Misrepresenting the Truth about Minorities*

Charles Breton†
Gregory Eady°
Yannick Dufresne•
Jennifer Hove•
Clifton van der Linden•

August 26, 2014

Abstract

In this article we ask when and why the public would 'lie' about their attitudes toward some ethno-cultural groups more than others. We argue that the underlying anti-prejudice norm that leads to social desirability bias does not apply equally to all groups, but varies in theoretically predictable ways that are open to empirical investigation. To test our hypotheses, we field a large-scale experiment using the unmatched count technique (n = 25,790) and vary the target group along two characteristics. First, we compare differences in support for a more 'abstract' statement regarding ethno-cultural minorities to those specifying a minority group by name. Second, we vary the ethno-cultural group of interest in terms of its historical relationship with the country. Using Canada as a a case, we find that an abstract notion of cultural groups receives more explicit support than Aboriginals and French-Canadians but not more than Asian-Canadians. We then show that this result, rather than representing genuine support for Asian-Canadians, is due to the fact that respondents who have negative views of Asian-Canadians are more likely to lie about it.

Keywords: multiculturalism, unmatched count technique, sensitive questions, public opinion, social desirability bias, ethnicity, race, Canada

*Corresponding authors: Charles Breton, cbreton@alumni.ubc.ca, and Gregory Eady, gregory.eady@mail.utoronto.ca
†University of British Columbia
°University of Toronto

It is well-known in survey research that respondents who are asked directly about their attitudes toward sensitive issues are more likely to provide answers that conform with social expectations (Edwards, 1953, 1957; Warner, 1965; Sigall and Page, 1971; Bradburn et al., 1978; Himmelfarb and Lickteig, 1982; Fisher, 1993; Berinsky, 1999; Johnson and Van de Vijver, 2002). This form of measurement bias — social desirability bias — can be quite large. Heerwig and McCabe (2009) found, for example, that only 70% of Americans surveyed in 2007 indicated support for having a black Presidential candidate when they used an indirect measurement technique. This compares with 92% who were asked the question directly during the 1990s.[1] The possibility of social desirability bias may lead to skepticism about the prevalence of socially undesirable attitudes. If this bias is ignored, it could result in the uncritical belief that direct questions elicit truthful answers. Biased estimates of the proportion of the population who hold socially undesirable attitudes may also have large effects on policy-making.

Researchers have designed and employed a variety of indirect measurement survey techniques to minimize this bias in an effort to estimate the proportion of those who genuinely hold an attitude or have truly engaged in a behavior. Indirect measurement techniques have been used to examine a wide range of attitudes and behaviors. For example, they have been employed to investigate support for affirmative action (Gilens, Sniderman and Kuklinski, 1998), attitudes toward a female or black presidential candidate (Streb et al., 2008; Redlawsk, Tolbert and Franko, 2010), voting (Karp and Brockington, 2005; Holbrook and Krosnick, 2010), vote buying (Gonzalez-Ocantos et al., 2012), and support for harsh immigration restrictions (Janus, 2010), among others.

One relatively unappreciated benefit of these techniques is that they permit one to estimate the degree of social desirability bias: the proportion of those who misrepresent an attitude or behavior when asked about it directly. Typically, the primary quantity of interest in studies that use indirect measurement strategies is the proportion of those who hold a sensitive attitude or engage in a socially unaccepted behavior. For example, one might ask what proportion of those in the United States would be upset if an African-American family moved next door? (Gilens, Sniderman and Kuklinski, 1998); How many of those in underdeveloped countries have been offered a gift in return for their vote? (Gonzalez-Ocantos et al., 2012); Or what proportion of the public is unwilling to vote for a female presidential candidate (Streb et al., 2008; Redlawsk, Tolbert and Franko, 2010)? In this article, we flip the current approach to social desirability bias on its head. Rather than treat social desirability bias as measurement error in need of correction, we treat it as the quantity of interest itself: we ask whether and why people might be more likely to 'lie' about their attitudes toward some groups more than others.

To the authors' knowledge, no studies have advanced hypotheses and used indirect

[1]Support for a black President could have declined significantly since the 1990s, although the trend in General Social Survey data until the question was dropped in 1992 was decidedly positive.

measurement techniques to test whether the public is likely to 'lie' about their attitude toward some groups or individuals more than others. A variety of related research questions are open to investigation. For example, in the United States, does the public feel more pressure to misrepresent prejudiced attitudes toward African-Americans and Hispanics than toward those of Middle Eastern origin? Are partisans more likely to misrepresent the negative characteristics of opposition candidates than the positive characteristics of their preferred political candidate? In highly multi-ethnic countries, are people more likely to 'lie' about their attitudes toward newly prominent ethnic groups than historically established ones? This is a potentially fruitful, yet unexplored, avenue of research, and one in the present article that we investigate systematically.

Here, we seek to answer the last of the above questions experimentally. Specifically, we ask whether the public in one of the most multicultural countries in the world, Canada, is more likely to 'lie' about their support for ethno-cultural groups that have more recent historical ties to the country than those with deeper historical ties. We argue that those living in highly multicultural countries, particularly those where the inclusiveness of new minorities is a strongly and publicly stated value, will experience more social pressure to express supportive attitudes toward newer ethno-cultural groups than well-established ones. Evidence in support of this argument would be particularly consequential for countries with increasing levels of ethnic diversity, which now make up much of the developed world. It would suggest that fostering genuine inclusiveness toward new ethno-cultural groups represents a substantial challenge: first, because support for new ethno-cultural groups is more expressed than real relative to traditional minority groups, and second, because genuine support for these groups is substantially hidden from everyday view.

To test our hypotheses, we field a survey in a highly multi-ethnic part of the country that includes a series of four list experiments to probe attitudes toward multicultural acceptance. We pair these experiments with direct questions. Doing so permits us to estimate the degree of social desirability bias for each sensitive item. A common practical limit of the study of social desirability bias is that reasonably precise measures from indirect questioning techniques require much larger sample sizes than traditional direct survey questions. This is due, in part, to the fact that social desirability bias is a quantity calculated from multiple survey responses, each measured with uncertainty. We address this problem directly by including the list experiments in a unique large-scale survey (n = 25,790) to provide sufficient power to test our hypotheses.

Our findings suggest that the public is more sensitive to expressing negative attitudes toward ethno-cultural groups whose presence in the country is historically more recent. In particular, when respondents are asked directly, the expressed level of support for Asian Canadians — a rapidly growing and prominent ethno-cultural group in Canada — is not substantively different from support for the more general principle of multiculturalism. Results from the list experiments show, however, that the driver of this similarity in support is due to differences in the degree of social

desirability bias: those surveyed are more likely to misrepresent their attitudes toward Asians, a group stemming from immigration, than to 'lie' about their support for the multiculturalist principle in general or two other ethno-cultural groups with deep historical ties to the country in particular.

Does the Anti-prejudice Norm Vary Across Minority Groups?

Recent studies show that overt prejudice and racism toward ethno-cultural groups in the developed world has decreased over time (e.g. Mendelberg, 2001; Sears and Henry, 2005; Huddy and Feldman, 2009). The strengthening of the anti-prejudice norm that accompanies this development suggests that attitudes incompatible with this norm will be more difficult to measure: people will be more likely to misrepresent socially unacceptable attitudes in an effort to appear consistent with societal expectations. From a public-policy policy perspective, this is unfortunate because multiculturalist policies are partially determined by data on attitudes toward specific ethno-cultural groups and minority groups in general. Unpacking the differences in these attitudes and determining differences in misrepresentation across groups is essential to understanding whether, how, and why populations react positively or negatively to increasing ethno-cultural diversity.

Efforts in political science to accurately measure these attitudes through the use of indirect measurement strategies, particularly toward ethno-cultural and racial groups, have grown steadily since the seminal research of Kuklinski, Cobb and Gilens (1997) and Kuklinski et al. (1997). Much of this work examines attitudes toward a single group of interest. Research on prejudice and social desirability bias toward ethno-cultural and racial groups is commonly conducted in an American setting where the focus is frequently on obtaining valid measures of prejudices toward blacks (e.g. Gilens, Sniderman and Kuklinski, 1998; Berinsky, 1999; Huddy and Feldman, 2009; Weber et al., 2013). Research into beliefs concerning other groups, such as immigrants and ethnic minorities, is similarly confined to one group (Janus, 2010). A drawback of this focus on single groups is that differences in social desirability bias *across* groups remains understudied.[2] The absence of cross-group comparisons may be due to the fact that researchers who are interested in ethnic and racial bias have research agendas specific to one minority group. Furthermore, practical concerns related to statistical power may also lead researchers to confine themselves to a single group: indirect measurement strategies often require larger sample sizes than typical direct measures to achieve similarly precise estimates. These limitations mean that researchers are in the relative empirical dark when it comes to whether and how social desirability bias

[2]To the authors' knowledge, studies in political science have neither sought to develop this avenue of research theoretically nor to test differences in social desirability bias empirically.

varies across minority groups.

