The Origins and Direction of Measurement Error in the "Vote Compass" on the CBC

Christopher Cochrane*
The University of Toronto

v.1.2 April 25, 2011†

Abstract

This methodological note advances two arguments. The first argument is that random error in the measure of the two underlying dimensions in the Vote Compass is converted through the process of averaging into systematic error in the measure of a user's position relative to the political parties on the two-dimensional plane. This argument can be demonstrated via simulation. The second argument is that the two-dimensional Vote Compass graphic is likely to "mis-classify" many users in terms of their actual proximity to political parties on the questions that the Vote Compass asks about. By "mis-classify" I mean that the two-dimensional graphic generated by the Vote Compass indicates to some users that they are "closest" to one political party when in fact there was some other political party with which these users agreed more often and more closely on the issues that the Vote Compass asked about. The possibility persists whether the users answer the questions randomly, ideologically, or some combination of both. My substantive conclusion is that the Vote Compass' two-dimensional graphic is misleading for many users.

*Christopher Cochrane, Assistant Professor, Department of Political Science, University of Toronto / Department of Social Sciences, University of Toronto Scarborough. This is a working paper. Please forward comments and corrections to christopher.cochrane@utoronto.ca. I am grateful to the creators of the Kieskompas, particularly André Krouwel, Matthew Wall, and Jan Kleinnijenhuis, who proposed a number of criticisms to an earlier draft of this document, which I have tried to integrate. I am grateful to my colleagues on the Vote Compass team, including Cliff van der Linden, Peter Loewen, Yannick Dufresne, Gregory Eady, and members of the advisory team, who have actively encouraged an open and public discussion of these concerns, and have addressed them in the current iteration of the Vote Compass. Although the views expressed in this document have been enhanced considerably by my conversations with colleagues, errors or omissions are mine alone.

†This version alters the title, adjusts the syntax somewhat to take account of a more specific understanding of the VC algorithm, and updates Figure 3 and the discussion to reflect a minor change in those results.

1 Introduction

The wildly popular Vote Compass, which is promoted by CBC and has attracted more than 1.5 million completions, purports to show its users "whose views are most like yours?," or, more recently, "It's election time! Take a few minutes to find out if your party is right for you." But what does it actually calculate? This methodological note advances two arguments. The first argument is that the Vote Compass algorithm converts, through the process of "averaging," random error in the measure of the underlying dimensions into systematic error in the measure of the distance between the user and the parties on these dimensions. This systematic error works massively in the direction of the Liberal party. All of the other parties are negatively affected by this, but especially the NDP. The second argument is that the Vote Compass graphic is likely to show a good number of users that they are "closest" to a political party when in fact they agree more often and more closely on these issues with some other political party.[1] In other words, the two-dimensional picture is "incorrectly" placing some users vis-a-vis their positions relative to the political parties on these issues. By "incorrectly" I mean simply that the picture shows a user that she is "closer" to Party A than Party B even though the Vote Compass itself records that she agrees less often and less intensely on these issues with Party A than with Party B. This note will demonstrate this possibility in the Vote Compass algorithm. My substantive conclusion is that the two-dimensional image distorts the information that the Vote Compass provides, and, as a result, it is likely to mislead many Vote Compass users.

2 Theory

The Vote Compass algorithm, whose basic parameters can be derived inductively by experimenting with the instrument, calculates the distance between a user and a party as the distance between, in effect, the average position of the user and the average position of the party on two "dimensions" - an "economic dimension" and a "social dimension" - dimensions which are each comprised of multiple loosely connected issues.[2]. This is a key point, because the distance of the averages is not the same as the average of the distances. A more effective way of calculating distance between a user and a party on these specific issues is to calculate the average distance between the user the party across the entire range of issues (Markus and Converse, 1979; Adams et al., 2005). For example, in a standard utility model, the distance (inverse utility) between voter i with positions x and party p with positions s across issues $j_{1...30}$ is

[1] The "Vote Compass graphic" refers to the two-dimensional Cartesian plane generated for the user after the user answers the items in the questionnaire.

