

Who Is “On Welfare”?

Validating the Use of Conjoint Experiments to Measure Stereotype Content

Abstract

We use the case of welfare recipients to validate conjoint experiments as a measure of stereotype content. Stereotypes are politically consequential, but their content can be difficult to measure. The conjoint measure of stereotype content, in which respondents see profiles describing hypothetical persons and rate these persons’ degree of belonging to the target group, offers several advantages over existing measures. However, no existing work evaluates the validity of this new measure. We evaluate this measurement technique using the case of welfare recipients. Stereotypes of welfare recipients are politically important and extensively studied, providing strong *a priori* expectations for portions of the stereotype, especially race, gender, and “deservingness.” At the same time, scholars disagree about the importance of another attribute with important political implications: immigration status. We find that aggregate stereotypes, measured via a conjoint experiment, match the strong *a priori* expectations: white Americans see welfare recipients as black, female, and violating the norms of work ethic. Individual-level stereotypes also predict welfare policy support—even when other demographic and ideological factors are accounted for. We also find that immigration status is not a part of the welfare recipient stereotype for most Americans, but support for welfare is lower among those who do stereotype welfare recipients as undocumented immigrants. Finally, we suggest an improvement in the conjoint task instructions. Overall, we confirm that conjoint experiments provide a valid measure of stereotypes.

Keywords: conjoint experiments, public opinion, stereotypes, survey methods, welfare

Politics often forces individuals to develop conceptions of objects—concepts, groups, categories—that are distant, abstract, and outside of their everyday experience. To develop opinions or attitudes that involve these objects, they must rely on mental representations of these objects, famously referred to as the “pictures in our heads” that guide public opinion (Lippman 1922). These *stereotypes* may or may not correspond to reality, though their common cultural sources mean that they are likely to be shared, at least among people who share a culture (Stangor and Schaller 2000). They nevertheless have great political import. Stereotypes of social groups strongly and significantly predict attitudes towards partisans (Ahler and Sood 2018), immigrants (Zhirkov 2021c), welfare recipients (Gilens 1999), and candidates (M. Schneider and Bos 2011). On the societal level, politically relevant populations that are negatively stereotyped by the public are disadvantaged in terms of policy outcomes (A. Schneider and Ingram 1993).

Understanding the content of these mental representations is essential to understanding the role that they play in political processes. Yet, doing so is difficult. Stereotypes can contain multiple attributes of various kinds. For example, stereotypes of welfare recipients, the substantive case we study, include demographic characteristics (e.g., race, gender), personality traits (e.g., being lazy), and behaviors (e.g., drug use). Further, the attributes making up a stereotype vary in importance, or centrality, to the stereotype (Garcia-Marques, Santos, and Mackie 2006). Finally, individuals may not have conscious access to the content of their own stereotypes, and if they do, may be motivated to describe them in a socially desirable way. These challenges present problems for measuring stereotype content using traditional methods that generally rely on direct questioning (Eagly and Mladinic 1989).

In this paper, we evaluate the validity of an alternative approach that measures stereotypes by asking respondents to categorize individuals in the context of a conjoint

experiment (Flores and Schachter 2018; Goggin, Henderson, and Theodoridis 2020). Such experiments show respondents a series of “profiles,” each describing a person using several attributes, and then asks subjects to judge that person’s group membership. In theory, this task—categorization—has the advantage of directly mirroring the psychological function of stereotypes, ascertaining the fit between individuals and groups. We interpret an attribute’s effect, as measured by the conjoint experiment, as the contribution of that attribute to the group stereotype. Since these effects are measured on the same scale, we can also compare the relative importance of attributes to the stereotype.

However, despite the novel use of the conjoint-based stereotype measure, existing studies avoid any discussion of validity.¹ Here, we evaluate the validity of using conjoint experiments to measure stereotypes by measuring the stereotypes white Americans have about welfare recipients. The extensive literature on these stereotypes gives us strong *a priori* expectations for their central elements, which we use to generate a series of validity tests. Using data from a demographically representative survey of white Americans, we find support for face, concurrent, and predictive validity of the measure.

At the same time, recent work disagrees about the importance of another potential element of the welfare recipient stereotype: immigration status. Some scholars claim that conservative rhetoric tying undocumented immigrants to welfare has “immigrationized” welfare—that is, made immigration status a central part of the welfare recipient stereotype (Garand, Xu, and Davis 2017). Others argue that this rhetoric has had little effect in the face of more general ethnocentric attitudes (Levy 2021). We find evidence that most Americans do not associate immigrants—undocumented or documented—with welfare, though those Americans

¹ While some work validates the use of conjoint experiments to measure preferences (Hainmueller, Hangartner, and Yamamoto 2015), none examines its validity as a measure of stereotypes.

who do stereotype welfare recipients as being undocumented immigrants are much less supportive of welfare.

We also introduce two innovations to the conjoint measurement of stereotype content. One weakness of existing uses of the conjoint measure of stereotypes is that they show only aggregate stereotypes, limiting researchers' ability to examine the individual-level correlates of stereotype content. We apply recent advances in conjoint methodology to obtain individual-level estimates of stereotype content (Zhirkov 2021a), that can be used as covariates in inferential analyses. Second, we use an embedded question-wording experiment to compare the phrasing of conjoint tasks used in recent published work, which asks how *likely* it is that profiles belong to the target group, to an alternative drawn from an older psychological literature on categorization, which asks how *typical* profiles are of the target group. We find that the latter is a better measure of stereotype content.

Stereotypes: Nature and Measurement

Stereotypes are “beliefs about the features and characteristics that are associated with social groups” (Bodenhausen, Kang, and Peery 2012, 320), that help people make sense of the social world. Stereotypes allow individuals to classify others into groups and subsequently draw inferences about what attributes to expect from them. The set of attributes forming the stereotype might include traits, demographic features, or behaviors associated with the group. Among these attributes, some are more important or central to the stereotype than others, and these play a larger role in categorization decisions. Exactly how stereotypes are stored and retrieved is a subject of debate (Sherman 1996), but all models share these basic aspects: abstraction, generalization across individuals, and the ability to accommodate a wide range of different attributes. Stereotype content may come from a variety of sources, including interactions with

group members, factual knowledge about the characteristics of group members, or the media (Brosius 2003). Thus, stereotypes may be “true” or “false” in that they accurately or inaccurately depict the characteristics of the target group, but they almost always reflect the biases embedded in the culture that produces them (Entman and Rojecki 2000).

These core features of stereotypes imply that a measure of stereotype content should have several qualities. First, it should recognize the multi-dimensional nature of stereotypes: a complex mix of attributes that can include positive, negative, and neutral qualities. Second, the measure should be able to accommodate a variety of kinds of attributes: demographic characteristics, traits, beliefs the group is thought to hold, behaviors the group is thought to engage in, and so on. Third, it should measure not only whether each attribute is part of a stereotype, but also the importance or centrality of each attribute to that stereotype. Fourth, certain stereotypes are subject to social disapproval, so an ideal measure would not rely on direct reports that are susceptible to social desirability bias.