Heterogeneity in the Strength of the Anti-Prejudice Norm

Social norms affect public expressions when individuals perceive them as applicable to their context and recognize the degree to which expressed prejudiced attitudes are deemed socially unacceptable. Recent studies show that individuals indeed perceive large differences in the applicability of social norms against expressing prejudice. Context and group characteristics influence when and where these norms are deemed applicable and the degree to which they are followed.

In a well-known study in social psychology, Crandall, Eshleman and O'Brien (2002) ask their research subjects to rate the social appropriateness of expressions of prejudice toward 105 social groups. The resulting ranked list is as one might expect: the anti-prejudice norm is perceived as weakest for prejudice against sexual predators, terrorists, and racists; it is strongest for expressions of prejudice toward the disabled, social service professionals, and various ethnic and racial groups (Crandall, Eshleman and O'Brien, 2002, 362). The authors also measure the degree of prejudice that people openly admit, asking a separate set of respondents directly about their attitudes toward the same 105 social groups. These measures of admitted prejudice correlate nearly perfectly with the measures of perceived social appropriateness ($r = 0.96$). The findings suggest, first, that there is large variation in the perceived anti-prejudice norm and, second, that individuals appear to respond to this norm by admitting to their own prejudices in ways consistent. What the results do not show, however, is the difference between the perceived anti-prejudice norm for each group and the level of genuine attitudes.

This more general finding is complemented by recent research in political science that demonstrates that context can affect the strength of the anti-prejudice norm (Blinder, Ford and Ivarsflaten, 2013; Weber et al., 2013). Weber et al. (2013) show in the United States,[3] for example, that whites who are particularly sensitive to social norms (high 'self-monitors') and live in neighborhoods with a large proportion of blacks are less likely than low self-monitors to indicate that they hold negative stereotypes of African-Americans. In neighborhoods with a small proportion of blacks, however, low and high self-monitors do not substantially differ in their indicated prejudices. Why is this? The authors argue that areas with a large proportion of blacks are more likely to have strong norms of egalitarianism. Those who are highly sensitive to social norms, consequently, are more likely to misrepresent their racial attitudes in neighborhoods where many residents are black. The authors' implicit empirical claim is therefore that if a measure of genuine racial attitudes among whites in such areas was available,

[3]The authors' theoretical claims concern low- and high-proportion black neighborhoods in the United States more generally, but their data are specific to New York state.

there would be little to no difference in observed prejudiced attitudes between low and high self-monitors. Context, here, is critical: the strength of the anti-prejudice norm and therefore the degree of attitude misrepresentation is conditional on the racial geographical landscape.

Similar work suggests that the applicability and strength of the anti-prejudice norm varies with the information provided about minority groups. For example, Blinder, Ford and Ivarsflaten (2013) argue that specifying group characteristics can trigger an awareness of the need to apply the anti-prejudice norm. Their claim is that a strong social norm dictates treating all citizens equally, but is much weaker for non-citizens. Specifying that a stigmatized group has citizenship status should therefore increase the group's support among fellow nationals. Using a framing experiment, they show that the strength of support for asylum seekers — a heavily stigmatized group — is conditional on specification of their citizenship status. The results show that when asylum seekers are specified as citizens, people are more likely to agree that these newcomers should have equal employment opportunities. Priming citizenship status, in other words, makes the anti-prejudice norm stronger.

The general conclusion from these studies is twofold. First, the results suggest that the strength of the anti-prejudice norm varies both across contexts and according to group characteristics. Second, they suggest that people are more likely to express prejudices toward groups that are weakly or not at all encompassed by the anti-prejudice norm. Results relevant to the anti-prejudice norm and differences in the levels of attitude misrepresentation are typically based on indirect measures of this norm however. The authors cited above rely on evidence across contexts and within groups that imply differences in the anti-prejudice norm, but do not measure these differences with direct measures of the social norm. It is not clear, for example, that support for asylum seekers with citizenship is due to the social norm being stronger when asylum seekers are specified as having citizenship. An alternative explanation is that support for asylum seekers with citizenship is greater than that for asylum seekers whose citizenship status is left unstated because citizenship status is a signal of asylum seeker 'quality': having passed through the institutional vetting process necessary to attain citizenship, asylum seekers with citizenship may be perceived as less likely to have a fraudulent asylum claim and more likely to provide greater economic and social benefits to the country. Lower support for asylum seekers without citizenship could, therefore, derive from principled opposition rather than differences in the strength of the anti-prejudice norm, or a combination of the two.

By measuring social desirability bias directly, we can get at the anti-prejudice norm and its strength: the stronger the norm, the more likely it is that people with the socially undesirable attitude will lie about it. For example, we can reasonably hypothesize that given the prevalence of the anti-prejudice norm in developed countries, those who have prejudiced opinions toward a minority group are more likely to lie about this attitude than those who hold prejudiced views of racist, extremist, or otherwise fringe groups. Social desirability bias is thus likely to be stronger when

people express attitudes toward ethno-cultural minorities than toward racist groups. The difference in the strength of the norm between these two groups is likely to be large. Past research is unclear as to what we should expect when the groups of interest are different ethno-cultural groups. Although research in social psychology seems to indicate that there is little reason to think that the underlying anti-prejudice norm will apply uniformly across all groups, it remains to be tested whether variation in the strength of the norm will generate heterogeneity in social desirability bias.

To investigate how the anti-prejudice norm differs across ethno-cultural groups, we take advantage of the apparent strength of this norm and of multiculturalism in Canada, one of the world's most multicultural countries, and the case to which we now turn.

The Anti-Prejudice Norm in a Highly Multicultural Country

Cross-country comparative studies generally find that residents of countries with multiculturalism policies tend to hold more inclusive attitudes toward ethnic minorities (e.g. Crepaz, 2006; Weldon, 2006; Wright, 2011). One question rarely considered however, is whether countries with multiculturalism policies are truly more inclusive or whether multiculturalism policies primarily increase the social pressure not to express prejudice without substantially influencing underlying attitudes. This latter possibility implies a potential dilemma for those living in highly multicultural countries who also hold relatively negative attitudes toward immigrants and ('newer') ethno-cultural groups: when asked about such attitudes, does one express the publicly defined and socially acceptable response that is consistent with multicultural acceptance and the national identity, or express one's personal belief openly? In the same vein, it is unclear whether this norm emphasizing the positive role of ethno-cultural groups is of equal strength for all groups.

The relationship between multiculturalism and reduced prejudice (e.g. Guimond et al., 2013; Richeson and Nussbaum, 2004) and that between multiculturalism and inclusive attitudes toward ethnic minorities (e.g. Crepaz, 2006; Weldon, 2006; Wright, 2011) is well-established in the literature. Guimond et al. (2013) show in a cross-country study that multiculturalism policy increases the strength of the perceived social norm against the expression of prejudice, which in turn influences individual attitudes. They find that, among their four countries of study (U.S., U.K., Germany, and Canada), Canadians perceive the anti-prejudice norm the strongest, a result they attribute to the importance attached to multiculturalism in the country.

From the level of formal policy [4] and from the political elite, therefore, the message

[4]Compared to other ethno-culturally diverse countries, Canada maintains an extensive set of multiculturalism policies. According to the Multiculturalism Policy Index, among Western countries, Canada is ranked first with respect to the scope of its multicultural policies directed at indigenous

is that the Canadian nation is constituted by manifold ethno-cultural groups and that these groups represent an essential and positive component of the country and the national identity. In Canada, the anti-prejudice norm is heavily strengthened by a recognition, belief, and respect for the idea that a diversity of ethno-cultural groups positively contributes to and defines society.