[2] The Vote Compass assigns a score of -2 to the most "left-wing" answers and +2 to the most "right-wing" answers, and it then sums the positions of the users and the parties across the range of questions and compares the results

$$D_i(p) = \sum_{j=1}^{30} \sqrt{(x_{ij} - s_{pj})^2} \quad (1)$$

In other words, distance is the sum of the absolute value of the differences between the party and the user across the range of issues.[3] Thus, if the NDP, Liberals and Conservatives are positioned respectively at locations -2,0, and 2 on issues A and B, and if a voter is at position -2 on issue A and at position 2 on issue B, then their "average" position across both issues is "0", the Liberal position, but the voter's actual distance from the political parties, using Equation 1, is 4 for the NDP (0+4), 4 for the Liberals (2+2), 4 for the Conservatives (4+0). The voter would not be closest to the "centrist" party; they would be equally close to all of the parties.

3 Distance in the Vote Compass and the Utility Algorithms

The Vote Compass' algorithm does not allow for this possibility in the two-dimensional picture. In this picture, a party with a position that is extremely to the left of the user on half of the issues and extremely to the right of the user on the other issues would be scored the same as a party that was exactly on the user's position on all issues. In fact, this first party with extreme right-left positions vis-a-vis the user would be scored more closely than a party that was say, one point left of the user on all issues, even though that latter party would be closer to the user on average, and, indeed, on every single issue. This point is illustrated in Figure 1.

In Figure 1, Voter 1 (v_1) and Voter 2 (v_2) are positioned relative to party P on two issues. The standard utility model in Equation 1 would situate v_2 closer to Party P on these issues. Compared to v_1, v_2 is much closer to party P on both issues. Indeed, v_1 is in both cases all the way over on the opposite side of Party P, and v_2 is right beside Party P. In the Vote Compass algorithm, however, v_1 would be considered to be closer than v_2 to party P. The reason for this is that Vote Compass compares the distance of the averages. Thus, when positions on these issues are averaged, v_2 is slightly to the left of center, whereas both v_1 and party P are exactly in the center. According to Vote Compass, v_1 is closer to party P in Figure 1 than is v_2. This extreme scenario illustrates a basic point: the Vote Compass may show a user that she is "closest" to one political

[3] A term for issue salience is easily integrated into this model (e.g., Adams, Merrill and Groffman 2005, p.17). If z represents the salience of issue j for voter i, then the utility model, integrating issue salience, is: $D_i(p) = \sum_{j=1}^{30} z_{ij} \sqrt{(x_{ij} - s_{pj})^2}$

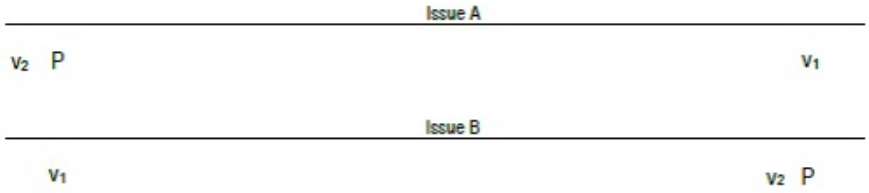

Figure 1: A Two-Voter/One-Party Hypothetical on Issues A and B

party when in fact she disagrees more often and more intensely with precisely that party than with any other party on the issues that the Vote Compass asks about.

