Maps in People's Heads: Assessing A New Measure of Context

Cara Wong* Jake Bowers Daniel Rubenson Mark Fredrickson

Ashlea Rundlett

August 24, 2017

Abstract

To understand the relationship between place and politics, we must measure *both* political attitudes *and* the ways in which place is represented in the minds of individuals. In this paper, we assess a new measure of mental-representation of geography, in which survey respondents draw their own local communities on maps and describe them. This mapping measure has been used in Canada, the UK, Denmark, and the U.S. so far. We use a panel study in Canada to present evidence that these maps are both valid and reliable measures of a personally relevant geographic area, laying the measurement groundwork for the growing number of studies using this technology. We hope to set efforts to measure 'place' for the study of context and politics on firmer footing. Our validity assessments show that individuals are thinking about people and places with which they have regular contact when asked to draw their communities. Our reliability assessments show that people can draw more or less the same map twice, even when the exercise is repeated months later. Finally, we provide evidence that the concept of community is a tangible consideration in the minds of ordinary citizens and is not simply a normative aspiration or motivation.

*Associate Professor, Department of Political Science, University of Illinois @ Urbana-Champaign (carawong@illinois.edu). *Acknowledgements:* This work was supported by the Social Science and Humanities Research Council of Canada.

Theories explaining how a local area can structure, catalyze, or inhibit political behavior and attitudes tend to rely on two types of causal stories: either a person has to perceive something about the place in order for it to matter (e.g., the racial composition of a community inspires fear) or something about the place structures action without the need for perception (e.g., lack of public space inhibits interaction). A common approach to studying context is to combine individual level data on behavior or attitudes with administrative data about place, e.g. a census tract, zipcode or city. However, we know that people experience these places differently and we suspect that these differences have political consequences.

In this paper we assess an alternative measure that overcomes this challenge and measures context at the individual level. A recent development to measure perceptions of contextual boundaries and content directly has caught the attention of a growing number of research teams around the world (Wong et al. 2012)[1]; the basic idea is to measure 'place' at the level of the individual rather than using aggregated, administrative data, asking respondents to draw a map of their 'local community.'

Asking online respondents to draw, rather than click buttons or type, involves a new task. Survey researchers have long asked people to engage in new tasks; the idea that a stranger would come into your home and ask you to express your opinions using scales was once a very novel thought, and it required adjustment and learning by the public as the study of public opinion grew (Groves 2011). We see the map drawing task as following in the same line of innovation in the measurement of public opinion as pioneered by Gallup and others. This paper aims to investigate the consequences of asking people to do this new task, so that as scholars of mass opinion, we can evaluate and improve efforts to measure context at the individual level.

The validity question is, roughly speaking, whether hand drawn maps and their descriptions bear some clear relationship with what we think of as a "local community." The reliability question is whether people can draw and describe similar maps if asked to do so at different points in time. We are also able to add to the conceptual literature on "community" by examining whether it is a tangible

[1]Similar measures were included in the 2014–5 British Election Study, the 2014 and 2016 Cooperative Congressional Election Studies in the U.S., in a 2016 national survey in Denmark, and in a 2012 survey of English Canada we analyze here.

construct in the minds of ordinary people, not simply a normative aspiration or motivation.

1 Data: The Measuring Local Communities Canada (MLCC) Survey

We use data from the MLCC survey—an online panel survey of roughly 7800 respondents in English Canada conducted in April–July 2012 (wave 1) and August–October 2012 (wave 2)—based on a convenience sample.[2] Our respondents' perceptions of their environments are not meant to be generalizable to the nation, but our measure of context is broadly applicable, as are the questions our measure raises about standard practices of using administrative units as measures of context.

In addition to answering traditional survey questions, our respondents created and referred to a couple of maps. They first gave their postal code and checked a Google Map centered on it to make sure that we had correctly located them. Then, after answering a few questions about their residence, they proceeded to a screen with a new map centered again on where they lived.[3] At that point, the questionnaire asked them to draw their "local community." The map-drawing task was one of the first in the survey, so the respondents were not primed to think about particular issues by other survey questions. They could zoom in and out, and they could draw any shape or shapes they liked.

Figure 1: Example hand drawn "local communities" from the MLCC Survey 2012.

Figure 1 shows four examples of the "local communities" drawn: the top row shows examples

[2]See Appendix A and Appendix B for more details about the sampling frame, sample, question wordings, and auxiliary analyses for this paper, including more details about the results of the map drawing.

[3]Previous research has shown that most Canadians are comfortable giving their postal codes. However, using IP addresses and the now common automatic location identification capabilities of web-browsers when available may ease the process of locating respondents.

of small and large communities, one focusing on a single street and the other containing most of the habitable land of Canada; the bottom left map represents a median size community of 12 square kilometers; and the bottom right is a community composed of multiple polygons. While the distribution in their sizes is unimodal, people's "local communities" vary a great deal—ranging from a single street to multiple continents—and about 16% of the communities drawn were composed of multiple polygons. The zoom-level of the Google map presented to the respondents was randomized, so the clustering of interpretations of "local community" as a place containing about 12 square kilometers was not, therefore, an artifact of cognitive anchoring.[4]

2 Validity of the Measure: Do the maps capture individualized context?

Can we interpret the drawings as telling us something about the boundaries of "local community" as they exist in people's minds? And, does asking for respondents' descriptions of their environments give us information above and beyond what Census numbers could provide?

2.1 Content and Convergent Validity

Scholars tend to conceptualize "community" as place, space, and networks of social ties and allegiances (Keller 2003; Agnew 2011; Wong 2010). Nevertheless, Hillery (1955) identified 95 definitions of "community," and the count can only have increased dramatically over the past half century. We provided no definition of "local community" in the survey, since part of our purpose was to discover its diverse meanings for ordinary people (which may not reflect what social scientists theorize).[5] The instructions did state that there was no right answer and that we had not included a definition because we wanted to know what was important to the respondents.[6] We included a follow-up question for 10% of the respondents in the initial survey and for all respondents in the second wave, asking them what they were thinking about as they drew their "local community." The options were derived from past theoretical and empirical research on community (e.g. Keller 2003; Agnew 2011; Wong 2010; Hillery 1955). As Table 1 shows, regular contact matters: neighbourhood was mentioned by 75% of

[4]See Appendix B.2 and Appendix B.3 for more information on the size of communities and time spent on the drawing task.

[5]This does make the traditional test of face validity—asking academic experts whether the question looks like what it is supposed to measure—moot.

[6]The terminology was chosen to emphasize a personally relevant grouping of people (a "community") that is spatially interdependent ("local").

respondents, 65% answered people or places seen on a weekly basis, and 68% chose places visited on a regular basis. Politics was much less salient: only 28% mentioned people or places you think about when you go vote in an election.[7]

Table 1: Considerations During "Local Community" Drawing

	% Respondents, Survey 1	% Respondents, Survey 2
Your neighbourhood	75	83
Places you visited regularly	68	71
People/places seen weekly	65	71
Family and friends	49	47
Voting in election	28	30
People like you	28	27
Other	19	15
Newspapers	16	16
Television or internet	5	4
N	809	3076

Valid measures of individualized context should reveal information about the salient people and places physically near a person. Respondents reported that they were thinking about people with whom they were spatially interdependent, and both frequency of interactions and proximity to one's home help determine where people draw the boundaries of their communities. These self-reports support the content validity of the map-drawing measure, eliciting the types of responses expected from operationalization of the concept of "local community."