Most existing measures of stereotypes lack one or more of these qualities. The stereotype battery in the American National Election Studies measures traits (e.g., “hardworking” or “intelligent”), but not other attributes that may go into a stereotype (e.g., demographic characteristics). Moreover, ratings on these traits tend to be highly correlated, even when asking about seemingly distinct qualities, and are often treated as a measure of general intergroup affect instead of a specific stereotype content (Kinder and Kam 2010). Other methods measure more specific content but come with their own weaknesses. For instance, the adjective-identification paradigm can accommodate a large number and range of attributes (Katz and Braley 1933), but does not measure their relative importance (though for an attempt to do so, see Schneider and Bos 2011). Open-ended responses allow for the reporting of a wide range of traits (Garcia-

Marques, Santos, and Mackie 2006), but make it hard to measure traits' relative importance and depend on coding procedures. Several recent studies of partisan stereotypes attempt to measure perceptions about the demographic make-up of parties (Claassen et al. 2021; Mason and Wronski 2018), but do not incorporate non-demographic attributes and rely on individuals' questionable ability to translate stereotypes into proportions (Landy, Guay, and Marghetis 2018). Importantly, all these measures rely on self-reports that may be affected by social desirability bias. Some work uses implicit measures of stereotypes based on time-response (Winter 2010; Zhirkov 2021b), but they are difficult to implement and require substantial cognitive effort.

Conjoint Experiments as a Measure of Stereotypes

We argue that conjoint experiments offer a better way to measure stereotype content because they have the qualities needed to ideally measure stereotypes. Conjoint experiments present participants with profiles composed of a set of attributes, with each attribute taking on one of several possible values. For each profile, the value of each attribute is independently randomized. Participants are then asked to choose from pairs of profiles or rate each profile according to some instruction. The independent randomization of these attribute values can, under certain assumptions, be used to calculate the independent effect of each attribute value on the rating or selection of profiles (Hainmueller, Hopkins, and Yamamoto 2014). These estimates, termed average marginal component effects (AMCEs), provide the effect of a given level of an attribute relative to other possible levels of the attribute. Recent work has extended the conjoint methodology to allow for measuring the effects of attributes at the respondent level, or individual marginal component effects (IMCEs; Zhirkov 2021a).

Conjoint designs in political science are mostly used to measure the effect of attribute values on choice, such as which candidate to vote for. However, some recent work uses conjoint

experiments to measure mental associations between categories of people and attributes; in short, stereotypes (Flores and Schacter 2018; Goggin, Henderson, and Theodoridis 2020).² In these studies, participants are presented with profiles describing hypothetical persons and are asked to choose more likely group members or rate the profiles' likelihood of belonging to the group. The resulting AMCEs are interpreted as measures of the centrality of each attribute to the commonly held stereotype of the group in question.

This use of a conjoint experiment appears to provide a good measurement of stereotype content because it mirrors the real-life categorization task that stereotypes are used for. Moreover, the conjoint task does not rely on individuals' ability and willingness to access and report the content of their stereotypes—it simply asks them to recognize group members when they appear. Additionally, the conjoint measure can accommodate many attributes, including different kinds of attributes, without the human coding decisions required by open-ended measures. Since the AMCEs and IMCEs for these attributes refer to the same outcome variable, researchers can use them to measure the relative importance or centrality of these attributes. Finally, some evidence suggests that conjoint experiments reduce social desirability bias because the large number of attributes make respondents feel that they can “conceal” socially undesirable motivations behind any given decision (Horiuchi, Markovich, and Yamamoto 2021). The conjoint experiment has practical advantages as well: it can be easily placed into online surveys, takes less survey time than measures like the adjective identification task, and requires less researcher involvement to interpret than open-ended responses.

Another important benefit of the conjoint measure concerns its resilience to respondents applying substantive knowledge, if not plain cheating, to answer the questions. For instance,

² Notably, these authors do not use the term, though their designs measure stereotypes: clusters of attributes associated with social groups.

when answering standard composition questions (“what percentage of group X have attribute Y?”), respondents can simply know the right answer—or even look it up online (Clifford and Jerit 2016). Conjoint experiments are not subject to this problem: their design requires respondents to guess the *conditional* probability of someone belonging to group X given *multiple* attributes—which is almost impossible for an individual to know or estimate in an objective manner.

Yet, despite the strong *prima facie* case for using conjoint experiments to measure stereotypes, no existing work validates this use. Conjoint experiments have been shown to match real-world choices in some areas (Hainmueller, Hangartner, and Yamamoto 2015), a fact cited as evidence of validity of the conjoint measurement of stereotypes (Flores and Schacter 2018, 845). However, even if results of conjoint experiments accurately represent preferences, this does not necessarily mean that respondents classify conjoint profiles in the same way they apply stereotypes in the real world. Thus, our first goal is to validate the conjoint measure as a measure of stereotype content.

In addition, we embed a question-wording experiment to evaluate two different ways to phrase the conjoint task instruction. Existing work in political science and sociology asks participants whether or how likely it is that the person described in a conjoint profile is a member a group (Flores and Schacter 2018; Goggin, Henderson, and Theodoridis 2020); we refer to this as the “likelihood” instruction. However, psychological studies of object categorization and stereotype content generally ask subjects how typical the object or person is of a category or group (Heit and Nicholson 2010); we refer to this as the “typicality” instruction. One advantage of the typicality instruction is that an individual can clearly be a group member while also being

a poor example of the group's characteristics,³ and soliciting typicality ratings avoids classifying these atypical group members as being highly characteristic of the target group.

Substantive Case: Stereotypes of Welfare Recipients

We evaluate the use of conjoint experiments to measure stereotypes using an important substantive case: stereotypes of welfare recipients. Welfare has long been the most controversial governmental assistance program in the US (Goren 2008; Shapiro et al. 1987), and attitudes on this issue have been driven by stereotypes of who is on welfare, which often diverge from reality (Gilens 1999). Moreover, these stereotypes have been consequential, driving public opinion and policy decisions like the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (Weaver 2000), that famously “ended welfare as we [knew] it.” Consequently, literature on welfare recipient stereotypes and their consequences is extensive in political science.

A central part of the welfare recipient stereotype is “undeservingness”: an unwillingness to work hard, instead opting to receive resources without putting in effort. Deservingness may be an evolved heuristic meant to prevent freeriding behavior and cheaters, which, translated into modern times, often relies on shortcuts such as employment status (Petersen 2012). Americans therefore distinguish between “undeserving” welfare recipients and the deserving poor (Appelbaum 2001; Goren 2003), with the former purportedly taking advantage of government “handouts” and becoming dependent on them rather than actively searching for employment. Conversely, those who are seeking employment or who have contributed to society financially are viewed as deserving. For instance, redistributive programs delivered via tax credits enjoy greater levels of support because their recipients are perceived as more deserving (Ellis and Faricy 2020). By extension, Americans also tend to believe that welfare recipients—and the poor

³ For example, most PhDs are Democrats, but the typical Democrat does not have a PhD.

generally—are lazy or freeriding, as well as more likely to abuse substances, be violent, or have a criminal record (Amundson, Zajicek, and Kerr 2015; Clawson and Trice 2000). The perception that welfare recipients do not work hard enough to achieve economic independence violates American values of individuality and hard work (DeSante 2013). As a result, Americans who score high on individualism are less supportive of social welfare programs (Federico 2006).