On the surface, this message appears to have found support among the larger public. Public opinion data in Canada frequently show that a large proportion of the population holds favorable attitudes toward immigration and ethnic diversity (e.g. Berry and Kalin, 1995; Adams, 2007), and that these favorable attitudes have been increasing over time (Wilkes and Corrigall-Brown, 2011). Recent research suggests, furthermore, that Canadians base their judgments about individual immigrants primarily on economic rather than ethnic criteria (Harell et al., 2012). Support for multiculturalism as a general principle, is consistently high. Not only are Canadians supportive of multiculturalism policies, but a majority identify multiculturalism as a critical element of the Canadian national identity (Jedwab, 2005; Bloemraad, 2006, 140).

Moreover, Canadian identity and national pride are often said to be intertwined with Canada's success in dealing with high levels of immigration and ethnic diversity (Joppke, 2004; Kymlicka, 2003; Bloemraad, 2006). Evidence at the individual level supports this link. Citrin, Johnston and Wright (2012) find, for example, that national pride is positively associated with support for multiculturalism and immigration in Canada. By contrast, in the United States national pride is a predictor of more anti-multicultural and anti-immigration attitudes.[5] There is general consensus in the literature, therefore, that in Canada support for and the relative success of Canadian multiculturalism is a strong foundation for the country.

Canada is an interesting case to study social desirability with respect with different ethno-cultural groups because of the presence of important national minorities. In theory, the anti-prejudice norm applies equally to all these groups and presents them as being worthy of respect. In discourse however, the focus of multiculturalism has been on recent waves of immigration and 'new ethno-cultural groups' as oppose to 'traditional' groups namely French-Canadians and Aboriginals.[6] A good illustration of this insistence of multiculturalism as being about these new ethno-cultural groups

peoples; first for policies directed at national minorities; and second for policies directed at immigrant minorities (Banting and Kymlicka, 2006). Canada also counts itself among only five countries to have formally defined multiculturalism's place in politics and society (Banting and Kymlicka, 2013)

[5]An interesting point however is that although pride in Canada was associated with more inclusive attitudes, these attitudes in themselves were not more inclusive in Canada than those in the U.S. overall. Citrin, Johnston and Wright (2012) conclude that "Canadian identity may foster an inclusive orientation but in doing so it starts from a less inclusive base".

[6]Politicians in the province of Quebec were quick to identify this as a problem. Their original opposition to a national policy of multiculturalism was not about the consequences of the policy for the integration of immigrants. Elites in Quebec were opposed to the policy because they perceived it as a strategy to diminish the role of the province by rejecting a binational vision of the country. For these elites, instead of the *Québécois* being one the two founding peoples, the policy made them just an ethnocultural group like any other (see Bouchard and Taylor, 2008, 122)

is the fact that the responsibility for the Multiculturalism program was shifted in 2008 from Canadian Heritage, the ministry in charge of culture, to Citizenship and Immigration Canada (Citizenship and Immigration Canada, 2012). This is not to say that the norm promotes exclusionary or prejudiced attitudes toward Aboriginals and French-Canadians. Rather, this focus on new groups sends the message that the "Canadian thing to do" is to welcome immigrants and their contribution to the country, it is less clear on what is expected vis-à-vis traditional groups.

Hypotheses

In the following, we pursue different objectives: (1) we are interested in finding unbiased measures of attitudes toward different ethno-cultural groups, (2) look for the presence of social desirability bias in a setting that is supposedly highly inclusive, and (3) investigate whether there is heterogeneity in this bias across target groups as a function of the strength of the anti-prejudice norm.

By all appearances the multicultural vision of Canada seems to be well-accepted by Canadians. Consequently, a question about this particular vision should elicit overwhelming support. This is especially true if prejudice toward a given group does not come into play. In other words, asking Canadians if they support a vision of Canada where every cultural group's contribution deserves respect, without referring to a particular group, is asking a question about the norm itself and should be supported by a large majority. This support should also be genuine. Again, when survey respondents are asked sensitive questions they are likely to lie to conform with the social norm. In the case of a question on an abstract group (e.g. "every cultural group"), the social norm is strong but it is also stripped of the prejudice that might taint attitudes toward a particular group. However, questions mentionning a precise group should not produce the same level of support. Here, the social norm will also be strong but personal prejudice might come into play. In other words, support might by high—though lower than for an abstract group—but it will come with a greater likelihood of lying. Consequently:

> **Hypothesis 1:** *An abstract notion of ethno-cultural groups will have higher explicit support than the particular groups that constitute the whole.*
>
> **Hypothesis 2:** *'Newer' ethno-cultural groups stemming from immigration will have higher explicit support than 'traditional' groups.*

Based on extant literature we do not have clear expectations about the difference in explicit support between the abstract notion and 'newer' ethno-cultural groups, but we do expect support for the abstract group to rest on a genuine attitude. However, research on the role and impact of multiculturalism in Canada leads us to expect that the anti-prejudice norm will be stronger for ethno-cultural groups originating from recent immigration. A stronger anti-prejudice norm means that:

> **Hypothesis 3:** *Respondents are more likely to misrepresent ('lie' about) their beliefs about 'new' ethno-cultural groups compared to 'traditional' ones.*

We test these hypotheses by using a combination of list experiments and direct questions. We describe this design and the data below.

Data and Methods

To minimize social desirability bias, researchers have increasingly relied on indirect survey techniques to elicit truthful answers from survey respondents. In line with these studies, we use an indirect measurement technique called a list experiment, which enables researchers to obtain more valid estimates of attitudes toward sensitive questions and behaviors. Specifically, we use the list experiment to estimate the proportion of respondents who agree or disagree that ethno-cultural groups in general, and three well-represented groups in particular, contribute to society. We combine the list experiments with equivalent direct questions. The comparison of indirect to direct questions allows us to estimate the proportion of respondents who misrepresent their attitudes when asked a sensitive question directly. This, in turn, allows us to determine whether the level of misrepresentation differs by the ethno-cultural group specified in each condition. In the following section, we explain the basic reasoning behind and set up of the list experiments. This is followed by a description of the survey questions and data we use to test our hypotheses

The List Experiment

The logic underlying list experiments is both simple and powerful. Respondents are randomly divided into groups and assigned a baseline or treatment list of items. Respondents in the baseline group receive a list of J items, typically 3 or 4 in total. Those in a treatment group receive the same J items, plus an additional item. This $J+1^{th}$ item is the item of interest: the sensitive attitude or behavior. Respondents are asked how many items in the list they agree or disagree with, how many make them upset, or a similar question consistent with a 'how many' formulation.

In the present study, respondents in the baseline group were presented with the following question and list:

> How many of the following do you *agree* with?
>
> - There should be more funding for the arts
> - The government should raise taxes on gasoline
> - Corporations are taxed too much
> - Unions have too much power

Those in a treatment group were presented with the same list, in addition to a fifth randomly selected item:

- `[Name of ethno-cultural group]'s contribution to Canada deserves respect`

We explain the theoretical reasoning behind the additional statement further below.

List experiments work for researchers because typical quantities of interest are (conditional) averages rather than individual responses. This is important because although we cannot know the answer to the sensitive question for any given individual we can nonetheless determine the proportion of those answering affirmatively to the sensitive item. This proportion is estimated by taking the difference of the mean response in the baseline group and mean response in a treatment group. Because the average answer to the J items for the baseline and treatment groups will be the same (in expectation) in both groups, the difference in means between baseline and treatment groups represents the proportion of respondents who answered affirmatively to the $J{+}1^{th}$ item.

The power of the list experiment resides in the fact that respondents are not asked about the items individually, but about *how many* of them. Unless a respondent answers affirmatively to all or none of the items, it is impossible to know how a respondent answered to the statements individually.[7] Knowing that the interviewer or analyst is blind to their answers per item provides respondents with the opportunity to express their attitudes safe from pressure to provide the socially desirable response. A complementary possibility is that attitudes queried in list experiments may differ from those in direct questions due to difference in the priming of social desirability: reactions to a sensitive statement asked about alone may elicit a stronger normative reflection or emotional response than that toward the same statement listed among many. In sum, list experiments help minimize the measurement error attributable to social desirability bias by relying on respondents' understanding that their answers cannot be deconstructed into individual responses and by minimizing the social desirability prime that arises from direct questions more generally.