Convergent validity captures the extent to which variables that theoretically should be related are, in fact, related. For example, one might expect that people who commute greater distances to work would include their work colleagues in their community and therefore stretch their community boundaries. A valid measure should capture this. We also expect that people living in cities would draw smaller communities than individuals living in more rural areas. For example, if one thinks only of neighbors as potential community members, a resident of Toronto could have hundreds of neighbors in her block, whereas a resident of Grise Fiord has about 130 residents in her entire hamlet.

We find evidence of a relationship between commuting distance and community size among our

[7] In the follow-up survey that we describe later, when asked the same question after drawing their local community again, respondents were thinking of very similar things. The order of the response options was randomized in both waves.

respondents in the workforce: the greater the commuting distance (measured as the distance between the centroids of the postal codes of respondents' home and workplaces), the larger the drawn community. In a robust regression model using the log of commuting distance as the predictor and the log of map size is the outcome, the coefficient is 0.37 ($p < .05$).[8]

We also find a negative relationship between population size and community size: the greater the number of residents in a respondent's Census subdivision (CSD), the smaller the community she draws. The coefficient for the log of CSD population is -.27 ($p < .00$) from a robust regression model.[9]

Community size and commuting distance and population size should be and are related, providing evidence of convergent validity (Campbell and Fiske 1959).

2.2 Discriminant Validity

If the drawings of "local communities" are very sensitive to the conditions under which the task is done, we might worry we are capturing a survey artifact rather than a personally relevant "pseudoenvironment" (Schwarz 1995; Schwarz and Sudman 1992; Lippmann 1991). Nevertheless, drawings should relate in some intelligible way to the survey conditions. We randomized the zoom-level of the map to assess anchoring effects, and preceded the map with a randomized geographic cue to assess priming effects. A valid measure would show patterns of responsiveness (say, bigger zoom levels should increase the size of maps drawn) but it shouldn't completely change the typical map if our measure of local communities had a focal geographic meaning in the minds of the survey respondents.

Anchoring Effects

The resolution of the map initially presented to the respondent could influence the size of people's maps. In other words, a respondent who sees a map encompassing her neighborhood might draw a smaller community than if she had seen a map encompassing her city. If there were large anchoring

[8] See Figure 6 in the appendix for more details and a plot of the relationship on both log and original scales. We use "robust regression" to refer to an influential-point resistant linear model fit using MM-estimation (For recent overviews of this approach see Maronna, Martin, and Yohai 2006; Huber 2011) implemented in the `rlm` command for R (Venables and Ripley 2002). We present results using these models in addition to OLS and non-parametric loess models because the size of maps drawn varied greatly and was quite skewed: OLS results were very sensitive to the presence of a few extremely large maps.

[9] See Figure 7 in the appendix for more details and a plot of the relationship on both log and original scales.

effects, we would be concerned about both the validity and reliability of the map-drawing measure (Wilson et al. 1996; Furnham and Boo 2011). To test this possibility, we randomized the resolution of the maps initially presented, ranging from several blocks to an entire metropolitan area. Respondents could zoom in and out using the Google Maps interface, but we chose to limit the resolutions *initially* presented so that it would be a relatively quick task for respondents to locate their own home on the map.

While there is greater variance in the size of the communities drawn for respondents shown maps with the greatest area, there is no consistent effect on the median size of the communities (See Appendix C.1). When we estimate the effect of randomly assigning the map resolution on the area of the "local communities", the results largely confirm these findings: for 3 of the 5 zoom levels, the initial zoom level of the maps people see first has no statistically or substantively significant effect on the size of their "local communities." Respondents who saw maps at the level of the city did draw maps that were distinct, but the means were only about 5 square km larger than people shown maps zoomed into the level of the street or buildings. The absence of large anchoring effects provide support in favor of the discriminant validity of the measure (Campbell and Fiske 1959), while the pattern of increasing variance is convergent validity evidence.[10]

Priming Effects

Social scientists have long been aware that the context of the survey—particularly questions that preceded one's variables of interest—could influence respondents' judgments and answers (Strack, Martin, and Schwarz 1988; Zaller and Feldman 1992; Tourangeau and Rasinski 1988). Because researchers cannot always control where in a survey their questions appear, we want to test whether our map-drawing measure is particularly susceptible to primes about geography. Could explicit cues about particular levels of aggregation before the map-drawing measure change the typical size of maps drawn? To assess this possibility, we randomized whether respondents were asked two questions immediately before the map drawing measure.[11] Four-fifths of the respondents were primed to think of

[10]Because of the greater variance for respondents shown larger areas, our practical advice to researchers would be to start with resolutions that show areas smaller than a city.

[11]The questions were modeled on items used in the General Social Survey (See,e.g., Furia 2005), and are not unusual in their format or substance.

one particular context with the following questions: "Thinking about where you live, how close do you feel to your [neighbourhood/ city/ province/ Canada]?" And, "If you could improve your work or living conditions, how willing or unwilling would you be to move to another [neighbourhood/ city/ province/ Canada]?" A control group received no treatment and proceeded from a question about home ownership directly to the map-drawing.

We find that priming respondents to think of different salient contexts immediately before the map-drawing measure has no substantively large effect on the typical size of maps drawn, except that the average size of the communities drawn by respondents receiving the "neighbourhood" prime is smaller than that of the control, and this difference is statistically significant. However, it is only 1.6 square kilometers smaller on average using a robust regression estimator. The variance of the map sizes drawn did increase with the size of the geographic cue, providing evidence that the treatment was received (see Appendix C.2).

The fact that the typical size of the maps drawn is not strongly or substantively influenced by anchoring or priming effects is evidence in favor of its discriminant validity, while the fact that the variance of the map sizes changes sensibly supports the convergent validity analysis. This is particularly encouraging for its future applications, since researchers cannot always control exactly where a question is placed in a questionnaire.[12]

2.3 Content Within Geographic Boundaries: Administrative vs. Individualized Context

What people say about their local communities (and also about maps with fixed bureaucratic units we showed them) is not identical to what the Census itself would say about those places. In other words, census numbers are not good proxies for people's perceptions. The mapping measure captures something different from what administrative units contain, although, as one might expect, people's subjective contexts are correlated with their "objective" contexts. We focus on the case of ethnic context here because of the large literature on inter-ethnic conflict in which context plays a

[12]For example, the map-drawing measure was part of the very last module of the September 2014 wave of the British Election Study, after respondents had answered many questions about both national and local politics. Given the greater variance in community size following national primes, our general advice is, if possible, avoid placing the map-drawing measure immediately following questions that could prime the national context.

major explanatory role.[13]

Measures of Objective Context

For our objective context measures, we use data from the 2006 Canadian Census. We created measures of ethnic context for five different geographic units: the dissemination area (DA), Census subdivision (CSD), federal electoral district (FED), forward sortation area (FSA), and provinces/territories in which our respondents live. By using these different geographic units, we are able to compare the perceptions of contexts varying in size and political salience. We created an index of the percentage of visible minorities following the Statistics Canada definition, which includes "persons who are non-Caucasian in race or non-white in colour and who do not report being Aboriginal." For the analyses, the "objective" context for respondents' local communities is the average of any dissemination area (DA) that is included in their drawing. (See Appendix D.1.)

