

July 2011

## **Social Welfare as Small-Scale Help: Evolutionary Psychology and the Deservingness Heuristic**

Michael Bang Petersen

Department of Political Science and Government, Aarhus University

Bartholins Allé 7, DK-8000 Aarhus C, Denmark

Email: [michael@ps.au.dk](mailto:michael@ps.au.dk)

Phone: (+45) 89 42 54 26

**Running head:** Social Welfare as Small-Scale Help

**Keywords:** Welfare Opinions; Deservingness; Heuristics; Evolutionary Psychology; ‘Who Said What?’ Paradigm

### **Acknowledgements**

I thank Lene Aarøe, Leda Cosmides, Pete Hatemi, Rob Kurzban, Andreas Roepstorff, Rune Slothuus, Svend-Erik Skaaning, Rune Stubager, Kim Mannemar Sønderskov, Jesper Sørensen, four anonymous reviewers and Editor Rick K. Wilson for thorough advice and comments in the preparation of this article.

Accepted for publication in *American Journal of Political Science*

*Public opinion concerning social welfare is largely driven by perceptions of recipient deservingness. Extant research has argued that this heuristic is learned from a variety of cultural, institutional, and ideological sources. The present article provides evidence supporting a different view: that the deservingness heuristic is rooted in psychological categories that evolved over the course of human evolution to regulate small-scale exchanges of help. To test predictions made on the basis of this view, a method designed to measure social categorization is embedded in nationally representative surveys conducted in different countries. Across the national- and individual-level differences that extant research has used to explain the heuristic, people categorize welfare recipients on the basis of whether they are lazy or unlucky. This mode of categorization furthermore induces people to think about large-scale welfare politics as its presumed ancestral equivalent: small-scale help-giving. The general implications for research on heuristics are discussed.*

Welfare spending constitutes a major part of the government budget of developed democracies, and voters consider many of the core responsibilities of the welfare state, such as education, health care, and care for the elderly, as being highly important (Gilens 2000).<sup>1</sup> Given the political and fiscal importance of welfare, one could perhaps expect people to form opinions on welfare policies via a detailed examination of their content and economic effectiveness. Yet research shows that people predominantly rely on a simple heuristic: Do the recipients deserve the benefits or not (e.g., Gilens 2000; Oorschot 2000; Petersen et al. 2011a; Sniderman, Brody, and Tetlock 1991)? This deservingness heuristic prompts people to categorize recipients as deserving or undeserving on the basis of recipient effort. People strongly oppose welfare spending benefiting individuals they regard as lazy, while they are quite supportive of benefits to the unlucky – i.e., those who try but fail due to circumstances beyond their control.

---

<sup>1</sup> This research was funded by a grant from the Danish Research Foundation to the author (grant #275-07-0068). Replication data is stored at the Danish Data Archive (<http://samfund.dda.dk/default-en.asp>).

The first step when using the deservingness heuristic to treat lazy and unlucky welfare recipients differently is a psychological act of categorization: The recipients must be categorized as the one or the other. Political science research has provided clear insights into the proximate factors governing this process (e.g., Oorschot 2000) and illuminated important national and individual differences in people's default tendencies to view welfare recipients as lazy or unlucky. For example, individuals living in residual welfare states such as the US (Larsen 2006) or individuals on the ideological right-wing are much more prone to think of welfare recipients as lazy (Skitka and Tetlock 1993). At the same time, systematic studies examining the ultimate roots of this categorization process have been lacking. That is, while we know much about the sources of individual differences in perceptions of recipient effort, we do not know why these perceptions of recipient effort are, in the first place, so psychologically important during the formation of welfare opinions.

One prominent researcher simply concedes that this is “an open question” (Oorschot 2006). Other researchers have argued that the very same sources that create variations in the perceptions of recipient effort are themselves responsible for their psychological importance and have pointed to, for example, individualism, policy feedback from specific welfare institutions, and ideological concerns (Gilens 2000; Rothstein 1998; Skitka and Tetlock 1993). The common thread for all of these explanations is that the deservingness heuristic is viewed as “learned” in the sense of being acquired through exposure to a certain culture, a certain set of institutions, or a specific ideology.

In this article, I argue for a different view: that the deservingness heuristic is rooted in psychological categories that evolved over the course of human evolution to regulate small-scale exchanges of help. That is, beneath the real and substantial variation in individual perceptions of the effort of welfare recipients, a species-typical set of psychological categories exists in us all,

motivating us to extract and respond to information about the effort of those in need of help. In making this argument, I follow an increasing number of political scientists and draw on interdisciplinary research concerning the biological foundations of behavior (e.g., Alford and Hibbing 2004; Fowler and Schreiber 2009; Hatemi et al. 2010; McDermott, Fowler, and Smirnov 2008). A review of research on help-giving from evolutionary psychology and related disciplines suggests that categorizing needy individuals on the basis of their effort is a biologically adapted feature of human psychology (Cosmides and Tooby 2005). It is universal, occurs automatically and draws attention towards the factors that were relevant in ancestral small-scale social groups and away from factors that were irrelevant. In the context of modern welfare politics, these insights are used to predict that people universally categorize welfare recipients on the basis of whether their effort is high or low if clear cues in this regard are directly provided. This is predicted to occur even across the sources of variation that generate individual differences in baseline stereotypes about recipient effort. For example, even those who generally believe that welfare recipients are lazy are predicted to pick up cues accurately suggesting otherwise in specific cases. A further prediction is that in the face of such cues, people come to view the evolutionarily recent phenomenon of large-scale welfare decisions as decisions about small-scale help-giving. These predictions are tested in three experimental studies utilizing a method from social psychology to measure implicit and explicit social categorization: the “Who Said What?” design. This design is embedded in nationally representative surveys conducted in two countries: the United States and Denmark. Within the realm of Western developed democracies, these countries are Most Different Systems (Przeworski and Teune 1970) in terms of national culture and type of welfare state. Furthermore, individual-level measures of the ideological positions and political sophistication of the respondents are obtained. Across all of these sources, the predictions are strongly supported.

## **The Deservingness Heuristic: Conceptualization and Origins**

In political science, researchers have defined heuristics as “judgmental shortcuts” (Sniderman, Brody, and Tetlock 1991) and focused on how such shortcuts allow people to form opinions in the absence of substantive political knowledge (e.g., Redlawsk and Lau 2006; Popkin 1991). Our attention to deservingness-relevant information constitutes a heuristic in this sense, and perceptions of whether welfare recipients are lazy or unlucky have been proven to facilitate the formation of opinions towards the welfare state in general (e.g., Gilens 2000), specific welfare programs (e.g., Larsen 2006; Oorschot 2000), and specific recipients within each program (e.g., Petersen et al. 2011a). In psychology, the concept of heuristics has been more rigorously defined as a combination of (1) a fast information processing mode and (2) a frugal decision-rule, where particular judgments follow from the presence of a narrowly defined and limited set of cues (Gigerenzer, Todd, and The ABC Research Group 2001). Deservingness judgments also reflect a heuristic in the more strict psychological sense. A recent study explicitly demonstrates how deservingness judgments occur rapidly and without effort (Petersen et al. 2011a). Specifically, the study reveals how the availability of deservingness-relevant information about welfare recipients both accelerates opinion formation in itself and is most influential when opinions are formed rapidly (i.e., based on a heuristic processing mode). Other studies have outlined the frugal nature of the deservingness heuristic. Hence, deservingness-based welfare opinions are not influenced by all types of positive or negative perceptions of welfare recipients but, in particular, by perceptions of recipient effort. Gilens (2000), for example, shows that of three racial stereotypes (that blacks are lazy, unintelligent, and violent), the American opposition to welfare is driven by laziness alone. Similarly, when examining Danish and American individuals, Petersen et al. (2011b) show that of two stereotypes about welfare recipients in general (that they are lazy and unintelligent), only perceptions of laziness contribute to feelings of anger towards welfare recipients.

In the study of heuristics, it is pivotal to distinguish the heuristic mechanism itself from the input on which the mechanism operates. As demonstrated by a long and important line of research, the input to the deservingness heuristic in the form of perceptions of recipient effort are learned from a large variety of sources (see, e.g., Gilens 2000; Larsen 2006). Recent research even suggests that genetic factors also contribute to variations in how specific individuals perceive others (Sturgis et al. 2010). The importance of this variation notwithstanding, the aim here is not to further illuminate the proximate causes of variation in input to the deservingness heuristic, but to illuminate the ultimate causes of the heuristic mechanism itself. In that regard, psychologists have predominantly referred to learning for many years. “Heuristics (...) are learned knowledge structures,” as Chaiken, Liberman, and Eagly (1989, 213) asserted in their influential work. This learning perspective has influenced how political scientists explain the preoccupation with recipient deservingness in the context of social welfare. Gilens (2000), for one, argues that the heuristic’s use of effort cues is an example of cultural learning and emerges in highly individualistic cultures such as the US. Rothstein (1998) argues that the heuristic is learned from certain institutional arrangements. Specifically, it is argued that liberal welfare states with extensive means-testing make people think about welfare in terms of whether recipients fulfill the necessary criteria and, hence, deserve the benefits. Finally, Skitka, and Tetlock (1993) point to the hypothesis that the heuristic could be part of a conservative ideological script and, hence, linked to an individual’s political ideology.

Yet, while people acquire information about whether welfare recipients are lazy or unlucky from their environment, this does not necessarily imply that the psychological motivation and capacity to seek out this information is in itself a product of learning (Tooby and Cosmides 1992). In fact, in recent years, the psychological literature on heuristics has proposed and provided evidence that some heuristics—or parts of them—are rooted in human nature rather than

representing the product of specific cultural, institutional, or ideological environments (Gigerenzer, Todd, and The ABC Research Group 2001).

Here, I apply this strand of research to illuminate the ultimate roots of the deservingness heuristic. When people utilize this heuristic, the fundamental step is an act of categorization: Recipients must be categorized as either low in effort (i.e., “lazy”) or high in effort (i.e., “unlucky”). On the basis of evidence drawn from the biological sciences, I argue that this propensity to categorize needy individuals on the basis of their effort is not learned. Rather, it seems to be an evolved solution to problems relating to the provision of assistance in the small social groups of our Pleistocene ancestors (see also Petersen et al. 2011b).

### **The Evolution of Help-Giving**

Homo sapiens evolved as social animals. Over the course of evolutionary history, they had to successfully navigate a range of complex social situations. One such situation related to the exchange of help (Cosmides and Tooby 2005). The fossil record, for example, suggests that our ancestors have cared for the crippled and injured as early as 1.77 million years ago (Hublin 2009). Similarly, the evidence suggests that they hunted cooperatively and shared meat in groups for at least 400,000 years (Stiner, Barkai, and Gopher 2009).