4 The Distortion of Random Error in the Vote Compass Algorithm

Another issue to consider is that comparing the averages across dimensions that are each constructed by a number of questions will render onto the center the random error, the statistical unreliability, in the measures of the underlying dimensions–the amount of error in this context is reflected in the degree to which the items in the dimensions are uncorrelated with each other. The extent to which this error averages to the center is proportional to the number of items in the measures and inversely proportional to the magnitude of the correlation between the items in the measure. Even very high correlations between items -much higher than the average inter-item correlations which would emerge between the items in each of the Vote Compass dimensions-will move the results toward the center. If the Vote Compass were simply measuring the underlying dimension, this would not be a major problem. The random error would simply reduce the accuracy of our predictions about how a user or a party would answer some other question on this dimension. But the Vote Compass is not simply measuring an underlying dimension. Its principal output is a measure of the distance between a user and party on that dimension. Thus, the fact that all of the random error in the measure of the underlying dimension is concentrated in the center when the positions on this dimension are averaged means that the random error in the measure of the dimension becomes systematic error in the measure of the distance between a user and a party on that dimension–it is "systematic" or "directional" error because it points in one direction: in this case, that direction happens to be Liberal. This is not because these users agree most with the Liberal Party (see above), rather, it is simply because *when averaging* responses that are weakly or not at all correlated to each other, there are many, many times more ways to generate a Liberal average randomly than the Conservative, Green, BQ or NDP average.

4.1 A Simulation

One way to assess the extent to which the algorithm shapes the results is to assume that a number of users responded to each question randomly. This allows us, in effect, to trace whether and how the error may be distorted in any particular direction. There are three ways in which a party may still come out ahead when respondents answer randomly. First, the party may be centrist on the largest number of issues, and thus the users can never really get "that far away" from the party on any issue. As a result of this centrist position, the party enjoys a strategic advantage vis-a-vis other parties. All things being equal, random respondents would be more likely to end up close to the centrist party than they would with parties at the extremes.

Yet a second way that a party can enjoy a strategic advantage is if the party occupies space that is unoccupied by other parties. Thus, this party picks up a higher proportion of the respondents who happen by random chance to wind up close to that party. In effect, there are no other parties competing for the users who end up in this part of the opinion distribution. Imagine a scenario with two left-wing parties and one right-wing party. By random chance alone, we would expect equal numbers of users to wind up on the left and the right. Yet, those who end up on the left are split between two parties and those who end up on the right are concentrated in one party. In this way, opinion space that is unoccupied by other parties provides a strategic benefit for a political party.

These scenarios are not inherently distortions or biases – they are genuine strategic advantages for political parties. One way that Vote Compass could shape the results toward a political party, therefore, would be to position a particular party in favorable strategic space. It is this decision that would be taken up in debates about whether the Liberals should be centrist on so many issues, and so on. But the Liberals do not actually enjoy this kind of positional advantage in the Vote Compass. Figure 3 plots the positions of 10000 simulated respondents vis-a-vis the actual positions of the political parties in the Vote Compass. Distance, in this case, is calculated using the typical formula for distance in Equation 1. Again, the positions of the users are simulated and random, but the positions of the parties are the actual positions, as recorded by Vote Compass and available on the CBC website.

Notice in Figure 2 that all of the parties do reasonably well, save for the NDP and the Greens who occupy similar ideological space. Yet, no single party enjoys a huge strategic advantage in terms of how Vote Compass positions the parties relative to each other on the issues. The party with the greatest situational disadvantage is the Green Party, in part because it so rarely occupies a centrist position on the issues (it is the party that is on average the farthest away from the center, according to Vote Compass), and in part because it has so much competition in the opinion space that it occupies. The party with the greatest situational advantage is Liberals, but only slightly. In the Vote Compass coding, the Liberals are closest to the center on average, but only slightly more so than the BQ and the Conservatives.

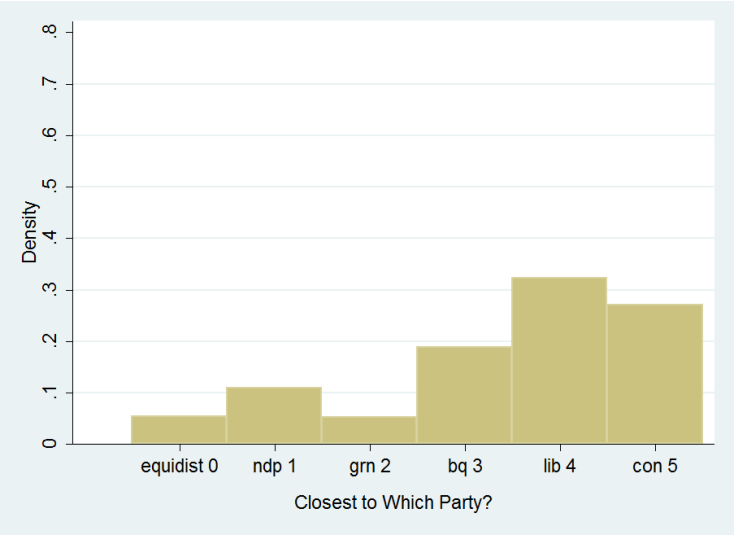

Figure 2: Proportion of Random Distribution of Simulated Users that is Closest to Each Party