Measures of Subjective Context

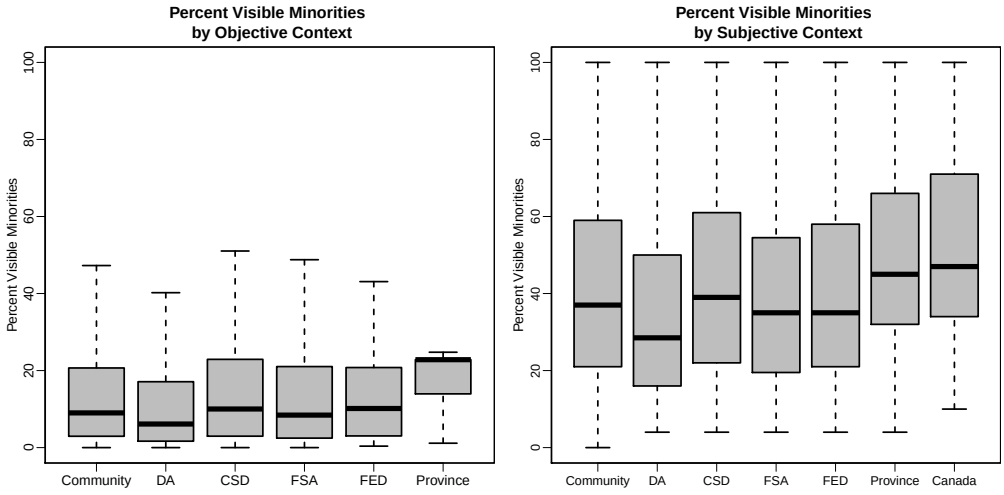
After respondents drew and described their local community, they then answered a battery of questions about their perceptions of the relative size of ethnic/racial groups in their community. The list of groups included the following: Blacks, Canadian Aboriginals, Whites, Chinese, Latin Americans, South Asians (East Indian, Pakistani, Sri Lankan, etc.), and Other Asians (Korean, Japanese, Filipino, etc.) The percentage visible minority in a context was an index adding together responses for Blacks, Chinese, Latin Americans, South Asians, and other Asians. We then showed respondents a map with one of six geographic areas highlighted at random with equal probability: the respondent's DA, CSD, FSA, FED, province, or Canada as a whole.[14] They were again asked to describe the ethnic make-up of this context. Thus, we were able to gauge perceptions of multiple politically-relevant characteristics of a range of geographic contexts.

Figure 2(a) shows a boxplot of the percentage of visible minorities as reported by the Census. The median percentage of visible minorities for the contextual units in our survey is 9% for respondents' local communities, 6% in DAs, 8% in FSAs, 10% in CSDs, 10% in FEDs, and 23% in provinces. In

[13]We did ask respondents to report on other perceptions in the study, including the percentages of partisans and unemployed in the various areas but we do not use that information here.

[14]Because respondents are asked to provide their postal code, we are able to geocode where they live and immediately situate respondents within these different geographic units so as to randomly assign them a correct administrative Census polygon.

Canada overall, visible minorities make up 16 % of the population.



(a) Objective (b) Subjective

Figure 2: Distribution of visible minorities recorded by the Canadian Census 2006 by Census geographies containing the MLCC respondents (left panel) & Perceived by respondents in their own hand drawn Communities, as well as in the official Census polygons pertaining to their postcode (right panel). In the right panel the boxplot labeled "Community" is the perception of the hand-drawn map. The other boxplots show the range of perceptions reported by respondents seeing one of those official polygons overlaid on a Google map.

In contrast, our respondents have different perceptions of the racial/ethnic diversity across the contexts in which they live.[15] Figure 2(b) shows a summary of the distributions of perceptions of visible minorities across the multiple levels of context. The median is 37% for their own local communities (labeled "Commmunity" on the plot), 30% for DAs, 35% for FSAs, 39% for CSDs, 35% for FEDs, 45% for provinces, and 47% for Canada; the degree of misperception tends to increase with the size of the contextual unit, regardless of whether one measures misperception using the actual numbers given by respondents or recalculated as a proportion of their overall responses.[16] Respondents are clearly overestimating the percentages of visible minorities at all levels, which means (1) Census numbers are not not good proxies for what people perceive in their contexts, (2) the relationships between contextual predictors and policy preferences can vary a great deal, depending on whether objective or

[15]Previous research repeatedly has shown innumeracy at the level of the nation (Nadeau, Niemi, and Levine 1993; Highton and Wolfinger 1991; Alba, Rumbaut, and Marotz 2005; Wong 2007), but we know of only one pilot study before now that looks at innumeracy at more local levels (Wong et al. 2012).

[16]Correlations between objective and subjective context tell a similar story of misperception: they range from .36 at the national level to .59 at the level of the CSD.

subjective measures are used, and (3) people's misperceptions of the levels of ethnic diversity in their individually-drawn communities are relatively similar to their misperceptions of diversity in other bureaucratic units. In other words, respondents are not more knowledgeable or accurate about a context that they have defined as being personally relevant, relative to other geographic units. Furthermore, the fact that respondents describe their "local community" in similarly distorted ways as administrative units in which they live provides more evidence that for ordinary people, "local community" is a tangible space and place (and not simply an abstract construct).[17]

3 Test-Retest Reliability

The same person asked to draw a map twice might end up drawing communities that differ, not because of changes in the internal understanding of local context, but because new and difficult tasks are likely to produce noisy outcomes. To address this question, we conducted a short follow-up survey 4 to 5 months later. We showed respondents the same street map and asked them to draw their "local community." We showed them the same resolution of the map they were initially shown in the first survey: if a respondent zoomed out from Google resolution 16 to 10 at time 1 to draw a large community, she would have to do it again at time 2 in order to draw a similar map. This is a difficult task and a high bar for reliability. While we do not believe anyone will draw an identical map from time 1 to time 2, a reliable measure would have communities drawn roughly in the same place and of the same size. After all, if respondents are thinking of people they know or places they see on a regular basis as they are drawing their "local communities," these should not change dramatically within a few months' time. However, if the maps are not at all similar, it raises questions about whether the term "local community" is actually substantively meaningful to ordinary people—i.e. it would also impugn the validity of the measure even as it would cast doubt on reliability. Assessing the test-retest reliability in this measure highlighted the novelty of this measurement assessment: A map is a two dimensional object and there are many more ways for such objects to differ from each other than there are for one dimensional objects like scale ratings.

[17] In Appendix D.3 and Appendix D.4 we engage further with the alternative explanations for these patterns, such as innumeracy.

3.1 Centroid Distance and Size of Communities

There are multiple ways to determine the reliability of the measure over time. As shown in Table 1, what respondents were thinking about while drawing their communities was very similar at both time 1 and 2. We also looked at whether the community maps drawn at time 1 and 2 overlapped and found that 95% did. Of course, even if maps intersect, are they generally centered in the same location? To capture this difference in centers, we measure the distance of the centroids of the polygons drawn. The median distance between centroids of maps drawn between four and five months apart is 1.3 kilometers, and the mean is 20 kilometers. Although some communities differed by large distances (such as those for five respondents who volunteered that they had moved), thus pulling the mean upwards, half of all maps differed by less than 1.3 kilometers in centroid. To see this kind of difference on a Google Map would require that one be zoomed in to about level 15, where one clearly can see streets. This is a surprisingly reliable method, if test-retest reliability is measured using differences between centroids.[18]

If respondents' maps at time 1 and 2 overlap and have relatively similar centroids, they still may differ a great deal in size. We were surprised to find that the median difference in the size of respondents' communities drawn over the two waves is 0 square kilometers, although the median of the absolute value of the difference is 13 square kilometers. This provides additional evidence that cues from neither previous questions from the survey nor from respondents' personal experiences grossly affect respondents' local community sizes.

3.2 Overlap of Maps

Finally, we looked at the percentage of the communities drawn that overlap. Because these numbers will differ, depending on whether the map drawn at time 1 or time 2 is used as the baseline, we calculated the percentages in both ways. The median percent of map 1 that is in map 2 is 66% (mean=60%). The median percent of map 2 that is included in map 1 is 68 (mean=60%). As an illustration of map overlap, Figure 3 shows an example where 63% of the map from time 1 is in that of

[18] The maps at time 1 and 2 were centered on the same point, which could influence this reliability. However, the centroids of people's communities were not always the centroids of the Google map presented, and many respondents moved the maps around, irrespective of whether they changed zoom levels.

time 2, and 98% of the map in time 2 is in the map of time 1.