Race also plays an important role in stereotypes of welfare recipients. People are more likely to say that a person is black if the image of that person is labeled “welfare recipient” (Brown-Iannuzzi et al. 2017). Blacks are most prominent in the most negative and stereotypical stories on poverty, the underclass, urban problems, and welfare reform and pregnancy (Clawson and Jett 2019, Gilens 1999). Conversely, the working poor are overwhelmingly portrayed as white. The role of race, however, may vary with economic conditions, with welfare stereotypes growing less racialized during economic downturns (Gilens 1999, 123–27). Stereotypes of welfare recipients as nonwhite translate to lower support for redistributive programs among whites. Those higher in racial resentment tend to be less supportive of welfare (Rabinowitz et al. 2009), particularly when they are given contextual racial cues (Banks 2014). However, some scholars argue that whites’ opposition to welfare stems from individualist values and preferences for limited government, not stereotypes of welfare recipients as nonwhite (Peffley, Hurwitz, and Sniderman 1997).

Gender and parental status are also prominent in stereotypes of welfare recipients. In conservative rhetoric, single mothers embody the undeserving welfare recipient via the “welfare queen” stereotype (H. Brown 2013). At the extreme, women are accused of having more children solely for the purpose of receiving more welfare money. This view was sufficiently mainstream to drive policies like the “family cap,” which denied additional cash assistance to families that

had new children while on welfare (Pierson-Balik 2003). Furthermore, Americans tend to perceive women with children on welfare as also being unmarried; in fact, many Americans believe that being on welfare encourages single motherhood (Hancock 2004).

Therefore, these characteristics—deservingness (work ethic), race, gender, marital status, number of children, and criminal record—are well-established and longstanding stereotypes of welfare recipients among Americans. For this reason, they are ideal for validating the use of conjoint experiments to measure stereotypes. Further, our study also contributes to this literature by comparing the relative centrality of these attributes to the stereotypes of welfare recipients.

More recently, some scholars have argued that immigrants, particularly undocumented immigrants from Latin America, have become more central to welfare stereotypes. At least since the battles over California’s Proposition 187 in the early 1990s, conservative elites have attempted to portray immigrants as “freeloaders” who come to the US not to work, but to receive welfare benefits (J. Brown 2021; Calavita 1996). While recent immigrants have been ineligible for most government benefits since 1996, conservative politicians have continued to link welfare programs and illegal immigration, as when former President Trump said: “I am tired of seeing our taxpayers paying for people to come into the country and immediately go onto welfare and various other things.”⁴

However, the success of this rhetoric is unclear. Some studies find a correlation between negative stereotypes of immigrants and negative attitudes towards welfare and conclude that “the American welfare state today has become more ‘immigrationized’ than ‘racialized’” (Garand, Xu, and Davis 2017, 147). But other recent work argues that most Americans have not incorporated immigration status into their stereotypes of welfare recipients, and the correlation

⁴ <https://www.nytimes.com/2019/08/14/us/immigration-public-charge-welfare.html>

between immigration and welfare attitudes is driven by general conservative ideology rather than an association of immigrants with welfare (Levy 2021). By comparing the importance of immigration status to other attributes, the conjoint measurement can shed light on this debate.

Validity Tests

A lengthy literature establishes that stereotypes of welfare recipients among white Americans cover a wide range of attributes: “deservingness” (often as a willingness to work and no criminal record), race, and gender or gendered characteristics (marital status and number of children).⁵ Further, the literature provides strong expectations for how individual characteristics should relate to the stereotype (e.g., those with negative attitudes toward blacks should associate welfare with blacks) and how these stereotypes should be associated with welfare support (e.g., those who see welfare recipients as lazy should be less supportive of welfare). These strong existing expectations of welfare recipient stereotypes provide a useful case for validating the conjoint measure. Moreover, the complexity of the welfare stereotype demonstrates a key strength of the conjoint-based measure: the ability to assess the role of multiple characteristics and compare the relative importance of different characteristics to the stereotypes.

Based on this literature, we develop tests for the face, concurrent, and predictive validity of the measure. Our hypotheses and tests focus on non-Hispanic white Americans because stereotypes about welfare recipients have been best studied within this group.

Tests: Face Validity

⁵ It is worth noting that these stereotypes differ from the average characteristics of actual recipients of means-tested government programs; only 25% of recipients are black, and most are also employed, with 85% of recipient families having at least one employed adult (Foster and Rojas 2013). Recipient families are more likely to be headed by a single woman, but there appears to be no relationship between welfare receipt and increased birth rate for women of child-bearing age (for a review, see Schettini Kearney 2004).

First, the literature establishes an expectation that welfare recipients are perceived to be black, female, and “undeserving,” in the sense of violating American norms of hard work. Thus, we can judge the face validity of a measure by whether it finds welfare receipt associated with race, gender, and deservingness cues.

- Black profiles should be more likely to be categorized as welfare recipients than white profiles.
- Female profiles should be more likely to be categorized as welfare recipients than male profiles.
- Profiles with cues that suggest low deservingness should be more likely to be categorized as welfare recipients than those with cues that suggest high deservingness.

Tests: Concurrent Validity

Second, we evaluate the concurrent validity of the measure by examining whether race plays a larger role in the stereotype of welfare recipients among those who have conservative racial attitudes. Similarly, work ethic should play a larger role in welfare stereotypes among those high on individualism.

- There should be a positive correlation between conservative racial attitudes and IMCEs for the black attribute.
- There should be a positive correlation between individualism and IMCEs for the work ethic attributes.

Tests: Predictive Validity

Finally, we evaluate predictive validity by testing whether white respondents who see welfare recipients as black and “undeserving,” as measured by IMCEs for the respective attributes, express lower support for welfare.

- Respondents with a high IMCEs for the black attribute exhibit lower support for welfare.
- Respondents with high IMCEs for the deservingness attribute exhibit lower support for welfare.

Tests: Question Wording

Finally, we evaluate the typicality vs. likelihood instructions by comparing the results on the above validity tests. As detailed above, we expect that the typicality instruction will produce a measure with better face, concurrent, and predictive validity than the likelihood instruction.

Research Design

Participants, Procedure, and Measures

We recruited participants through Lucid, an online source of representative convenience samples of U.S. adults, in January 2021. Lucid samples are close to the national demographic benchmarks and replicate experiments conducted on probability samples (Coppock and McClellan 2019). Lucid also provides participants' demographics.

A total of 1,295 non-Hispanic white respondents completed the survey.⁶ Sample characteristics were as follows. Mean age was 49.6 years, and 53.2% were female. Median income was between \$45,000 and \$49,999, and 46% reported having college degrees. The sample was balanced by partisanship with 37.5% of respondents being Democrats, 35.3% Republicans, and 27.2% independents.