Providing assistance in the form of shared resources most likely evolved to buffer variations in the access to resources that are nutritiously important but difficult to acquire, e.g., hunted game (Kaplan and Gurven 2005). In fitness terms, however, providing help is a risky strategy posing non-trivial adaptive problems for the helper. As with the free-rider problem (Olson 1971), providing help is only adaptive to the extent it is reciprocal, i.e., if those who receive also give (Axelrod 1984; Trivers 1971). The challenge, in this respect, is that adaptive help-giving requires one to avoid helping *cheaters*, i.e., organisms that strategically receive more than they give,

instead providing direct help towards *reciprocators*, i.e., those who will return the favor when the shoe is on the other foot. In this manner, a fundamental necessity for the evolution of adaptive help-giving has been the evolution of cognitive categories enabling the individuals to represent and discriminate between “cheaters” and “reciprocators” (Cosmides and Tooby 2005). Detailed psychological experiments have demonstrated how individuals do in fact have such specialized abilities for categorizing others as “cheaters” and “reciprocators” in exchange situations (Brown and Moore 2000; Cosmides and Tooby 2005). These abilities are found across cultures (the success rates on these tasks for Amazonian Indians and Harvard students are indistinguishable) (Sugiyama, Tooby, and Cosmides 2002), reliably develop in children as young as three years (Harris, Nunez, and Brett 2001), rest on the activation of specific regions in the brain, and are independent of general cognitive abilities (Stone et al. 2002).

Evolutionary analysis suggests that evolution has structured the cognitive categories of “cheater” and “reciprocator” such that the major criteria of discrimination relates to the exact cues of effort, i.e., the displayed motivation to accrue and exchange resources, that the deservingness heuristic takes as input (Cosmides and Tooby 2005; Cosmides, Tooby, and Barrett 2010). Neuroscience has provided further evidence supporting a biological foundation for this reliance on effort cues as categorization criteria. Research has revealed that the human brain contains dedicated systems for reading the motivations of others (Baron-Cohen 1997), that these systems are activated in situations of social exchange (Ermer et al. 2006), that intentional—but not unintentional—cheating activates brain regions associated with anger and disgust (Sanfey et al. 2003; de Quervain et al. 2004), and that intentional—but not unintentional—help-giving releases a hormone (oxytocin) found to regulate pro-social behavior (Zak 2007).

Many evolved cognitive categories operate automatically and implicitly, require little effort to use, enter decision-making early in the process, and influence subsequent processing (Price

2008). Conversely, most learned categories would enter the decision-making process at later stages, require explicit and more effortful reasoning, and feel less intuitive and compelling (Boyer 2001). Supporting the evolutionary origins of the categories “cheater” and “reciprocator,” the available evidence suggests that they operate automatically to a significant extent. For example, individuals often cannot explain why they help when they do, and this seems linked to the fact that the hormonal responses to “reciprocators” are regulated by neural systems operating on subconscious levels (Zak 2007). Other strands of research also support this notion of automaticity and have shown that individuals solve tasks involving cheater-detection faster than other comparable tasks (Ermer et al. 2006), that initial attributions of responsibility require substantial effort to override (Skitka et al. 2002), and that the presence of effort cues induces individuals to rely less on explicit opinion formation processes (Petersen et al. 2011a).

### **Welfare Recipients in the Evolved Mind: Predictions**

Building on this evidence from evolutionary psychology, I argue that it is the operation of evolved cognitive categories—rather than learned responses—that draw our attention to deservingness-relevant information in the context of social welfare and enable them to process this information rapidly and effortlessly. In this perspective, modern citizens represent lazy welfare recipients using the evolved category “cheater,” while unlucky welfare recipients are represented as “reciprocators.” This argument entails a number of predictions about how people regard welfare recipients in the presence and absence of cues regarding their effort.

Modern welfare politics should arouse our help-giving psychology, as it basically revolves around whether benefits should be transferred to needy individuals. This being the case, the first and rather obvious prediction is that the structure of the evolved cognitive categories designed to regulate help-giving ought to motivate people to pick up on any available cues

concerning efforts made by welfare recipients and categorize recipients as different on the basis of such cues. In other words, whenever clear cues are available, people should categorize recipients low in effort (“cheaters”) as distinct from recipients who are high in effort (“reciprocators”). In many ways, the validity of this prediction is evident in the existing literature on deservingness. Yet the evolutionary perspective parts with the existing literature when it comes to, first, *how* this categorization operates and, second, *who* engages in it.

As described above, the automaticity of many evolved categories implies that they enter decision-making early in the process and influence the subsequent processing. In order to facilitate adaptive behavior, evolved categories are particularly expected to direct attention towards information that would be relevant to consider in the environment in which it evolved and away from evolutionarily irrelevant information (Cosmides and Tooby 2006). In this regard, it is important that while modern social welfare can be viewed as an abstract form of help-giving, it is played out in a radically different context than the context in which the relevant evolved cognitive categories were selected for. Modern mass society consists of millions of inhabitants, most of whom do not know and will never meet each other. Yet mass society is a recent cultural development. A few thousand years is too little time for natural selection to engineer complex cognitive categories, and reasoning about mass society and its macro dynamics most likely rests on explicit and learned categories (Cosmides and Tooby 2006). Conversely, the design of our psychology of help-giving appears to have evolved over millions of years and is adapted to life in small-scale social groups most likely consisting of 25–200 individuals and characterized by intense face-to-face interaction (Dunbar 1998; Kelly 1995). In other words, the “cheater” and “reciprocator” categories should “presume” a small-scale context and direct attention away from factors that would be irrelevant to consider here.

In the context of modern welfare politics, one such factor is the large-scale nature of the welfare state itself. Upon activation, our evolved cognitive categories should consequently draw attention away from the differences between this mass phenomenon and the small-scale exchanges of help that the categories evolved to regulate. In essence, the second prediction is that people should be induced to think of large-scale welfare in terms of small-scale help-giving when cues concerning the efforts of welfare recipients are present; not because people are not explicitly aware that the problems of state-sponsored welfare are different from help-giving in the family, between friends, or colleagues in everyday life, but because the operation of these automatic, evolved categories should swamp people's decision-making process and make the differences seem irrelevant and uninteresting. Importantly, if it is specifically the activation of these categories that induces people to think about welfare politics as small-scale help-giving, people should be quite able to differentiate between the two in the absence of cheater-relevant information (and, hence, these evolved categories are not activated). Further along these lines, several researchers have argued that people usually regard mass politics as something very removed from their everyday lives (e.g., Carpini and Keeter 1996; Lippmann 1922).

Moreover, according to this hypothesis, the cognitive categories of “cheater” and “reciprocator” are parts of our universal species-typical psychology. In contrast, previous learning-oriented explanations suggest that the tendency to treat unlucky and lazy welfare recipients differently arises from specific welfare institutions, political cultures, or ideologies and predicts variation between individuals in terms of the extent to which they categorize welfare recipients according to their respective efforts. It is therefore possible to test the evolutionary explanation against previous arguments by investigating whether people categorize welfare recipients according to their efforts *across* those cultural factors that previous explanations have indicated to be important. Building on the evolutionary perspective, my third prediction is that all people—even

those living in universal welfare states, in collectivistic cultures, and those who hold left-leaning political convictions—categorize welfare recipients on the basis of effort cues just as easily as citizens in residual welfare states, individualistic cultures, and those with right-leaning political convictions. A final factor to consider here is political sophistication or awareness. While the literature on deservingness judgments has focused on other factors, the more general literature on public opinion has focused on political sophistication as the basis of political learning. As a number of studies demonstrate, political sophistication regulates the extent to which we learn “what goes with what” regarding the communications of political elites (Converse 1964; Zaller 1992). If categorization on the basis of effort cues is learned from political elites, it should be most outspoken among the politically sophisticated. Again, Prediction 3 entails that categorization is independent of this individual-level difference.

### **Research Design and Data**

Three studies were conducted in order to test how people categorize welfare recipients in the presence of effort cues. These studies were based on a method from social psychology designed to measure how people categorize others—the “Who Said What?” method. For the first time, this method was embedded in web surveys administered to nationally representative samples in the United States (study 1) and Denmark (study 2), allowing us to investigate whether categorization is uniform across these two countries despite their differences. The US has a residual welfare state and a highly individualistic culture (Gilens 2001). Conversely, Denmark has a universal welfare state (Esping-Andersen 1990) and a culture characterized by numerous collectivistic elements (Nelson and Shavitt 2002). In this sense, Denmark and the US constitute Most Different Systems (Przeworski and Teune 1970), and observations of uniform responses across these different populations would strengthen the argument that the responses emerge from evolved, species-typical

psychological categories. Using nationally representative samples also provided us with large, individual-level variation in the political ideology and political sophistication of the respondents. This allowed us to test the effect of these potential individual-level moderators of categorization. Study 3 was conducted in Denmark and designed to investigate the specificity of the effects uncovered in the two first studies: Do people stop distinguishing between situations of social welfare and situations of small-scale help-giving when any kind of evolved categories are activated? Or is it specifically the availability of cues to the efforts of recipients that prompts people to regard social welfare as a form of small-scale help-giving?

### *Participants*

In November and December 2010, respondents were recruited for Study 1 (US) and Study 2 (Denmark) by the YouGov survey agency. Samples were drawn from the agency's standing web panels in the countries. To approximate national representativeness, subjects were selected on the basis of their gender, age (> 18), and geography (state in the US case, region in the Danish case). Quote sampling was applied to ensure that the sample corresponded to national proportions on these three dimensions. Because the paradigm used is memory-intensive (see below), it was important to avoid people with age-related working memory deficits. The sample was therefore limited to people under age 50 (Salthouse 1994). There were 1006 subjects in the Danish sample (497 females, 509 males, mean age = 34.90) and 1014 in the US sample (508 females, 506 males, mean age = 33.95).

In April 2010, respondents were recruited for Study 3. Based on the sampling criteria described above, a nationally representative sample of Danish citizens was drawn from the YouGov survey agency's standing web panel. Again, the sample was limited to people under age 50. The total number of subjects was 204 (101 females, 103 males, mean age = 35.23).

## *Procedure*

Study 1 was administered in English, Study 2 in Danish. Other than that, the two studies were identical. Study 3 was also administered in Danish and only varied slightly from the other studies. Given this, I describe the basic setup of the three studies at the same time. In the results section, I expand on how the variation between the first two studies and the third study allows us to illuminate important further aspects of how people categorize welfare recipients.

After a general introduction, subjects were presented with the following information:

*Eight different people will be introduced to you. These individuals will be presented one at a time. We will ask you to look at pictures of these persons and to read what each one says. Try to gain an impression of each individual. Later, you will be asked to answer questions about them.*

On the next eight screens, respondents were presented with pictures of eight different individuals, one at a time, and read a short statement from each. Other than simply looking at the picture and reading the associated statement, the subjects were not supposed to do anything. Each individual was presented on separate screens for 20 seconds, and the screen then automatically proceeded to the next individual. In the first two studies, all of the individuals were young males. In Study 3, there were four males and four females. This presence of female pictures is the only difference between Study 3 and the first two studies (see results section for further discussion). To picture the individuals, eight neutralized pictures of young faces were taken from the Center for Vital Longevity Face Database (Minear and Park 2004) and randomly paired with the statements for each subject.

Each study involved two experimental conditions. The experimental manipulations related to the content of the statements expressed by the individuals as illustrated in Figure 1. In the baseline condition (cf. the sentences in ordinary type in Figure 1), the statements revealed that each of the individuals were either (1) receiving welfare benefits from the state because of

unemployment or (2) had some kind of everyday need that a friend helped them with. Four individuals were presented as welfare recipients and four others as receiving small-scale assistance. In the treatment condition, a statement was added to the baseline presentations which depicted each individual as either being low or high in effort (cf. the sentences in italics in Figure 1). Specifically, a sentence was added to the statements from condition 1, indicating either (1) that the individual was not making an effort to alleviate the need himself (i.e., a “cheater”) or (2) that he was making an effort (i.e., a “reciprocator”). In the treatment condition, an equal number of welfare recipients and everyday help-receivers were depicted as high and low in effort (specifically, two of each). All of the statements and further information on the procedure can be found in the online appendix.