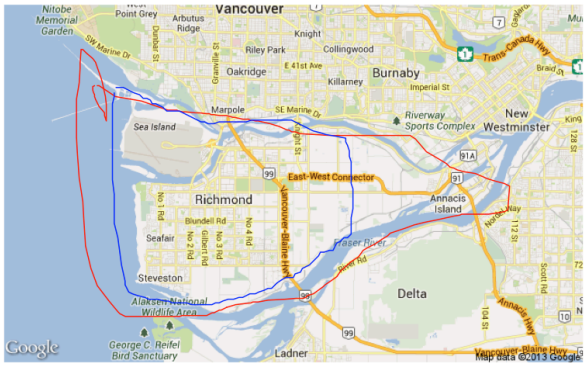


Figure 3: An Example of One Respondent's "Local Communities" Drawn in the Initial Survey (red, larger) and the Follow-up (blue, slightly smaller).

Overall, the results indicate that the map-drawing measure is surprisingly reliable despite the novel task of asking respondents to (1) draw on Google maps online and (2) describe their communities. The evidence of reliability is important because it confirms that the concept "local community" is meaningful to respondents, that the task of drawing one's local community is not so challenging that it cannot be replicated, and that the measure can capture similar information over time. For its future use as a potential outcome measure for program evaluation, this evidence is especially promising.

4 Discussion and Conclusion

Ordinary citizens are capable of picturing local contexts, drawing them using a Google Map interface on a computer, and describing them. We show these measures are valid and reliable; the maps drawn are neither non-attitudes nor ephemeral entities. We also show that what people report about their maps does not directly reflect government statistics—at least when it comes to racial/ethnic characteristics—and that researchers can measure ordinary citizens' perceptions of where they live.

The environments in which people live may affect them in numerous ways. The accuracy of people's perceptions of the partisanship and unemployment levels of where people reside may also vary, and scholars will need to look at how people learn about where they live. Should we think of "context" broadly speaking having an impact on political judgments, or are different facets of it seen and interpreted in a variety of ways? By assessing a novel measure of Lippmann (1991)'s "pictures in their heads" focused on the subjective nature of both the boundaries of communities and their content, we hope to support new approaches to studying the mechanisms involved in linking context with

racial threat, sociotropic retrospective voting, and partisan sorting, among other important political phenomenon.

Finally, we speculate that these maps may help scholars assess the impact of policy interventions aimed at improving the lives of residents living in a particular locale. For example, does the introduction of ideas of private property change how people see the land where they live? Does a change in policing policies affect how people would draw the area in which they feel safe? Or, does the presence of city bikes expand the meaning of "local," as they explore new areas in the city where they live with a different mode of transportation? By measuring changes in maps over time, scholars could help policy makers better track the impact of these types of interventions and understand the mechanisms by which context affects attitude and actions.

References

Agnew, John. 2011. "Space and place". *Handbook of geographical knowledge* 2011:316–331.

Alba, R., R.G. Rumbaut, and K. Marotz. 2005. "Distorted Nation: Perceptions of Racial/Ethnic Group Sizes and Attitudes toward Immigrants and Other Minorities". *Social Forces* 84:901.

Campbell, Donald T, and Donald W Fiske. 1959. "Convergent and discriminant validation by the multitrait-multimethod matrix." *Psychological bulletin* 56 (2): 81.

Furia, Peter A. 2005. "Global citizenship, anyone? Cosmopolitanism, privilege and public opinion". *Global Society* 19 (4): 331–359.

Furnham, Adrian, and Hua Chu Boo. 2011. "A literature review of the anchoring effect". *The Journal of Socio-Economics* 40 (1): 35–42.

Groves, Robert M. 2011. "Three Eras of Survey Research". *Public Opinion Quarterly* 75 (1): 861–871.

Highton, B., and R.E. Wolfinger. 1991. *Estimating the Size of Minority Groups*. Tech. rep. Memo to Board of Overseers, National Election Studies.

Hillery, G.A. 1955. "Definitions of community: Areas of agreement". *Rural sociology* 20 (2): 111–123.

Huber, Peter J. 2011. "Robust statistics". In *International Encyclopedia of Statistical Science*, 1248–1251. Springer.

Keller, Suzanne. 2003. *Community: Pursuing the dream, living the reality*. Princeton University Press.

Lippmann, W. 1991. *Public opinion*. Transaction Publishers.

Maronna, RARD, R Douglas Martin, and Victor Yohai. 2006. *Robust statistics*. John Wiley & Sons, Chichester. ISBN.

Nadeau, R., R.G. Niemi, and J. Levine. 1993. "Innumeracy about minority populations". *Public Opinion Quarterly* 57 (3): 332–347.

Schwarz, Norbert. 1995. "What respondents learn from questionnaires: The survey interview and the logic of conversation". *International Statistical Review/Revue Internationale de Statistique*: 153–168.

Schwarz, Norbert, and Seymour Sudman, eds. 1992. *Context effects in social and psychological research*. Springer Science & Business Media.

Strack, Fritz, Leonard L Martin, and Norbert Schwarz. 1988. "Priming and communication: Social determinants of information use in judgments of life satisfaction". *European journal of social psychology* 18 (5): 429–442.

Tourangeau, Roger, and Kenneth A Rasinski. 1988. "Cognitive processes underlying context effects in attitude measurement." *Psychological bulletin* 103 (3): 299.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. Fourth. ISBN 0-387-95457-0. New York: Springer.

Wilson, Timothy D, et al. 1996. "A new look at anchoring effects: basic anchoring and its antecedents." *Journal of Experimental Psychology: General* 125 (4): 387.

Wong, Cara. 2010. *Boundaries of Obligation in American Politics: Geographic, National, and Racial Communities*. Cambridge University Press.

— . 2007. "" Little" and" Big" Pictures in Our Heads: Race, Local Context, and Innumeracy about Racial Groups in the United States". *Public Opinion Quarterly* 71 (3): 392.

Wong, Cara, et al. 2012. "Bringing the Person Back In: Boundaries, Perceptions, and the Measurement of Racial Context". *Journal of Politics* 74, no. 4 ().

Zaller, John, and Stanley Feldman. 1992. "A simple theory of the survey response: Answering questions versus revealing preferences". *American journal of political science*: 579–616.

Appendix A Information about the Sampling Frame and Sample

Our sampling frame came from Vote Compass, a website sponsored by the Canadian Broadcasting Corporation (CBC). The Vote Compass website allowed visitors to answer about twenty policy questions and place themselves in a policy space relative to the major political parties (which had also completed the survey). Over 1 million Canadians participated in the Vote Compass project surrounding the May 2011 federal election. We emailed all of the 80,000 or so respondents of the 2011 Vote Compass who agreed to be contacted for future academic studies, and about 10 percent agreed to take our survey; 7817 respondents completed the first survey, and 3076 completed the second survey. The different AAPOR response rates range between 11 and 15 percent for wave 1, and 43 and 50 percent for wave 2.

This sample is not representative of Canada as a whole, since respondents in the sampling frame had to have been interested in the Vote Compass website and also interested in further research on politics using the web. Compared to the 2006 Census, our respondents are older, better educated, wealthier, and more likely to be men and white. Furthermore, our online survey was only conducted in English (whereas Vote Compass was conducted in both English and French). Nevertheless, we do have respondents living across many parts of Canada, including both urban and rural areas, and in all 13 provinces and territories. They live in CSDs with populations ranging from 75 to 2,477,000; 25 percent of all respondents lived in CSDs with 19,080 residents or fewer, half lived in CSDs with 126,800 residents or fewer, and the mean number of residents was 523,000.