The survey proceeded in the following order. Respondents started by reporting their opinions on welfare policies using four questions about support for redistribution and attitudes towards welfare recipients (Gilens 1999), that have been used in several studies on the topic

⁶ We excluded two respondents who failed attention checks and 33 respondents who used the exact same rating scores for all profiles in the conjoint experiment.

(Brown-Iannuzzi et al. 2019; Levy 2021). Next, they completed the conjoint task described below. Finally, respondents answered questions on racial attitudes and individualism.

We measured racial attitudes using a scale designed to capture their affective and cognitive aspects: fear, institutionalized racism, and empathy (FIRE; DeSante and Smith 2020). Just as the racial resentment scale was introduced to capture the evolving shape of American racial attitudes in the post-civil rights period, FIRE was designed to measure contemporary manifestations of racial animus characterized by “colorblind racism” and has been used in a range of recent studies of racial attitudes (Green and McElwee 2019; Jardina, Kalmoe, and Gross 2020). The items in the FIRE scale have the added benefit of not invoking conceptions of deservingness, as the racial resentment scale does.⁷ Like racial resentment, though, they are predictive of attitudes towards welfare (Desante and Smith 2020, 644–45). Individualism was measured using a four-item battery asking respondents about their support for the ethos of self-reliance (Feldman et al. 2020). See Appendix A in Supplementary Material for question wording.

Conjoint Design

Based on the existing literature on stereotypes of welfare recipients, we included the attributes and values in the conjoint design listed in Table 1. These include race, gender, marital status, number of children, and immigration status. For immigration status, we used the term “undocumented/illegal,” as either “undocumented” or “illegal” by itself could signal political affiliation (for a similar use, see Flores and Schachter 2018). In addition to these, we also included two attributes that signaled deservingness. First, we used the person’s current employment status, including, for those unemployed, whether they were looking for a job. Second, we included the person’s criminal record. There are seven different values for this

⁷ For example, the racial resentment scale includes the statement “It’s really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites.”

attribute including three substance abuse crimes, three violent crimes, and no criminal record. We used three different crimes in each category to maintain realism; all analyses group these together.

Table 1. Attributes for profiles in conjoint experiment

Attribute	Values
Race/Ethnicity	White Black Hispanic
Gender	Male Female
Marital Status	Married Not Married
Has Children	<i>No: Zero</i> <i>Yes: One, Two, Three</i>
Immigration Status	U.S. Citizen Green-Card Holder Undocumented/Illegal Immigrant
Employment Status	Employed Unemployed, Seeking Employment Unemployed, Not Seeking Employment
Criminal Record	No Criminal Record <i>Yes, Drug-related: DUI, Heroin Possession, Drug Sales</i> <i>Yes, Violent: Aggravated Assault, Robbery, Threatening with a Weapon</i>

Note. Collapsed attribute values are in italics.

For race, gender, marital status, number of children, immigration status, and employment status we drew attribute levels from uniform distributions. For criminal record, we assigned a one-third probability to each kind of criminal record (no criminal record, violent crime, substance-related crimes), and within the kinds of criminal records an equal chance of each specific crime. For each profile, attribute values were drawn randomly and independently.⁸ The

⁸ Due to randomization, two respondents who never saw a profile with no children and were excluded from the analysis.

order of attributes was random across participants but held constant across all tasks for each individual participant.

To evaluate the likelihood vs. typicality instructions in the conjoint task, we randomly assigned participants to one of two instructions. Participants in the likelihood condition were asked to rate “how likely” it is that the person is on welfare as the outcome variable, with the response scale from 0 = *Extremely unlikely* to 10 = *Extremely likely*. Participants in the typicality condition were instead asked to rate how much the person is “typical of people on welfare” with the response scale from 0 = *Extremely nontypical* to 10 = *Extremely typical*. All participants were presented with 30 conjoint tasks;⁹ each conjoint task showed a single profile.¹⁰ Figure 1 shows an example of the two conditions as presented to respondents.

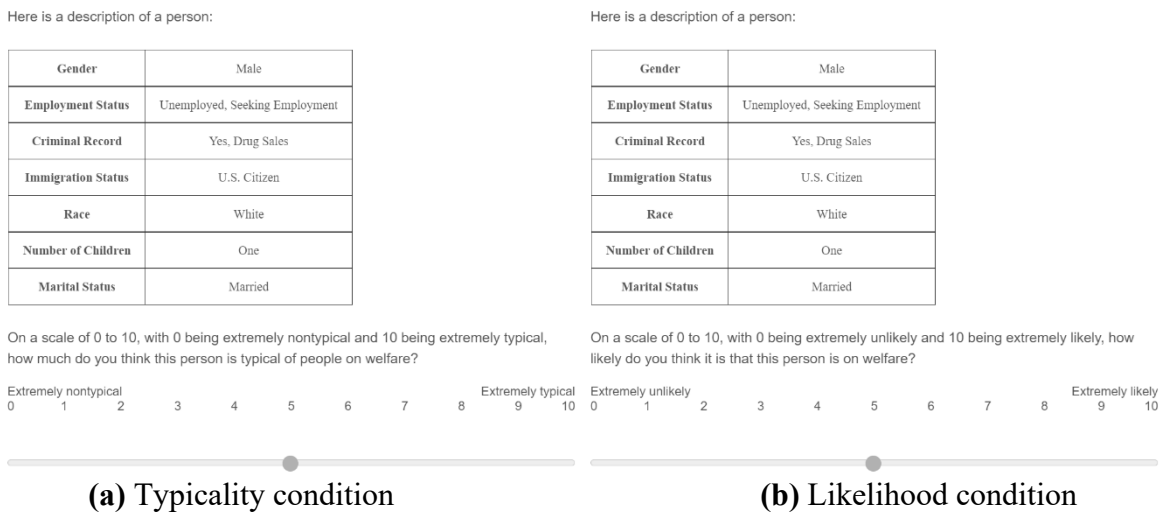


Figure 1. Examples of conjoint profiles by condition

Results

⁹ Respondents can complete this number of conjoint tasks without evidence of satisficing (Bansak et al. 2018). See Appendix B in Supplementary Material for results broken down by earlier and later rounds.

¹⁰ When completing the conjoint experiment, 55 respondents ended up rating less than 30 profiles (the lowest number of rated profiles was 27 for three respondents). These respondents were kept in the analysis.