After the respondents had viewed the presentation of all eight individuals, they were confronted with a short distracter task before being introduced to the next task: a surprise recall task. They were told that they would be presented with the statements made by the eight individuals and the images of all eight individuals. They were then asked to mark the image of the individual who made the relevant statement. The order in which the statements from the presentation phase were presented was randomized. The subjects’ answers were recorded.

After obtaining the recall data, data on the two outlined individual-level variables—political ideology and political sophistication—were obtained. To measure political ideology, subjects were asked to place themselves on the political left-right spectrum. Answers were obtained on an 11-point scale ranging from 0 to 10 with the endpoints labeled “Left” and “Right,” respectively. To measure political sophistication, I followed the standard method and relied on factual questions about domestic politics (Zaller 1992). The measure of sophistication is simply the number of correct answers to such questions. All item wordings appear in the online appendix.

### *Analyses*

The procedure outlined above reflects one of the most widely used methods to examine categorization in social psychology: Taylor et al.'s (1978) "Who Said What?" (WSW) method (cf. Sani, Bennett, and Soutar 2005). As is apparent from the above description, the paradigm involves two phases: First, subjects are told that they are to be presented with a series of individuals and asked to form impressions about these individuals. The second phase is a surprise recall phase in which subjects are asked to pair the statements with the individuals, i.e., recall who said what. The crux of the method is that the analysis of the subjects' answers in the recall phase allows one to make inferences about how they have categorized the targets in the presentation phase.

The analysis of the WSW paradigm is oriented towards analyzing patterns in the errors subjects make during recall. If subjects in the recall phase of the present studies mistakenly attribute the statement of one welfare recipient to another welfare recipient, it demonstrates that they have noticed this similarity, mentally categorized both targets as "welfare recipients," and because of this joint categorization mistake one for the other during recall (i.e., either implicitly or explicitly recall that the statement was made by a welfare recipient and then pull out the wrong recipient from memory). In other words, the fundamental premise of the paradigm is that subjects during recall more readily confuse individuals they have categorized together, and psychologists have provided evidence for the correctness of this premise outside of the laboratory (Fiske, Haslam, and Fiske 1991) and in other, more ecologically valid settings (Sani, Bennett, and Soutar 2005). Importantly, given that the paradigm relies on errors in recall, the measures obtained are fully unobtrusive and, therefore, allow one to gauge both implicit and explicit categorization (see Klauer and Wegener 1998). Correct responses are not analyzed due to the uncertainty as to whether they are due to accurate memory, category confusion, or mere guessing.

In the baseline conditions of the present experiments, respondents were presented with individuals that varied along one of the dimensions of the relevant categories: Welfare recipients versus receivers of small-scale help from their friends. The extent to which respondents during recall confuse welfare recipients with other welfare recipients and receivers of help from a friend with other receivers of help from a friend provides us with a baseline measure of the respondents' tendency to think about welfare politics as being different from small-scale exchanges of help. Another dimension of categories was added in the treatment condition: Whether the individuals were low in effort (i.e., "cheaters") or high in effort (i.e., "reciprocators"). The treatment condition allows us to test our core predictions. First, if subjects confuse "cheaters" with "cheaters" and "reciprocators" with "reciprocators" during recall, this would support Prediction 1: people categorize on the basis of effort cues. Second, by comparing the baseline and treatment conditions, we are able to test Prediction 2: that the activation of the categories of "cheater" and "reciprocator" crowds out our attention to the differences between mass welfare politics and small-scale help-giving. In other words, categorization on the basis of that dimension should show a sharp drop in the treatment condition relative to the baseline condition. Finally, by comparing categorization across the US (Study 1) and Denmark (Studies 2 and 3) as well as across the described individual-level moderators, we are able to test Prediction 3 concerning the universality of the predicted categorization patterns.

Exact measures of categorization are calculated by comparing the number of within-category errors (e.g., when the statement of one "cheater" is attributed to another) with the number of between-category errors (e.g., when the statement of a "cheater" is attributed to a "reciprocator"). The higher the number of within-category errors relative to the number of between-category errors, the more strongly subjects confuse individuals along the lines of this category. This provides support for the inference that subjects have categorized the presented individuals using that

category. Statistically, the measure of categorization is basically calculated by subtracting the number of between-category errors from the number of within-category errors (although it should be noted that – as is standard – different correction procedures are used prior to subtraction; all details are provided in the appendix). Two further calculations are made in the presentation of the results (see Table A1 in the appendix): First, to test whether these differences are significant, one uses a t-test of the difference between this difference and 0. From this, p-values are obtained for the difference. Second, the issue is not just whether categorization is significantly different from zero, but how large the tendency to categorize along the dimension is. The magnitude of the raw difference is difficult to interpret, as the exact interpretation also depends on the associated standard deviations. Therefore, I follow Kurzban, Tooby, and Cosmides (2001) and report a standardized measure of the difference between the number of within- and between-category errors by transforming the t-value of this difference into an effect size (cf. Rosenthal 1991). This effect size provides the final measure of categorization and is interpretable as a Pearson's correlation. It varies between  $-1$  and  $1$ , with higher values indicating stronger within-category confusion. A value of  $0$  indicates a lack of categorization in the sense that subjects are no more likely to make a within-category recall error relative to a between-category recall error.

## **Results**

Figure 2 reports the results from the first two experiments conducted in the US (Study 1) and Denmark (Study 2). Panel A reports the findings from the US, while Panel B reports the Danish findings. The displayed coefficients are the effect sizes ( $r$ ) of the difference between the within-category and between-category errors. That is, they reflect how strongly subjects have categorized the presented individuals along the specific dimension. The baseline condition of the experiments allows us to gain a baseline measure of the extent to which respondents categorize on the basis of

whether individuals receive welfare or everyday help in the absence of any effort cues. The white bar on the left side of each panel in Figure 2 (“Baseline Condition”) shows how subjects categorize the individuals in this condition. As can be seen, subjects across the two countries strongly categorize the individuals on the basis of whether they are welfare recipients or receivers of small-scale help (US:  $r = .43$ ,  $p < .001$ ; Denmark:  $r = .56$ ,  $p < .001$ ). In other words, in the condition without effort cues, the subjects in both samples made more within-category than between-category errors in recall and therefore confused welfare recipients with other welfare recipients and receivers of small-scale help with other receivers of everyday help—but not welfare recipients and receivers of everyday help with each other. The reported effect sizes are high. By implication, in the absence of effort cues, receiving social welfare from the state and receiving help from a friend are perceived as two worlds apart for people in both Denmark and the US.

The right side of panels A and B in Figure 2 reports the categorization results obtained in the treatment condition of the studies (“Treatment Condition”). In this second condition, the individuals presented to the subjects varied not only in terms of whether they received social welfare or small-scale help from a friend but also in the effort they exerted towards alleviating their own need. In the treatment condition, in other words, subjects were presented with information that should activate the presumed evolved categories of “cheater” and “reciprocator.” The first prediction entails subjects picking up this information and categorizing the individuals on the basis of their respective efforts. The black bars in Figure 2 show that this is the case. In both the US and Danish experiments, the subjects confuse individuals low in effort with each other (“cheaters”) and individuals high in effort with each other (“reciprocators”) (US:  $r = .31$ ,  $p < .001$ ; Denmark:  $r = .41$ ,  $p < .001$ ). In support of Prediction 2, people’s attention is drawn to cues about the effort made by the recipients of help and they mentally discriminate between low- and high-effort recipients.

According to the evolutionary argument, the psychological categories of “cheater” and “reciprocator” evolved to function within ancestral small-scale environments and are designed to operate on the assumptions relevant in this environment. When these semi-automatic categories are activated, they are expected to swamp the opinion formation process of modern citizens and crowd out attention to information that was irrelevant in the environment in which it evolved.<sup>2</sup> Given this, Prediction 2 entails the attention of the subjects to the distinction between large-scale social welfare provisioning and small-scale help-giving being markedly smaller in the treatment condition. The white bars in Figure 2 display the subjects’ categorization of the scale of the situation, and Prediction 3 can be tested by comparing the white bars in the panels of Figure 2 for the two conditions. As can be seen, the subjects cause the categorization of whether the individuals receive social welfare or small-scale help to drop substantially as compared to the situation without effort cues. In both the US and Denmark, further tests reveal this drop to be highly significant (correlation between experimental manipulation and categorization: US:  $r = -.12$ ,  $p < .001$ ; Denmark:  $r = -.20$ ,  $p < .001$ ). While categorization on the basis of this dimension is still significant in the presence of effort cues, the categorization is not only weak (US:  $r = .19$ ,  $p < .001$ ; DK:  $r = .23$ ,  $p < .001$ ), but also significantly weaker than categorization on the basis of effort (paired t-test of difference in

---

<sup>2</sup> It is important to note that the WSW paradigm measures both implicit and explicit categorization. Nevertheless, the present study does add further support to the existing evidence that categorization on the basis of effort occurs effortlessly. In Studies 1 and 2, after completing the experiment, subjects were directly asked how much effort they put into the recall task. If discriminating between “cheaters” and “reciprocators” reflects an explicit, effort-involving process, we should expect the level of self-reported effort to correlate strongly with this type of categorization. There is no such correlation in neither experiment (US:  $r = 0.04$ ,  $p = 0.27$ ; Denmark:  $r = 0.01$ ,  $p = 0.76$ ). In contrast, when examining a measure that must tap an explicit process, namely the number of correct recalls that subjects have made, we do find an effect of self-reported effort. Hence, in both Denmark and the US, the correlation is highly significant and, in Denmark particularly, of substantial importance (US:  $r = 0.15$ ,  $p < 0.001$ ; Denmark:  $r = 0.34$ ,  $p < 0.001$ ). See the appendix for further discussion.

categorization of the two dimensions: US:  $t = 3.12$ ,  $df = 708$ ,  $p = .002$ ; Denmark:  $t = 5.19$ ,  $df = 700$ ,  $p < .001$ ). In terms of effect sizes, the categorization of “cheaters” versus “reciprocators” is 1.6 and 1.8 times higher than the categorization on of social welfare versus small-scale help-giving in Denmark and the US, respectively. These results support Prediction 2.

### *Assessing National- and Individual-Level Moderators*

Given that these effects are expected to emerge from the operation of evolved, species-typical psychological categories, Prediction 3 entails that they are present across the two countries as well as across individual differences. The analyses above have shown that similar results were obtained in both the US and Denmark. A formal t-test corroborate that we with great confidence can reject the learning-oriented expectation that people in residual welfare states pay more attention to effort cues than people in universal welfare states (t-test of the prediction that categorization is larger in the US than in Denmark:  $t = 2.01$ ,  $df = 1408$ ,  $p = .04$ ).<sup>3</sup> Despite the marked cross-national differences in terms of national culture and welfare state regime, categorization on the basis of effort cues is as good as indistinguishable. Furthermore, the categorization of the effort cues did not correlate with ideology or political sophistication in either of the two countries (US:  $r_{\text{ideology}} = -.07$ ,  $p = .09$ ;  $r_{\text{sophistication}} = .04$ ,  $p = .26$ ; Denmark:  $r_{\text{ideology}} = .04$ ,  $p = .32$ ;  $r_{\text{sophistication}} = -.01$ ,  $p = .78$ ). Consistent with Prediction 3, not only across countries but also across political ideologies and levels

---

<sup>3</sup> If anything, there is a slightly higher level of categorization among Danes (corresponding to a correlation between respondent nationality and categorization of  $r = 0.05$ ;  $p = .04$ ). While this difference just reaches significance, its size is trivial and it most likely reflects minor wording differences between the original and the translated stimuli material (the stimuli was originally developed and pretested in Denmark).

of political sophistication, subjects spontaneously process the efforts of other individuals and, in effect, cheaters and non-cheaters are encoded as distinct.