Appendix B Details on the Mapping Measure of Context Boundaries

Appendix B.1 Instructions

The instructions for our measure of subjective context were fairly long because map drawing is a new task for survey respondents and we discovered a need for more instruction during pilot testing. The following are the exact instructions given:[19]

> Please draw what you think of as your Local Community on the map.
>
> 1. Click "Start Drawing" to begin drawing your community.
> 2. Hold down the mouse button to make a line. Open shapes will be enclosed if you lift your mouse for a few seconds.
> 3. You may make as many shapes as you like.
> 4. Even though we have centered the map on your postal code, you should feel free to zoom in or out and to move the map if it does not contain the areas that you consider your local community.
> 5. If you need to zoom in or out, or move the map, after you have begun drawing, click Stop Drawing, adjust the map, and then click Start Drawing to finish drawing your community.
> 6. If you left click on a completed shape you will have the chance to delete that community.
> 7. You may Reset Map to delete all of your shapes to start over.

[19]Note: The map functions were also designed to allow drawing using track pads even though we refer to a "mouse" in the instructions.

Appendix B.2 Map Sizes

Figure 4 shows the distribution in the map sizes, and this includes maps with multiple polygons. The median size community is 12 square kilometers. The right panel of the figure shows that only about 10% drew maps that were greater than 250 square kilometers. It also shows that most people drew maps that were smaller than about 50 square kilometers, and 20 percent drew maps that were smaller than 1 square kilometer.

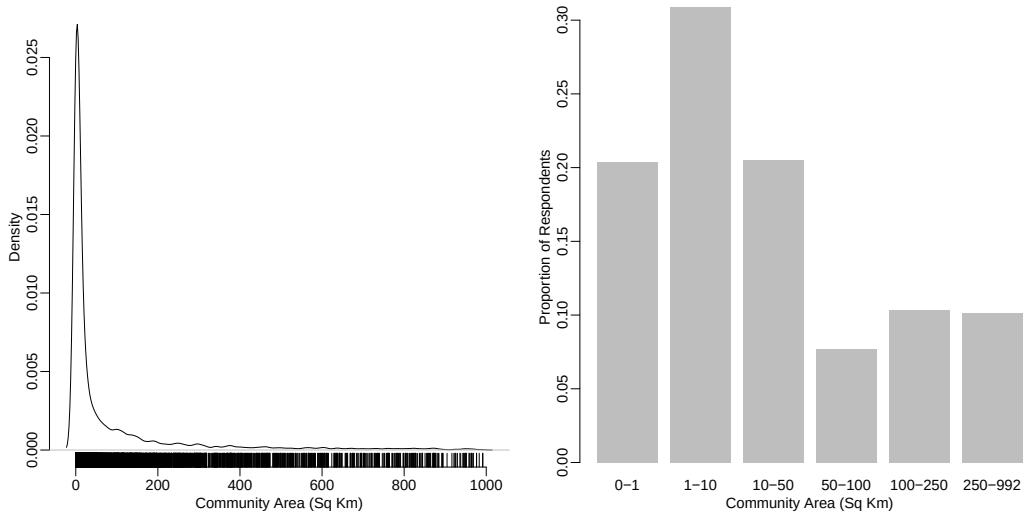


Figure 4: The distribution of sizes of maps drawn.

Appendix B.3 Map Drawing Time

Figure 5 shows how long people spent on the drawing task. If respondents were not able to complete the map-drawing task in a reasonable period, we would worry about the quality of the information gathered. Some time was devoted to reading the instructions and considering what constitutes their communities, but we also assume that even for younger, more tech-savvy respondents, drawing on a map may include some trial and error; therefore, we allowed respondents to delete maps drawn and start over again. For the initial survey, the median response time was 2.2 minutes for this map-drawing measure, indicating that it was neither a simple nor a Herculean task for respondents. The right panel shows that few people rushed through the task (completing it in less than 30 seconds), but that also few people spent more than 480/60=8 minutes on the task.

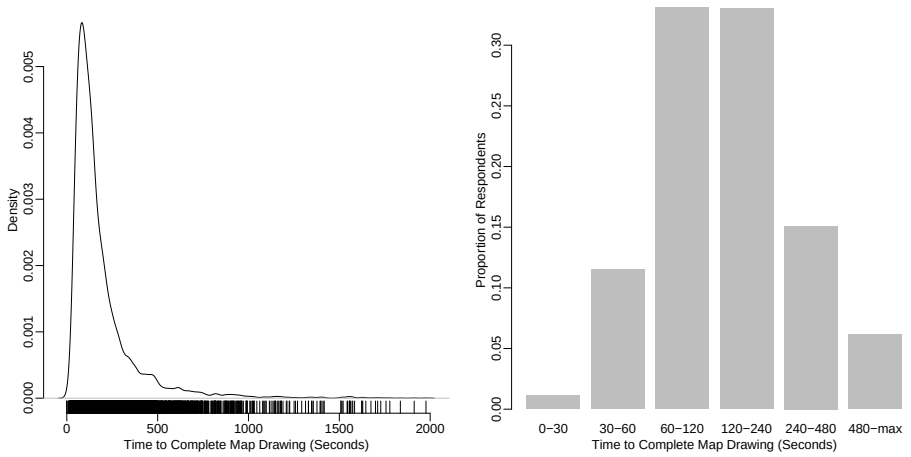


Figure 5: The distribution of the time respondents took to draw their "local communities."

By Wave 2, the respondents had learned how to answer the community measure, leading to much shorter response times. The median response time for drawing their local community in the follow-up survey was 1.2 minutes (compared to 2.2 minutes in Wave 1). About 10 percent included multiple polygons in their self-drawn communities, compared to 16 percent that were non-contiguous in the initial survey.

Appendix C Validity Assessment

Appendix C.1 Convergent Validity Tests: Commuting Distance and Population Size of CSD

Commute Distance and Map Size

We expected that longer commutes would be associated with larger maps. In Figure 6 we see that this is so: the coefficient for the robust linear model in the left plot is .024 ($t = 2.24$) (shown by the dotted line), and the coefficient for the same model but on the log scale on the right panel is .366 ($t = 7.83$).

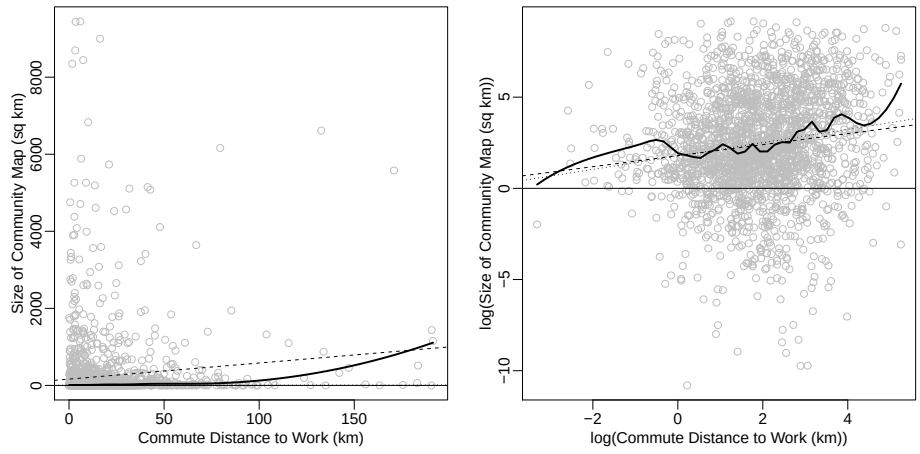


Figure 6: Longer commutes predict larger maps (length of commute measured as distance between centroids of home postal code areas and workplaces). In each panel we plot an OLS model (dashed line), the robust regression (dotted line), a smoothed nonparametric loess model (dark black solid line), and a zero-line (thin horizontal black line). The right panel uses a log scale to reveal details among the majority with short commutes and small maps.

