Evaluating the Use of Conjoint Experiments to Measure Stereotype Content:

The Example of Welfare Recipients

C. Daniel Myers[1], Kirill Zhirkov[2], and Kristin Lunz Trujillo[1,3]

[1] University of Minnesota

[2] University of Virginia

[3] Carleton College

[Draft. Please don't cite or circulate]

Abstract

In this paper, we validate conjoint experiments as a measure of the content of stereotypes about politically relevant groups. 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, which rely on respondents to directly report the content of their stereotypes. However, no existing work evaluates the validity of this new measure. We evaluate this measurement technique using a case where an extensive existing literature gives us strong *a priori* expectations: stereotypes of welfare recipients. We find that aggregate stereotypes, measured via a conjoint experiment, match 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 demonstrate the advantages of the conjoint measure by using it to compare the relative importance of different elements of the stereotype, shedding light on longstanding disputes about the relative importance of race and deservingness as well as on more recent debates about the role of immigration status in the welfare recipient stereotype. Finally, we suggest an improvement in the wording of the conjoint task. 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, what Lippman (1922) famously referred to as the "pictures in our heads" that guide public opinion. These *stereotypes* do not necessarily correspond to reality, though their common cultural sources mean that they are likely to be shared, at least among people who share a culture. They nevertheless have great political import. Stereotypes of social groups strongly and significantly predict attitudes towards partisans (Ahler and Sood 2018), immigrants (Blinder 2015; Zhirkov 2021b), welfare recipients (Gilens 1999), and candidates (Karl and Ryan 2016; Schneider and Bos 2011). On the societal level, politically relevant populations that are negatively stereotyped in the eyes of the mass public are disadvantaged in terms of policy outcomes (Schneider and Ingram 1993).

Exploring the content of these mental representations is essential to understanding the role that they play in political processes. Yet, doing so is difficult. Stereotypes are multidimensional, such that measuring their content requires measuring multiple attributes of various kinds.[1] For example, stereotypes of welfare recipients, the substantive case we study, include demographic characteristics (e.g., race and 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. Finally, individuals may not have conscious access to the content of the stereotypes they hold, and even if they do, may be motivated to describe them in a socially

---

[1] Our approach differs from those that attempt to reduce stereotype content to one or two dimensions (e.g., Fiske et al. 2002), because studying the implications of specific stereotype components—such as, for instance, race vs. "deservingness" in the case of welfare recipients—is often of substantive interest to political scientists.

desirable way—creating challenges for traditional methods of measuring stereotype content that rely on direct questioning.

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; Schachter, Flores, and Maghbouleh 2021). Such experiments show respondents a series of "profiles" that each describes a person in terms of a variety of attributes and then ask subjects to guess that person's category membership. In theory, this task—categorization—has the advantage of directly mirroring the psychological function of stereotypes: sorting individuals into groups. Effects obtained from such conjoint experiments are interpreted as the contributions of the respective attributes to the stereotype of the target group. Since these effects are measured on the same scale, this method can also compare the relative importance of attributes to the stereotype.

However, despite the novelty of the conjoint-based stereotype measure, the existing studies avoid any discussion of validity. While some work validates the use of conjoint experiments to measure preferences (Hainmueller, Hangartner, and Yamamoto 2015; Jenke et al. 2021), none examines its validity as a measure of stereotypes. In this paper, we evaluate the validity of this measure by using it to measure stereotypes about welfare recipients among white Americans. The extensive literature on these stereotypes (e.g., Brown-Iannuzzi et al. 2017; DeSante 2013; Gilens 1999) provides us strong *a priori* expectations for stereotype content, which we use to generate a series of validity tests for the conjoint measure of stereotypes. Using data from a demographically representative survey of white Americans, we find support for face, concurrent, and predictive validity.

In addition to validating the measure, we 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 researcher's ability to examine the individual-level correlates of stereotype content. We apply recent methodological 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 (how likely it is that profiles belong to the target group) to an alternative drawn from an older psychological literature on categorization (how typical profiles are of the target group). We find that the latter produces a better measure of stereotype content.

Finally, the results of this study also have important substantive implications for the politics of welfare in the United States. One advantage of the conjoint measure of stereotype content is that it allows researchers to compare the relative importance of different attributes to the stereotype. In the context of stereotypes about welfare recipients, we find that while race, gender, and deservingness all contribute to the stereotype, gender and deservingness play a larger role than race. We also find that, on average, welfare has not become associated with immigrants in the minds of Americans; however, those who do stereotype welfare recipients as undocumented immigrants are less supportive of welfare policies.

Stereotypes: Nature and Measurement

Stereotypes are "beliefs about the features and characteristics that are associated with social groups" (Bodenhausen, Kang, and Peery 2012, 320), or "pictures in our heads" (Lippmann 1922), that allow us to make sense of the social world. Stereotypes allow us to classify other individuals into groups and, having so classified them, draw inferences about what attributes to

5

expect from them (Hilton and Von Hippel 1996). Rather than learn about each individual one meets, one can save mental energy by categorizing that individual as a member of a group based on salient features and then applying the content of the group stereotypes to the individual. This content, the set of attributes forming the stereotype, might include traits, demographic features, or behaviors that are associated with the group. Among these attributes, some are perceived as more important or central to the stereotype than others. Individuals are categorized as group members, and as more or less prototypical group members, depending on how well they match the stereotyped set of attributes, with more central attributes playing a larger role in this decision than less central attributes. Exactly how stereotypes are stored and retrieved is a subject of debate (Garcia-Marques, Santos, and Mackie 2006; Sherman 1996), but all models share these basic aspects: abstraction, generalization across individuals, and the ability to accommodate a wide range of different attributes.