Results: Face Validity

Figure 2 shows the AMCEs describing aggregate stereotype content for the typicality vs. likelihood response conditions, estimated via OLS with standard errors clustered by respondent.¹¹ For a sense of scale, the standard deviation of conjoint ratings in our data is approximately three on the 0–10 scale. As expected in the face validity hypotheses, the gender, marital status, number of children, and deservingness (both employment status and criminal record) attributes are significantly associated with judging a profile to be a welfare recipient in both typicality and likelihood conditions. These results suggest that the conjoint measure has face validity—the characteristics that existing literature expects to be part of the welfare recipient stereotype (female, not married, parent, unemployed, and rule-breaker) have positive AMCEs. Interestingly, the two types of criminal record have similar effects, suggesting that respondents categorize profiles as norm-violating or not, as opposed to making a distinction between substance abuse crimes, which might be expected to relate more to the low work ethic stereotype, and violent crimes.¹²

--- Figure2.eps ---

Figure 2. Effects of profile attribute values on stereotype ratings

Note. $N = 38,721$ (total number of profiles rated by all respondents). Response scales are 0–10 in both conditions. 95% confidence intervals shown. Numerical results in Table S1 (Appendix D, Supplementary Material).

However, there are important differences between the typicality and likelihood conditions. First, race/ethnicity is significantly associated with welfare recipients in the typicality condition but not in the likelihood condition, though the difference between these conditions is

¹¹ Replication materials for all presented analyses are available at: <https://doi.org/10.7910/DVN/6ECD1D>

¹² Appendix C in Supplementary Material reports the AMCEs for nonwhite respondents. Having children and employment-seeking appear to play a smaller role in black and Hispanic respondents' stereotypes of welfare recipients, though our samples of these groups are small.

not statistically significant. Second, unauthorized immigrants are judged as less likely to be welfare recipients than U.S. citizens—but not as less typical welfare recipients. Third, being unemployed makes a profile much more “likely” to be judged as a welfare recipient but has a significantly smaller, though still quite large, effect on judgments of typicality. Notably, immigration and employment status relate to what many believe to be formal rules for receiving welfare.¹³ This suggests that the likelihood instruction may lead people to rely on attributes that objectively define group membership instead of their more subjective sense of how “good” a member of the group the individual is.

Figure 2 also compares the relative importance of these attributes to whites’ stereotypes of welfare recipients. The results suggest that the most important attribute is related to conceptions of deservingness: whether the person is employed and, if not, whether they are searching for work. Perceptions of criminality also play a role, with either type of criminal record being associated with welfare receipt. Demographic characteristics, on the other hand, play a significant but smaller role with gender (as well as the gendered characteristics of being unmarried and having children) being more important than race. In contrast to expectations that welfare has become “immigrationized,” we find no evidence that being an immigrant is a part of the welfare recipient stereotype among whites, on average.¹⁴

Results: Concurrent Validity

Figure 3 tests concurrent validity by estimating the relationships between FIRE and individualism on the one hand and IMCEs for being black and “undeserving” (being unemployed and not seeking employment) on the other, controlling for partisanship and demographics. For a

¹³ In the case of employment, this is a misperception, as welfare recipients are often employed.

¹⁴ Appendix E in Supplementary Material (Figure S3) investigates whether race interacts with other attributes. No notable interaction effects found.

sense of scale, IMCEs can theoretically range from -10 to 10 while observed standard deviations are approximately 1.5 for “Black” and 2.5 for “Not Seek Job.” For comparability, all predictors are recoded to the same 0--1 range. As expected, we find a positive association between FIRE and individual-level stereotypes of welfare recipients as black (although it is significant on the conventional 95% level only in the typicality condition). Furthermore, racial attitudes are the only significant predictor of the racialized welfare stereotype. We also find a positive association between individualism and stereotyping welfare recipients as not seeking employment (significant in the likelihood condition whereas in the typicality condition this effect seems to be partly suppressed by partisanship). Additionally, strong Republicans and older respondents have more pronounced stereotypes of welfare recipients as undeserving.¹⁵

--- Figure3.eps ---

Figure 3. Predictors of stereotypes: race and deservingness dimensions

Note. $N = 1,256$. 95% confidence intervals shown. Full results in Table S2 (Appendix D, Supplementary Material).

Results: Predictive Validity

Finally, we turn to using individual-level stereotypes about welfare recipients to predict welfare policy support. Results are presented in Figure 4. Recall that all IMCEs are on the same scale, so their coefficients can be directly compared. Model without controls shows less support for welfare when its beneficiaries are seen as unauthorized immigrants and unemployed people not seeking employment. This also adds nuance to our earlier finding that, on average, welfare has not been “immigrationized.” While on average respondents have not incorporated immigrants

¹⁵ We re-estimate these effects accounting for uncertainty in the estimated IMCEs (Figure S4 in Appendix F, Supplementary Material). The substantive interpretation of the results is unchanged, though the estimated effect of FIRE on “Black” IMCE in the typicality condition is no longer significant.

into their stereotypes of welfare recipients, those who do stereotype welfare recipients as unauthorized immigrants are significantly less supportive of welfare.

--- Figure4.eps ---

Figure 4. Effects of stereotypes on support for welfare

Note. Controls: FIRE, individualism, partisanship, age, gender, education, income. 95% confidence intervals shown. Full results in Table S3 (Appendix D, Supplementary Material).

When demographic controls, as well as the FIRE scale and individualism, are added to the model, these coefficients decrease but retain statistical significance in the typicality condition (while turning insignificant in the likelihood condition). Race is significantly related to welfare support in the typicality condition only in the model without controls,¹⁶ suggesting that some of the effect of racial stereotypes may be captured in the FIRE scale or partisanship. Overall, in agreement with some recent findings (Fang and Huber 2020; Gross and Wronski 2021), the effect of deservingness-related attributes on welfare support is the strongest and most consistent.¹⁷ These results support the predictive validity of the method.¹⁸

Conclusion

Stereotypes of Welfare Recipients

¹⁶ IMCEs for “Black” and “Hispanic” attribute values are still insignificant when we control only for demographics and partisanship but not for FIRE and individualism (see Figure S6 in Appendix G, Supplementary Material).

¹⁷ We thank an anonymous reviewer for pointing out that this result could be a function of our use of measure of welfare support that includes several questions that might implicate the deservingness of welfare recipients. To evaluate this possibility, Figure S7 (Appendix H, Supplementary Material) replicates the analysis in Figure 4 using the two items from the welfare support scale that are least connected to recipients’ deservingness (“The high cost of welfare...” and “When people can’t support themselves...”). The coefficient of “deservingness” indeed decreases. This can indicate sensitivity of the results to how support for welfare is measured, but also lower reliability of a two-item scale.

¹⁸ We re-estimate these effects accounting for uncertainty in the estimated IMCEs (Figure S5 in Appendix F, Supplementary Material). The substantive interpretation of the results is unchanged, though the estimated effect for “Black” IMCE in the typicality condition without controls is no longer significant.

Consistent with the existing literature, we have found that race, gender, and deservingness-related attributes all play a role in stereotypes about welfare recipients.¹⁹ In addition, an important strength of the conjoint measure is the ability to directly compare the independent contributions of these attributes to the stereotype. We found that deservingness, operationalized as employment seeking behavior, as well as gender and gendered characteristics such as marital status and family size, are as or more important than race. At the individual level, stereotyping welfare recipients as black is associated with lower support for welfare, but this relationship is not statistically significant once controls for demographics, individualism, and racial attitudes are added. Stereotyping welfare recipients as not seeking work, on the other hand, is negatively associated with welfare support even in models that include these controls. This is somewhat surprising given the prominent literature on the link between race and welfare among white Americans (Gilens 1999), though the effect of race may work through deservingness perceptions. Our design randomized race and deservingness independently, so we cannot directly test this conjecture, but it may persist outside controlled experimental settings.