Population Size/Density and Map Size

We expected that living in places with a larger population (like living in cities) would be associated with smaller maps. In Figure 7 we see that this is so: The coefficient for the robust linear model in the left panel is .00 ($t = -4.14$), and the coefficient for the model in the right panel is -.359 ($t = -22.55$).

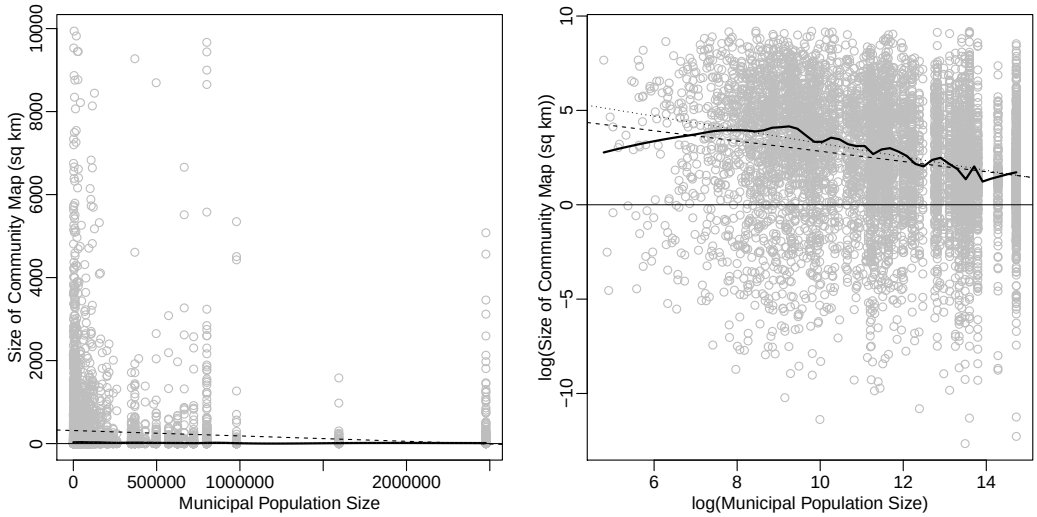


Figure 7: Population size and density (Canadian Census Subdivision or municipal population size) predicts smaller maps. In each panel we plot an OLS model (dashed line), the robust regression (dotted line), a smoothed nonparametric loess model (dark black solid line), and a zero-line (thin horizontal black line). The right panel uses a log scale to reveal details among the majority with short commutes and small maps.

Cognitive Bias of Anchoring: The Effect of Initial Resolution of the Map

Did the size of the initial map influence or anchor the size of the community drawn? While the definition of someone's local community may be influenced by a variety of factors that should differ across individuals and time, its meaning (and political relevance) should not vary depending on the resolution of a map shown. If the context of the survey itself strongly influences what people conceptualize and draw as their community, we should not see it relate strongly to attitudes outside of this context (e.g., policy preferences). It also would be very unlikely that respondents would draw similar communities over time, since the "context" was contingent on the survey context.

There were 5 different resolutions that respondents could have been shown at the start of the question about their local community, ranging between Google maps zoom-level 10 and 17.[20] In the end, 14 different resolutions were chosen by the respondents for their communities; only 8 percent of respondents' communities were drawn at a map resolution greater than any of the ones shown initially (e.g., a resolution zoomed out to the North American continent), and the rest fell within the range of resolutions initially shown. About a third (34 percent) of respondents did not move from the initial randomly assigned resolution, with less movement from those shown a greater area—like one's metropolitan area—and more zooming out from the smaller areas. Regardless of whether respondents

[20]For a point of comparison, at resolution 10, 1 screen centimeter is 5.4 kilometers. At resolution 17, 1 screen centimeter is 42.3 meters. Or, at 96 dpi, the scale for resolution 10 is about 1: 500,000, and for resolution 17 it is about 1:4000. (See `http://wiki.openstreetmap.org/wiki/Zoom_levels` for more details.) Another way to think of the scale is that zoom level 0 is the whole world, 10 is a city, level 16 is a small road, and 19 a building. The scale is roughly doubled as the zoom level increases by one step.

changed the zoom level of the map, they could still move the center of the map.

Figure 8 and table 2 show the relationship between size of map drawn and this zoom-resolution randomization. We can see that variance of the map sizes differs even if the medians do not appear to differ according to zoom level.

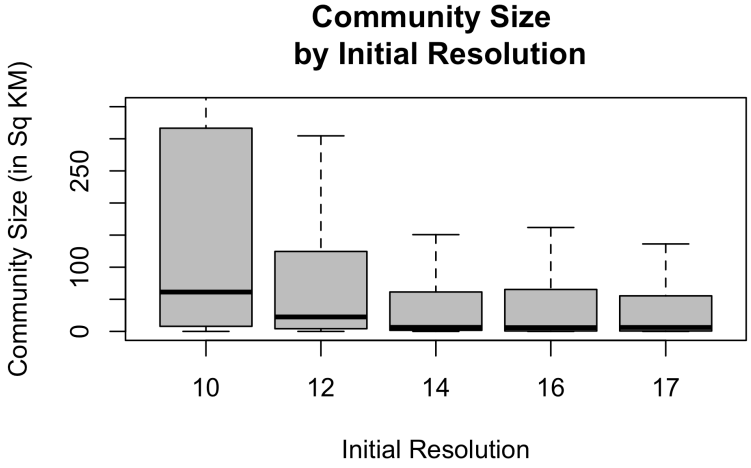


Figure 8: The Potential Effects of Anchoring: Box plot of the size of "Local Communities" drawn, by initial resolution shown: 10 represents a map showing the largest area, comparable to a metropolitan area.

The typical map size differed by zoom-level, but only for the larger (province and country-level) maps as compared to the very zoomed-in maps. For example, among those seeing the country level zoom, their maps were on average about 5 sq km larger than those seeing the smallest zoom, below the neighborhood level. These differences are statistically detectable although they are substantively small (5 sq km differences).

Table 2: Effect of Zoom Level of Community Size (Robust Regression Coefficients)

	Coefficient	T value
Zoom Level 10	5.18	10.45
Zoom Level 12	4.91	9.88
Zoom Level 14	0.37	0.75
Zoom Level 16	-0.36	-0.70
Intercept	6.11	17.24

Appendix C.2 Cognitive Bias of Priming: The Effect of Geographic Priming

Four-fifths of the respondents were primed to think of one particular context with the following 2 questions asked just before the map drawing task at the start of the first survey: "Thinking about where you live, how close do you feel to your [neighbourhood/ city/ province/ Canada]?" And, "If you could improve your work or living conditions, how willing or unwilling would you be to move to another [neighbourhood/ city/ province/ Canada]?" A control group received no treatment and proceeded from a question about home ownership directly to the map-drawing.[21]

[21] We did not assess policy-issue based geographic priming because we worried that the geographic cue would not be as clear for respondents and therefore could be difficult for us to interpret. For example, would the average respondent know that health care policy is decided at the federal level,

While there is no large substantive effect of geographic primes on typical drawing size when we regress local community size on the treatments using a robust linear model, the box plots in figure 9 make clear that priming does affect the variance of the size of communities drawn. We also see in table 3 that respondents who received the "neighbourhood" prime did draw communities that were statistically distinguishable from the control, but they were only 1.6 square kilometers smaller on average (again using a robust regression to handle the out-sized influence of a few very large maps).

