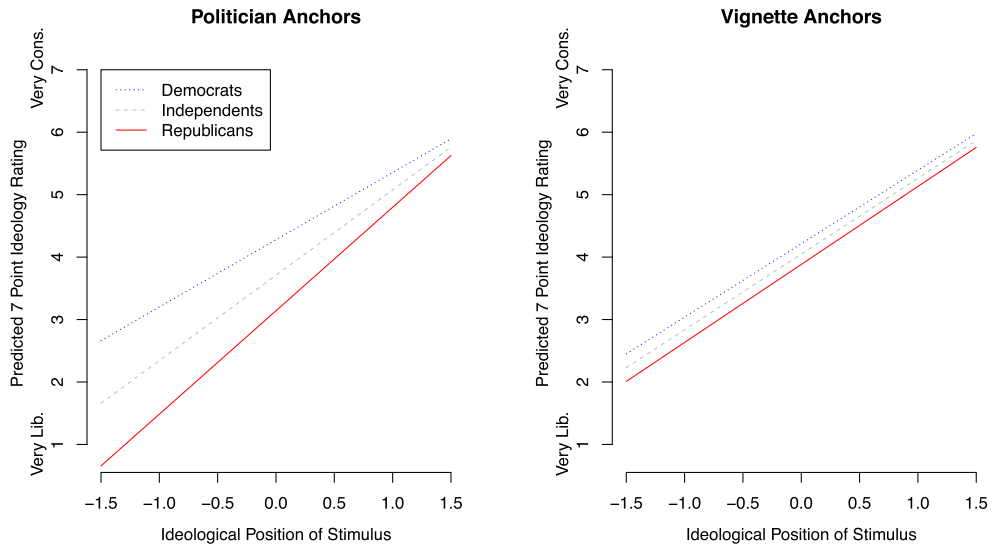


Figure 5. Predicted seven-point scale ratings of stimuli as a function of underlying ideology, by respondent PID, from politician anchors (left pane) and vignette anchors (right pane) estimation. The left panel plots predicted ratings of a politician as a function of the politician’s underlying ideology based on the average respondent-specific intercept and slope estimates from the politician anchors setup (where α_i and β_i are estimated based on respondents’ ratings of politicians). The right panel plots predicted ratings of a vignette as a function of the vignette’s underlying ideology based on the average respondent-specific intercept and slope estimates from the vignette anchors setup (where α_i and β_i are estimated based on respondents’ ratings of vignettes).

difference in ratings of conservative stimuli is less than one-fourth the size of the difference for a more liberal one.

As discussed above, however, the standard politician anchors setup of the A–M model assumes that all respondents view the ideology of a given political actor in the same way. This results in the

shift and stretch parameters capturing not only DIF in its traditional form—people interpreting and using the seven-point scale values differently—but also any partisan differences in perceptions of the actors' ideological positions. In other words, the large partisan differences between the seven-point scale ratings of a liberal stimulus could be due to Democrats thinking of terms like “very liberal” or “somewhat liberal” in different ways than Republicans do. But this same pattern could also result from partisan respondents understanding the meaning of the response scale identically but perceiving politicians' ideologies differently.

By contrast, the vignette anchors setup estimates the DIF parameters (α_i and β_j) for each respondent based on the assumption that they perceive the hypothetical vignettes, rather than real politicians or parties, similarly. As Figures 3 and 4 as well as Table 1 showed, the partisan differences in the shift and stretch terms for the vignette anchors setup are much smaller than that in the politician anchors setup. The right panel of Figure 5 plots the expected ratings under the vignette anchors setup for strong Democrats, strong Republicans, and independents as a function of a stimulus's underlying ideology (ζ_j). While some partisan difference exists, these differences are much smaller than those estimated under the assumption that all respondents perceive politicians identically.