These core features of stereotypes imply several qualities that a measure of stereotype content should have. First, the ideal measure should recognize the multi-dimensional nature of stereotypes; stereotypes are not merely "positive" or "negative," but a complex mix of attributes that can include positive, negative, and neutral qualities. Second, the measure should be able to accommodate a variety of attributes: demographic characteristics, personal traits, beliefs the group is thought to hold, behaviors the group is thought to engage in, etc. Third, it should be able to 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, and thus measures that rely on direct reports may be particularly susceptible to social desirability bias.

Most existing measures of stereotypes used in behavioral political science 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., behaviors or demographic characteristics). Moreover, ratings on these traits tend to be highly correlated, even when they ask about seemingly distinct qualities. As a result, prominent uses of these items combine them into scales to measure general inter-group affect, such as ethnocentrism (Kinder and Kam 2010). Thus, while useful as measures of general inter-group attitudes, they do not provide a measure of the specific content of group stereotypes.

Other methods allow for the measurement of more specific content, but come with their own weaknesses. The adjective-identification paradigm can accommodate a large number and range of attributes (Katz and Braley 1933), but it 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; Rothschild et al. 2019), but also 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 (Ahler and Sood 2018; Mason and Wronski 2018), but do not incorporate non-demographic attributes and rely on individuals' estimates of proportions that can be biased (Landy, Guay, and Marghetis 2018). Importantly, all of these measures rely on explicit reports that may be affected by social desirability bias. Implicit measures of stereotypes based on time-response tasks have been used as well (e.g., Winter 2010), but they are difficult to implement outside of the lab and require substantial cognitive effort on respondents' part.

Conjoint Experiments as a Measure of Stereotypes

We argue that conjoint experiments offer a better way to measure stereotype content because they have many of the qualities needed to ideally measure stereotypes. As noted above, 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.

One downside to AMCEs is that these represent the average effect of each attribute value across the entire sample—or specific subsamples. However, more recent work has extended this framework to allow for measuring the effects of attributes at the respondent level, or individual marginal component effects (IMCEs; Zhirkov 2021a). IMCE estimation takes advantage of the fact that respondents in conjoint experiments usually rate multiple profiles with randomized attribute values. As a result, it becomes technically possible to estimate marginal component effects for each individual respondent. Such estimation does not rely on any assumptions other than those already required for the standard conjoint analysis. However, feasible estimation of IMCEs adds some requirements to the conjoint design: an interval outcome scale, as well as maximizing the number of rated profiles and minimizing the number of values per attribute.

Conjoint designs in political science are mostly used to measure the effect of attribute values on choice, such as what immigrant to admit or 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; Schachter, Flores, and Maghbouleh 2021).[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 measuring the commonly held stereotypes about the group in question.

This use of a conjoint experiment appears to provide a good measurement of stereotype content because, as closely as is possible in the survey context, it models the real-life categorization task that stereotypes are used for (Macrae and Bodenhausen 2000). 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 a large number of attributes as well as a variety of 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 2020). 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.

---

[2]   Notably, these authors do not use the term "stereotypes." Some say they are measuring the "social construction" of the target groups (Flores and Schachter 2018), while others describe their analysis as measuring "associative networks" between attributes and groups (Goggin, Henderson, and Theodoridis 2020). Stereotypes, however, are defined in the psychological literature exactly as clusters of attributes associated with groups that are to a large degree socially constructed.

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 arenas (e.g., Hainmueller, Hangartner, and Yamamoto 2015), a fact cited as evidence of validity of the conjoint measurement of stereotypes (Flores and Schachter 2018, 845). However, translating the results of conjoint experiments into choices is straightforward since these conjoint experiments literally ask individuals to make choices. Translating conjoint results into stereotypes is not as straightforward: it relies on inferring the content of group stereotypes from the ratings of specific profiles, and depends on how profiles presenting rare combinations of traits are perceived,[3] and assumes that respondents classify conjoint profiles in the same way they classify actual humans in the real world. None of this is to suggest that the conjoint experiment cannot be a valid measure of stereotype content, but merely that its validity should not simply be assumed, as existing studies do.

In addition to validating this measure, 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 of the target group (e.g., 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 to rate the typicality of an object or person for the target category or group (Garcia-Marques, Santos, and Mackie 2006; Heit and Nicholson 2010); we refer to this as the "typicality"

---

[3] As Flores and Schachter (2018, 846) note, "these uncommon combinations—the result of randomizing all the treatments—are what allow us to untangle effects of characteristics that commonly co-occur." However, the authors do not explicitly discuss this issue as a threat to measurement validity. For a point on the importance of real-life attribute values' distributions for conjoint experiments measuring preferences, see De la Cuesta, Egami, and Imai (2021). It is unclear, however, whether the same concern applies to conjoint experiments measuring stereotypes since their goal is to explore the distributions of attributes *in the minds of voters*.

instruction.[4] This literature provides reasons to believe that the typicality instruction will produce a better measure of stereotype content. A key feature of stereotypes is that group members vary in the degree to which they are seen as "good," or prototypical, exemplars of the group (Garcia-Marques et al. 2006), depending on how well their attributes match the content of the group stereotype. An individual can clearly be a group member while also being poor example of the group's characteristics.[5] Asking respondents to assess typicality may produce a more accurate measure by avoiding classifying atypical persons 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 United States (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). As a result, the study of stereotypes of welfare recipients and their consequences has a long history in political science. Moreover, these stereotypes have been consequential, driving public opinion and policy decisions such as the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (Weaver 2000), that famously "ended welfare as we [knew] it."