The distortion of the Vote-Compass toward the Liberal Party therefore has little to do with the way that the parties are positioned on these issues – the Liberal positions offer some strategic advantage vis-a-vis the other parties, but it is not a major advantage. The problem, rather, stems from the "averaging" of the party positions on each "dimension." Figure 3 illustrates this effect. When positions are averaged, the distribution of voters across the parties in Figure 2 is changed into something entirely different: fully 87 percent of the simulated respondents wind up as Liberals, even though only 32 percent of the users agree most with the Liberal Party across this range of issues. The simple fact of averaging increases nearly three-fold the proportion of the simulated random respondents who are "closest" to the Liberal Party. This is the origin of systematic error in the Vote Compass. It is not "bias" in the intentional sense of partisan manipulation; it is systematic error in the statistical sense of directional error due to averaging.

As a result of the averaging in Vote Compass, the Vote Compass reports that a number of users are "closest" to a political party, when in fact there is one or many other parties with which that user agrees more often and more closely across these 30 issues. Figure 4 displays the results from the above simulation. More than 60 percent of the simulated users were classified by the Vote Compass algorithm as being closest to one political party when in fact there was at least one other party with which that user agreed more closely across these issues. Of those who were misclassified in terms of their proximity to the parties on these 30 issues, about 6 percent ended up Conservative, and 91 percent ended

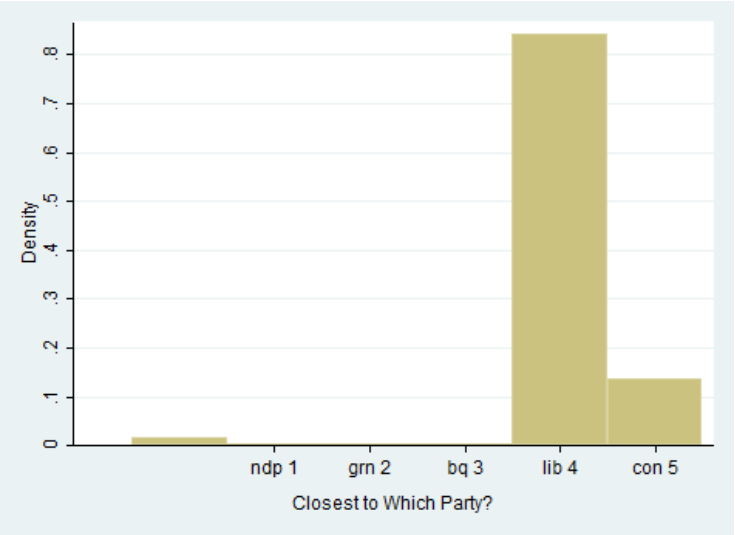

Figure 3: Proportion of Random Distribution that is "Closest" to Each Party According to the Vote Compass Algorithm

up Liberal. These are not actual user data-which I have not seen and which are not publicly available-but this is what we would expect by chance alone.

5 Discussion and Conclusion

This is a simulation. Again, these are not actual user-data from the Vote Compass, but they are the actual party positions that the Vote Compass uses. Thus, following the path of random respondents is a useful way to identify where and to what extent a statistical distortion occurs. Of course, Vote Compass respondents do not answer entirely at random – for many users, at least some of their opinions will be structured on each dimension by partisan loyalties or ideologies. But this is not the point. Indeed, even if these questions were perfect measures of a single underlying ideological dimension (which they are not), and even if every voter was perfectly constrained by left/right ideology and answered every single question in exactly the same category (which they do not), the Vote Compass would still improperly classify many of its respondents.