Appendix section A3 presents the results of additional regression specifications predicting these DIF terms. Most notably, the strong relationship between the shift terms and PID remains when controlling for political interest, education, and a policy-based ideology measure. Ideology does not appear to predict either of the DIF terms and does not meaningfully change the estimated coefficient on partisanship under any of these new specifications. The relationship between partisanship and these α_i terms is estimated to be stronger for respondents who report being more politically interested. It is also shown that the shift terms are larger in magnitude for more politically interested respondents, although this relationship is strongest in magnitude for Democrats. The estimated coefficient for strong Republicans on political interest for predicting these β_j terms is much smaller in magnitude and not statistically significant at the 0.05 level. While the findings presented in the Appendix are exploratory in nature, they suggest potential for future study, including looking at whether the sources of political information (e.g. *Fox News* vs. *MSNBC*) may impact the interpretation or use of these ideology rating scales.

These results suggest that the vast majority of what was estimated as partisan DIF in the politician anchors setup is actually due to partisan differences in the perceptions of politicians and parties, particularly when evaluating more liberal stimuli, rather than differing interpretations of the rating scale itself. To investigate this more directly, we can examine the estimated perceptions of politicians, calculated based on the formula in equation (2) above, including whether they differ significantly by respondent partisanship.

Table 2 presents the results of six regression models, each one predicting respondents' estimated perceptions of a given stimulus under the vignette anchors setup with seven-point PID.¹¹ Most notably, the coefficients on PID are negative, highly statistically significant, and relatively large in magnitude for all of the stimuli except Trump and the Republican Party. For example, each unit increase in a respondent's PID (e.g. from “strong Democrat” to “weak Democrat”) is estimated to move their perception of Obama by -0.16 . This translates in a difference between strong Democrats and strong Republicans of nearly one unit on the latent scale or, roughly, the distance between the position of the most liberal vignette (someone taking five consistently liberal policy positions) and the third most liberal one (someone taking liberal positions on three out of five specific policies). The relationship between perceptions and PID are similar for Clinton and the Democratic Party and, to a somewhat lesser extent, for perceptions of the Supreme Court.

11 Because of the division by β_j in equation (2), estimates can vary wildly for respondents with β_j having a nontrivial amount posterior density very close to zero. Therefore, following Hare *et al.* (2015), I use posterior medians, rather than means, as estimates of these DIF-adjusted perceptions.

Table 2. Relationship between estimated perception of politician ideologies from vignette anchors model and party identification. Columns show results of linear regression models for respondents' estimated perceptions (posterior medians) for each political actor as predicted by seven-point party identification.

	Obama	Clinton	Trump	Dem. Party	Rep. Party	Sup. Court
Intercept	-0.41 (0.12)	-0.23 (0.12)	0.72 (0.16)	-0.39 (0.11)	1.08 (0.13)	0.49 (0.10)
Party ID	-0.16 (0.03)	-0.19 (0.03)	-0.01 (0.04)	-0.16 (0.03)	-0.00 (0.03)	-0.11 (0.02)
<i>n</i>	809	804	705	805	796	751
Residual SE	1.63	1.64	2.01	1.61	1.78	1.32
<i>R</i> ²	0.04	0.06	0.00	0.05	0.00	0.03

Interestingly, perceptions of Trump and of the Republican Party show little, if any, partisan difference, with coefficients on PID in both of these models being estimated quite close to zero and not coming near statistical significance.

These regression results are consistent with the results in Figure 5 and they confirm that much of the so-called partisan DIF estimated by the A–M model under the standard politician anchors setup is not actually driven by differential use of the ideological rating scale but, instead, by differential perceptions of politicians, parties, and other political stimuli. Republican respondents do tend to utilize the seven-point ideological rating scale differently from independents or Democrats. But these scale use differences are much smaller in magnitude than the partisan differences in perceptions of the ideology of politicians or parties. To put it differently, the right panel of Figure 5 suggests that DIF is unlikely to be a major driver of partisan differences in ideological ratings of politicians given that scale use differences should show up in the ratings of the hypothetical individuals' ideologies, but only very small partisan differences are observed for vignette ratings.

The estimates from the vignette anchors setup demonstrate that the common perceptions assumption that is implicit in the standard A–M setup does not hold for Americans evaluating the ideology of politicians, political parties, and other stimuli. While relatively small differences observed between Democratic, independent, and Republican respondents' ratings of the ideology of Trump or the Republican Party appear to be driven mainly by differing use of the response scale (i.e. standard DIF), the large partisan divergence in ratings of more liberal stimuli are not principally the result of partisan DIF but, instead, are overwhelmingly driven by large partisan divergence in actual perceptions of the ideology of these stimuli. Republican respondents perceive these stimuli, such as Democratic politicians, the Democratic Party, and even the Supreme Court, as being much more liberal than do Democratic respondents.