---

[4]   Studies in the object-categorization literature generally solicit typicality ratings by asking how "good" an exemplar is of a category. Following Garcia-Marques et al. (2006), we do not ask how "good" each exemplar is because this may invite normative judgement, which is particularly problematic when measuring a normatively loaded category like "welfare recipients." While Garcia-Marques et al. (2006)'s studies with Portuguese participants use ask how "characteristic" exemplars are, we simply ask how "typical" they are because "typical" is the more commonly used word in English.

[5]   For example, people with Ph.Ds. are very likely to be Democrats but also atypical of the category "Democrats."

A central part of the stereotype of welfare recipients is that they are undeserving: unwilling to work hard and instead opting to receive resources without putting in effort. Deservingness may be an ingrained 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 that 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 tend to also believe that welfare recipients—and the poor generally—are lazy or freeriding, as well as more likely to abuse substances, be violent, or have a criminal record (Admundson, Zajicek, and Kerr 2015; Clawson and Trice 2000). This 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, and of welfare in particular (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," compared to an image of someone not labeled as being on welfare (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). 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). Further, more recent studies have found that white participants' racial attitudes predict welfare attitudes even when they are not exposed to racial cues (Tesler 2016; Valentino, Neuner, and Vandenbroek 2018). At the same time, some scholars argue that whites' opposition to welfare is a result of individualist values and ideological beliefs about limited government, not stereotypes of welfare recipients as nonwhite (Peffley, Hurwitz, and Sniderman 1997).

Gender and parentage are also prominent in stereotypes of welfare recipients. In conservative rhetoric, single mothers embody the undeserving welfare recipient in the minds of Americans via the "welfare queen" stereotype (H. Brown 2013). At the extreme, women are accused of having more children solely for the purpose of increasing the amount of welfare money they receive. 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 family breakup and single motherhood (Hancock 2004).

These characteristics—deservingness (work ethic), race, gender, marital status, number of children, and criminal record—are all therefore well-established and longstanding stereotypes of welfare recipients among Americans. For this reason, they are ideal for validating the use of conjoint experiments in measuring stereotypes. Further, our study also contributes to this large

literature by directly comparing the centrality of these attributes to the stereotypes of welfare recipients; many studies claim that one is more important than others, and this method can help directly compare them.

More recently, some scholars have argued that immigrants, particularly undocumented immigrants from Latin American countries, have come to play a central role in 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 United States 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 taxpayer paying for people to come into the country and immediately go onto welfare and various other things."[6]

However, the success of this rhetoric is less clear. Some studies find a correlation between negative stereotypes of Hispanics or immigrants and negative attitudes towards welfare (H. Brown 2013; Reese 2005), with one recent study concluding that "the American welfare state today has become more 'immigrationalized' than 'racialized'" (Garand, Xu, and Davis 2017, 147; also see Haselswerdt 2021). But other recent work argues that most Americans have not incorporated immigration status into their stereotypes of welfare recipients, and that the correlation between immigration and welfare attitudes is driven by general conservative ideology, not by a growing association of immigrants and Hispanics with welfare (Levy 2020). By comparing the importance of immigration status to other attributes, the conjoint measurement can shed light on this debate.

---

[6]   https://www.nytimes.com/2019/08/14/us/immigration-public-charge-welfare.html

Validity Tests

A lengthy literature establishes that stereotypes of welfare recipients among white Americans cover a wide range of attributes: "deservingness" (generally characterized in terms of willingness to work and criminal record), race/ethnicity, and gender or gendered characteristics (marital status and number of children).[7] It makes stereotypes of welfare recipients a useful case for validating the conjoint measure by providing strong *a priori* expectations. These include the overall content of the stereotype (e.g., deservingness, race, and gender should play a significant role in the stereotype) as well as strong expectations for how individual characteristics should be related to the stereotype (e.g., those with negative attitudes towards blacks should associate welfare more strongly with blacks), and how these stereotypes should be associated with support for welfare (e.g., people who stereotype welfare recipients as lacking work ethic should show less support for welfare). Moreover, the complexity of the welfare stereotype allows us to demonstrate the key strengths of the measure: its ability to assess the role of multiple characteristics and compare the importance of different characteristics to the stereotypes.

Based on this literature, we develop tests for the face, concurrent, and 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.

Face Validity

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

---

[7]   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 is likely no relationship between welfare receipt and increased birth rate for women of child-bearing age (for a review, see Schettini Kearney 2004).

judge the face validity of a measure by whether it finds an association between race, gender, and deservingness cues on the one hand and welfare receipt on the other hand.

- 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.

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 anti-black attitudes and deservingness cues play a larger role in welfare stereotypes among those high on individualism.

- There should be a positive correlation between anti-black attitudes and IMCEs for the black attribute.

- There should be a positive correlation between individualism and IMCEs for the deservingness attributes.

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 spending.

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

Question Wording

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

<div align="center">

Research Design

</div>

Participants, Procedure, and Measures

We recruited participants through the Lucid platform, an online source of demographically representative convenience samples of U.S. adults, in January 2021. Lucid samples are close to the demographic benchmarks from the American National Election Studies and replicate effects from experiments conducted on national probability samples (Coppock and McClellan 2019). Lucid also provides participants' demographics: age, gender, education, income, race/ethnicity, and partisanship (on the 7-point scale from 1 = *Strong Democrat* to 7 = *Strong Republican*).

