Taking a Step Back
Things we don't see, things we don't like.

Elizabeth Penttila
Department of Anthropology, McGill University

Cite as: Penttila, Elizabeth. 2008. Taking a Step Back: Things we don't see, things we don't like. *Stop: Toutes Directions*, report 10. http://stoptoutesdirections.org/report10.pdf.

Abstract: With a wide variety of mostly inexperienced research teams, an observational bias on continuous traits is certainly possible. The Stop *project displays no such bias on its four point scales for vandalism, wear and damage. Could student attitudes have affected this? An opinion survey throws light on what the students are really thinking- and the diverse results support a Portfolio or Averaging effect on the data, or a need for greater sample size.*

Introduction

Stop signs- those invested in this project have probed so many angles of these everyday objects, that many of us feel almost blind to them. But in our effort to obtain such a tremendous quantity of data- almost 3,000 signs, spread throughout six municipalities of Montreal- we may have lost our ability to correctly discern the very information we sought. We may have been blinded to the signs.

Another point of interest is the attitude of students, and their willingness and desire to complete the project. Anecdotally, it is easy to see how these things might affect information recovery, but what can we tell from the data itself? Students asked to respond to an email survey on their feelings on the project did so with a 92% response rate. This data provided the basis for correlations between project appreciation and data, as well as interesting internal correlations; for example, the correlation between number of zones, and scaled response to "I liked this project." With this data, it should be possible to estimate internal bias on various parameters of the project due to states of opinion.

Methods

Most of the data necessary for this project came from the fieldwork the class participated in, taking various qualitative and quantitative data about stop signs in Montreal. The data taken include language present, damage, wear, vandalism, presence of horizontal supports, luminescence, and the presence of related signs (all way signs, specifically). Additionally, the students had a comment region in which to add observations, though this was not always utilized. In order to establish any in-team bias, I first determined four groups of data to examine. First, data from similar zones (contiguous zones with similar densities of intersections) collected by the same teams. All the values for damage, wear, and front and back vandalism was averaged for each zone, then the sites compared by regression. The same process was undertaken for similar zones with different teams analyzing them, different zones (sites with at least 4 zones between them) undertaken by the same teams, and finally different zones undertaken by different teams. To undertake this analysis, I must assume that nearby, similarly dense sites will in reality have similar values; if this is not the case then only by sampling the areas again by some means could a result possibly be found.

Teams were determined by arbitrarily assigning letters to groups of people who surveyed areas together. Similarly, zone numbers were assigned based on time of assignment and are essentially arbitrary. Most teams were consistently made up of the same members. Some switching did occur, however, and so teams were considered to be those made up of a majority of members from another team. Single member teams were considered separate teams, as single person input would represent individual bias. The teams which assessed only one zone were excluded from these analyses.

The second portion of this analysis involved a survey, sent to all students (See appendix for original message, object 2.0). Questions were designed to both establish facts about individuals (e.g., number of zones studied) and allow qualitative responses (comment space), as well as provide some scaled responses to statements of attitude (agree/disagree, "I had a lot of choice in projects."). Questions were not intentionally leading, but in the interest of brevity and higher return rate I did not employ reverse-statements ("I like my project," I regret the project I chose."). Responder's results were transferred to a master document, omitting names except initials used to check accuracy. Some qualitative factors were quantified. This was then subject to the various statistical pokings and proddings necessary. Some of the responses had to be dropped eventually, when it became apparent that the question was poorly worded (transportation) or were not specific enough to yield relevant data (work expected) (Bertrand, 68).

In addition to this, the role of gender in determining participant activity was examined by simple mean-percent analyses.

The largest problem with all of this is the limited sample size. With only 23 individuals to analyze (myself and non-responders excluded) and 9 teams (myself included, as when I was surveying I did not know this would become my project), results may seem meaningful while being in fact an artefact of the limited sample. Seeing as this is the only data available, however, the results will have to stand as they are: correlations and other tools within the bounds of a single repetition of a project, carried out by a small group.

Results

"Similar sites" Analysis (For relevant original data tables, see appendix, tables 1.0-1.3)

Figure 1

R^2	0.8282
Slope	1.0536
P Value	0.0017

Figure 2

Figure 3

Figure 4

Figure 5

R^2	0.2457
Slope	0.5262
P Value	0.2108

Figure 6

R^2	0.6517
Slope	0.8795
P Value	0.0154

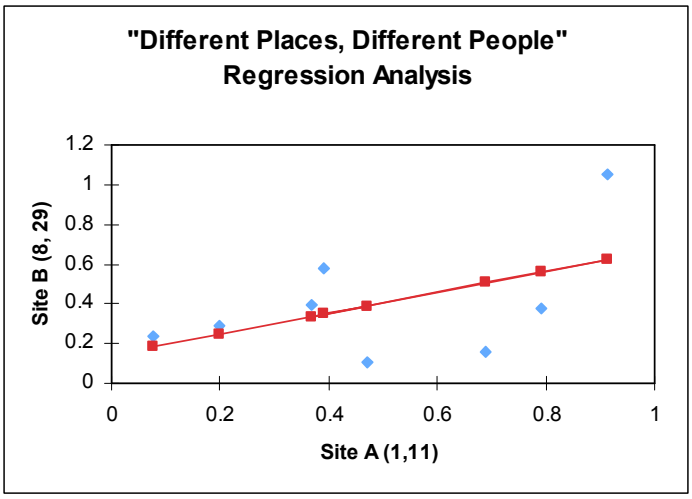

Figure 7

R^2	0.2526
Slope	0.5242
P Value	0.2043

Figure 8

Survey Analysis (original response table in Appendix, table 2.1)

Abbreviations are as follows: SQ for Scaled response Questions. CWC for Comment Word Count.