Prediction 3 is also sustained when we analyze the drop in the categorization of social welfare versus small-scale help-giving in the face of effort cues. This drop is statistically unrelated to the subjects' country, political ideology, and political sophistication. Hence, regressing processing on models with the two-way interactions between the experimental treatment (absence versus presence of effort cues), on the one hand, and these three sources of cultural variation, on the other hand, does not return significant effects (US:  $F_{\text{sophistication*condition}} = .65$ ,  $p = .43$ ;  $F_{\text{ideology*condition}} = 1.44$ ,  $p = .23$ ;  $F_{\text{country*condition}} = 3.83$ ;  $p > .05$ ; DK:  $F_{\text{sophistication*condition}} = .02$ ,  $p = .89$ ;  $F_{\text{ideology*condition}} = .30$ ,  $p = .58$ ; all models appear in Table A2 in the appendix). Across countries, ideologies, and levels of political sophistication, the presence of effort cues crowds out the sense of the provisioning of social welfare as something distinct from small-scale help-giving.

### *Assessing Alternative Explanations: Results from Study 3*

Studies 1 and 2 support the specified predictions. Study 3 tests whether an alternative explanation can account for these effects. The argument drawn from evolutionary psychology is that the availability of effort cues in the context of social welfare causes people to think about social welfare as small-scale help-giving, because other cues associated with social welfare (most prominently, a need and the granting of benefits) enable it to be mentally construed as such. As argued in theories of bounded rationality (e.g., Jones 2001), however, the attention capacity of the individual is quite limited. The drop in the categorization of social welfare versus small-scale help-giving in the treatment condition relative to the baseline condition could therefore simply reflect a zero-sum relationship between categorization on the basis of distinct cues; that is, when effort cues are present, general attention constraints weaken categorization on the basis of other cues.

Study 3 was devised to investigate whether the findings obtained in Studies 1 and 2 are simply driven by general attention constraints. To test this, Study 3 added an additional dimension of variation to the presented individuals: Their sex. In half of the cases (and orthogonally to the other dimensions), the pictures used in the first two experiments were substituted with images of young females. Sex as a cue has a number of properties that are important here. Sex is an extremely important evolutionary attribute and organisms without cognitive mechanisms designed to identify what constitutes the pool of potential mating partners would not have done well over the course of evolution (Buss 2005). In line with this, previous studies using the memory confusion protocol demonstrate that categorization on the basis of sex is very strong (Kurzban, Tooby, and Cosmides 2001; Lieberman et al. 2008) and proceeds automatically (Spears et al. 1999). In this way, it is expected that categorization on the basis of sex shares fundamental properties with categorization on the basis of cheater-relevant information in terms of evolutionary status and automaticity. If a decrease in categorization on the basis of social welfare versus small-scale help results from general processing constraints, the availability of sex cues should trigger it. If the effect is specifically triggered by cheater-relevant cues, however, the addition of sex as a dimension of categorization should leave any predicted categorization patterns from Studies 1 and 2 unchanged.

Figure 3 reports the results from this third experiment. The left side of the Figure reports the categorization of the first condition in which the individuals now vary both in terms of sex and the scale of the situation. Again, the white bars reflect the categorization of social welfare versus small-scale help. The grey bars show the categorization of males versus females. As can be seen, the subjects categorize strongly on the basis of both dimensions (social welfare versus small-scale help;  $r = .56$ ,  $p < .001$ ; Females versus males:  $r = .40$ ,  $p < .001$ ). If we compare the categorization of social welfare versus small-scale help in Study 3 to Study 2 (the two experiments

with Danes), we see that they are exactly the same with and without sex cues (t-test of difference in categorization across the studies:  $t = -.04$ ,  $df = 365$ ,  $p = .97$ ).

In the treatment condition in Study 3, all three sets of cues are present. The results are presented in the right side of Figure 3. First, as revealed by the grey bar, the subjects continue to categorize the individuals on the basis of their sex in the presence of effort cues ( $r = .46$ ,  $p < .001$ ). In fact, the categorization on the basis of sex is slightly (but insignificantly) higher when effort cues are also present (t-test of difference in categorization across the conditions:  $t = .29$ ,  $df = 202$ ,  $p = .77$ ). In addition to categorizing on the basis of sex, subjects again categorize on the basis of effort cues and, hence, treat “cheaters” and “reciprocators” as distinct ( $r = .27$ ,  $p = .001$ ). Although we see a slight drop in categorization on the basis of effort compared to the condition without sex cues (black bars in Figure 2), categorization remains substantial. Third, the presence of effort cues facilitates a very strong drop in the tendency of subjects to categorize social welfare receivers as being different from the receivers of small-scale assistance (correlation between experimental manipulation and categorization:  $r = -.37$ ,  $p < .001$ ). In fact, in this condition, categorization on the basis of the scale of the situation disappears entirely ( $r = -.02$ ,  $p = .78$ ). Here, subjects do not in any way attend to whether people receive social welfare or whether they receive everyday help from a friend. Psychologically speaking, they are treated as the same.

The Study 3 results show how subjects are quite able to categorize on the basis of several uncorrelated dimensions at once. It is not the presence of any alternative cue—even one that is processed automatically—but rather the specific presence of effort cues that prompts people to conflate social welfare and small-scale help-giving. It is also worth noting how the effect of effort cues on the categorization on the basis of other cues was specific: effort cues removed subjects’ tendency to use the presumed learned distinction between large-scale and small-scale situations, but not the evolutionarily important category of sex.

## **Conclusions and Implications**

The reported experiments have uncovered two important features of how people categorize welfare recipients on the basis of effort cues: First, it has been demonstrated that people categorize welfare recipients on the basis of available effort cues across the cultural and individual-level variations previously thought to moderate attention to the efforts of recipients. In this manner, the present results speak against the hypothesis that the deservingness heuristic as a mechanism is a product of feedback processes in welfare states with widespread means-testing, individualistic national cultures, conservative ideological scripts, or lack of political sophistication. While these and other environmental factors—together with genetic differences across individuals—create substantial variation in individuals' default propensity to see welfare recipients as lazy or unlucky, these factors do not in themselves explain why people pay close attention to the effort of needy individuals. While the input to the mechanism varies, the mechanism itself seems to exist in us all.

Second, the presence of cues regarding recipients' efforts has been demonstrated to induce people to think about welfare provisioning as being no different from small-scale exchanges of help between friends. It is not that people do not initially recognize the differences between these situations. In the absence of effort cues, people attend strongly to their difference, but this difference seems to be considered irrelevant in the presence of effort cues. It was also observed that the automaticity of effort categorization in itself does not crowd out attention to the large-scale nature of welfare provisioning. The activation of another set of highly automatic categories, male versus female, did not have similar effects. It is only when cues relevant to help-giving (a need, a transfer of resources, etc.) is specifically paired with cues regarding the efforts of the recipient that political opinion formation is swamped and people construe large-scale welfare provision as small-scale help-giving.

These present results provide support for the hypothesis that the deservingness heuristic is rooted in evolved cognitive categories designed to detect and represent “cheaters” and “reciprocators,” respectively. The cross-cultural nature of the present evidence supports the view that these categories are not learned. Rather, they are something we as humans naturally come equipped with. Americans and Danes are obviously not representative of humans across time and space. Still, these populations are maximally different on factors pinpointed by learning-oriented accounts as responsible for the attention to effort cues. Furthermore, the fact that the activation of these categories crowds out attention to the large-scale nature of modern welfare states supports the view that these categories evolved to solve adaptive problems in ancestral small-scale environments and therefore direct attention to the features that were relevant in this environment. In this perspective, when people apply the deservingness heuristic to modern issues of welfare, they are applying a species-typical mode of thought that evolved as a response to the specific adaptive problems posed by help-giving in the course of human ancestral history.

Arguing for the evolved roots of the deservingness heuristic constitutes in many ways a most likely case. For thirty years, evolutionary theorists have demonstrated the cross-cultural existence of highly sophisticated abilities to detect cheaters and cooperators in help-giving situations (Cosmides and Tooby 2005). At the same time, there is reason to believe that the deservingness heuristic is not the only political heuristic influenced by our ancestral past. As group-living animals, a range of proto-political problems such as sharing decisions, the formation of collective action, punishment of free-riders, management of intergroup relations, and hierarchy formation would have impinged directly on the survival and reproduction of our ancestors, and evolutionary psychologists have produced evidence indicating the existence of a number of evolved decision-rules in each of these domains (for an overview, see Buss 2005). In this manner, the

arguments in the present article might have more general implications for the political science literature on heuristics.

In particular, the notion of evolved heuristics sheds novel light on a much debated topic within the political science literature on heuristics: Do heuristics produce opinions that are biased or “good enough” relative to more effortful processes (see Lau and Redlawsk 2006; Drukman, Kuklinski, and Sigelman 2009; Kuklinski and Quirk 2000)? From an evolutionary perspective, heuristics do not produce either one or the other. Instead, evolved heuristics are *ecologically rational*; that is, they evolved by natural selection to operate effectively within a particular environment (Gigerenzer, Todd, and The ABC Research Group 2001; Cosmides and Tooby 2006). The effectiveness of heuristics is a matter of the match between the heuristic and the environment in which it is deployed, not merely the heuristic in itself. In the case of mass politics, however, this match will often be poor because evolved heuristics are designed to operate in small-scale rather than large-scale situations. Opinions concerning mass politics arising from the operation of evolved heuristics are therefore likely to be biased in specific respects relating to the differences between these kinds of situations. Here, welfare politics is illustrative. In small-scale environments, helping is a simple matter of transferring resources to a needy individual. In large-scale environments, however, the effectiveness of any social solutions is also affected by macro dynamics that do not arise in small-scale environments. Because evolved heuristics induce people to think about large-scale politics as small-scale sociality, they will miss such dynamics altogether when forming opinions. Extant research has identified several examples consistent with such a lack of attention to macro dynamics. One such example relates to rent control (cf. Cosmides and Tooby 2006). Rent control as a form of assistance makes sense from a small-scale perspective: Resources are transferred from the landlords to the poor. Yet macro-economic dynamics may cause rent control to have adverse consequences that may in fact neutralize and possibly exceed any positive

effects. Hence, some studies have uncovered a significant relationship between whether a city has adopted rent control policies and the level of homelessness such that cities *with* rent control have higher homeless rates than cities without it (e.g., Grimes and Chressanthis 1997). Given macro-economic dynamics, rent control reduces the supply of rental units, thereby offsetting some of the potential positive effects of rent control. Although economists disagree to some extent over the size of the negative macro dynamics (Jenkins 2009), people will fail to even consider them if the activation of evolved categories causes them to think in purely small-scale terms. Another example relates to targeted versus universal welfare programs. From a small-scale perspective, targeted programs should be most effective in helping the needy, as they bring resources specifically to those in need. Yet this small-scale logic cannot be easily applied to large-scale welfare provisioning, where comparative studies have provided compelling evidence that the most redistributive programs are programs that are not targeted but universal. This is again due to macro dynamics—here, among the electorate—which drive down the benefit levels in targeted programs (Korpi and Palme 1998; Rothstein 1998). Politically, it has been shown to be possible to sustain high levels of benefits from a welfare program only when the electorally important middle class are among those benefiting from the program. When they do not benefit, most voters can easily be persuaded to support scaling down the benefit levels. Because of this electoral macro dynamic, universal programs are on balance more redistributive than targeted programs. This net result arises from (1) their high benefit rates and (2) that higher income groups contribute more to the program via taxation than low income groups. Again, thinking about large-scale welfare provisioning as small-scale help-giving would make people miss such macro dynamics when forming welfare opinions.