A total of 1,309 non-Hispanic white respondents completed the survey.[8] Sample characteristics are as follows. Mean age is 49.5 years and the gender composition is 53.2% female. Median income is between $45,000 and $49,999. As many as 45.8% of respondents report having college degrees. The sample is balanced by partisanship with 35.7% of respondents being Democrats, 38.4% Republicans, and 25.9% independents.

The survey proceeded in the following order. Respondents started from reporting their opinions on welfare policies using four questions about support for redistribution and attitudes to

---

[8] This number excludes two respondents who failed attention checks and 33 respondents who used the exact same rating scores for all profiles in the conjoint experiment.

welfare recipients (see Gilens 1999; Levy 2020). Then, they completed the conjoint task (described in more detail in the section below). After the conjoint task, respondents answered the questions on anti-black attitudes and individualism.

We operationalized negative attitudes towards blacks using the Fear, Institutionalized Racism, and Empathy (FIRE) scale designed to capture both affective and cognitive aspects of modern racial attitudes (DeSante and Smith 2020a). The FIRE scale may better measure racial attitudes among the younger generations (DeSante and Smith 2020b), is not explicitly related to conceptions of deservingness, and has been used in other recent studies (e.g., Green and McElwee 2019; Jardina, Kalmoe, and Gross 2020). Individualism was measured using a four-item battery asking respondents about their support for the ethos of self-reliance (Feldman et al. 2020). See Supplementary Material for the question wordings and response options.

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/ethnicity, gender, marital status, number of children, and immigration status that existing literature identified as elements of stereotypes about welfare recipients. 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 attribute including no criminal record, three violent crimes, and three substance-related crimes. Our expectation was that the effect of the three substance-related crimes would have a similar

18

effect, as would the three violent crimes, so we grouped them together; we used three different crimes in each category to maintain realism.

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	*No*: Zero
	Yes: One, Two, Three
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
	Yes, Drug-related: DUI, Heroin Possession, Drug Sales
	Yes, Violent: Aggravated Assault, Robbery, Threatening with a Weapon

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

---

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

To evaluate the "likelihood" vs. "typicality" wording of the conjoint task instruction, 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 with the response scale from 0 = *Extremely unlikely* to 10 = *Extremely likely*. Participants in the typicality condition were 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 then presented with 30 conjoint tasks; each conjoint task showed a single profile.[10] Figure 1 shows an example of the conjoint task (likelihood condition) as presented to respondents.

Here is a description of a person:

Gender	Male
Employment Status	Unemployed, Seeking Employment
Criminal Record	Yes, Drug Sales (small amount of marijuana)
Immigration Status	U.S. Citizen
Race	White
Number of Children	One
Marital Status	Married

On a scale of 0 to 10, with 0 being extremely unlikely and 10 being extremely likely, how likely do you think it is that this person is on welfare?

Extremely unlikely   Extremely likely
0   1   2   3   4   5   6   7   8   9   10

Figure 1. Example conjoint design profile.

---

[10]   When completing the conjoint experiment, 56 respondents ended up rating less than 30 profiles (the lowest number of rated profiles was 27 for two respondents). These respondents were kept in the analysis.

Results

Face Validity

Figure 2 shows the AMCEs describing aggregate stereotype content contrasting the typicality vs. likelihood response conditions. 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 the existing literature expects to be part of the welfare recipient stereotype (female, not married, parent, unemployed, and rule-breaker) have positive AMCEs. Yet, there are important differences between the typicality and likelihood conditions. First, blacks and Hispanics are significantly associated with welfare recipients in the typicality condition but not in the likelihood condition, though the difference between these conditions is 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 employed makes a profile much less "likely" to be judged as a welfare recipient but has a significantly smaller, though still quite large, effect on judgments of typicality.

Figure 2 also allows comparing 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 characteristic of being unmarried and having children) being more important than race. In contrast to expectations that

21

welfare has become "immigrationized," we find no evidence that being an immigrant is a part of