In any case, from the perspective of the underlying latent dimensions that Vote Compass uses and on which the averaging takes place, user responses are "random" vis-a-vis these dimensions when they answer some of the questions on a dimension in a manner that has nothing to do with how they answered the other questions on that dimension. That random element, which is certainly plentiful on these dimensions, is by definition equally likely to end up in any response category for any single user on any single question, and, for this reason,

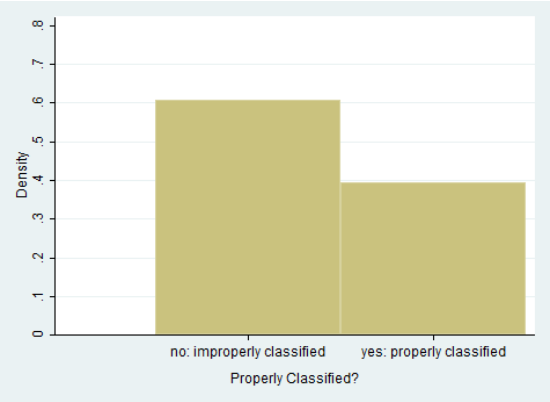

Figure 4: Proportion of Simulated Users Classified by Vote Compass as Being Closer to a Party when in Fact there was at least One Other Party with which that Simulated User Agreed More Closely Across the 30 Issues in the Vote Compass Questionnaire

it concentrates in the center for each user when it is averaged. In some cases the error will benefit the left, in other cases the error will benefit the right, and in still other cases it will benefit the center. It is therefore evenly distributed across the political spectrum. When it is averaged, however, it becomes far more likely to benefit the center–the proportion on the right and the left "cancel each other out", in effect, and shapes the distribution toward the center. Thus, when the Vote Compass shifts from measuring an underlying dimension to measuring the distance between a party and a user on that dimension (i.e., the two-dimensional output that the Vote Compass reports to the user), the process of averaging converts the random error in the measure of the former into systematic error in the measure of the latter. This does not mean that the user is in fact closest to the Liberal Party (see Equation 1 and Figure 1). It simply means that the process of averaging distorts systematically the error in the measure of the underling dimensions, and this distortion just so happens to work in the direction of the Liberal party.

The purpose of this analysis is to trace the measurement error by highlighting the magnitude, direction, and origins of systematic error in the Vote Compass algorithm. As I have shown, this distortion emerges not from how the researchers have positioned the political parties on these issues, nor from the issues that were chosen, but instead from how the Vote Compass algorithm "averages" the positions before it calculates distance on the Cartesian plane. The process of averaging has other implications has well. If a user answers the Vote Compass by disagreeing as strongly as possible with the Liberal party position on every single issue, then, depending on how that user answers the handful of questions where the Liberals are in the middle (i.e., whether the user strongly

agrees, strongly disagrees, some combination of both, or doesn't know), the user will be scored by the Vote Compass as being closest or second closest to the Liberal Party, even though the user is further away from the Liberals on these issues, by a considerable extent, than from any other party.[4] Some users will find this confusing. Indeed, it is my contention that the averaging in the Vote Compass algorithm jeopardizes the educational value of this tool.

One response to the objections listed here may well take the form of the empirical claim that many respondents did, in fact, answer the questions on these dimensions in a manner that would allow us to predict their answer to one question by knowing their answer to some other question. There is no question that this is the case – user responses are partly random, but not wholly random. Of course, even if a user does answer ideologically in left/right terms, the Vote Compass is still not guaranteed to position them correctly vis-a-vis their distance from the parties on these thirty issues. But there is one group for whom the Vote Compass is likely to work properly. Hardcore and highly engaged partisans are likely to answer each question in a way that aligns with the policy positions of their favorite political party (Green et al., 2002; Goren, 2005). NDP partisans take-up the NDP position on every issue, Conservative partisans take up the Conservative positions on every issue, and so on. The answers of these users on each dimension are not only non-random in the sense that they are predictable, but they also align with the positions of at least one political party. For these users, therefore, the result that they get from the Vote