The power of evolutionary psychology lies in fleshing out the structure of the species-typical psychological mechanisms that have ensured the survival and reproduction of our species

over the course of human evolution. Given this focus on mechanisms, evolutionary psychology provides excellent tools for studying the structure of heuristics. At the same time, this focus implies that evolutionary psychology can easily be integrated with other levels of analysis. In the present study, for example, I have argued that the importance of responding differently to “cheaters” and “reciprocators” in help-giving situations over the course of human evolution causes people across cultures, ideologies, and levels of sophistication to think about modern welfare in terms of whether recipients are lazy or unlucky. The *mechanism* underlying such categorizations seems to be rather universal, as would be expected given their importance for the evolution of human cooperation. Yet manifest differences in welfare opinions will readily occur if the input to these mechanisms varies across individuals. Such variations can occur between countries where, for example, the level of ethnic heterogeneity (Alesina et al. 2001) or level of unemployment (Gilens 2000) seem to influence perceptions of the efforts of welfare recipients. Variations can also occur between individuals within the same country. While ideological predispositions do not influence individuals’ tendencies to discern between sloth and misfortune in the presence of clear effort cues, liberals and conservatives still differ in their baseline stereotypes about welfare recipients. Liberals are likely to see most welfare recipients as “reciprocators,” while conservatives likely view them as “cheaters.” Such individual-level variations can also emerge from genetic differences as is increasingly emphasized in applications of behavior genetics to the study of politics (e.g., Alford, Funk, and Hibbing 2005; Hatemi et al. 2010). While species-typical mechanisms are the focus of evolutionary psychological theory, the present results do not preclude the importance of genetic differences capable of influencing people’s propensity to, e.g., view others as “cheaters.” Instead, these results provide directions for the search for points around which these two major biological approaches to behavior, evolutionary psychology, and behavior genetics can converge. First, the present results lend support to the assertion of both approaches that most between-population genetic differences

are small (see, e.g., Boyd and Silk 2003). Hence, Americans and Danes did not differ in their sensitivity to effort cues, and any potential genetic differences of importance should therefore be located within each of these populations. Second, the present results suggest that heritable differences in welfare attitudes can emerge as downstream effects of heritable differences in the tendency of people to view others as either low or high in effort, in effect serving as input to a universal susceptibility to effort cues (see Tooby and Cosmides 1990). While it is increasingly acknowledged that manifest behavior is an interaction between nature and nurture, this opens for the possibility that the contribution of nature is in itself an interaction—an interaction between genetic differences and genetic similarities in the form of evolved universal mechanisms.

## References

- Alesina, Alberto, Edward Glaeser & Bruce Sacerdote (2001). “Why Doesn’t the US Have a European-Style Welfare System?” *Brookings Papers on Economic Activity* 2: 187-254.
- Alford, John, and John Hibbing. 2004. “The Origins of Politics: An Evolutionary Theory of Political Behavior.” *Perspectives on Politics* 2: 707–23.
- Alford, John R., Carolyn L. Funk, and John R. Hibbing. 2005. “Are Political Orientations Genetically Transmitted?” *American Political Science Review* 99(2): 153–67.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Baron-Cohen, Simon. 1997. *Mindblindness*. Cambridge, MA: The MIT Press.
- Boyer, Pascal. 2001. *Religion Explained. The Evolutionary Origins of Religious Thought*. New York: Basic Books.
- Boyd, Robert, and Joan B. Silk. 2003. *How Humans Evolved*. 3<sup>rd</sup> ed. New York & London: W.W. Norton & Company.
- Brown, William Michael, and Chris Moore. 2000. “Is Prospective Altruist-Detection an Evolved

- Solution to the Adaptive Problem of Subtle Cheating in Cooperative Ventures? Supportive Evidence Using the Wason Selection Task.” *Evolution & Human Behavior* 21: 25–37.
- Buss, David M. (ed.). 2005. *The Handbook of Evolutionary Psychology*. Hoboken, NJ: John Wiley and Sons, Inc.
- Chaiken, Shelly, Akiva Liberman, and Alice Eagly. 1989. “Heuristic and Systematic Processing within and beyond the Persuasion Context.” In *Unintended Thought*, eds James S. Uleman and John A. Bargh. New York: The Guilford Press, 212–52.
- Converse, Philip E. 1964. “The Nature of Belief Systems in Mass Publics.” In *Ideology and Discontent*, ed. David E. Apter. New York: The Free Press, 206–261.
- Cosmides, Leda, and John Tooby. 2005. In *The Handbook of Evolutionary Psychology*, ed. David M. Buss. Hoboken, NJ: John Wiley and Sons. 584–627.
- Cosmides, Leda, and John Tooby. 2006. “Evolutionary Psychology, Moral Heuristics, and the Law.” In *Heuristics and the Law*, eds. Gerd Gigerenzer and Christoph Engel. Cambridge: MIT Press, 182–212.
- Cosmides, Leda, H. Clark Barrett, and John Tooby. 2010. “Adaptive Specializations, Social Exchange, and the Evolution of Human Intelligence.” *Proceedings of the National Academy of Sciences* 107: 9007–14.
- de Quervain, Dominique J. F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. “The Neural Basis of Altruistic Punishment.” *Science* 305: 1254–8.
- Druckman, James N., James H. Kuklinski, and Lee Sigelman. 2009. “The Unmet Potential of Interdisciplinary Research: Political Psychological Approaches to Voting and Public Opinion.” *Political Behavior* 31: 485-510.
- Dunbar, Robin I. M. 1998. “The Social Brain Hypothesis.” *Evol Anthro* 6: 178–90.

- Ermer, Elsa, Scott Guerin, Leda Cosmides, John Tooby, and Michael Miller. 2006. "Theory of Mind Broad and Narrow: Reasoning about Social Exchange Engages TOM Areas, Precautionary Reasoning Does Not." *Social Neuroscience* 1: 196–219.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Polity Press.
- Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90: 980–94.
- Fiske, Alan P., Nick Haslam, and Susan E. Fiske. 1991. "Confusing One Person with Another: What Errors Reveal about the Elementary Forms of Social Relations." *Journal of Personality and Social Psychology* 60: 656–74.
- Fowler, James H., and Darren Schreiber. 2009. "Biology, Politics and the Emerging Science of Human Nature." *Science* 322: 912–4.
- Grimes, Paul W., and George A. Chressanthis. 1997. "Assessing the Effect of Rent Control on Homelessness." *Journal of Urban Economics* 41(1): 23–37.
- Gigerenzer, Gerd, Peter M. Todd, and The ABC Research Group. 2001. *Simple Heuristics that Make Us Smart*. New York: Oxford University Press.
- Gilens, Martin. 2000. *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*. Chicago: University of Chicago Press.
- Harris, Paul L., Maria Nunez, and Caroline Brett. 2001. "Let's Swap: Early Understanding of Social Exchange by British and Nepali Children." *Memory & Cognition* 29(5): 757–64.
- Hatemi, Peter K., John Alford, John Hibbing, Matthew Keller, Nicholas Martin, Sarah Medland, and Lindon Eaves. 2010. "Not by Twins Alone: Using the Extended Twin Family Designed to Investigate the Genetic Basis of Political Beliefs." *American Journal of Political Science* 54(3): 798–814.

- Hublin, Jean-Jacques. 2009. "The Prehistory of Compassion." *Proceedings of the National Academy of Sciences USA* 106: 6429–30.
- Jenkins, Blair. 2009. "Rent Control: Do Economists Agree?" *Econ Journal Watch* 6(1): 73–112.
- Jones, Bryan. 2001. *Politics and the Architecture of Choice*. Chicago: University of Chicago Press.
- Kaplan, Hillard, and Michael Gurven. 2005. "The Natural History of Human Food Sharing and Cooperation." In *Moral Sentiments and Material Interests*, eds Herbert Gintis, Samuel Bowles, Robert T. Boyd, and Ernst Fehr. New York: MIT Press, 75–113.
- Kaplan, Hillard, and Kim Hill. 1985. "Food Sharing Among Ache Foragers: Tests of Explanatory Hypotheses." *Current Anthropology* 26(2): 223–46.
- Kelly, Robert L. 1995. *The Foraging Spectrum. Diversity in Hunter-Gatherer Lifeways*. Washington: Smithsonian Institution Press.
- Klauer, Karl Christoph, and Ingo Wegener. 1998. "Unraveling Social Categorization in the 'Who Said What?' Paradigm." *Journal of Personality and Social Psychology* 75, 1155–78.
- Korpi, Walter, and Joakim Palme. 1998. "The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries." *American Sociological Review* 63: 661–87.
- Kuklinski, James H., and Paul J. Quirk. 2000. "Reconsidering the Rational Public: Cognition, Heuristics, and Mass Opinion." In *Elements of Reason*, eds Arthur Lupia, Mathew McCubbins, and Samuel Popkin. Cambridge: Cambridge University Press.
- Kurzban, Robert, John Tooby, and Leda Cosmides. 2001. "Can Race be Erased? Coalitional Computation and Social Categorization." *Proc Natl Acad Sci USA* 98: 15387–92.
- Lau, Richard R., and David P. Redlawsk. 2006. *How Voters Decide*. Cambridge, MA: Cambridge University Press.