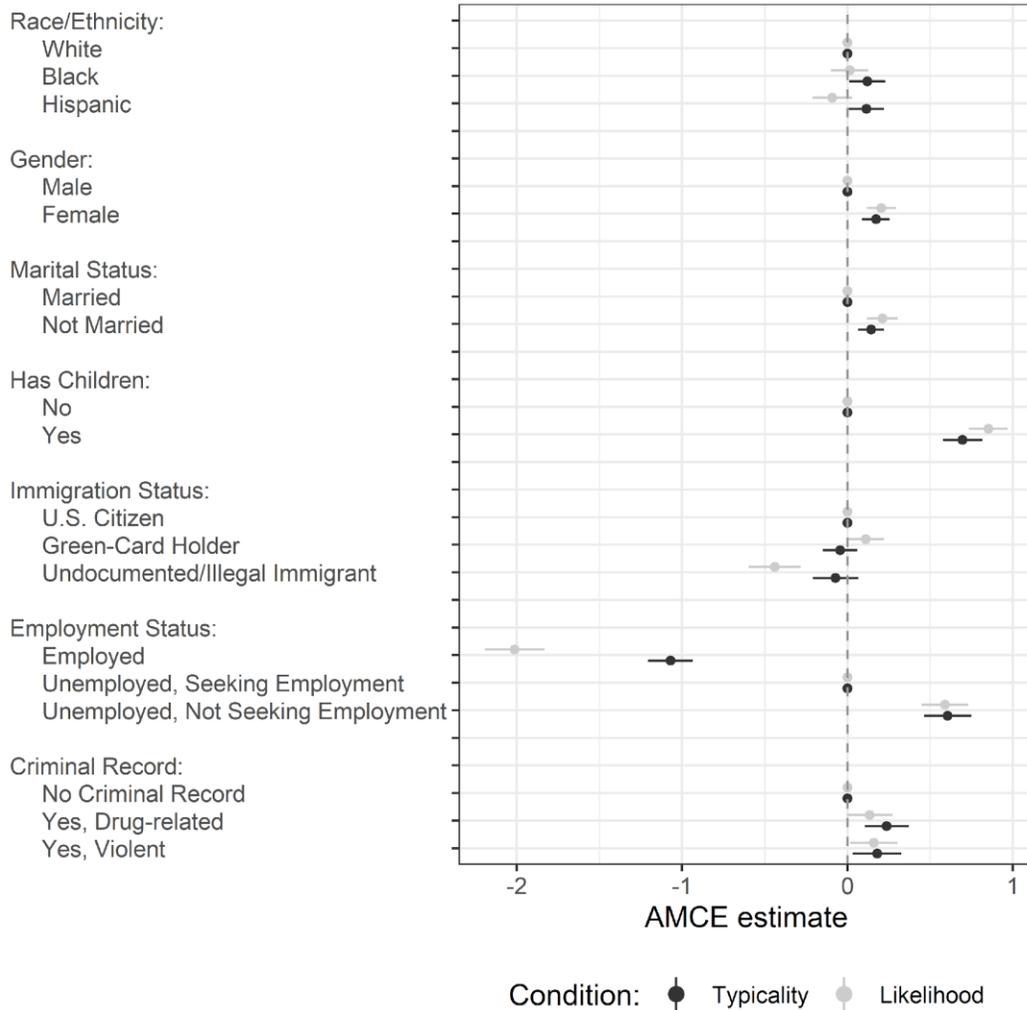the welfare recipient stereotype for whites, on average.

Figure 2. Effects of profile attribute values on stereotype ratings.
Note. N = 38,721 (total number rated profiles by all respondents). Response scales are 0–10 in both conditions. 95% confidence intervals shown. See Table S3 in Supplementary material for full results.

Concurrent Validity

Figure 3 tests concurrent validity by estimating the relationships between FIRE scores and

individualism on the one hand and IMCEs for being black and "undeserving" (being unemployed

and not seeking employment) on the other hand, controlling for partisanship and demographics.

22

As expected, we find a positive association between FIRE scores and individual-level stereotypes of welfare recipients as black (although it is significant on the conventional 95% level only in the typicality condition). Furthermore, anti-black attitudes are the only significant predictor of the racialized welfare stereotype. Also as expected, we find a positive association between individualism and stereotyping welfare recipients as not seeking employment in both the typicality and likelihood conditions. Other interesting effects are that Republicans and older respondents have stronger stereotypes of welfare recipients as undeserving above and beyond the effect of individualism.
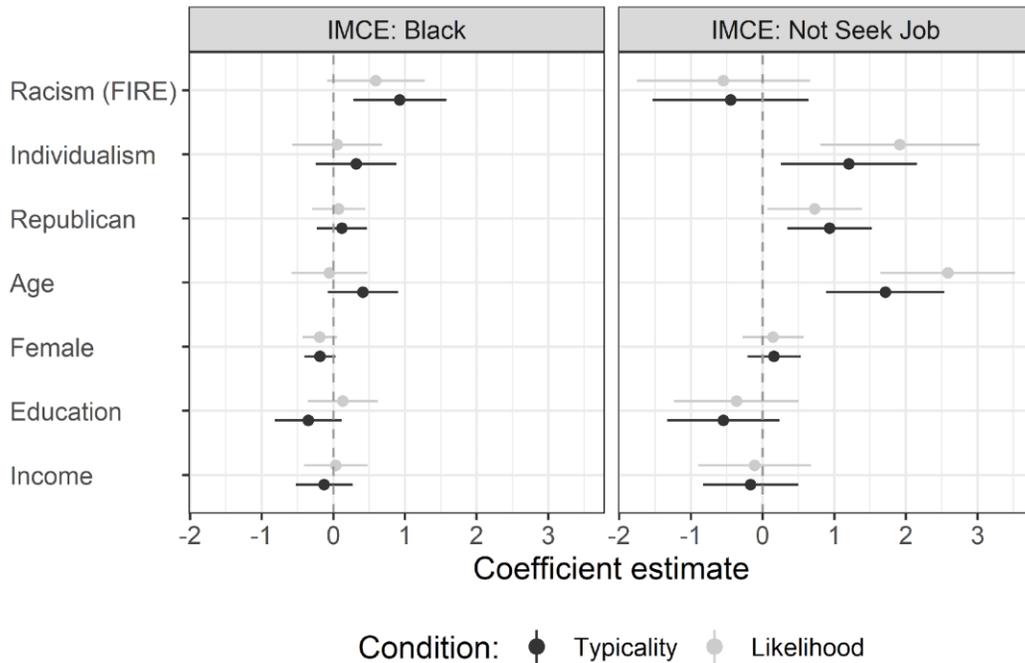

Figure 3. Predictors of stereotypes: race and deservingness dimensions.
Note. N = 1,256. 95% confidence intervals shown. See Table S3 in Supplementary material for full results.