There are several other possible substantive explanations for the relatively small role of race in the welfare recipient stereotype. Race *per se* may play a smaller role in the welfare stereotype than previously thought. This is consistent with explanations suggesting that perceptions of deservingness are the key determinants of welfare attitudes, with race functioning as a culturally constructed heuristic for deservingness (Peffley, Hurwitz, and Sniderman 1997). Welfare stereotypes may have also grown less racialized over time. This may be the result of

¹⁹ Compared to welfare recipient data referenced in footnote 5 (Forster and Rojas 2013; Schettini Kearney 2004), respondents accurately assess non-citizen immigrants as less typical due to eligibility, as well as mothers and unmarried individuals being more typical. But they are inaccurate in recipient race, employment status, and number of children. However, typicality scores are difficult to compare with objective data on welfare recipients, therefore our results do not allow claiming presence of certain (in)correct stereotypes in the public.

long-term changes to the politics of welfare, as the Trump-era Republican Party appears relatively friendlier toward redistributive programs directed at poor whites (Smith and King 2021), or short-term economic factors, specifically the pandemic-induced recession (Bridgman et al. 2021). That said, it is important to note that racial attitudes still strongly and significantly predict welfare opposition among whites.²⁰ Potentially, this effect can be explained by some other mechanism—for instance, beliefs that the government allocates money unfairly across racial groups (Krimmel and Rader 2021).

In addition, purely methodological factors may affect the estimated relationship between racialized stereotypes about welfare recipients and support for redistributive policies. Conjoint experiments may be subject to social desirability bias, though this can only account for the small role of race if associating welfare with race is still perceived as socially undesirable (see Valentino, Neuner, and Vandebroek 2018). In addition, knowledge about formal eligibility rules for certain welfare programs (that sometimes include income/employment) may have increased the estimated effect of “deservingness” vis-a-vis race. Finally, our measure of support for welfare partially implicates deservingness and may have boosted the corresponding effects.²¹ In any case, more studies are needed to determine the nature of welfare attitudes and racial perceptions in contemporary American society. For instance, researchers can use the conjoint-based method proposed here to directly study whether and to what extent whites may stereotype nonwhite groups as “undeserving.”

The ability of the conjoint measure to directly compare the importance of different attributes also allows us to speak to whether welfare has been “immigrationized,” or associated

²⁰ Racial stereotypes can impact support for welfare via racial conservatism and partisanship (although reverse causality is also possible; see Goren 2021), and thus our multivariate model may underestimate the effect of “Black” IMCE. We are grateful to an anonymous reviewer for suggesting this possibility.

²¹ We are grateful to anonymous reviewers for suggesting these possibilities.

with immigrants (Garand, Xu, and Davis 2017; cf. Levy 2021). We find that non-citizen immigrants, regardless of legal status, are no more likely to be seen as welfare recipients than U.S. citizens.²² We also show that the immigration–welfare stereotype measured via conjoint IMCEs has a negative association with welfare support: those who think of welfare recipients as unauthorized immigrants are less supportive of redistributive policies. These results suggest one possible way to reconcile competing findings: on average, attempts to associate undocumented immigrants with welfare have failed, but among those for whom such rhetoric has been successful, it has lowered support for welfare.

The Validity of the Conjoint Measure

In addition to our substantive findings about welfare recipient stereotype content, we validate the conjoint measure of stereotype content. Using the extensive literature on welfare stereotypes, we put forward three core tests that concern, respectively, face, concurrent, and predictive validity of conjoint-based stereotype measures. Our results support all three types of measurement validity. First, consistent with existing literature, respondents see welfare recipients as more likely to be nonwhite, female, not married, parents, unemployed, and lawbreakers. Second, stereotype content is associated with relevant individual-level covariates: for instance, those with individualist values are more likely to stereotype welfare recipients as not seeking employment. Third, some aspects of individuals’ stereotypes about welfare recipients measured via conjoint IMCEs predict welfare policy opinions—even controlling for demographics, partisanship, racial attitudes, and individualism.

Our results also suggest an improvement to the instructions in conjoint experiments measuring stereotypes. Existing uses of conjoint experiments measuring stereotypes ask

²² Importantly, we do not explore whether naturalized immigrants are stereotyped as welfare recipients, which could be a welfare stereotype that some Americans hold.

respondents how likely it is that the person described by the profile is a member of the group (the likelihood instruction). In contrast, studies in the psychological literature on categorization ask respondents to judge how typical a person is of the group (the typicality instruction). Our results suggest that the typicality instruction better measures stereotypes. The typicality instruction passes face, concurrent, and predictive validity checks that we have put forward in the paper setup. The likelihood instruction, in turn, demonstrates notably weaker results in terms of predictive validity. Likelihood instructions may also lead respondents to rely on information that is objectively predictive of group membership, such as (perceptions of) formal eligibility rules, rather than more subjective fit with the group stereotype. While our results do not invalidate previous findings obtained in conjoint experiments with likelihood or probability instructions, they suggest alternative interpretations of some results.

Stereotypes play an important role in political cognition, and conjoint experiments can unlock new advances on studying them. The standard conjoint experiment can measure aggregate-level stereotypes in the society as a whole or the specific sub-populations. Combined with new methods of deriving individual-level estimates, conjoint experiments can provide a measure of stereotype content that can then be used as either an explanatory or dependent variable in subsequent inferential analyses.

That said, this measurement strategy has some downsides. First, unlike measures that rely on universal stereotype dimensions (for instance, “warmth” and “competence” in the stereotype content model; Fiske et al. 2002), or open-ended responses, the attributes included in the conjoint-based measure must be adjusted to the specific research question. Therefore, it requires that investigators have solid *a priori* expectations about possible elements of the stereotype. In this paper we draw on existing literature on the content of welfare stereotypes; in the absence of

such literature, the investigator may start with more qualitative, open-ended methods to first explore stereotype content (Rothschild et al. 2019). Second, researchers should decide whether they are interested in studying aggregate- or individual-level stereotypes before designing the conjoint experiment. If they are only interested in aggregate stereotypes, researchers can use fewer rated profiles and a wider range of attribute values; measuring individual stereotypes, in turn, requires minimizing the number of attribute values and using at least 20 rated profiles per respondent (Zhirkov 2021a).

Stereotypes are central to political cognition; their contents impact how individuals view political actors and government policies, how they relate to other groups in society, and how group affiliations shape beliefs and behaviors via self-stereotyping. The importance of stereotypes is a result of their richness—their ability to encode a variety of kinds of attributes—but this richness also complicates attempts to measure their content. In this paper, we have shown that conjoint experiments provide a valid measure of stereotype content at both aggregate and individual level. We have demonstrated how it can be used to describe a substantively important stereotype, that of welfare recipients, and to examine the relationship between stereotypes, other psychological constructs, and policy opinions. As political scientists increasingly adopt a group-based view of politics, we validate an important tool for understanding how individuals and collectives perceive politically relevant groups, and how those perceptions shape political attitudes and behaviors.