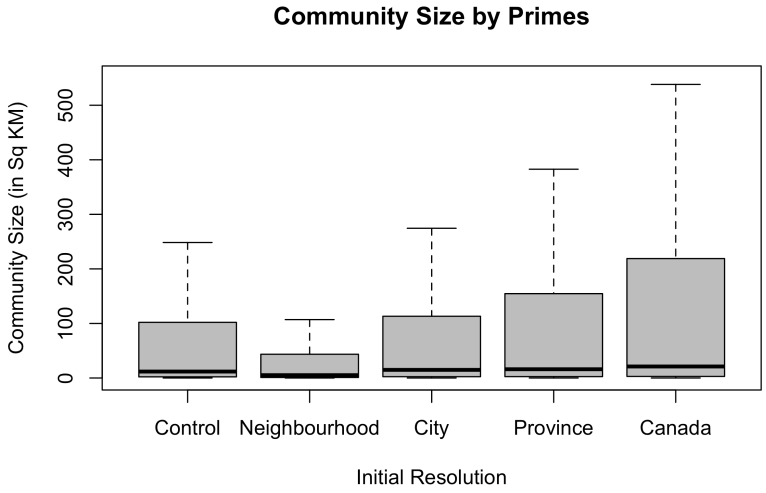


Figure 9: The Potential Effects of Priming: Box plot of the size of "Local Communities" drawn. Respondents were randomized either to be in the control group or receive one of four geographic primes that made salient their neighborhood, city, province, or country.

Table 3: Effect of Zoom Level of Community Size (Robust Regression Coefficients)

	Coefficient	T value
Canada	0.25	0.49
Province	0.04	0.09
City	0.76	1.50
Neighbourhood	-1.61	-3.18
Intercept	8.02	22.37

In other analyses, not shown here, but in the reproduction materials associated with this article, we find no detectable interactions between anchoring and priming treatments.

Appendix C.3 Limitations of the validity assessment: Criterion Validity

In theory, the introduction of a new measure should show that the measure relates to another measure of the concept in a different context. For example, to assess the validity of a math test, we might observe whether that test predicts the ability to calculate a tip at a restaurant. This form of

even if they are thinking about their own local doctor? Or, would they know federal election districts do not decide whether there should be multilingual ballots, even if they are the most salient geography during an election? In contrast, the question we used is based on an item included in the General Social Survey, and Wong (2010) has shown that the different variables predict different policy-related responses.

validity, often called criterion validity, is especially convincing: the pressure gauge on the car's tires should not show high pressure when the tires look flat to the eye, and reading scores should predict comprehension and speed of reading outside the test. In the field of public opinion, this kind of validity test tends to be ignored because, after all, the study of opinion at a mass level implies a trade-off between close physical observation of what a person says during political discussions with friends and their score on an ideology scale, or their score on a stereotyping measure and their everyday interactions with members of different groups. Our measure is in the same boat as the rest of the measures of public opinion, although one could imagine leveraging advances in technology beyond Google maps to track actual movement through space and/or interactions with people. At this stage in the research program, however, our evidence in favor of the validity of this measure is of the same standard as that used to validate other measures of public opinion.

Appendix D Content of Contexts: Variables and Coding Decisions

Appendix D.1 Objective Ethnic Context: Census Data

To create our measure of objective ethnic context, we use data from the following two Census items:

1. Is this person an Aboriginal person, that is, First Nations (North American Indian), Metis or Inuk (Inuit)?

2. Is this person: White, South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.), Chinese, Black, Filipino, Latin American, Arab, Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, Laotian, etc.), West Asian (e.g., Iranian, Afghan, etc.), Korean, Japanese, Other (specify).[22]

We measure context across five different administrative levels: Dissemination Areas (DAs), Census Subdivisions (CSDs), Forward Sortation Areas (FSAs), Federal Election Districts (FEDs), and Provinces/Territories. DAs and CSDs are both Census units: DAs are composed of one or more blocks with 400 to 700 individuals and are the smallest Census unit for which all information is disseminated, and CSDs represent cities, towns, and municipalities. FSAs are contextual units where the first three digits of the postal code are the same (and there were roughly 1600 in Canada in 2006). There are about 300 FEDs (with 60,000 to 120,000 people in each), ten provinces, and three territories in Canada. In our survey, we have 6370 DAs, 1158 FSAs, 984 CSDs, 289 FEDs, and 13 provinces/territories represented.

In order to calculate the objective ethnic context within the boundaries of a respondent's hand-drawn "local community," we tried multiple approaches. In the end, none seemed superior to the one we used in the paper, which averages the context of any DA that is included in their drawing. For example, if a "local community" was drawn to overlap with three different DAs, then the "objective" point of comparison for racial context used is the average of the Census-reported percentages of VM for those three DAs. Using only DAs that were entirely enclosed within a map is problematic because some maps were drawn so small as to not include a single entire DA. We also considered calculating the fraction of the area of a DA contained within a map and including only that proportion in the demographics; however, this approach assumed that individuals are evenly distributed across a DA,

[22] Respondents are also given the following information: "Mark more than one or specify, if applicable. This information is collected in accordance with the Employment Equity Act and its Regulations and Guidelines to support programs that promote equal opportunity for everyone to share in the social, cultural, and economic life of Canada."

which is obviously false. The approach we present here probably overstates the visible minorities one would find if household level data were available. As such it attenuates the evidence of misperception shown below. Population density information within a DA is not publicly or readily available and its precision and reliability (usually based on satellite photos of night light or on the ground surveys of wifi hotspots) varies greatly once one is outside of a city. For those reasons, as well as cost, we chose to work with errors of interpolation and extrapolation with a known direction of bias and also known precision and reliability from the Canadian Census.

We had originally planned to look at both the effect of change in diversity as well as the level. However, in 2011, the long-form of the Census—which is where Canadians are asked about their ethnicity and race—became voluntary in the newly renamed National Household Survey (Thompson 2010). In contrast to a 94 percent response rate to the mandatory 2006 form, Statistics Canada reported a 69 percent response rate for the 2011 data released in May 2013. While we had originally planned to use the more recent data and also examine the effects of *changes* in diversity in locales, as the former Chief Statistician of Canada Sheikh explains, "Given the magnitude of change from the 2006 census, it is not clear whether the NHS data reflect a real change in outcomes or simply a statistical artifact due to the change in methodology" (Sheikh 2013). What is particularly challenging for our research is that the data are most compromised and problematic for smaller geographic levels. In 2016, the Census reinstated its mandatory long form, and the response rate went back to about 98 percent; however, the data will not be available until late 2017 at the earliest.

Appendix D.2 Subjective Ethnic Context: Perceptions of Visible Minorities

Respondents answered questions about the composition of their subject maps using a slider, so any response between 0 and 100 was possible. We chose to use sliders instead of pie charts because past research has shown respondents have greater difficulty in calculating angles and areas in pie charts, and are more accurate in discriminating proportions when aligning positions on a common scale (Hollands and Spence 2001; Schonlau and Peters 2012). Because people commonly overestimate the size of these groups, the index often exceeded 100 percent.[23] Thus, we have subjective perception measures for the hand drawn maps for all of the respondents and also for each of six administrative contextual units for about 1/6 of our sample. Using previously available data, it is surprisingly difficult to test the accuracy of people's perceptions of their contexts at anything besides the national level. There are multiple surveys that ask respondents to describe the demographic make-up of their country. However, for most sub-national perceptions, the boundaries of the contexts are often unclear. For example, surveys sometimes ask respondents to describe their local community or neighborhood (for example in the GSS), but past research has shown that both terms vary a great deal in people's minds, even between neighbors who are engaged in community activism. In our survey, we not only ask about fixed administrative units common across multiple respondents, we also show our respondents administrative maps of these areas; we believe most people do not know the borders of their federal election district, for example, so the maps at least provide a common picture across respondents that they can then describe.