An important benefit of list experiments is that they can be used in concert with direct questions to estimate the degree of social desirability bias attached to a sensitive statement. This is accomplished by asking respondents in the baseline condition the sensitive directly.[8] In the present study, respondents in the baseline group were randomly assigned a sensitive item after being asked the list question, and asked whether they agreed with the statement:

`Do you agree or disagree with the following statement?`

[7]Statements in the list are designed to avoid, as much as possible, cases in which respondents agree with all or none of the statements.

[8]Direct questions are only asked to those in the baseline group to avoid consistency bias.

```
[Name of ethno-cultural group]'s contribution to Canada
deserves respect
```

List experiments require some assumptions. We assume first that answers to the baseline items are not affected by the addition of the sensitive item. Recent methodological advances permit this assumption to be tested (Imai, 2011; Blair and Imai, 2012), the result of which we provide in the results section. Second, we assume that answers to the sensitive item in the list experiment are truthful. Although this assumption is not testable, studies commonly find differences between responses to the direct questions and list experiment questions (Gilens, Sniderman and Kuklinski, 1998; Janus, 2010) that, at minimum, suggest that list experiments improve validity, even if social desirability bias is not entirely eliminated.

Previous studies using list experiments relied on difference-in-means tests to compare answers to the list experiment from various groups of interest, such as those by gender, age, and region. These estimates are inefficient however, and do not permit one to control for potential confounders. To remedy this, Imai (2011) and Blair and Imai (2013) propose a set of methods that allow list experiments to be analyzed within a regression framework. These methods also permit researchers to more easily test hypotheses regarding differences in social desirability bias across questions and groups. Where applicable, we use these methods as implemented in Blair and Imai's (2013) `list` package in R.

Measuring Social Desirability Bias

Social desirability bias is typically measured as the proportion of the population that misrepresents that they hold a socially unacceptable attitude or engage in a socially unacceptable behavior (e.g. Streb et al., 2008; Heerwig and McCabe, 2009; Janus, 2010; Redlawsk, Tolbert and Franko, 2010).[9] In this section, we argue that this measure is inconsistent with the definition of social desirability bias however, and does not permit meaningful comparisons of social desirability bias across groups. We argue, furthermore, that for a measure of social desirability bias to be meaningful theoretically and comparatively, it must measure the strength of the social norm itself. We provide a simple reformulation of the measure to ensure consistency with an intuitive definition of social desirability bias.

In social psychology, where the concept of social desirability was first advanced (Edwards, 1957), social desirability bias is frequently defined as the "the tendency to answer in a manner that is socially desirable instead of expressing one's true feeling" (Crowne and Marlowe, 1960).[10] In political science, definitions are similar. Berinsky

[9] Gilens, Sniderman and Kuklinski (1998) is an important exception. They present both the total proportion who misrepresent their attitude in addition to the measure that we define in this section and use as our measure of social desirability bias in our empirical section.

[10] Other definitions in social psychology are similar: the "[tendency] to distort self-reports in a

(1999, 1211) defines social desirability bias as the "gap between private opinion and public utterance [that arises from] a desire to cloak attitudes that society as a whole might deem unacceptable for fear of social sanctions."

To operationalize these definitions, researchers typically measure the degree of social desirability bias by the difference in proportions between responses to direct and indirect questions (Blair and Imai, 2012, 49).[11] This difference can be thought of equivalently as the total proportion of the population that misrepresents their response. However, this proportion is not a measure of the strength of the social norm. This is because the norm, theoretically and by assumption, only has an effect on responses from those who hold the socially unacceptable attitude[12]. The strength of the social norm can only be observed among individuals who hold an attitude that they have reason to misrepresent. For the measure to be meaningful it should therefore represent the pressure that one feels to 'lie' if one holds a socially unacceptable attitude—the strength of the social norm that drives misrepresentation. To measure this norm requires focusing empirically on the behavior of those affected it.

We explain this first by examining two hypothetical cases. Imagine first a case in which 10% of the public hold a socially undesirable attitude, all of whom lie about it. Second, imagine a case in which 20% of the public hold the socially undesirable attitude, only half of whom lie about it. By the typical measure — the total proportion of 'liars' in the population — the degree of social desirability in both cases is 10%. It is clear, however, that the degree of social desirability bias is not equal in both cases: the pressure to respond in accordance with social expectations is stronger in the former case, where lying is absolute. Figure 1 illustrates these two hypothetical cases graphically.

We therefore measure social desirability as the probability that one who holds a socially undesirable attitude will misrepresent it. In other words, it is the probability that someone who has reason to lie about an attitude or behavior actually does so. In our hypothetical example, therefore, social desirability bias is 1 in the first case because the proportion of those 'lying' among those who hold the undesirable belief is 100%, but 0.5 in the second case, where only half of those who hold the undesirable belief 'lie' about it. More formally, we measure social desirability bias, θ, as,

$$\theta = \frac{|\pi_{direct} - \pi_{list}|}{\pi_{undesirable}} \tag{1}$$

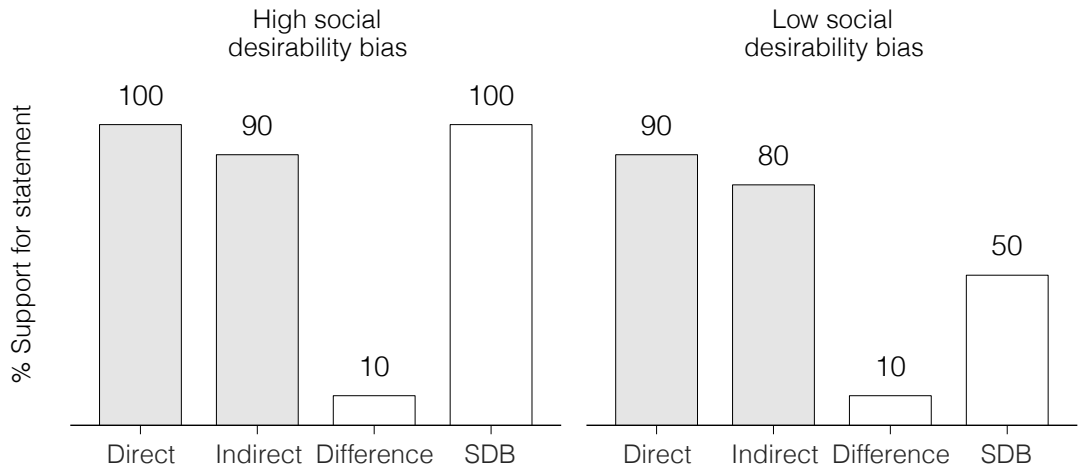
where π_{direct} is the proportion who answer affirmatively to the statement asked directly,

favourable direction" (Furnham, 1986, 350); the "tendency of individuals to present themselves favorably with respect to current social norms and standards" (Zerbe and Paulhus, 1987).

[11] As Blair and Imai (2012, 49) write, the "difference between responses to direct and indirect questioning has been interpreted as a measure of social desirability bias in the list experiment literature."

[12] Here, we leave aside the possibility that the social norm has *had* an effect on people's attitude over time so that they have now changed their true position. This possible effect of the norm is outside the scope of this paper

FIGURE 1: MEASURING SOCIAL DESIRABILITY BIAS: TWO HYPOTHETICAL CASES



This bar graph shows the proportion of support for two hypothetical statements. The bar labeled 'Direct' refers to the level of support to a direct statement, while 'Indirect' refers to the level of support with the same statement when asked indirectly. 'Difference' is the quantity often referred to as social desirability bias. The 'SDB' bar is the actual level of social desirability bias, the proportion of people with the socially undesirable attitude who misrepresent it.

π_{list} is the proportion who answer affirmatively to the statement in the list experiment, and $\pi_{undesirable}$ is the proportion who hold the socially undesirable belief. This latter proportion is equivalent to π_{list} or $1-\pi_{list}$ depending on whether answering affirmatively to the statement indicates a socially desirable or undesirable response. By assumption, $|\pi_{direct} - \pi_{list}| \leqslant \pi_{undesirable}$.[13]

SURVEY QUESTION DESIGN

National identity and multiculturalism policies in Canada do not center around the passive acceptance of all ethno-cultural groups, but champion these groups' contributions to Canadian society. To probe support for the multicultural project, we therefore design questions around whether Canadians believe that these groups' contributions to the country deserve respect. Specifically, we ask whether respondents agree or disagree with the following statements, one of which was randomly assigned to each respondent:

[13]This is the 'no liars' assumption (Blair and Imai, 2013, 52): that the proportion of 'liars' in the list experiment is less than or equal to that in the direct question.