References

Ahler, Douglas J., and Gaurav Sood. 2018. “The Parties in Our Heads: Misperceptions about Party Composition and Their Consequences.” *Journal of Politics* 80 (3): 964–81.

- Amundson, Kalynn, Anna M. Zajicek, and Brinck Kerr. 2015. "A Social Metamorphosis: Constructing Drug Addicts From the Poor." *Sociological Spectrum* 35 (5): 442–64.
- Appelbaum, Lauren. 2001. "The Influence of Perceived Deservingness on Policy Decisions Regarding Aid to the Poor." *Political Psychology* 22 (3): 419–42.
- Banks, Antoine. 2014. *Anger and Racial Politics: The Emotional Foundation of Racial Attitudes in America*. New York: Cambridge University Press.
- Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2018. "The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments." *Political Analysis* 26 (1): 112–19.
- Bodenhausen, Galen V., Sonia K. Kang, and Destiny Peery. 2012. "Social Categorization and the Perception of Social Groups." In *The SAGE Handbook of Social Cognition*, edited by Susan T. Fiske and C. Neil Macrae, 311–29. Los Angeles: SAGE.
- Bridgman, Aengus, Eric Merkley, Peter John Loewen, Taylor Owen, and Derek Ruths. 2021. "All in This Together? A Preregistered Report on Deservingness of Government Aid during the COVID-19 Pandemic." *Journal of Experimental Political Science*. Published ahead of print. <https://doi.org/10.1017/XPS.2021.10>
- Brosius, Hans-Bernd. 2003. "Exemplars in the News: A Theory of the Effects of Political Communication." In *Communication and Emotion: Essays in Honor of Dolf Zillmann*, edited by Jennings Bryant, David R. Roskos-Ewoldsen, and Joanne Cantor, 179–94. Mahwah, NJ: Routledge.
- Brown, Hana. 2013. "The New Racial Politics of Welfare: Ethno-Racial Diversity, Immigration, and Welfare Discourse Variation." *Social Service Review* 87 (3): 586–612.

- Brown, Jessica. 2021. “‘Not Sending Us Their Best’: Media Framings of Immigrants as the ‘Parasitic Poor.’” In *The Routledge Companion to Media and Poverty*, edited by Sandra L. Borden, 277–87. New York: Routledge.
- Brown-Iannuzzi, Jazmin L., Ron Dotsch, Erin Cooley, and B. Keith Payne. 2017. “The Relationship between Mental Representations of Welfare Recipients and Attitudes toward Welfare.” *Psychological Science* 28 (1): 92–103.
- Brown-Iannuzzi, Jazmin L., Erin Cooley, Stephanie E. McKee, and Charly Hyden. 2019. “Wealthy Whites and Poor Blacks: Implicit Associations between Racial Groups and Wealth Predict Explicit Opposition toward Helping the Poor.” *Journal of Experimental Social Psychology* 82: 26–34.
- Calavita, Kitty. 1996. “The New Politics of Immigration: ‘Balanced-Budget Conservatism’ and the Symbolism of Proposition 187.” *Social Problems* 43 (3): 284–305.
- Claassen, Ryan L., Paul A. Djupe, Andrew R. Lewis, and Jacob R. Neiheisel. 2021. “Which Party Represents My Group? The Group Foundations of Partisan Choice and Polarization.” *Political Behavior* 43: 615–36.
- Clawson, Rosalee A., and Rakuya Trice. 2000. “Poverty as We Know It: Media Portrayals of the Poor.” *Public Opinion Quarterly* 64 (1): 53–64.
- Clawson, Rosalee A., and Janell Jett. 2019. “The Media Whiteness of Social Security and Medicare.” *Politics, Groups, and Identities* 7 (1): 207–18.
- Clifford, Scott, and Jennifer Jerit. 2016. “Cheating on Political Knowledge Questions in Online Surveys: An Assessment of the Problem and Solutions.” *Public Opinion Quarterly* 80 (4): 858–87.

- Coppock, Alexander, and Oliver A. McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research and Politics* 6 (1). <https://doi.org/10.1177/2053168018822174>
- DeSante, Christopher D. 2013. "Working Twice as Hard to Get Half as Far: Race, Work Ethic, and America's Deserving Poor." *American Journal of Political Science* 57 (2): 342–56.
- DeSante, Christopher D., and Candis Watts Smith. 2020. "Fear, Institutionalized Racism, and Empathy: The Underlying Dimensions of Whites' Racial Attitudes." *PS: Political Science and Politics* 53 (4): 639–45.
- Eagly, Alice H., and Antonio Mladinic. 1989. "Gender Stereotypes and Attitudes Toward Women and Men." *Personality and Social Psychology Bulletin* 15 (4): 543–58.
- Ellis, Christopher, and Christopher Faricy. 2020. "Race, "Deservingness," and Social Spending Attitudes: The Role of Policy Delivery Mechanism." *Political Behavior* 42: 819–43.
- Entman, Robert M., and Andrew Rojecki. 2000. *The Black Image in the White Mind: Media and Race in America*. Chicago: University of Chicago Press.
- Fang, Albert H., and Gregory A. Huber. 2020. "Perceptions of Deservingness and the Politicization of Social Insurance: Evidence From Disability Insurance in the United States." *American Politics Research* 48 (5): 543–59.
- Federico, Christopher. 2006. "Race, Education, and Individualism Revisited." *Journal of Politics* 68 (3): 600–10.
- Feldman, Stanley, Leonie Huddy, Julie Wronski, and Patrick Lown. 2020. "The Interplay of Empathy and Individualism in Support for Social Welfare Policies." *Political Psychology* 41 (2): 343–62.

- Fiske, Susan T., Amy J. C. Cuddy, Peter Glick, and Jun Xu. 2002. "A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition." *Journal of Personality and Social Psychology* 82 (6): 878–902.
- Flores, Rene D., and Ariela Schachter. 2018. "Who Are the 'Illegals'? The Social Construction of Illegality in the United States." *American Sociological Review* 83 (5): 839–68.
- Foster, Ann, and Arcenis Rojas. 2013. "Program Participation and Spending Patterns of Families Receiving Government Means-Tested Assistance." U.S. Bureau of Labor Statistics. <https://www.bls.gov/opub/btn/volume-2/pdf/spending-patterns-of-families-receiving-means-tested-government-assistance.pdf>
- Garand, James C., Ping Xu, and Belinda C. Davis. 2017. "Immigration Attitudes and Support for the Welfare State in the American Mass Public." *American Journal of Political Science* 61 (1): 146–62.
- Garcia-Marques, Leonel, A. Sofia C. Santos, and Diane M. Mackie. 2006. "Stereotypes: Static Abstractions or Dynamic Knowledge Structures?" *Journal of Personality and Social Psychology* 91 (5): 814–31.
- Gilens, Martin. 1999. *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*. Chicago: University of Chicago Press.
- Goggin, Stephen N., John A. Henderson, and Alexander G. Theodoridis. 2020. "What Goes with Red and Blue? Mapping Partisan and Ideological Associations in the Minds of Voters." *Political Behavior* 42: 985–1013.
- Goren, Paul. 2003. "Race, Sophistication, and White Opinion on Government Spending." *Political Behavior* 25 (3): 201–20.