- Lieberman, Debra, Robert E. Oum, and Robert Kurzban. 2008. "The Family of Fundamental Social Categories includes Kinship: Evidence from the Memory Confusion Paradigm." *European Journal of Social Psychology* 38: 998-1012.
- McDermott, Rose, James H. Fowler, and Oleg Smirnov. 2008. "On the Evolutionary Origin of Prospect Theory Preferences." *The Journal of Politics* 70(2): 335–50.
- Miner, Meredith, and Denise C. Park. 2004. "A Lifespan Database of Adult Facial Stimuli." *Behavior Research Methods, Instruments, and Computers* 36: 630–33.
- Nelson, Michelle R., and Sharon Shavitt. 2002. "Horizontal and Vertical Individualism and Achievement Values: A Multimethod Examination of Denmark and the United States." *Journal of Cross-Cultural Psychology* 33: 439–58.
- Olson, Mancur. 1971. *The Logic of Collective Action*. Cambridge: Harvard University Press.
- Oorschot, Wim van. 2000. "Who Should Get What, and Why? On Deservingness Criteria and the Conditionality of Solidarity Among the Public." *Policy and Politics* 28(1): 33–48.
- Oorschot, Wim van. 2006. "Making the Difference in Social Europe: Deservingness Perceptions among Citizens of European Welfare States." *Journal of European Social Policy* 16(1): 23–42.
- Petersen, Michael Bang, Rune Slothuus, Rune Stubager, and Lise Togeby. 2011a. "Deservingness versus Values in Public Opinion on Welfare: The Automaticity of the Deservingness Heuristic." *European Journal of Political Research* 50: 25–52.
- Petersen, Michael Bang, Daniel Sznycer, Leda Cosmides, and John Tooby. 2011b. "Who Deserves Help? Evolutionary Psychology, Social Emotions, and Public Opinion about Welfare." *Political Psychology*, forthcoming.
- Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

- Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Price, Michael E. 2008. "The Resurrection of Group Selection as a Theory of Human Cooperation." *Social Justice Research* 21: 228–40.
- Rosenthal, Robert. 1991. *Meta-Analytic Procedures for Social Research*. Newbury Park: Sage.
- Rothstein, Bo. 1998. *Just Institutions Matter*. Cambridge University Press.
- Salthouse, Timothy A. 1994. "The Ageing of Working Memory." *Neuropsych* 8: 535–43.
- Sanfey, Alan G., James K. Rilling, Jessica A. Aronson, Leigh E. Nystrom, Jonathan D. Cohen. 2003. "The Neural Basis of Economic Decision-Making in the Ultimatum Game." *Science* 300: 1755–8.
- Sani, Fabio, Mark Bennett, and Ashley Urie Soutar. 2005. "The Ecological Validity of 'Who Said What?' Technique: An Examination of the Role of Self-Involvement, Cognitive Interference and Acquaintanceship." *Scandinavian Journal of Psychology* 46: 83–90.
- Skitka, Linda J., and Philip E. Tetlock. 1993. "Providing Public Assistance: Cognitive and Motivational Processes Underlying Liberal and Conservative Policy Preferences." *Journal of Personality and Social Psychology* 65: 1205–23.
- Skitka, Linda J., Elizabeth Mullen, Thomas Griffin, Susan Hutchinson, and Brian Chamberlin. 2002. "Dispositions, Ideological Scripts, or Motivated Correction? Understanding Ideological Differences in Attributions for Social Problems." *Journal of Personality and Social Psychology* 83: 470–87.
- Sniderman, Paul M. 2000. "Taking Sides: A Fixed Choice Theory of Political Reasoning." In *Elements of Reason. Cognition, Choice, and the Bounds of Rationality*, eds. Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin. Cambridge, MA: Cambridge University Press, 67–84.

- Sniderman, Paul M., Richard A. Brody, and Philip E. Tetlock. 1991. *Reasoning and Choice. Explorations in Political Psychology*. Cambridge, MA: Cambridge University Press.
- Spears, Russell, Alexander S. Haslam, and Ruurd Jansen. 1999. "The Effect of Cognitive Load on Social Categorization in the Category Confusion Paradigm." *European Journal of Social Psychology* 29: 621-639.
- Stiner, Mary C., Ran Barkai, and Avi Gopher. 2009. "Cooperative Hunting and Meat Sharing 400–200 kya at Qesem Cave, Israel." *Proceedings of the National Academy of Sciences USA* 106: 13207–12.
- Sturgis, Patrick, Sanna Read, Peter Hatemi, Gu Zhu, Tim Trull, Margaret Wright, and Nicholas Martin. 2010. "A Genetic Basis for Social Trust?" *Political Behavior* 32: 205–30.
- Stone, Valerie, Leda Cosmides, John Tooby, Neal Kroll, and Robert Knight. 2002. "Selective Impairment of Reasoning about Social Exchange in a Patient with Bilateral Limbic System Damage." *Proceedings of the National Academy of Sciences* 99(17): 11531–6.
- Sugiyama, Lawrence S., John Tooby, and Leda Cosmides. 2002. "Cross-Cultural Evidence of Cognitive Adaptations for Social Exchange among the Shiwiar of Ecuadorian Amazonia." *Proceedings of the National Academy of Sciences* 99(17): 11537–42.
- Taylor, Shelley E., Susan T. Fiske, Nancy Etcoff, and Ann Ruderman. 1978. "Categorical and Contextual Bases of Person Memory and Stereotyping." *Journal of Personality and Social Psychology* 36: 778–93.
- Tooby, John, and Leda Cosmides. 1992. "The Psychological Foundations of Culture." In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds. Jerome Barkow, Leda Cosmides, and John Tooby. New York: Oxford University Press, 19–136.
- Tooby, John, and Leda Cosmides. 1990. "On the Universality of Human Nature and the Uniqueness of the Individual: The Role of Genetics and Adaptation." *Journal of Personality* 58: 17–67.

Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46: 35–57.

Zak, Paul 2007. "The Neuroeconomics of Trust." In *Renaissance in Behavioral Economics*, ed. Frantz. Oxon: Routledge, 17–33.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

**FIGURE 1.** Illustrations of experimental stimuli used in Studies 1 and 2. Each experiment contains two conditions. The Words in ordinary type are present in both conditions, whereas the words in italics only appear in condition 2.

“Cheater” versus “Reciprocator” dimension present in Treatment Condition only (in italics)

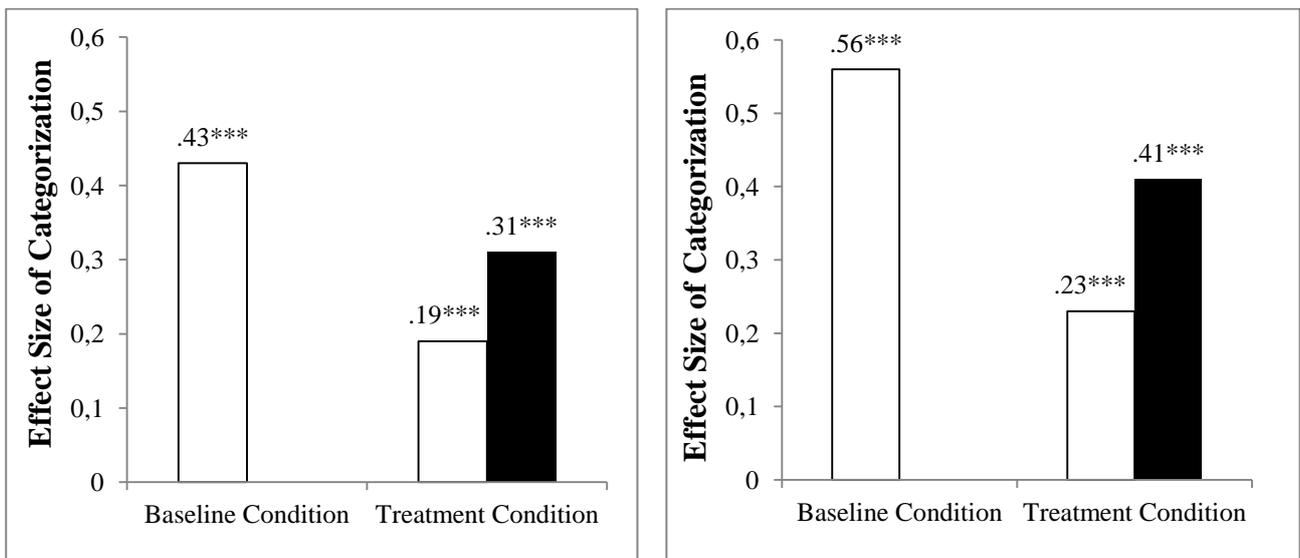
		“Cheater”	“Reciprocator”
Social Welfare vs. Small-Scale Help dimension present in both conditions (in ordinary type)	<b>Social Welfare</b>	 <p>I get support from the government because I am out of work. <i>I could easily find a job, but to tell the truth, I prefer to do without all the hassle when money lands on my account anyway.</i></p>	 <p>I am on social welfare, because I don't have a job. <i>I lost my work when I was injured on the job a year ago. But I am doing all I can to get back in the labor market.</i></p>
	<b>Small-Scale Help</b>	 <p>I have exams coming up and borrowed the notes that one of my classmates took at school. <i>I didn't feel like going to class – I might as well sleep late, when I just as easily can bum off the others and use the notes they took in class.</i></p>	 <p>I am using my friend's bike, because mine is broken. <i>The bike was damaged when a car struck me. I took it to the bicycle repair shop right away, so my friend will have his bike back again soon.</i></p>

Notes. The images are taken from the Center for Vital Longevity Face Database (Minear and Park 2004).

**FIGURE 2.** Consistent with an evolutionary perspective, the presence of cues about the efforts of welfare recipients (i) prompts people across cultures to categorize the recipients on the basis of their efforts and (ii) impedes categorization on the basis of whether recipients receive large-scale social welfare or small-scale help from a friend. Effect sizes reported.

A. Study 1 (US)

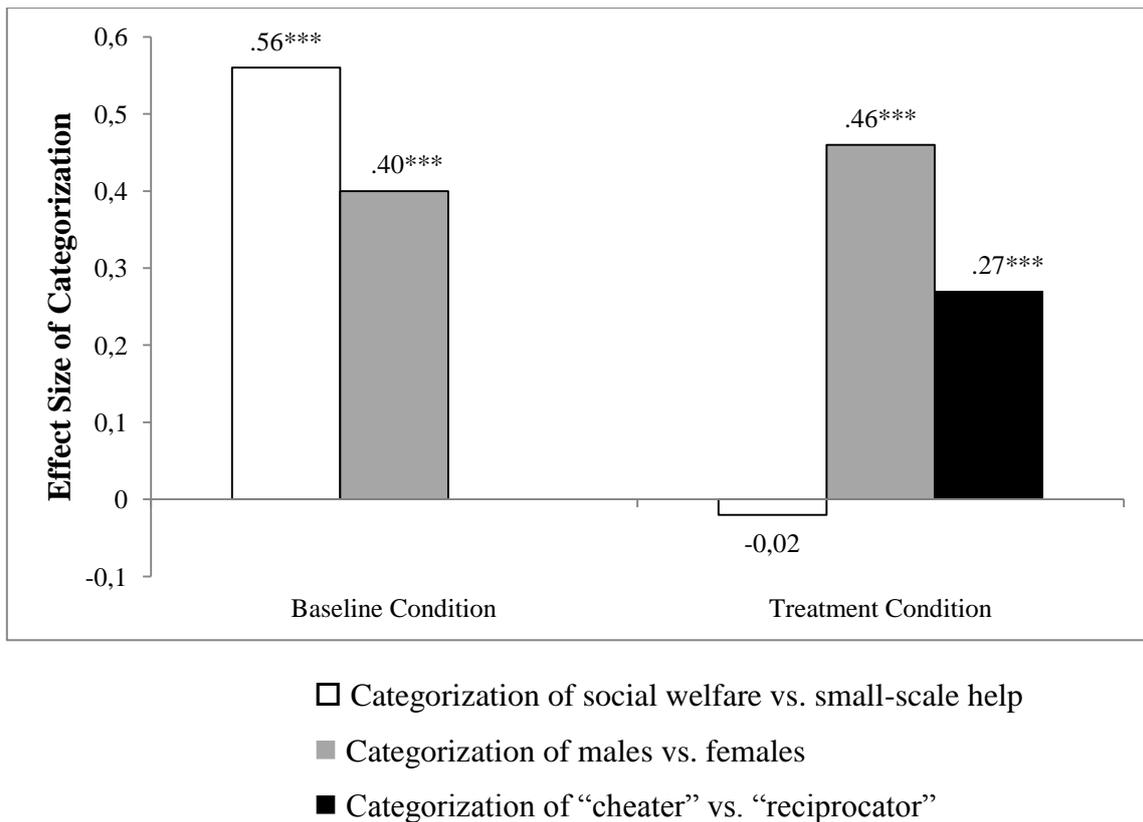
B. Study 2 (Denmark)



□ Categorization of social welfare vs. small-scale help  
 ■ Categorization of “cheaters” vs. “reciprocators”

*Notes.* The reported numbers are effect sizes ( $r$ ) calculated from independent samples  $t$ -tests. Panel A reports categorization among a representative sample of US citizens ( $N = 1014$ ), and Panel B reports categorization among a representative sample of Danish citizens ( $N = 1006$ ). Additional analyses reveal that this categorization pattern is unaffected by the political ideology and political sophistication of the subject.