4 The Relationship Between Respondent Partisanship and Ideological Ratings of Parties Over Time

The above analyses provide strong findings about partisan differences in ideological perceptions. But the 2016 presidential election campaign, during which these data were gathered, was a unique political environment to say the least. In this section, I examine whether the patterns identified in 2016 appear similar in previous years going back nearly a half century. In other words, is it the case that in past years, Republicans tended to perceive the Democratic Party as much more liberal than did Democratic and independent identifiers, while perceptions of the Republican Party differed minimally by respondent partnership?

Obviously, previous surveys like the ANES do not typically include vignette questions of the type required to estimate the model used above. But by looking at basic associations between

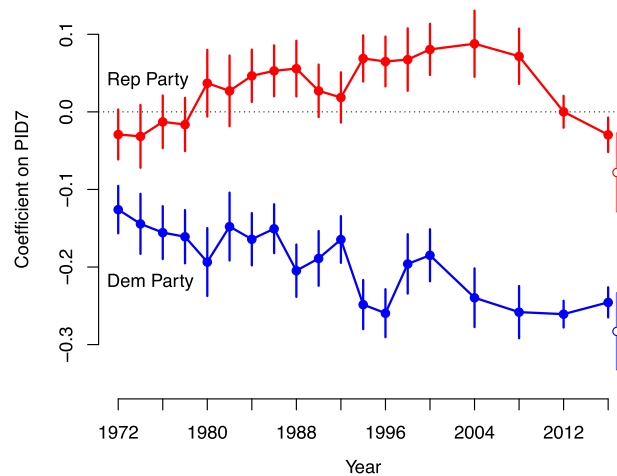


Figure 6. Coefficients on party identification predicting ideological ratings of Democratic and Republican Party by year of ANES. Dots indicate estimated coefficients and vertical line segments show 95% confidence intervals. The rightmost point shows corresponding estimates from University of Texas module of 2016 CCES used in the previous sections.

respondents' partisanship and their ideological ratings of the parties, we can see whether these patterns were similar in previous years to the ones identified above for 2016. If this is the case, it might imply that these findings are more general. Furthermore, if one is willing to assume that the findings above of very small partisan differences in scale in 2016 also hold in previous years (a strong assumption to be sure), then the raw seven-point ideology ratings of each of the parties can be viewed as a reasonably good measure of respondents' actual underlying perceptions—ones that can meaningfully be compared across respondents. In other words, if partisan differences in interpretation and use of the rating scale are minimal, looking at these raw ratings should produce similar findings to what would have been found if we were able to estimate the full vignette anchors adjustments used above. Obviously it is possible that in previous years, partisan differences in scale use were much larger, so the findings in this section should be interpreted with some caution. But even if one does not want to maintain “minimal partisan DIF” assumption over time, it may still be interesting to examine the characteristics of the raw ideological ratings of the two parties for their own sake.

Here I present the results of simple regression analyses using the ANES time series cumulative file. Since 1972, respondents have been asked to rate the ideologies of the Democratic and Republican parties on a seven-point ideology scale.¹² In each of these years, I estimate linear regressions predicting respondents' ideological ratings of the Democratic Party and, separately, the Republican Party, each using PID as the only predictor. Figure 6 plots the estimated slope coefficients from each of these separate bivariate regressions. These coefficients represent the change in perceptions of the party (Democrat or Republican) that would be expected to result from moving up the seven-point PID scale one unit based on a given year's data.