- `Every cultural group's contribution to Canada deserves respect`
- `Aboriginals' contribution to Canada deserves respect`
- `Asian Canadians' contribution to Canada deserves respect`
- `French Canadians' contribution to Canada deserves respect`

These questions enable us to test the hypotheses as follows. First, the statement concerning respect for 'Every cultural group' probes attitudes toward the abstract notion of ethno-cultural group stripped of prejudice. It provides a useful contrast to the remaining three statements, which refer specifically to ethno-cultural groups by name. Comparisons between the 'Every cultural group' condition and the three group-specific conditions test whether respect for the abstract notion differs from that of the groups that make this abstract a social reality. Ideally, one would have a treatment condition for all possible ethno-cultural groups, and average over their proportions in population. This is not possible for practical reasons. We therefore choose three groups—Aboriginals, Asians, and French Canadians—that form a large proportion of ethno-cultural minorities in Canada,[14] and which therefore provide much of the ethno-cultural 'content' implicit in the 'Every cultural group' statement.

Second, the choice of these three groups permits a test of the hypothesis that respondents will misrepresent their attitudes more toward 'new' ethno-cultural groups than 'traditional' ones. Aboriginal and French Canadians are the two major groups traditionally and historically tied to Canada and its culture: the majority of aboriginals are considered the 'First Nations' of Canada and French Canadians are formally one of Canada's founding peoples. Asian Canadians, on the other hand, have formed a substantial proportion of the immigration to Canada in the last 50 years or so. Asian Canadian immigration to Canada began in the late eighteenth century, primarily in response to the gold rush in British Columbia and to work on the pan-Canadian railroad. As with Aboriginals, Asians suffered heavily from highly discriminatory government policies (see for example Ward, 1978).[15] Based on the 2006 census, they constitute roughly 11% of the Canadian population.

The Survey Instrument

To test the hypotheses, we embed the list experiment in a unique survey instrument fielded during the 2013 British Columbian provincial election. The instrument, called

[14]For example, Canadians of South Asian and Chinese origin constitute 46% of the visible minority population of Canada. As stated earlier, French-Canadians form roughly 20% of the Canadian population while 4.3% of Canadians report having an Aboriginal identity (Statistics Canada, 2013).

[15]To give just three historically important examples of discrimination toward Asians: Between 1923 ans 1947, the Chinese Immigration Act put a stop to Chinese immigration, already subject to a prohibitive head tax, by closing the border to all but a few Chinese immigrants. Japanese Canadians on their part were subject to internment inland and away from the West Coast during the Second World War. Finally, in 1914, the Komagata Maru, a vessel transporting more than 300 would-be immigrants from India was not allowed to dock in Vancouver and had to sail back to India.

Vote Compass, is an electoral literacy and engagement tool, which asks respondents a series of election issue-related and socio-demographic questions. It was fielded through the Canadian Broadcasting Corporation (CBC), Canada's public broadcaster, and widely advertised throughout British Columbia. The list experiment itself ran from May 4 through May 14, 2013.

British Columbia is an interesting case because it includes an important aboriginal population (5.4% compared to 4.3% nationally) as well as the highest proportion of Asians Canadians in the country. People of Chinese origin make up 10.1% of the provincial population compared to 4% for Canada as a whole.[16]

To ensure the validity of responses, we make conservative choices about those included in the sample for analysis. We implement a wide variety of technical measures, and capture a series of survey-specific variables to prevent invalid responses. We include only those respondents who filled out the majority of socio-demographic questions, with the assumption that this optional section would not be filled out by those inputting invalid responses.[17] In addition, the analysis only uses non-visible minorities respondents, which means that from our original set of 44,131 respondents, our sample includes 25,790 respondents.

Results

We begin by testing the first two hypotheses. The first hypothesis predicts that support for the principle of multiculturalism ('respect for every cultural group') will be greater than support for each of the ethno-cultural groups named specifically. We expect that this difference in support will hold irrespective of whether the question is asked directly or in the list experiment. The second hypothesis predicts that among the direct questions, support for 'new' ethno-cultural groups ('Asians') will be greater than support for 'traditional' groups ('Aboriginals' and 'French Canadians'). This second hypothesis is predicated on the assumption that there will be larger social desirability bias for questions concerning 'Asians' than for the other three groups. We explore this possibility — our primary hypothesis — further below.

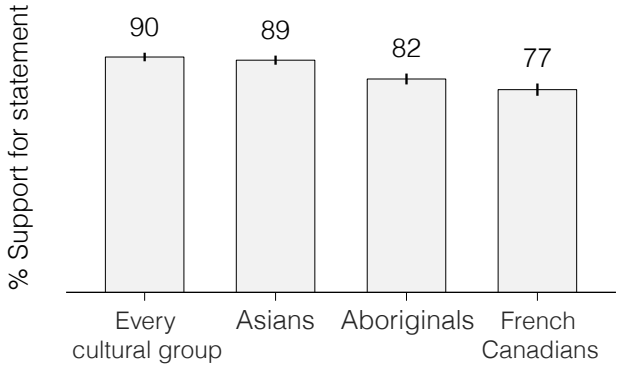
Direct Questions

We first examine the direct questions. Recall that respondents who received the baseline list experiment question were also presented with one of the four statements, selected at random, and asked whether they agree with it. This question was asked

[16]When combined, people of Chinese, South Asian and Filipino origin represent 47% of the Vancouver metropolitan area population (Statistics Canada, 2013)

[17]For the integrity of the use of the application in the future, we do not discuss the technical details of the tool's ability to detect invalid responses.

FIGURE 2: PROPORTION SUPPORT FOR ETHNO-CULTURAL GROUPS (DIRECT QUESTION)



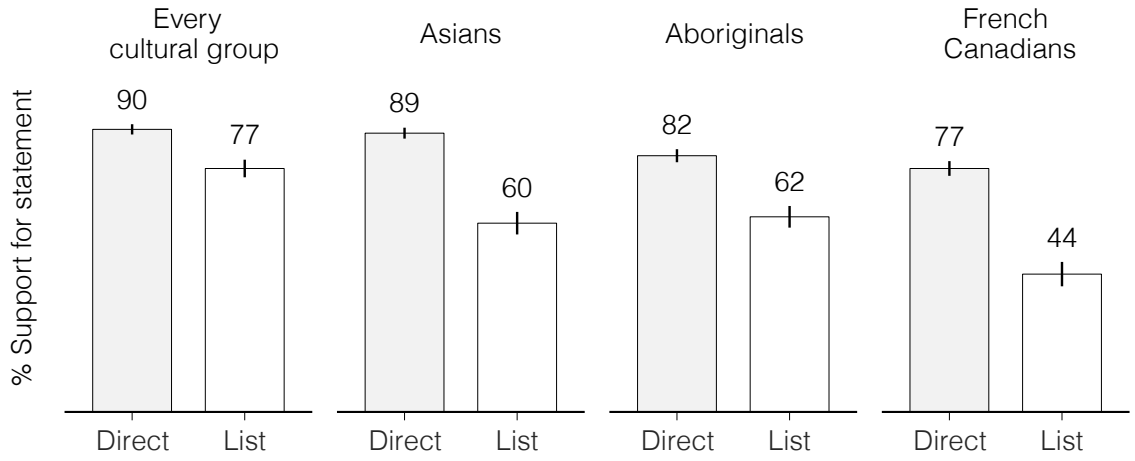
These bar graphs show the proportion of support for each statement among non-visible minorities, as estimated for the direct questions. Vertical lines indicate 95% confidence intervals around each proportion.

immediately after baseline respondents received the list experiment. We present estimates of the proportions of support for each of these statements in Figure 2.

Surprisingly, the results in support of the first hypothesis are mixed. As expected, the proportion of respondents agreeing with the 'every cultural group' statement (90%) is significantly ($p < 0.01$) and substantively different from agreement with both the 'aboriginal' (82%) and 'French Canadians' (77%) statements. Support for the 'every cultural group' statement does not, however, differ substantively (1.2 percentage points) from support for the 'Asian' statement (89%), nor is this difference statistically significant ($p = 0.36$). The results are therefore mixed: the direct question responses suggest that support for the statement regarding respect for Asians' contributions to the country has roughly the same amount of support as that for the more abstract statement regarding respect for multiculturalism in general. This is surprising since this abstract notion is stripped of possible prejudice toward a specific group. The data clearly support the second hypothesis however: respect for 'Asian' contributions to the country has more support in the direct questions than that for the 'Aboriginals" and 'French Canadians" statements ($p < 0.01$).