Predictive Validity

Finally, we turn to using individual-level stereotypes about welfare recipients as predictors of welfare policy. Results of the corresponding OLS regressions are presented in Supplemental

Table S2. Model 1, without any controls, shows that welfare spending is supported less when its beneficiaries are seen as unauthorized immigrants (an attribute that likely incorporates both racial and normative considerations) and unemployed people not seeking employment. 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). Overall, in agreement with some recent findings (e.g., Gross and Wronski 2021), the effect of deservingness on welfare policy support is the strongest and most consistent. These results provide support for the predictive validity of the method.
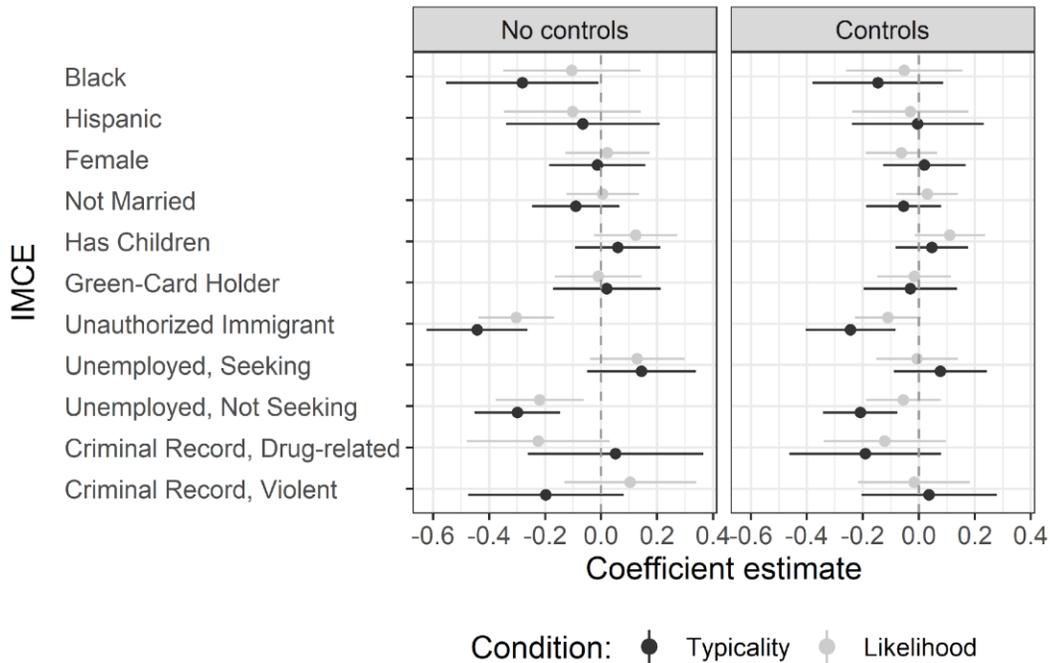

Figure 4. Effects of stereotypes on support for welfare policies.
Note. Controls: racism (FIRE), individualism, partisanship, age, gender, education, income. 95% confidence intervals shown. See Table S3 in Supplementary material for full results.

24

<div align="center">**Conclusion**</div>

Stereotypes of Welfare Recipients

Consistent with the existing literature, we find that race, gender, and deservingness all play a role in stereotypes about welfare recipients. 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 find 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 (e.g., Gilens 1999), although in line with other work that emphasizes perceptions of deservingness as the key determinant of support for redistributive policies (Peterson 2012; Fang and Huber 2020).

There are several possible explanations for the relatively small role of race in the welfare recipient stereotype. First, the conjoint measure may still 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 Vandenbroek 2018). Second, race *per se* may play a smaller role in the welfare stereotype than previously thought. This is consistent with explanations of the race–welfare attitude link that suggests 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; Petersen 2012).

<div align="center">25</div>

Third, welfare stereotypes may have grown less racialized over time. This may be the result of 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.[11] At the same time, it is important to note that anti-black attitudes still strongly and significantly predict opposition to welfare among whites. 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).

The ability of the conjoint measure to directly compare the importance of different kinds of attributes also allows us to speak to the debate about whether welfare has been "immigrationalized," or associated with immigrants (Garand, Xu, and Davis 2017; but cf. Levy 2020). We find that immigrants, regardless of legal status, are no more likely to be seen as welfare recipients than U.S. citizens. 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 the content of the welfare recipient stereotype, we validate the conjoint measure of stereotype content. Using extensive literature in the field, we have put forward three core tests that concern, respectively, face, concurrent, and predictive

---

[11] There is evidence that the COVID-19 pandemic increased overall support for redistributive policies suppressing discriminatory attitudes (Bridgman et al. 2021).

validity of conjoint-based stereotype measures. Our results show strong support for all three types of measurement validity. First, we have found that, consistent with existing literature, respondents see welfare recipients as more likely to be nonwhite, female, not married, parents, unemployed, and law-breakers. Second, stereotype contents are associated with relevant individual-level covariates: for instance, those with individualist values are more likely to see welfare recipients as "undeserving." Third, some aspects of individuals' stereotypes about welfare recipients measured via conjoint IMCEs predict welfare policy opinions—even controlling for demographics, partisanship, racism, and individualism.

Our results also suggest an improvement to the instructions in conjoint experiments measuring stereotypes. Existing uses of conjoint experiments to measure stereotypes ask respondents how likely it is that the person described by the profile is a member of the group (the "likelihood" instruction). This contrasts with studies in the psychological literature on categorization that ask respondents to judge how typical a person is of the group (the "typicality" instruction). Our results suggest that the typicality instruction makes a better measure of stereotypes. The typicality formulation passes face, concurrent, and predictive validity checks that we have put forward in the paper setup. The likelihood formulation, in turn, demonstrates notably weaker results in all three tests. First, results from the likelihood condition suggest that race does not play a significant role in stereotypes about welfare recipients. Second, in the likelihood condition we found no effect of anti-black attitudes on racialization of welfare in the likelihood condition. Third, conjoint IMCEs generated by the likelihood instruction have weak predictive power *vis-a-vis* support for welfare policies. These three results are in clear contradiction to the extensive literature on welfare politics in the United States—and to the results obtained in the typicality condition.