**FIGURE 3.** Consistent with the notion that the people’s conflation of large-scale welfare and small-scale help-giving is not caused by attentional constraints, the categorization patterns are substantially unaffected by the introduction of an additional evolutionary important cue (recipient sex). Effect sizes reported.



*Notes.* The reported numbers are effect sizes ( $r$ ) calculated from the t-tests of independent samples based on Study 3. The sample is a representative sample of Danish citizens ( $N = 204$ ).

## Supporting online appendix

for

### **Social Welfare as Small-Scale Help: On the Evolved Roots of the Deservingness Heuristic**

#### Contents:

1. Additional Information on Materials
2. Additional Information on Analyzing Recall Data
3. Additional Analyses
4. Additional References
5. Additional Tables

#### **Additional Information on Materials**

##### *“Who Said What?” with Representative Samples*

The “Who Said What?” (WSW) paradigm was originally designed for use in laboratory settings with limited samples. From the perspective of political science, approximately representative samples are important. To this end, the design of the WSW protocol was slightly revised to adapt it to online surveys with thousands of subjects.

A primary difference between the laboratory setting and online surveys with nationally representative samples relates to the level of attention that can be expected on behalf of the subjects (Sniderman and Grob 1996). In the laboratory version used by Kurzban et al. (2001), subjects were presented with eight targets, each appearing three times with three different statements. On the one hand, this consecutive reappearance of the targets enables subjects to rehearse their categorization and, hence, this presumably strengthens the effects. On the other hand, it requires several minutes of uninterrupted and intense concentration on behalf of the respondents as they must read through all of the statements. The version used in this study therefore only presented subjects with eight individuals appearing a single time. During the presentation phase, each individual—together with the associated statement—appeared on the screen for 20 seconds. After the presentation, the subjects completed a short word-based distracter task and were then introduced to the surprise recall task. Pilot studies were run to calibrate the design.

### *Introducing the WSW to the Subjects*

After a general introduction, the subjects were introduced to the task in the following manner:

Eight different people will be introduced to you. These individuals will be presented one at a time. We will ask you to look at pictures of these persons and to read what each one says. Try to gain an impression of each individual. Later, you will be asked to answer questions about them.

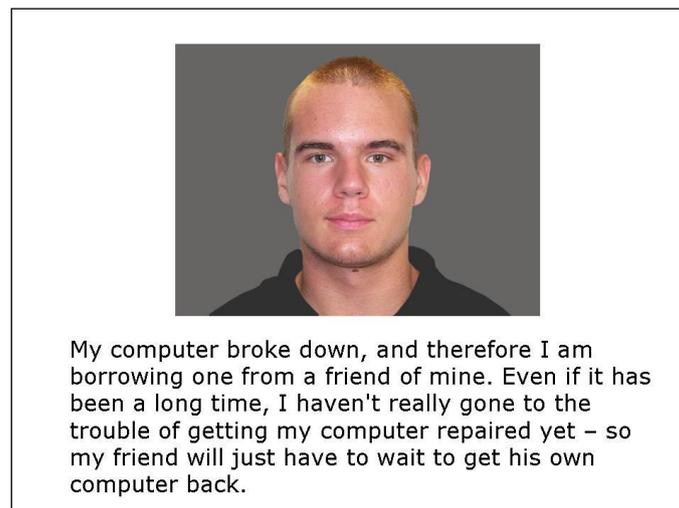
Each portrait will be shown for 20 seconds, and the next portrait will come up automatically. So you do not have to press any keys during the introduction of these people. Just let yourself form an impression of them.

**Click here when you are ready.**

Most importantly, in keeping with standard procedures, the subjects were not informed about the recall phase that followed the presentation phase.

### *The Presentation Phase*

The image below is the image that the survey programmers used as the template for setting up the presentation of the individuals to the subjects.



To picture the individuals in Studies 1 and 2, eight neutralized male pictures (i.e., all looked straight into the camera, all were only shown from the shoulders and up and all wore black polo shirts) were taken from the Center for Vital Longevity Face Database (Minear and Park 2004). No formal

criterion was used in the selection of these pictures—except that they should look “normal.” To picture the individuals in Study 3, four of the male pictures were replaced with four neutralized female pictures from the same database. Again, no formal criterion was used in the selection of these pictures except that they should look normal and be of ordinary physical attractiveness. In all of the studies, it was randomly determined for each individual subject which picture was paired with which statement.

The statements attributed to the individuals were used to experimentally manipulate the kind of information about the individuals that the subject had available. In condition 1, the statements indicated whether the individual received social welfare or some kind of help with an everyday problem. In condition 2, these statements were supplied with a statement indicating whether the individual was making an effort or not to alleviate their own need. The procedure that randomly determined whether a subject was presented with condition 1 or 2 was programmed such that 30% of the subjects were presented with the baseline condition, while 70% were presented with the treatment condition. This assured that enough statistical power would be present in the important second condition with effort cues present such that any unhypothesized effects of individual-level variables would be detected.

The following is a full list of statements in the order of presentation during the experiments. The statements in ordinary type are the statements present in both conditions 1 and 2, which indicate whether the individual is a social welfare recipient or recipient of small-scale help. The statements in italics are the statements added in condition 2, which indicate whether the individual is high or low in effort. The information in parentheses shows which categories each individual belongs to. Of course, this information was not available to the subjects.

1. I have exams coming up and borrowed the notes that one of my classmates took at school. *I didn't feel like going to class – I might as well sleep late, when I just as easily can bum off the others and use the notes they took in class.* (Small-Scale Help/Low Effort)
2. I am spending nights at a friend's place, because my partner kicked me out with nowhere to go. *My friend stepped in when my partner found a new boyfriend. But I am looking everywhere for a new apartment and have placed an ad in the newspaper.* (Small-Scale Help/High Effort)

3. I get support from the government because I am out of work. *I could easily find a job, but to tell the truth, I prefer to do without all the hassle when money lands on my account anyway.* (Social Welfare/Low Effort)
4. I am on social welfare, because I don't have a job. *I lost my work when I was injured on the job a year ago. But I am doing all I can to get back in the labor market.* (Social Welfare/High Effort)
5. I am using my friend's bike, because mine is broken. *The bike was damaged when a car struck me. I took it to the bicycle repair shop right away, so my friend will have his bike back again soon]* (Small-Scale Help/High Effort)
6. I do not have work and get welfare money from the government. *If I felt like it, I could without doubt hold a job, but I am actually pretty happy about doing nothing, while other people slave all day at work.* (Social Welfare/Low Effort)
7. My computer broke down, and therefore I am borrowing one from a friend of mine. *Even if it has been a long time, I haven't really gone to the trouble of getting my computer repaired yet – so my friend will just have to wait to get his own computer back.* (Small-Scale Help/Low Effort)
8. I am without a job and receive help from the government. *My old company went down in the finance crisis. But I have already applied for more training in continued education, and I look forward to starting that.* (Social Welfare/High Effort)

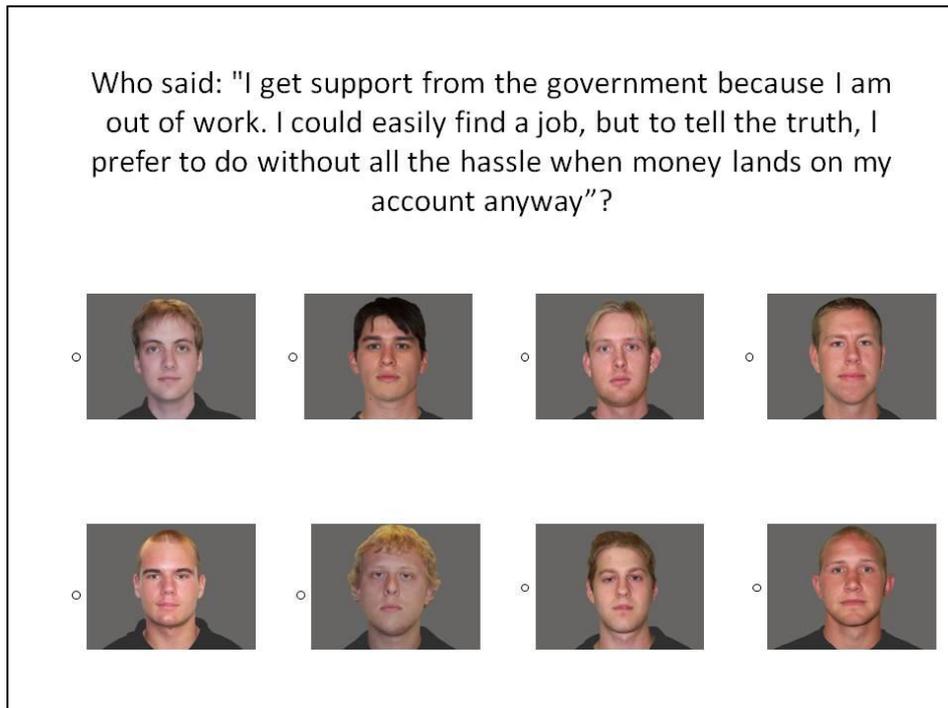
In Study 3, statements 1, 2, 4, and 6 were attributed to males while statements 3, 5, 7, and 8 were attributed to females. In this way, it was ensured that all three dimensions (scale, effort, and sex) were orthogonal to each other. Within the statements attributed to each sex, the specific picture used to represent the individual was randomly assigned to the statement for each subject.

### *The Recall Phase*

After the presentation phase, subjects were presented with a short distracter task, asking them to write down the names of 15 countries. The median time spent on the task was 104 seconds, 88 seconds and 79 seconds in Experiments 1, 2, and 3, respectively. The subjects were then introduced to the recall phase in the following manner:

The following screens will present the statements made by the 8 individuals shown to you previously. Please mark the image of the individual who said the relevant statement, in your opinion.