Appendix D.3 Recalculating Perceptions Using Proportions

We use the index of estimates of the different visible minority groups as our measure of subjective context. However, it is possible that a respondent thinks that groups are roughly the same size and therefore reports similar percentages, without paying attention to whether they sum to 100 percent.

[23] We created a few versions of this index—some restricted to go to 100, others merely summing the proportions reported—and we saw that our results are robust to different specifications.

Therefore, we also checked our results, recalculating perceptions as proportions. As an example of what we mean by proportion, someone who says that Canada is 50 percent white, 50 percent black, and 50 percent Asian would be coded to have the perception that Canada is roughly 67 percent VM, not 100 percent. We find similar results.

Appendix D.4 Innumeracy

Of course, if people provide answers that sum to more than 100, it is possible that ordinary citizens do not have a sophisticated understanding of percentages (Paulos 1988). The gap between objective and subjective context, in that case, could result from a simple lack of mathematical understanding. In order to determine whether this is the case (and test the validity and reliability of our measure of perceived context), the first wave of the survey included multiple measures of respondents' perceptions of their self-defined communities. In addition to being asked to use 0 to 100 sliders to report the percentage of various racial/ethnic groups, for example, living in their "local community," respondents were asked (later in the survey) to look again at the "local community" map they drew and say whether the place as drawn was "mostly white," "mostly racial and ethnic minorities' (after which they were asked to explain which groups),' "about half and half," or "some other mixture" (after which they were asked to explain what they mean).

In general, it seems that people have a pretty good understanding of percentages. Among those respondents who said their local community was "mostly racial and ethnic minorities," 74 percent gave estimates of different visible minority (VM) groups that totalled over 60 percent VM (only 11 percent gave estimates that totalled less than 40 percent VM). Of those respondents who described their communities as "mostly white," 69 percent gave estimates for separate VM groups that together totalled less than 40 percent. As one might expect, the "about half and half" response was more vague, and responses showed a greater variance as a result: of those respondents who chose this description of their local community, more than a half had given estimates of different VM groups that totalled over 60 percent, and about a third have estimates of groups that totalled between 40 and 60 percent VM.

Another way that we can account for possible innumeracy is by recalculating perceptions proportionally, as we mentioned in the previous section in the appendix. After all, if people understand that 50 percent is bigger than 30 percent, but not that percentages should add up to 100, our proportional measure of perceptions gives us a measure of subjective context that is less sensitive to these limitations arising from innumeracy. Using this proportional measure, we can also compare the responses of the same people about perceived VM in Canada in survey 1 (which asked about perceptions of each ethnic group separately) and the subsequent survey 2 (which asked respondents a single question about the percentage of visible minorities in Canada). Figure 10 shows the relationship between the perceived VM index from survey 1 and the perceived VM estimate in survey 2. As one might expect, the index led to higher numbers (and greater misperceptions), although there is also a clear relationship between the estimates.

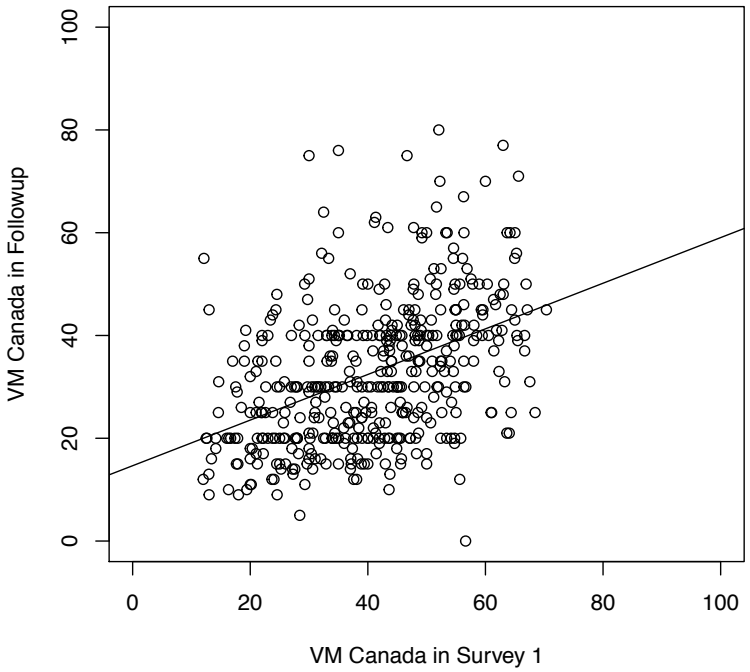


Figure 10: Scatter Plot of Perception of Visible Minorities at Time 1 (index) and Time 2 (single response) with OLS line.

Appendix D.5 Other Cognitive Biases

Another source for overestimation of minority groups is cognitive biases, rather than innumeracy. Burns, Chiu, and Wu (2010) explain that rare events tend to be overestimated because of the availability heuristic and anchoring (Tversky and Kahneman 1974); this would help explain why asking about multiple VM groups separately lead to greater misperceptions cumulatively than when asking about VM as a single group. This cognitive bias can even cycle, such that underestimation follows overestimation one or more times (Hollands and Dyre 2000). There is, of course, an altogether different explanation for over- or under-estimates of the size of groups, which is based in fear or threat. Herda (2010), for example, shows perceived threat from a group is related to innumeracy (i.e., overestimation), confirming previous research (Nadeau, Niemi, and Levine 1993; Gallagher 2003). However, that is a task for future research; in this paper, we are not focused on who in particular misperceives. We only want to address the question of whether the task of answering these questions about where respondents live is too difficult to ask in a survey, and therefore we only focus on innumeracy here. As we have shown, although we have followed in the tradition of survey research by asking respondents to engage in novel cognitive and social tasks, the respondents learned the tasks, and behaved as if the tasks were meaningful when we asked them to talk about the contents of their maps as "local communities".

References

Burns, Zach, Andrew Chiu, and George Wu. 2010. "Overweighting of small probabilities". *Wiley Encyclopedia of Operations Research and Management Science.*

Gallagher, Charles A. 2003. "Miscounting race: Explaining Whites' misperceptions of racial group size". *Sociological Perspectives* 46:381–396.

Herda, Daniel. 2010. "How many immigrants? Foreign-born population innumeracy in Europe". *Public opinion quarterly*: 674–695.

Hollands, JG, and Brian P Dyre. 2000. "Bias in proportion judgments: The cyclical power model". *Psychological Review* 107:500.

Hollands, J.G, and Ian Spence. 2001. "The Discrimination of Graphical Elements". *Applied Cognitive Psychology* 15:413–431.

Nadeau, R., R.G. Niemi, and J. Levine. 1993. "Innumeracy about minority populations". *Public Opinion Quarterly* 57 (3): 332–347.

Paulos, John Allen. 1988. *Innumeracy: Mathematical illiteracy and its consequences*. Macmillan.

Schonlau, Matthias, and Ellen Peters. 2012. "Comprehension of Graphs and Tables Depends on the Task". *Statistics, Poltiics and Policy* 3.

Sheikh, Munir. 2013. "Canada Has Lost Its Census Anchor". *The Globe and Mail* (). `https://goo.gl/KOd9Zh`.

Thompson, D. 2010. "The politics of the census: lessons from abroad". *Canadian Public Policy* 36 (3): 377–382.

Tversky, A., and D. Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases". *Science* 185 (4157): 1124–1131.

Wong, Cara. 2010. *Boundaries of Obligation in American Politics: Geographic, National, and Racial Communities*. Cambridge University Press.