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.

Researchers who plan on using this measure should follow some guidelines. First, unlike measures that rely on universal stereotype dimensions (such as, for instance, "warmth" and "competence" in the stereotype content model), the attributes included in the conjoint-based measure must be adjusted to the specific research question. Therefore, it requires that investigators have some idea 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 preliminarily explore the content of the stereotype. Second, researchers should decide on 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).[12]

---

[12] Researchers may want to minimize the number of tasks in order to conserve survey time or because of concerns about respondent fatigue. However, rating 20 conjoint tasks requires only slightly more than four minutes for the median respondent. Moreover, respondents completing as many as 30 conjoint tasks show few signs of satisficing (Bansak et al. 2018). Together, these suggest that the survey costs of including enough conjoint tasks to measure individual-level stereotypes may be less than appears at first glance relatively modest—especially given the benefits of the measure.

Group 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, one about welfare recipients, to examine the relationship between stereotypes, other psychological constructs, and policy opinions. As political scientists increasingly adopt a group-based view of politics, this measure provides an important tool for understanding how individuals and collectives perceive politically relevant groups, and how those perceptions shape political attitudes and behaviors.

References

Amundson, Kalynn, Anna M. Zajicek, and Brinck Kerr. 2015. "A Social Metamorphosis: Constructing Drug Addicts from the Poor." *Sociological Spectrum*, 35(5): 442–64.

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.

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.

Blinder, Scott. 2015. "Imagined Immigration: The Impact of Different Meanings of 'Immigrants' in Public Opinion and Policy Debates in Britain." *Political Studies* 63 (1): 80–100.

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

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, chap. 26. 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.

Calavita, Kitty. 1996. "The New Politics of Immigration: 'Balanced-Budget Conservatism' and the Symbolism of Proposition 187." *Social Problems* 43 (3): 284–305.

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.

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 & Politics* 6 (1). https://doi.org/10.1177/2053168018822174

De la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. 2021. "Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution." *Political Analysis*. Published ahead of print. https://doi.org/10.1017/pan.2020.40

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 Smith. 2020a. "Fear, Institutionalized Racism, and Empathy: The Underlying Dimensions of Whites' Racial Attitudes." *PS: Political Science and Politics* 53 (4): 639–45.

DeSante, Christopher D., and Candis Smith. 2020b. "Less Is More: A Cross-Generational Analysis of the Nature and Role of Racial Attitudes in the 21st Century." *Journal of Politics* 82 (3): 967–80.

Ellis, Christopher, and Christopher Faricy. 2020. "Race, "Deservingness," and Social Spending Attitudes: The Role of Policy Delivery Mechanism." *Political Behavior* 42: 819–43.

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, René 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.

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): 2395–2400.

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.

Haselswerdt, Jake. 2021. "Social Welfare Attitudes and Immigrants as a Target Population: Experimental Evidence." *Perspectives on Politics* 19 (2): 442–59.

Heit, Evan, and Stephen P. Nicholson. 2010. "The Opposite of Republican: Polarization and Political Categorization." *Cognitive Science* 34 (8): 1503–16.

Hilton, James L., and William von Hippel. 1996. "Stereotypes." *Annual Review of Psychology* 47: 237–71.

Horiuchi, Yusaku, Zachary D. Markovich, and Teppei Yamamoto. 2020. "Does Conjoint Analysis Mitigate Social Desirability Bias?" https://doi.org/10.2139/ssrn.3219323

Jardina, Ashley, Nathan Kalmoe, and Kimberly Gross. 2020. "Disavowing White Identity: How Social Disgust Can Change Social Identities." *Political Psychology*. Published ahead of print. https://doi.org/10.1111/pops.12717

Jenke, Libby, Kirk Bansak, Jens Hainmueller, and Dominik Hangartner. 2021. "Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments." *Political Analysis* 29 (1): 75–101.

Kinder, Donald R., and Cindy D. Kam. 2010. *Us against Them: Ethnocentric Foundations of American Opinion*. Chicago: University of Chicago Press.

Karl, Kristyn L., and Timothy J. Ryan. 2016. "When Are Stereotypes about Black Candidates Applied? An Experimental Test." *Journal of Race, Ethnicity and Politics* 1 (2): 253–79.

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. 2020. "Once Racialized, Now 'Immigrationized'? Explaining the Immigration-Welfare Link in American Public Opinion." *Journal of Politics*. Published ahead of print. https://doi.org/10.1086/711404

Lippmann, Walter. 1922. *Public Opinion*. New York: Macmillan, Harcourt, Brace.

Macrae, C. Neil, and Galen V. Bodenhausen. 2000. "Social Cognition: Thinking Categorically about Others." *Annual Review of Psychology* 51: 93–120.

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.

Reese, Ellen. 2005. *Backlash against Welfare Mothers: Past and Present*. Berkeley: University of California Press.

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.

Schachter, Ariela, Rene D. Flores, and Neda Maghbouleh. 2021. "Ancestry, Color, or Culture? How Whites Racially Classify Others in the U.S." *American Journal of Sociology* 126 (5): 1220–63.

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.

Tesler, Michael. 2016. *Post-Racial or Most-Racial? Race and Politics in the Obama Era*. Chicago: University of Chicago 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*. Published ahead of print. https://doi.org/10.1017/pan.2021.4

Zhirkov, Kirill. 2021b. "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.