- Goren, Paul. 2008. "The Two Faces of Government Spending." *Political Research Quarterly* 61 (1): 147–57.
- Goren, Paul. 2021. "Pliable Prejudice: The Case of Welfare." *American Journal of Political Science*. Published ahead of print. <https://doi.org/10.1111/ajps.12652>
- Green, Jon, and Sean McElwee. 2019. "The Differential Effects of Economic Conditions and Racial Attitudes in the Election of Donald Trump." *Perspectives on Politics* 17 (2): 358–79.
- Gross, Kimberly, and Julie Wronski. 2021. "Helping the Homeless: The Role of Empathy, Race and Deservingness in Motivating Policy Support and Charitable Giving." *Political Behavior* 43: 585–613.
- Hancock, Angie-Marie. 2004. *The Politics of Disgust: The Public Identity of the Welfare Queen*. New York: NYU Press.
- Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments against Real-World Behavior." *Proceedings of the National Academy of Sciences* 112 (8): 2395–400.
- Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22 (1): 1–30.
- Heit, Evan, and Stephen P. Nicholson. 2010. "The Opposite of Republican: Polarization and Political Categorization." *Cognitive Science* 34 (8): 1503–16.
- Horiuchi, Yusaku, Zachary Markovich, and Teppei Yamamoto. 2021. "Does Conjoint Analysis Mitigate Social Desirability Bias?" *Political Analysis*. Published ahead of print. <https://doi.org/10.1017/pan.2021.30>

- Jardina, Ashley, Nathan Kalmoe, and Kimberly Gross. 2020. "Disavowing White Identity: How Social Disgust Can Change Social Identities." *Political Psychology* 42 (4): 619–36.
- Kinder, Donald R., and Cindy D. Kam. 2010. *Us against Them: Ethnocentric Foundations of American Opinion*. Chicago: University of Chicago Press.
- Katz, Daniel, and Kenneth W. Braly. 1933. "Racial Stereotypes of One Hundred College Students." *Journal of Abnormal and Social Psychology* 28 (3): 280–90.
- Krimmel, Katherine, and Kelly Rader. 2021. "Racial Unfairness and Fiscal Politics." *American Politics Research* 49 (2): 143–56.
- Landy, David, Brian Guay, and Tyler Marghetis. 2018. "Bias and Ignorance in Demographic Perception." *Psychonomic Bulletin & Review* 25 (5): 1606–18.
- Levy, Morris E. 2021. "Once Racialized, Now 'Immigrationized'? Explaining the Immigration-Welfare Link in American Public Opinion." *Journal of Politics* 83 (4): 1275–91.
- Lippmann, Walter. 1922. *Public Opinion*. New York: Macmillan, Harcourt, Brace.
- Mason, Lilliana, and Julie Wronski. 2018. "One Tribe to Bind Them All: How Our Social Group Attachments Strengthen Partisanship." *Advances in Political Psychology* 39: 257–77.
- Peffley, Mark, Jon Hurwitz, and Paul M. Sniderman. 1997. "Racial Stereotypes and Whites' Political Views of Blacks in the Context of Welfare and Crime." *American Journal of Political Science* 41 (1): 30–60.
- Petersen, Michael Bang. 2012. "Social Welfare as Small-Scale Help: Evolutionary Psychology and the Deservingness Heuristic." *American Journal of Political Science* 56 (1): 1–16.
- Pierson-Balik, Denise A. 2003. "Race, Class, and Gender in Punitive Welfare Reform: Social Eugenics and Welfare Policy." *Race, Gender & Class* 10 (1): 11–30.

- Rabinowitz, Joshua L., David O. Sears, Jim Sidanius, and Jon A. Krosnick. 2009. "Why Do White Americans Oppose Race-Targeted Policies? Clarifying the Impact of Symbolic Racism." *Political Psychology* 30 (5): 805–28.
- Rothschild, Jacob E., Adam J. Howat, Richard M. Shafranek, and Ethan C. Busby. 2019. "Pigeonholing Partisans: Stereotypes of Party Supporters and Partisan Polarization." *Political Behavior* 41 (2): 423–43.
- Schettini Kearney, Melissa. 2004. "Is There an Effect of Incremental Welfare Benefits on Fertility Behavior? A Look at the Family Cap." *Journal of Human Resources* 39 (2): 295–325.
- Schneider, Anne, and Helen Ingram. 1993. "Social Construction of Target Populations: Implications for Politics and Policy." *American Political Science Review* 87 (2): 334–47.
- Schneider, Monica C., and Angela L. Bos. 2011. "An Exploration of the Content of Stereotypes of Black Politicians." *Political Psychology* 32 (2): 205–33.
- Shapiro, Robert Y., Kelly D. Patterson, Judith Russell, and John T. Young. 1987. "Public Assistance." *Public Opinion Quarterly* 51 (1): 120–30.
- Sherman, Jeffrey W. 1996. "Development and Mental Representation of Stereotypes." *Journal of Personality and Social Psychology* 70 (6): 1126–41.
- Smith, Rogers M., and Desmond King. 2021. "White Protectionism in America." *Perspectives on Politics* 19 (2): 460–78.
- Stangor, Charles, and Mark Schaller. 2000. "Stereotypes as Individual and Collective Representations." In *Stereotypes and Prejudice: Essential Readings*, edited by Charles Stangor, 64–82. Philadelphia, PA: Psychology Press.

- Valentino, Nicholas A., Fabian G. Neuner, and L. Matthew Vandenbroek. 2018. "The Changing Norms of Racial Political Rhetoric and the End of Racial Priming." *Journal of Politics* 80 (3): 757–71.
- Weaver, R. Kent. 2000. *Ending Welfare as We Know It*. Washington, DC: Brookings Press.
- Winter, Nicholas J. G. 2010. "Masculine Republicans and Feminine Democrats: Gender and Americans' Explicit and Implicit Images of the Political Parties." *Political Behavior* 32 (4): 587–618.
- Zhirkov, Kirill. 2021a. "Estimating and Using Individual Marginal Component Effects from Conjoint Experiments." *Political Analysis* 30 (2): 236–49.
- Zhirkov, Kirill. 2021b. "Perceptions of Stereotypically Immigrant Groups as Darker-Skinned and Politics of Immigration in the United States and Britain." *Politics, Groups, and Identities*. Published ahead of print. <https://doi.org/10.1080/21565503.2021.1992285>
- Zhirkov, Kirill. 2021c. "Who Are 'the Immigrants'? Beliefs about Immigrant Populations and Anti-Immigration Attitudes in the United States and Britain." *Social Science Quarterly* 102 (1): 228–37.