In random order, each statement from the presentation phase was the presented to the subjects together with all of the pictures from that phase. The following image illustrates the set-up:



### *Additional Measures*

In Studies 1 and 2, measures of ideology and political sophistication were obtained. To measure political sophistication, I relied on factual questions about politics. The level of sophistication was the number of correct answers to these questions. In the US, subjects were asked the following four questions: 1) Do you happen to know how many times an individual can be elected President of the United States under current laws? 2) How many U.S. Senators are there from each state? 3) For how many years is a member of the United States House of Representatives elected—that is, how many years are there in one full term of office for a U.S. House member? 4) What percentage vote of the House and the Senate is needed to override a Presidential veto? In Denmark, subjects were asked the following four questions: 1) Which of the following parties are members of the current Danish government? 2) To which party does Morten Bødskov [a Danish politician] belong? 3)

Some of the political parties are more favorable towards refugees and immigrants than others. Are the Social Liberals among the more or less favorable? 4) To which party does Eva Kjer Hansen [a Danish politician] belong?

To measure ideology, the subjects were asked: “In politics they often talk of the right and left—like a party or a politician on the right wing (conservative), left wing (liberal), or middle of the road (mainstream). Where would you place yourself from left to right on this political scale?” The scale was a 10-point scale with only the endpoints labeled “Left” (0) and “Right” (10), respectively.

### **Additional Information on Analyzing Recall Data**

Analyzing the recall data in the WSW protocol involves comparison of the within-category errors (i.e., when the statement of a low-effort person is attributed to another lazy person) with the between-category errors (i.e., when the statement of a lazy person is attributed to a hardworking person). If the number of within-category errors is higher than the number of between-category errors, subjects have encoded the specific category. In this respect, however, it is important to note that there are four targets in each category (i.e., four welfare recipients versus four receivers of everyday help and four low-effort individuals versus four high-effort individuals). This implies that there are four possibilities for a between-category error during recall (i.e., attributing an individual’s statement to an individual from the other category), but only three possibilities for a within-category error (as one of the four targets in the category is indeed the one who made the statement). As is standard in the literature, I correct for this difference in base rate probabilities of making a within-category error as opposed to a between-category error (for procedure, see Taylor et al. 1978; Kurzban et al. 2001).

To measure categorization of “cheaters” versus “reciprocators” (Studies 1-3) and males versus females (Study 3), we calculate the corrected number of within-category errors minus the number of between-category errors relative to the control condition (i.e., in the absence of effort and sex cues) and test whether this is greater than zero. This means that the effect sizes for the Danish Study 3 are calculated relative to the control condition in Danish Study 2. In this way, the experimental design allows us to correct for any conflation of targets that merely reflects artificial within-category similarities or differences in the baseline descriptions of the recipients (rather than, e.g., their displayed level of effort). This is not a real issue in Denmark, where the stimuli was pretested. In the control condition (i.e., absent effort and sex cues), however, US subjects weakly

confuse the baseline descriptions of the targets that belong to different effort categories in the effort condition. While this tendency is weak, it is nonetheless significant ( $r = .14$ ,  $p = .01$ ). If uncorrected, this apparent within-category difference in the wording of the stimuli would artificially lower the reported categorization on the basis of effort.

To measure the categorization of social welfare versus small-scale help (Studies 1–3), we calculate the corrected number of within-category errors minus the between-category errors and test whether this is greater than zero (in this case, we cannot compare with a control condition as these cues are present in all conditions).

To test the significance of categorization, these measures are subjected to independent sample t-tests against 0. From these t-tests, effect sizes are calculated. Calculations and further descriptions appear in Table A1.

## **Additional Analyses**

### *Some new evidence for the automaticity of effort categorization*

Theoretically, the implicit and automated nature of the cognitive categories of “cheater” and “reciprocator” played an important role in predicting that the attention to the large-scale nature of modern welfare provisioning would be crowded out in the face of effort cues. It was argued that this arises from downstream effects from the activation of these categories that enter the decision-making process early. That these categories do indeed operate implicitly was backed by evidence from political science, social psychology, and neuroscience (see the theoretical discussion). At the same time, it is important to note that the WSW paradigm measures both implicit and explicit categorization. In the current study, this particular strength of the WSW paradigm was important, as it allowed us to investigate how implicit categorization on the basis of effort interacted with explicit categorization on the basis of the scale of the helping situation. Of course, this also implies that finding evidence indicating that subjects categorize individuals along the effort dimension in the WSW paradigm does not in itself imply that this happens automatically. Still, the present study does add further support to the existing evidence that categorization on the basis of effort occurs effortlessly. First, political sophistication has been argued to be related to the degree to which people process political information through explicit memory-based processes (e.g., Druckman and Nelson 2003). Given that we did not find a moderating effect of sophistication on effort-based categorization, this supports the existing view. Second, in Studies 1 and 2, after completing the

experiment, subjects were directly asked how much effort they put into the recall task using the following question: “Generally speaking, how much effort did you put into remembering who said what? Not any effort at all, a little effort, some effort, much effort or very much effort?” If discriminating between “cheaters” and “reciprocators” reflects an explicit, effort-involving process, we should expect the level of self-reported effort to correlate strongly with this type of categorization. There is no such correlation in neither experiment (US:  $r = .04$ ,  $p = .27$ ; Denmark:  $r = .01$ ,  $p = .76$ ). In contrast, when examining a measure that must tap an explicit process, namely the number of correct recalls that subjects have made, we do find an effect of self-reported effort. Hence, in both Denmark and the US, the correlation is highly significant and, in Denmark particularly, of substantial importance (US:  $r = .15$ ,  $p < .001$ ; Denmark:  $r = .34$ ,  $p < .001$ ). In tandem, these findings provide additional support for the notion of the automaticity of the categorization of needy individuals on the basis of effort.

In this regard, it should also be noted that there is no reason to expect that these categories operate fully automatically in the sense that they are encapsulated and unable to draw on more explicit processes. In many cases, the adaptive choice in help-giving situations would have involved recalling past encounters with the individual, planning ahead, etc. In real life situations, the automatic components of our psychology of help-giving is in this way expected to interact with more explicit components (see also Barrett et al. 2006).

#### *Does the presence of sex cues interact with the presence of effort cues?*

In Study 3, where both male and female individuals are presented to the subjects, we see a drop in categorization of social welfare versus small-scale help in the treatment condition compared to this condition in Studies 1 and 2 (compare  $-.02$  with  $.19$  and  $.23$ , respectively). Subjects simply do not pay attention to this dimension when both sex and effort cues are present. We also see a very slight drop in categorization on the basis of effort (compare  $.27$  with  $.31$  and  $.41$ ). How should we interpret these drops? Here, it is important to note that these drops do not in any way suggest that general processing constraints explain the core finding, i.e., that the presence of effort cues crowds out attention to cues about the scale of the situation. Hence, in both the baseline condition and treatment condition of Study 3, we find that subjects can easily categorize along two dimensions at once. Rather, it is possible that in the presence of effort cues, sex cues provide additional input for our decision rules designed for small-scale help-giving. For example, from a mate-selection perspective, males might find it more beneficial to convey benefits to reciprocating females than to

males. In this way, the two cues could interact in facilitating the perception of welfare politics as small-scale help (although, the two-way interaction between the presence versus absence of effort cues, on the one hand, and presence versus absence of gender cues, on the other hand, is not significant:  $F = 2.47$ ,  $p = .12$ ).

#### *Replication of the core results in a pilot study*

A final replication of the basic findings from Studies 1, 2, and 3 is available from the pilot study used to calibrate the paradigm. It was conducted in Denmark by the YouGov agency with 56 subjects (33 females and 23 males; mean age = 36) and used a slightly different order and wording of the statements. Only the condition with effort cues present was used in the pilot study. In the presence of effort cues, subjects discriminate substantially between cheaters and non-cheaters ( $r = .33$ ,  $p = .01$ ) but only weakly between situations of welfare provisioning and situations of everyday help ( $r = .23$ ,  $p = .09$ ). Moreover, the categorization on the basis of neither dimension correlates with political sophistication (“cheaters” vs. “reciprocators:”  $r = -.07$ ,  $p = .58$ ; social welfare vs. small-scale help:  $r = -.15$ ,  $P = .26$ ; the political ideology of the subjects was not measured in the pilot study). Hence, the pilot study fully replicates the results, lending even stronger support to their robustness.

#### **Additional References**

- Sniderman, Paul M., and Douglas B. Grob. 1996. “Innovations in Experimental Design in Attitude Surveys.” *Annual Review of Sociology* 22: 377–99.
- Druckman, James N. and K. Nelson. 2003. “Framing and deliberation: How citizens’ conversations limit elite influence.” *American Journal of Political Science*, 47: 729–45.
- Barrett, H. Clark, David A. Frederick, Martie G. Haselton, and Robert Kurzban. 2006. “Can manipulations of cognitive load be used to test evolutionary hypotheses?” *Journal of Personality and Social Psychology* 91(3): 513–647.

## Additional Tables

**Table A1. Overview of Results and Calculations**

Study	Condition	Dimension	Within-minus between-category errors <sup>a</sup>	T-value of difference to 0 <sup>b</sup>	Degrees of Freedom	Effect size (correlation) <sup>c</sup>	Significance
1 (US)	Baseline	Social Welfare vs. Small-Scale Help	.62	8.28	304	.43	.000
	Treatment	Social Welfare vs. Small-Scale Help	.26	5.04	708	.19	.000
		“Cheater” vs. “Reciprocator”	.46	8.71	708	.31	.000
2 (Denmark)	Baseline	Social Welfare vs. Small-Scale Help	.89	11.67	304	.56	.000
	Treatment	Social Welfare vs. Small-Scale Help	.31	6.20	700	.23	.000
		“Cheater” vs. “Reciprocator”	.61	11.91	700	.41	.000
3 (Denmark)	Baseline	Social Welfare vs. Small-Scale Help	.90	5.27	61	.56	.000
		Male vs. Female	.88	3.39	61	.40	.001
	Treatment	Social Welfare vs. Small-Scale Help	-.02	-.28	141	-.02	.781
		“Cheater” vs. “Reciprocator”	.26	3.32	141	.27	.001
		Male vs. Female	.78	6.10	141	.48	.000

<sup>a</sup> This figure is calculated by subtracting the number of between-category errors from the number of within-category errors. Given that eight individuals are presented, the maximal number of possible errors in either direction is 8. Prior to subtraction, the errors have been corrected for differences in baseline probabilities of making a within- and between-category error (see text).

<sup>b</sup> A t-test is performed that tests whether the difference between the number of within-category and the number of between-category errors is significant. This corresponds to a paired t-test of the difference between the number of within-category errors and the number of between-category errors.

<sup>c</sup> Following Kurzban et al. (2001), a correlational effect size is calculated from the formulae offered by Rosenthal (1991):  $r = \sqrt{t^2/(t^2+df)}$ , where t is the t-value and df are the degrees of freedom.

**Table A2. Categorization of social welfare vs. small-scale help regressed on experimental treatment, respondent's ideology, political knowledge, country and interaction effects. Standardized coefficients.**

	<b>Study 1 (US)</b>	<b>Study 2 (Denmark)</b>	<b>Study 1 + 2</b>
Treatment	-.14	-.23***	-.122***
Ideology	-.06	-.05	-
Political knowledge	.05	.02	-
Treatment × Ideology	.11	.05	-
Treatment × Political Knowledge	-.08	-.01	-
Country	-	-	.088*
Country × Ideology	-	-	-.078
R2	.02	.04	.03

Notes. N (study 1) = 870, N (study 2) = 928. A high value on the ideology variable corresponds to left-leaning political conviction.

\*p < .05, \*\*p < .01, \*\*\*p < .001