What conclusions are one to draw from these results? The data appear to suggest that Asians enjoy widespread respect in general — as much as the principle of multiculturalism — and greater support than Aboriginals and French Canadians. One may be tempted to find explanations in Asians' higher economic status and their popular image as hard-working first- and second-generation immigrants. Conversely, one may argue that support for Aboriginals and French Canadians is lower than that for Asians because of the history of discrimination against Aboriginals and negative

Figure 3: Direct Question v. List Experiment



These bar graphs show the proportion of support for each statement among non-visible minorities, as estimated for the direct and list questions. List experiment estimates calculated using the R package `list` (Blair and Imai, 2013). Vertical lines indicate 95% confidence intervals around each proportion.

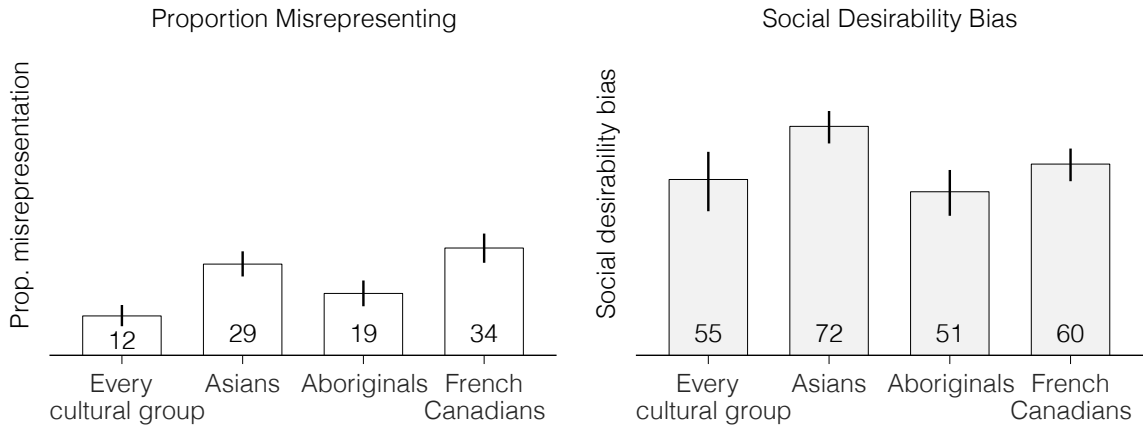
attitudes toward Quebec separatism respectively. If there is heterogeneity in social desirability bias, however, and this bias is strongest for the 'Asian' statement, then these results may be misleading. To explore this possibility, we turn to the results from the list experiment. As we will see, they paint a much different picture.

List Experiment

We present estimates of support for each statement from both the direct and list experiments in Figure 3. As one would expect where there is social desirability bias, the levels of support for each of the four statements in the list experiment are lower than in each of the direct questions. These differences in support between direct and list questions also appear to vary across the four statements. This suggests that there are differences in the magnitude of social desirability bias, which will we investigate in the next section.

Unlike the mixed results for the first hypothesis in the direct questions, the list experiment strongly supports the hypothesis that their is more support for the principle of multiculturalism than that for the specified groups individually. The difference to note here is that between the direct and list experiment estimates of support for the 'Asian' statement. Recall from the direct questions that support for the 'every cultural group' and 'Asian' statements did not differ substantially (90% v. 89%) nor significantly ($p = 0.36$). In the list experiment however, support for the 'Asian'

Figure 4: Degree of Social Desirability Bias by Group



This bar graph shows the difference in support for each respective statement between the direct and list question among non-visible minorities. Vertical lines indicate 95% confidence intervals around each proportion.

statement is much lower than that for the 'every cultural group' statement (77% v. 60%). In the direct questions, where the difference between support for the 'Every cultural group' and 'Asian' statements was 1 percentage point, in the list experiment, this difference is 17 percentage points ($p < 0.01$). Furthermore, support for the 'Asian' statement, which was higher than support for 'Aboriginals' in the direct questions, is 2 percentage points *lower* in the list experiment ($p = 0.059$).

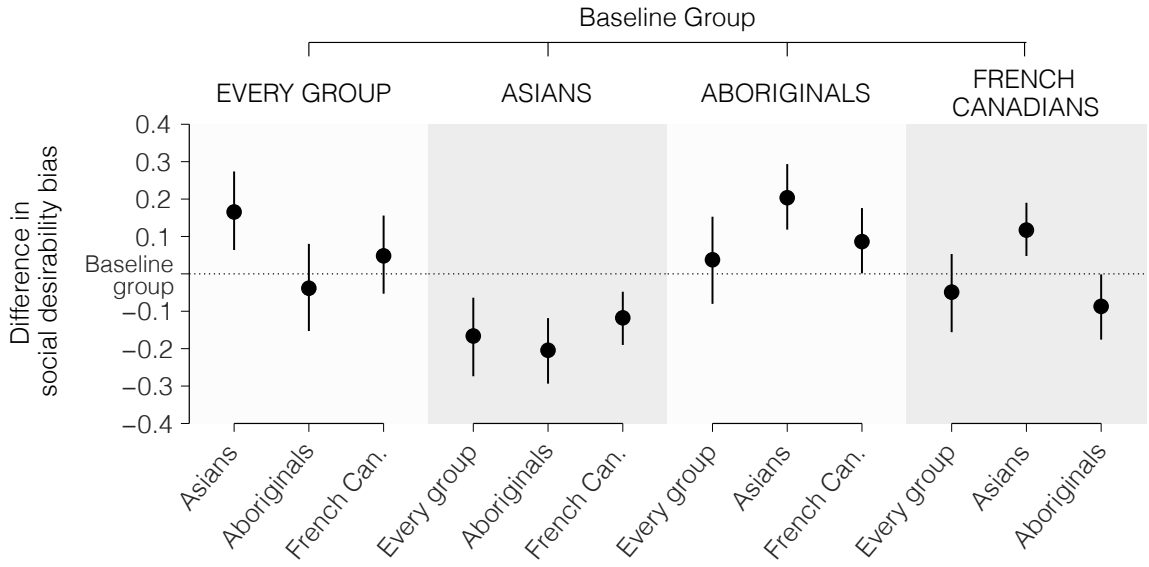
In sum, the results from the list experiment support the first hypothesis: the statement specifying the general principle of multiculturalism ('every cultural group') enjoys more (genuine) support than that for the ethno-cultural groups individually.

Heterogeneity in Social Desirability Bias

We now turn to the third hypothesis, which predicts that respondents will be more likely to misrepresent their attitudes toward 'new' ethno-cultural groups than 'traditional' groups and the more general principle of multiculturalism. To estimate the level of social desirability bias, we generate predicted probabilities through simulation for the direct and list experiment questions, and use the resulting probabilities to calculate the level of social desirability bias for each statement. We test the hypothesis by taking the differences in social desirability bias between groups and calculating the 95% confidence interval of these differences (Blair and Imai, 2013).

We begin by examining the estimates of the total proportion of respondents misrepresenting their attitude toward each statement. We present these graphically in

FIGURE 5: PREDICTED DIFFERENCES IN SOCIAL DESIRABILITY BIAS



These bar graphs show the estimated differences between each pair of social desirability bias estimates. Vertical lines indicate 95% confidence intervals around the difference.

the first panel of Figure 4. As a proportion of all respondents, 34% misrepresent their attitude toward the French Canadian statement; 29% toward the Asian statement; 19% toward the Aboriginal statement; and 12% toward the more general 'every cultural group' statement. The 95% confidence intervals of the differences in these proportions for each pairwise comparison (not shown) do not contain zero except for that between the Asian and French Canadian statements. As we noted earlier however, it is unclear what conclusions one should draw from these proportions: their size is a function of the total proportion of respondents who hold the socially undesirable attitude. Recall from Figure 3 that fewer respondents agreed with the Asians and French Canadian statements than with the other two statements. The results therefore show that their is more misrepresentation of attitudes toward the French Canadians and Asian statements in total, but is it unclear whether this is due to a greater pressure to conform with social expectations. We therefore now turn to our measure of social desirability bias.