[4]The Vote Compass team drew some attention to the issue of "acquiescence bias" in its response to Professor Kathy Brock's criticism of the instrument. Acquiescence bias refers to the tendency of some users to respond with the same response category (e.g., strongly agree), no matter the question. The threat of acquiescence bias is that a correlation between multiple items will emerge simply by virtue of the fact that some users respond to all questions with the same answer. Thus, when constructing multi-item measures of underlying opinions, social scientists commonly reverse the coding of the categories so that acquiescence bias does not work in the direction of the expected correlation between items–in other words, a potential source of systematic error is converted by reversing the order of the categories into a source of random error. This reflects the inherent conservatism of statistical methods. In the case of measuring underlying dimensions, the reversal of the categories in the Vote Compass has the intended effect. But when it comes to measuring the distance between a user and the parties on that dimension, which is what the Vote Compass displays to the users on the two-dimensional Cartesian plane, the reversal of the response categories has an altogether different effect. If the researchers had simply left all of the categories so that they ran in the same direction (e.g., left/right), then acquiescence bias would benefit different parties depending on which category the user selected. If "strongly disagree" was always the left-wing answer and "strongly agree" was always the right-wing answer, then acquiescence bias in the "strongly disagree" category would work in the direction of the most left-wing party, acquiescence bias in the "strongly agree" category would work in the direction of the most right-wing party, and acquiescence bias in the middle would work in the direction of the most centered party. By reversing the categories, however, all of the distortion from acquiescence was directed toward a single party in the measure of the user's proximity to that party: whether the user answers consistently "strongly agree", "strongly disagree", or "neither", one always ends up in the Liberal category. In other words, reversing the coding of the categories transformed systematic error into random error in the measure of the underlying latent dimension, as researchers are supposed to do, but, in so doing, it generated directional error in the measure of the user's position relative to the parties on that dimension. This particular problem, however, is a symptom rather than the cause of the directional error in the Vote Compass.

Compass will be accurate, at least insofar as they are told by the Vote Compass that they are, in fact, "closest" to the party with which they agreed most on the questions that they just answered. But although hardcore and deeply engaged partisans are likely to get the "right answer" from Vote Compass, they are also the least likely to have been influenced by the results that the Vote Compass revealed to them (Zaller, 1992). A person who is deeply committed and highly active in, say, the the New Democratic party, and who agrees with everything that the New Democratic party stands for, is unlikely to have been persuaded by the Vote Compass, no matter the result (e.g., if the result were Green).

The story may altogether different, however, for citizens who are not highly partisan and engaged politically. This group includes, for example, young people and new Canadians who are just learning their way around the Canadian political landscape, and it includes Canadians who care enough about politics to use the Vote Compass but who do not have enough time to follow politics closely on a day-to-day basis. These are precisely the citizens who are most likely to be influenced, many argue, by the information that they get from the media and from public opinion polls (Zaller, 1992; Daigle, 2010). At the same time, these are also the citizens who are very unlikely to answer all of the questions on each dimension in a manner that would allow us to predict their answer to one question in the Vote Compass by knowing their answer to some other question on that dimension (Converse, 1964).[5] In other words, these citizens are likely to answer randomly vis-a-vis the underlying dimension, and, as I have shown above, they are therefore unlikely to get the proper result when the Vote Compass shows them on the two dimensional plane the party to which they are closest on the questions that they just answered. Taken together, this may well generate an unfortunate irony for the Vote Compass' two dimensional representation of its results. This representation may be most likely to work properly for those that it is least likely to influence, and least likely to work properly for those that it is most likely to influence. Further research would be needed to test this possibility, but if this hypothesis is right, then Vote Compass' two dimensional picture is most likely to mislead precisely the citizens who depend on it the most for direction.