As found in the 2016 CCES data above, Republican respondents in each year of the ANES rate the Democratic Party as significantly more liberal than Democratic respondents do, as evidenced by the negative coefficient estimates on PID predicting perceptions of the Democratic Party in each year. There is also evidence of a general time trend, with this partisan divergence in perceptions of the Democratic Party growing larger in magnitude (more negative) in more recent years. In 1972, for example, the slope estimate for perceptions of the Democratic Party is -0.13 . This implies

¹² The ANES has also frequently asked respondents to rate specific politicians or other political stimuli, but given that these change frequently over time, I analyze only ratings of the two major parties here since these items are included in the vast majority of ANES waves.

that moving a respondent's PID from strong Democrat to strong Republican would be expected to make their rating of the Democratic Party roughly three quarters of a point more liberal. In 2016, the corresponding slope coefficient is estimated to be -0.25 —almost double the 1972 estimate—which translates to a nearly one and a half point difference in expected seven-point scale ratings. There are many potential explanations for this trend such as increased ideological sorting including the decline of the southern wing of the Democratic Party, which future research could explore. This difference could also be related to the fundamental asymmetry in the nature of ideology and group interest motivating the two parties in contemporary American politics, as discussed by Grossmann and Hopkins (2016).

Perceptions of the Republican Party, by contrast, show much smaller relationships with respondent partisanship. In many of the years analyzed, we cannot reject the null hypothesis of no relationship between PID and perceptions of the Republican Party at conventional significance levels. There is some evidence of time trends, with this coefficient being estimated as positive (and usually significant) between 1980 and 2008. Most notably, all of the estimates for perceptions of the Republican Party are much closer to zero than are any of those for perceptions of the Democratic Party, indicating that partisanship is a much weaker predictor of perceptions of the Republican Party than perceptions of the Democratic Party in all of these years.

The corresponding coefficient estimates using the same 2016 CCES data analyzed in the previous section instead of ANES data are plotted with hollow dots on the far right of Figure 6. Interestingly, these 2016 CCES estimates are similar to those based on the 2016 ANES, albeit with larger confidence intervals due to the smaller sample size. Of course, the results in Figure 6 are not adjusted for DIF in the same way that the previous sections' results are because the ANES study does not include the vignettes necessary to do so. To the extent that one believes that the results from the 2016 CCES presented above, which show minimal partisan differences in response scale use, apply to the ANES study, including those dating back several decades, this may not be problematic. But this is of course quite a strong assumption, so these results should be interpreted with some caution.

5 Discussion

The approach used here modifies the classic model of Aldrich and McKelvey in a way that results in several advantages. First, it is possible under this new vignette anchors framework to estimate each respondent's perceptions of the ideology of politicians, while also adjusting these perceptions to account for possible differences in the way that different respondents may interpret and use the response scale on which they report these ideological perceptions. Under the existing A–M approach, estimation of respondent-specific DIF was only possible by assuming that all respondents perceive a given politician in the same way. This approach used here thus provides a way of assessing the plausibility of this identical perceptions assumption. Applying the vignette anchors modeling framework to a dataset including respondents' ratings of the ideology of US political actors, I find strong and important differences in the perceptions of politicians by different respondents. This finding casts doubt on this key assumption of the A–M framework.

Furthermore, it is shown that the assumption that respondents perceive hypothetical individuals described in vignettes in the same way is much more plausible (even if not exactly true) than the assumption that respondents perceive actual politicians identically. This suggests that partisans tend to have similar understandings on average of the terms “liberal” and “conservative,” at least with respect to collections of policy positions described for hypothetical individuals. By contrast, their perceptions of prominent politicians and other political actors show notable differences by respondent PID. These findings can provide new nuance to the arguments of Converse (1964), as well as others such as Kinder and Kalmoe (2017), about the mass public's understanding of the ideological map of American politics. It may be too simple to say that

citizens do not hold realistic ideological understandings of the American political landscape, but rather that their understandings may be realistic in some respects, but importantly distorted by partisanship in others.

Most notably, the results presented above show that large differences exist between the perceptions of Democrats and Republicans. In particular, respondent partisanship is strongly related to perceptions of more liberal politicians and the Democratic Party as well as the Supreme Court. Perceptions of more conservative political actors, by contrast, show little or no partisan difference. Some have suggested that these sorts of partisan discrepancies are attributable to DIF, in which partisanship affects the way respondents perceive and use the survey response scales. But using novel survey data in which respondents rate the ideology of hypothetical individuals described by their policy positions, I find that partisan scale use differences are relatively small. Even after adjusting for these scale use differences, the vast majority of the large partisan differences in respondents' ideological perceptions remain.