In the second panel of Figure 4, we present the estimates of social desirability bias, defined as the proportion of those who misrepresent their attitude among those who hold the socially unacceptable attitude. As expected, the results strongly support the hypothesis. Among respondents who hold the socially undesirable attitude toward the Asian statement, 72% of them misrepresent it. This compares with 60% who misrepresent their attitude toward French Canadians; 55% toward the 'every cultural

group' statement; and 51% toward the Aboriginal statement. We present the differences in social desirability bias for each condition in Figure 5, along with their respective 95% confidence intervals. As one can see, the confidence intervals for the difference between the Asian statement and each of other statements do not cross zero. Although we did not hypothesize a difference in the social desirability of the other statements, the 95% confidence interval for the difference in social desirability bias between the French Canadian and Aboriginal statements also does not cross zero. This means that among the particular groups, the norm appears to be the strongest when it comes to Asian-Canadians and the weakest for Aboriginals. In other words, respondents who have a prejudiced opinion towards Aboriginals are less likely to lie about it than respondents who have similar views toward Asian-Canadians or French-Canadians.

Discussion and Conclusion

When governments promote diversity through policies such as multiculturalism, they hope both to decrease prejudiced behaviors and expressions, and to remove the attitudes that support them. Changes in behaviors and expressions do not necessarily reflect changes in attitudes however. Expressions of respect for minority groups may derive from truly held beliefs or from compliance with social norms. Obtaining reliable estimates of the proportion of individuals who hold socially sensitive attitudes toward ethno-cultural groups is therefore difficult: social pressure may lead to the self-censorship of attitudes. A great deal of effort has been devoted to finding ways to circumvent the biases produced by social desirability. In this article, instead of treating social desirability as a nuisance, we present it as a quantity of interest that reflects the strength of the social norm on a given issue. Doing so, we highlight a problem that arises when this norm does not apply equally to different minority groups. The consequence is one of heterogeneity in social desirability bias.

Our hypotheses were built on two intuitions. The first was that the anti-prejudice norm in an immigration country like Canada would be stronger when the question is about 'newer' ethno-cultural groups than more 'traditional' ones. The second intuition, inspired by research done on prejudice in social psychology, was that this difference in the strength of the norm would translate into a variation in lying. Because we expected the norm to be stronger for these groups, our hypothesis was that people with the undesirable opinion would be more likely to lie about it. Specifically, results from our large-scale survey experiments show that respondents lie more when the question is about Asian Canadians than when it is about Aboriginals or French Canadians. We also demonstrate that an abstract notion of 'every cultural group' receives more support than the specific groups that form this whole and that this support is more genuine.

The findings reported in this article have several potentially important implications both methodologically and politically.

Methodologically, to our knowledge this represents the first investigation of social desirability bias across a series of groups. The heterogeneity uncovered here raises doubt on past research such as Berry and Kalin (1995) and Coenders, Lubbers and Scheepers (2008) comparing direct measures of support for different ethnic groups. As our results show, a difference in level of support between two groups based on a direct measure can hide support levels that are in fact quite different, even the opposite. One could be led to conclude that a group enjoys higher support when in fact it is the underlying norm that makes people lie more about this group than others.

This finding means that future research wishing to compare attitudes toward different groups should either employ indirect measures or have clear expectations about the strength of the anti-prejudice norm for each group of interest. Another interesting possibility is to use measures of self-monitoring to control for the propensity to follow social norms at the individual level (Berinsky and Lavine, 2012; Hainmueller and Hopkins, 2012; Weber et al., 2013). For example, if the expectation is that there might be heterogeneity in social desirability bias, looking at results for respondents who have low scores on a scale such as the Self-Monitoring Scale (Gangestad and Snyder, 2000; Snyder and Gangestad, 1986, e.g.) would partially control for the social norm and mitigate heterogeneity.

We argue that it is the difference in the strength of the norm that explains this heterogeneity in social desirability bias across groups. For this reason, our results should also be taken into account when comparing attitudes across countries. If the strength of the anti-prejudice norm can vary across minorities in a single country, it might also vary across geographical contexts. Here again, measures of self-monitoring represent a potential solution. Testing these propositions— the possibility of heterogeneity in social desirability bias across countries and the use of self-monitoring as a correction—are interesting new directions for future research.

This also demonstrates that list experiment should be used for more than just obtaining unbiased estimates of opinions on sensitive issues. The combination of a direct question and list experiments, something that we are not the first to exploit (Gilens, Sniderman and Kuklinski, 1998; Janus, 2010; Blair and Imai, 2012), makes it possible to gain more insight in the dynamics of social desirability. More importantly, we specify the quantity that should be identified as social desirability bias and distinguish it from the proportion of respondents misrepresenting their true attitudes. Gilens, Sniderman and Kuklinski (1998) also reported this quantity but without identifying it as social desirability bias. Studies following in their footsteps often only focused on the difference between the two experimental conditions. More generally, and as we show here, adding a direct question to the baseline condition in a list experiment only has advantages. Of course, some questions might be too sensitive to be asked directly, but we believe that when possible, a direct question should be part of basic list experiment designs.

Politically, these results have implications both for our specific case, Canada, and for immigration countries in general. A country like Canada that prides itself on

its acceptance of immigration should want its citizens to hold these views genuinely. We asked respondents whether they agreed with the statement that different ethnic groups' contribution to Canada deserved respect, a rather easy statement to agree with. Although a large proportion agreed with the abstract notion that 'every cultural group' deserved respect, it was not the case for the actual groups that make up this abstract notion. Asians Canadians, for example, saw only 60% of respondents support the statement when asked indirectly. The result for Asian-Canadians is not the only worrisome finding about Canadians' acceptance of the multiculturalism idea. A question asking whether the contribution to the country of French Canadians, one of Canada's founding people, deserves respect received only 44% of support. With such a result, a multicultural vision of the country that celebrates the contribution of all ethno-cultural groups appears to be on less than solid grounds. The fact that we find these levels of support in a country that is often hailed as a success story of host-society and immigrants relations offers grim prospects for other countries of immigration.

Based on what we presented here, it is not surprising that a study such as Guimond et al. (2013) finds that a stronger anti-prejudice norm is associated with less prejudice toward a minority group—in their case Muslims. These scholars see this result as a demonstration of the positive effect of multiculturalism as an integration policy (Guimond et al., 2013, 955). However, our results suggest that this finding could actually rest on citizens lying more about their attitudes toward Muslims in countries that have a stronger anti-prejudice norm. Whether this still represents a positive effect of multiculturalism remains, empirically, an open question, and a research agenda which, we hope, will be explored more fully. ■

References

Adams, Michael. 2007. *Unlikely Utopia: The Surprising Triumph of Canadian Pluralism.* Toronto: Viking Canada.

Banting, Keith and Will Kymlicka. 2006. *Multiculturalism and the Welfare State: Recognition and redistribution in Contemporary Democracies.* Oxford University Press.

Banting, Keith and Will Kymlicka. 2013. "Multiculturalism Policy Index." `http://www.queensu.ca/mcp/index.html`.

Berinsky, Adam J. 1999. "The Two Faces of Public Opinion." *American Journal of Political Science* 43(4):1209–1230.

Berinsky, Adam J. and Howard Lavine. 2012. *Improving Public Opinion Surveys: Interdisciplinary Innovation and the America National Election Studies.* Princeton University Press chapter Self-Monitoring and Political Attitudes, pp. 27–45.

Berry, John W. and Rudolf Kalin. 1995. "Multicultural and Ethnic Attitudes in Canada: An Overview of the 1991 National Survey." *Canadian Journal of Behavioural Sciences* 27(3):301–320.

Blair, Graeme and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20(1):47–77.

Blair, Graeme and Kosuke Imai. 2013. "list: Statistical Methods for the Item Count Technique and List Experiment." The Comprehensive R Archive Network (CRAN).

Blinder, Scott, Robert Ford and Elisabeth Ivarsflaten. 2013. "The Better Angels of Our Nature: How the Antiprejudice Norm Affects Policy and Party Preference in Great Britain and Germany." *American Journal of Political Science* 0(0):0–0.

Bloemraad, Irene. 2006. *Becoming a Citizen: Incorporating Immigrants and Refugees in the United States and Canada*. University of California Press.