The intention of Vote Compass was to get people thinking about where they stand on the issues, and where the parties stand on those issues. There is clear educational value in the student-led research behind this tool. As it stands, however, that educational value is being undermined, in my opinion, by the two-dimensional picture that the Vote Compass presents to its users. This picture is telling Canadians that they are "closest" to a political party and it is advertised on the promise of revealing to Canadians "whose views are most like yours." A reasonable person could certainly be led to interpret "closest," "right

[5]There is no obvious logical connection, for example, between one's opinion about Canada's military presence in the Arctic and one's opinion about whether the federal government should play more of a role in Quebec culture and one's opinion about whether the environmental damage caused by the oil sands is exaggerated and one's opinion about whether the senate should be abolished, and so on. For citizens who are not deeply partisan, these are altogether separate issues.

for you" and "most like yours" to mean "most in agreement" on the issues in the survey that the person just answered. But the Vote Compass algorithm makes no such calculation. As a result, it is often showing people who agreed most often and most closely with one political party on these issues that they are in fact "closest" to some other party on the two-dimensional plane. I have been told that the two-dimensional picture is but an "abstract conceptualization of ideological space." But all measures contain some element of error (i.e., unreliability). And it is clear that the questions in the Canadian Vote Compass do not represent reliable measures of two "abstract ideological dimensions," even by the most generous standards of statistical reliability. Indeed, I do not believe that the intent of the Canadian researchers was to gear their questions for the narrow technical purpose of measuring two underlying dimensions. In any case, this particular point does not address the issue of how random error in the measure of these dimensions turns into systematic error in the measure of the user's position relative to the political parties on these dimensions. This is the methodological concern addressed in this paper.

It worth noting that the Vote Compass has been employed in a number of elections in other countries, and there are political scientists who see this tool as an opportunity to offset the rising levels of apathy, particularly among young people. Certainly, these goals should be acknowledged. Moreover, it is my understanding that the decision to represent distance in political space in abstract ideological terms, rather than issue-by-issue, is a part of a well-intentioned effort to offset the potential bias which may emerge from the selection on the part of the researchers of which specific issues and questions to include in the questionnaire. That decision would figure front and center in the typical utility measure (Equation 1), though I suspect that it would also be plainly obvious to the users that a standard utility algorithm would purport to tell them nothing beyond their location relative to the parties on the specific questions that they were just asked about. In any case, this is an election campaign, citizens use information shortcuts (Lau and Redlawsk, 2006; Sniderman et al., 1991), and the source credibility attached to this project may have a way of piercing the healthy guard of otherwise skeptical citizens (Petty and Cacioppo, 1986). Although it is important to acknowledge and respect the well-intentioned arguments of those who disagree with me, my substantive conclusion is nonetheless the same: Canadians who are interested in using Vote Compass to find out where they stand on these issues relative to the political parties should ignore altogether the potentially misleading two-dimensional picture and skip right to the back-end where the party positions are listed.

References

Adams, J. F., S. Merrill III, and B. Grofman (2005). *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. New York, NY: Cambridge University Press.

Converse, P. E. (1964). *Ideology and Discontent*, Chapter The Nature of Belief Systems in Mass Publics, pp. 206–261. New York, NY: The Free Press of Glencoe.

Daigle, D. T. (2010). *Catching the Big Wave: Public Opinion Polls and Bandwagons in US and Canadian Elections*. Ph. D. thesis, The Ohio State University.

Goren, P. (2005, October). Party identification and core political values. *American Journal of Political Science 49*(4), 881–896.

Green, D., B. Palmquist, and E. Schickler (2002). *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CO: Yale Univesity Press.

Lau, R. R. and D. P. Redlawsk (2006). *How Voters Decide: Information Processing During Election Campaigns*. New York, NY: Cambridge University Press.

Markus, G. B. and P. E. Converse (1979). A dynamic simultaneous equation model of electoral choice. *The American Political Science Review 73*(4), 1055–1070.

Petty, R. and J. Cacioppo (1986). *Communication and Persuastion: Central and Peripheral Routes to Attitude Change*. New York, NY: Springer-Verlag.

Sniderman, P. M., R. A. Brody, and P. E. Tetlock (1991). *Reasoning and Choice: Explorations in Political Psychology*. New York: Cambridge University Press.

Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.