Supplementary Material

Support for welfare

"In the next four screens, please indicate how much you agree or disagree with the statement."

- Most people on welfare could get by without it if they really tried. (reversed)

- The high cost of welfare puts too big a burden on the average taxpayer. (reversed)

- When people can't support themselves, the government should help by giving them enough money to meet their basic needs.

- Most people on welfare would rather be working than taking money from the government.

Question order randomized.

Answers coded from 1 = *Strongly disagree* to 7 = *Strongly agree*.

Individualism

"Please indicate how strongly you agree or disagree with the following statements."

- Even if people try hard, they often cannot reach their goals. (reversed)

- Any person who is willing to work hard has a good chance of succeeding.

- Even if people are ambitious, they often cannot succeed. (reversed)

- If people work hard, they almost always get what they want.

Question order randomized.

Answers coded from 1 = *Strongly disagree* to 7 = *Strongly agree*.

FIRE

"Please indicate how strongly you agree or disagree with the following statements."

- I am fearful of people of other races.

- White people in the U.S. have certain advantages because of the color of their skin. (reversed)

- Racial problems in the U.S. are rare, isolated situations.

- I am angry that racism exists. (reversed)

Question order randomized.

Answers coded from 1 = *Strongly disagree* to 7 = *Strongly agree*.

Table S1. Effects of profile attribute values on stereotype ratings by condition

	Typicality	Likelihood
Black	0.12*	0.01
	(0.06)	(0.06)
Hispanic	0.12*	−0.09
	(0.05)	(0.06)
Female	0.17***	0.21***
	(0.04)	(0.05)
Not Married	0.14***	0.21***
	(0.04)	(0.05)
Yes	0.70***	0.85***
	(0.06)	(0.06)
Green-Card Holder	−0.04	0.11*
	(0.05)	(0.06)
Undocumented/Illegal Immigrant	−0.07	−0.44***
	(0.07)	(0.08)
Employed	−1.07***	−2.01***
	(0.07)	(0.09)
Unemployed, Not Seeking Employment	0.61***	0.59***
	(0.07)	(0.07)
Yes, Drug-related	0.24***	0.14
	(0.07)	(0.07)
Yes, Violent	0.18*	0.16*
	(0.08)	(0.07)
N	18,990	19,731

Note. Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S2. Predictors of stereotypes: race and deservingness dimensions

	Typicality		Likelihood	
	IMCE: Black	IMCE: Not Seek Job	IMCE: Black	IMCE: Not Seek Job
Racism (FIRE)	0.93**	0.19	0.59	0.85*
	(0.33)	(0.42)	(0.35)	(0.39)
Individualism	0.32	0.79*	0.05	0.98**
	(0.29)	(0.36)	(0.32)	(0.35)
Age	0.41	1.09***	−0.06	0.52
	(0.25)	(0.32)	(0.27)	(0.30)
Female	−0.19	0.24	−0.19	0.16
	(0.11)	(0.14)	(0.12)	(0.14)
Income	−0.13	−0.27	0.04	0.15
	(0.20)	(0.25)	(0.23)	(0.25)
Education	−0.35	−0.23	0.13	−0.42
	(0.24)	(0.30)	(0.25)	(0.28)
Republican	0.12	0.45*	0.08	0.18
	(0.18)	(0.23)	(0.19)	(0.21)
N	616		640	

Note. Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S3. Effects of stereotypes on support for welfare policies by condition

	Typicality		Likelihood	
	No controls	Controls	No controls	Controls
IMCE: Black	−0.28*	−0.15	−0.11	−0.05
	(0.14)	(0.12)	(0.13)	(0.11)
IMCE: Hispanic	−0.06	−0.00	−0.10	−0.03
	(0.14)	(0.12)	(0.12)	(0.11)
IMCE: Female	−0.01	0.02	0.02	−0.06
	(0.09)	(0.07)	(0.08)	(0.07)
IMCE: Not Married	−0.09	−0.05	0.01	0.03
	(0.08)	(0.07)	(0.07)	(0.06)
IMCE: Has Children	0.06	0.05	0.12	0.11
	(0.08)	(0.07)	(0.08)	(0.06)
IMCE: Green-Card Holder	0.02	−0.03	−0.01	−0.02
	(0.10)	(0.08)	(0.08)	(0.07)
IMCE: Unauthorized Immigrant	−0.44***	−0.24**	−0.30***	−0.11
	(0.09)	(0.08)	(0.07)	(0.06)
IMCE: Employed	0.12	0.11	0.07	0.06
	(0.07)	(0.06)	(0.05)	(0.05)
IMCE: Unemployed, Not Seeking	−0.30***	−0.21**	−0.22**	−0.06
	(0.08)	(0.07)	(0.08)	(0.07)
IMCE: Criminal Record, Drug-related	0.05	−0.19	−0.22	−0.12
	(0.16)	(0.14)	(0.13)	(0.11)
IMCE: Criminal Record, Violent	−0.20	0.04	0.10	−0.02
	(0.14)	(0.12)	(0.12)	(0.10)
FIRE		−0.28***		−0.33***
		(0.04)		(0.04)
Individualism		−0.30***		−0.22***
		(0.04)		(0.04)
Age		−0.02		0.01
		(0.03)		(0.03)
Female		0.00		−0.02
		(0.01)		(0.01)
Income		−0.07**		−0.08**
		(0.03)		(0.03)
Education		−0.02		0.01
		(0.03)		(0.03)
Republican		−0.10***		−0.15***
		(0.02)		(0.02)
N	634	616	659	640