Bouchard, Gerard and Charles Taylor. 2008. *Fonder l'Avenir, Le Temps de la Conciliation: Rapport de la Commission Bouchard-Taylor*. Quebec: Quebec Government.

Bradburn, Norman M., Seymour Sudman, Ed Blair and Carol Stocking. 1978. "Question Threat and Response Bias." *Public Opinion Quarterly* 42(2):221–234.

Citizenship and Immigration Canada. 2012. "Our Mandate." `http://www.cic.gc.ca/english/department/mission.asp`.

Citrin, Jack, Richard Johnston and Matthew Wright. 2012. "Do Patriotism and Multiculturalism Collide? Competing Perspectives from Canada and the United States." *Canadian Journal of Political Science* 45(3):531–552.

Coenders, Marcel, Marcel Lubbers and Peer Scheepers. 2008. "More than two decades of changing ethnic attitudes in the Netherlands." *Journal of Social Issues* 64(2):269–285.

Crandall, Christian S., Amy Eshleman and Laurie O'Brien. 2002. "Social Norms and the Expression and Suppression of Prejudice: The Struggle for Internalization." *Journal of Personality and Social Psychology* 82(3):359–378.

Crepaz, Markus. 2006. *Multiculturalism and the Welfare State: Recognition and redistribution in Contemporary Democracies*. New York: Oxford University Press chapter Public Opinion on Multiculturalism, Trust, and the Welfare State.

Crowne, Douglas P. and David Marlowe. 1960. "A New Scale of Social Desirability Independent of Psychopathology." *Journal of Consulting Psychology* 24(4):349–354.

Edwards, Allen Louis. 1953. "The Relationship Between the Judged Desirability of a Trait and the Probability That the Trait Will be Endorsed." *The Journal of Applied Psychology* 37(2):90–93.

Edwards, Allen Louis. 1957. *The Social Desirability Variable in Personality Assessment and Research*. New York, NY: Dryden.

Fisher, Robert J. 1993. "Social Desirability Bias and the Validity of Indirect Questioning." *Journal of Consumer Research* 20(2):303–315.

Furnham, Adrian. 1986. "Response Bias, Social Desirability and Dissimulation." *Personality and Individual Differences* 7(3):385–400.

Gangestad, Steven W. and Mark Snyder. 2000. "Self-Monitoring: Appraisal and Reappraisal." *Psychological Bulletin* 126(4):530–555.

Gilens, Martin, Paul M. Sniderman and James H. Kuklinski. 1998. "Affirmative Action and the Politics of Realignment." *British Journal of Political Science* 28(1):159–183.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio and David W. Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56(1):202–217.

Guimond, Serge, Pierre De Oliveira, Nour Kteily, Richard N. Lalonde, Felicia Pratto, Jim Sidanius, Richard J. Crisp, Rodolphe Kamiejski, Beate Kuepper, Shana Levin, Francine Tougas and Andreas Zick. 2013. "Diversity Policy, Social Dominance, and Intergroup Relations: Predicting Prejudice in Changing Social and Political Contexts." *Journal of Personality and Social Psychology* 104(6):941–958.

Hainmueller, Jens and Daniel J. Hopkins. 2012. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes Toward Immigrants." *MIT Political Science Department Research Paper no.2012-22.* .

Harell, Allison, Stuart Soroka, Shanto Iyengar and Nicolas Valentino. 2012. "The Impact of Economic and Cultural Cues on Support for Immigration in Canada and the United States." *Canadian Journal of Political Science* 45(3):499–530.

Heerwig, Jennifer A. and Brian J. McCabe. 2009. "Education and Social Desirability Bias: The Case of a Black Presidential Candidate." *Social Science Quarterly* 90(3):674–686.

Himmelfarb, Samuel and Carl Lickteig. 1982. "Social Desirability and the Randomized Response Technique." *Journal of Personality and Social Psychology* 43(4):710–717.

Holbrook, Allyson L. and Jon A. Krosnick. 2010. "Social Desirability Bias in Voter Turnout Reports." *Public Opinion Quarterly* 74(1):37–67.

Huddy, Leonie and Stanley Feldman. 2009. "On Assessing the Political Effects of Racial Prejudice." *Annual Review of Political Science* 12:423–447.

Imai, Kosuke. 2011. "Multivariate Regression Analysis for the Item Count Technique." *Journal of the American Statistical Association* 106(494):407–416.

Janus, Alexander L. 2010. "The Influence of Social Desirability Pressures on Expressed Immigration Attitudes." *Social Science Quarterly* 91(4):928–946.

Jedwab, Jack. 2005. "Neither Finding nor Losing our Way: The Debate Over Canadian Multiculturalism." *Canadian Diversity* 4(1):95–102.

Johnson, Timothy P. and Fons J. R. Van de Vijver. 2002. Social Desirability in Cross-cultural Research. In *Cross-Cultural Survey Methods*, ed. Janet A. Harkness, Fons J. R. Van de Vijver and Peter Ph. Mohler. New York, NY: Wiley pp. 193–202.

Joppke, Christian. 2004. "The retreat of multiculturalism in the liberal state: theory and policy." *British Journal of Sociology* 55(2):237–257.

Karp, Jeffrey A. and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries." *Journal of Politics* 67(3):825–840.

Kuklinski, James H., Michael D. Cobb and Martin Gilens. 1997. "Racial Attitudes and the "New South"." *The Journal of Politics* 59(2):323–349.

Kuklinski, James H., Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence and Barbara Mellers. 1997. "Racial Prejudice and Attitudes Toward Affirmative Action." *American Journal of Political Science* 41(2):402–419.

Kymlicka, Will. 2003. "Being Canadian." *Government and Opposition* 38(3):357–385.

Mendelberg, Tali. 2001. *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton, NJ: Princeton University Press.

Redlawsk, David P., Caroline J. Tolbert and William Franko. 2010. "Voter, Emotions, and Race in 2008: Obama as the First Black President." *Political Research Quarterly* 63(4):875–889.

Richeson, Jennifer A. and Richard J. Nussbaum. 2004. "The IMpact of Multiculturalism versus Color-Blindness on Racial Bias." *Journal of Experimental Social Psychology* 40:417–423.

Sears, David and P.J. Henry. 2005. "Over Thirty Years Later: A Contemporary Look at Symbolic Racism and Its Critics." *Advances in Experimental Social Psychology* 37:95–150.

Sigall, Harold and Richard Page. 1971. "Current Stereotypes: A Little Fading, a Little Faking." *Journal of Personality and Social Psychology* 18(2):247–255.

Snyder, Mark and Steven W. Gangestad. 1986. "On the Nature of Self-Monitoring: Matters of Assessment, Matters of Validity." *Journal of Personality and Social Psychology* 51(1):125–139.

Statistics Canada. 2013. "Immigration and Ethnocultural Diversity in Canada." `http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001-eng.pdf`.

Streb, Matthew J., Barbara Burrell, Brian Frederick and Michael A. Genovese. 2008. "Social Desirability Effects and the Support for a Female American President." *Public Opinion Quarterly* 72(1):76–89.

Ward, W. Peter. 1978. *White Canada Forever: POpular Attitudes and Public Policy Toward Orientals in British Columbia*. Montreal: McGill-Queen's University Press.

Warner, Stanley L. 1965. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." *Journal of the American Statistical Association* 60(309):63–69.

Weber, Christopher R., Howard Lavine, Leonie Huddy and Christopher M. Federico. 2013. "Placing Racial Stereotypes in Context: Social Desirability and the Politics of Racial Hostility." *American Journal of Political Science* forthcoming.

Weldon, Steven. 2006. "The Institutional Context of Tolerance for Ethnic Minorities: A Comprative, Multilevel Analysis of Western Europe." *American Journal of Political Science* 50(2):331–349.

Wilkes, Rima and Catherine Corrigall-Brown. 2011. "Explaining Time Trends in Public Opinion: Attitudes Towards Immigration and Immigrants." *International Journal of Comparative Sociology* 52(1-2):79–99.

Wright, Matthew. 2011. "Policy Regimes and Normative Conceptions of Nationalism in Mass Public Opinion." *Comparative Political Studies* 44(5):598–624.

Zerbe, Wilfred J and Delroy L. Paulhus. 1987. "Socially Desirable Responding in Organizational Behavior: A Reconception." *The Academy of Management Review* 12(2):250–264.