The results presented beg the question of what mechanism produces these strong partisan differences as well as why there is little difference, if any, between Democratic and Republican respondents' perceptions of more conservative political actors. Although we do not have conclusive findings on this front, several possibilities seem plausible and future work should focus on arbitrating between these various explanations. First, there likely exists a fundamental asymmetry between the labels used for the two sides of the ideological scale. In recent decades, in particular, the term "liberal" has been assigned negative connotations, particularly by Republicans. For example, Schiffer (2000) refers to the "popular stigmatization of the word 'liberal'" (see also Neiheisel (2016) for related experimental results). This possibility is also related to the findings of Bullock *et al.* (2015) and Prior *et al.* (2015), who argue that much of the partisan divergence in factual beliefs about politics can be attributed to expressive actions—the so-called partisan "cheerleading" in which respondents derive utility from making statements that reflect positively on their own party or negatively on the opposing party, even if they may not believe these statements to be accurate.

In contrast to the perception-based perspective on ideology taken here, other research has also measured ideology using more objective policy stances taken by political actors (see e.g. Bafumi and Herron 2010, Jessee 2012, Shor and Rogowski 2016, Tausanovitch and Warshaw 2018). As shown in the Appendix (Section A2), similar estimates for politicians and vignettes are obtained whether one uses the perception-based approach of the vignette anchors model or a policy-based approach as estimated through a standard ideal point model. This suggests that, aside from the important partisan differences in perceptions that have been identified here, the average perceptions of political actors' ideologies correspond closely with the ideological positions implied by the actors' respective positions on specific policies. Although further work is needed to examine this more fully, this finding might be thought to suggest that on balance, the American public on average has a relatively accurate understanding of the positions taken by politicians, parties, and other political actors. But underlying these on average perceptions are important partisan differences. An interesting question would be whether certain types of political actors are perceived more similarly by Democrats and Republicans than others. Although the results above show that partisan differences for perceptions of Republican actors are smaller than those for others, determining whether other types of stimuli show larger or smaller partisan perceptual differences would require a survey in which respondents rate the ideology of many more political actors.

Going forward, future research can use the approach employed here to learn about how different people may perceive stimuli differently. In particular, this approach allows for the estimation of respondent-specific perceptions that are also adjusted based on each respondent's particular scale use characteristics. While the results presented here shows that the central

assumption of the A–M framework appears to be strongly violated when analyzing Americans' perceptions of prominent political actors (the specific context in which the A–M approach was developed), it is possible that this common perceptions assumption is more plausible in other contexts. It may also be the case that although DIF was found to be a relatively small issue in this application, it could be a much bigger factor in other applications. For example, King *et al.* (2004) use anchoring vignettes to allow for cross-cultural comparisons of concepts like political efficacy, finding that DIF-adjusted measures produce very different results than do raw (unadjusted) measures. One could imagine using this approach in the comparative context to, for example, explore cross-country or cross-party differences in perceptions related to European Union politics or in other international contexts. The vignette anchors framework introduced here presents a straightforward way of assessing this assumption's plausibility as well as learning about possible differences in the perceptions of different types of respondents, a quantity that is often of important theoretical interest.

Acknowledgements

I thank Chris Hare, Neil Malhotra, and Elizabeth Simas for helpful feedback.

Conflicts of interest

The author declares no conflicts of interest related to this project.

Data availability

Replication data and code for this paper can be found at Jessee (2019).

Supplementary material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2019.47>.

References

- Aldrich, J. H., and R. D. McKelvey. 1977. "A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections." *The American Political Science Review* 71(1):111–130.
- Bafumi, J., and M. C. Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and their Members in Congress." *American Political Science Review* 104(3):519–542.
- Bakker, R., S. Jolly, J. Polk, and K. Poole. 2014. "The European Common Space: Extending the Use of Anchoring Vignettes." *The Journal of Politics* 76(4):1089–1101.
- Brady, H. E. 1985. "The Perils of Survey Research: Inter-Personally Incomparable Responses." *Political Methodology* 11(3/4):269–291.
- Brady, H. E., and P. M. Sniderman. 1985. "Attitude Attribution: A Group Basis for Political Reasoning." *The American Political Science Review* 79(4):1061–1078.
- Brody, R. A., and B. I. Page. 1972. "Comment: The Assessment of Policy Voting." *American Political Science Review* 66(2):450–458.
- Bullock, J. G., A. S. Gerber, S. J. Hill, and G. A. Huber. 2015. "Partisan bias in factual beliefs about politics." *Quarterly Journal of Political Science* 10:519–578.
- Campbell, Angus, P. E. Converse, W. E. Miller, and D. E. Stokes. 1960. *The American Voter*. Chicago, IL: University of Chicago Press.
- Conover, P. J., and S. Feldman. 1982. "Projection and the Perception of Candidates' Issue Positions." *The Western Political Quarterly* 35(2):228–244.
- Converse, P. E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by R. Nieme and H. Weisberg, 322–357. New York: Free Press.
- Grossmann, M., and D. A. Hopkins. 2016. *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. Oxford: Oxford University Press.
- Hare, C., D. A. Armstrong, R. Bakker, R. Carroll, and K. T. Poole. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59(3):759–774.
- Jessee, S. A. 2012. *Ideology and Spatial Voting in American Elections*. Cambridge: Cambridge University Press.

- Jessee, S. A. 2019. "Replication Data for: Estimating Individuals' Political Perceptions While Adjusting for Differential Item Functioning." <https://doi.org/10.7910/DVN/MCXXBS>, Harvard Dataverse, V1.
- Kinder, D. R., and N. P. Kalmoe. 2017. *Neither Liberal nor Conservative: Ideological Innocence in the American Public*. University of Chicago Press.
- King, G., C. J. L. Murray, J. A. Salomon, and A. Tandon. 2004. "Enhancing the validity and cross-cultural comparability of measurement in survey research." *American Political Science Review* 98(1):191–207.
- King, G., and J. Wand. 2007. "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes." *Political Analysis* 15(1):46–66.
- Levendusky, M. S., and N. Malhotra. 2015. "(Mis) Perceptions of Partisan Polarization in the American Public." *Public Opinion Quarterly* 80(S1):378–391.
- Neiheisel, J. R. 2016. "The 'L' Word: Anti-liberal Campaign Rhetoric, Symbolic Ideology, and the Electoral Fortunes of Democratic Candidates." *Political Research Quarterly* 69(3):418–429.
- Plummer, M. et al. 2003. "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling." In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, vol. 124, edited by K. Hornik, F. Leisch, and A. Zeileis. Vienna, Austria.
- Prior, M., S. Gaurav, and K. Khanna et al. 2015. "You cannot be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10(4):489–518.
- Rasch, G. 1966. "An Item Analysis which Takes Individual Differences into Account." *British Journal of Mathematical and Statistical Psychology* 2:45–57.
- Schiffer, A. J. 2000. "I'm not that Liberal: Explaining Conservative Democratic Identification." *Political Behavior* 22(4):293–310.
- Shor, B., and J. C. Rogowski. 2016. "Ideology and the US Congressional Vote." *Political Science Research and Methods* 6(2):323–341.
- Simas, E. N. 2017. "Ideology Through the Partisan Lens: Applying Anchoring Vignettes to US Survey Research." *International Journal of Public Opinion Research* 30(3):343–364.
- Simas, E. N. 2018. "Perceptions of the Heterogeneity of Party Elites in the United States." *Party Politics* 24(4):444–454.
- Struthers, C. L, C. Hare, and R. Bakker. 2019. "Bridging the Pond: Measuring Policy Positions in the United States and Europe." *Political Science Research and Methods*, doi:10.1017/psrm.2019.22.
- Tausanovitch, C., and C. Warshaw. 2018. "Does the Ideological Proximity Between Candidates and Voters Affect Voting in US House Elections?" *Political Behavior* 40(1):223–245.
- Westfall, J., L. Van Boven, J. R. Chambers, and C. M. Judd. 2015. "Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide." *Perspectives on Psychological Science* 10(2):145–158.