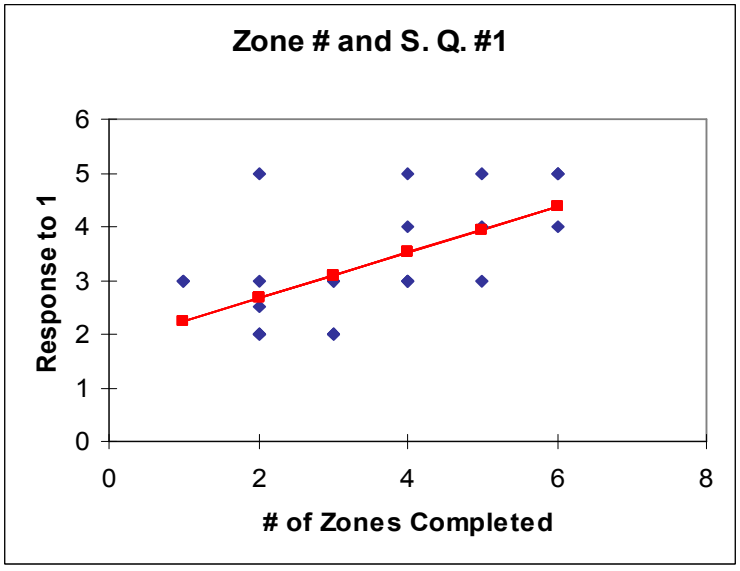

Figure 9

R^2	0.351
Slope	0.4234
P value	0.0023

Figure 10

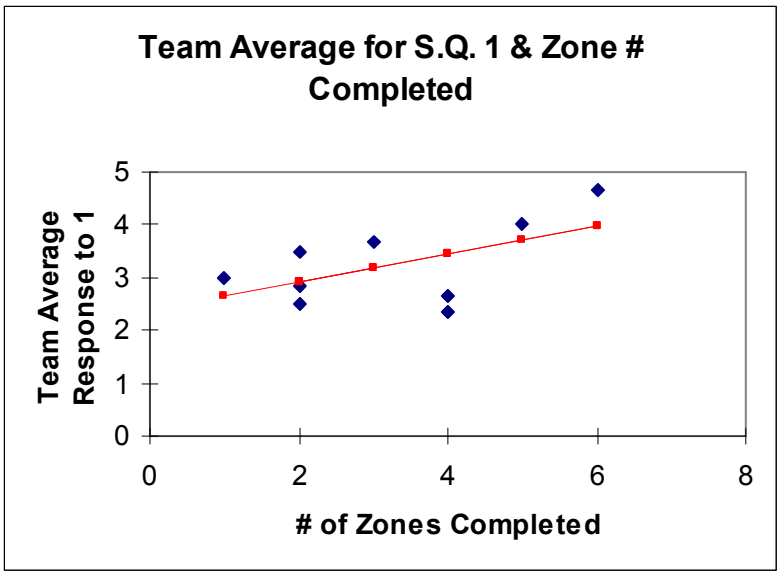

Figure 11

R^2	0.3134
Slope	0.2637
P value	0.117

Figure 12

S.Q. 1 & S.Q. 2	
R^2	0.0114
Slope	0.0836
P value	0.619

Figure 13

S.Q. 1 & S.Q. 3	
R^2	0.0712
Slope	0.3036
P value	0.2075

Figure 14

S.Q. 1 & S.Q. 4	
R^2	0.0248
Slope	-0.1153
P value	0.4623

Figure 15

S.Q. 1 & S.Q. 5	
R^2	0.0002
Slope	-0.01
P value	0.9533

Figure 16

S.Q. 1 & S.Q. 6	
R^2	0.0218
Slope	0.1147
P value	0.5062

Figure 17

S.Q. 1 & S.Q. 7	
R^2	0.0328
Slope	0.2321
P value	0.5286

Figure 18

S.Q. 1 & S.Q. 8	
R^2	0.0203
Slope	-0.128
P value	0.4909

Figure 19

S.Q. 1 & S.Q. 9	
R^2	0.0183
Slope	0.1964
P value	0.3974

Figure 20

S.Q. 1 & S.Q. 10	
R^2	0.0057
Slope	-0.1719
P value	0.7251

S.Q. 9 & S.Q. 10	
R^2	0.0022
Slope	-0.0625
P value	0.8265

Figure 21

Survey & S.Q. 2	
R^2	0.001
Slope	0.0625
P value	0.881

Figure 22

Figure 23

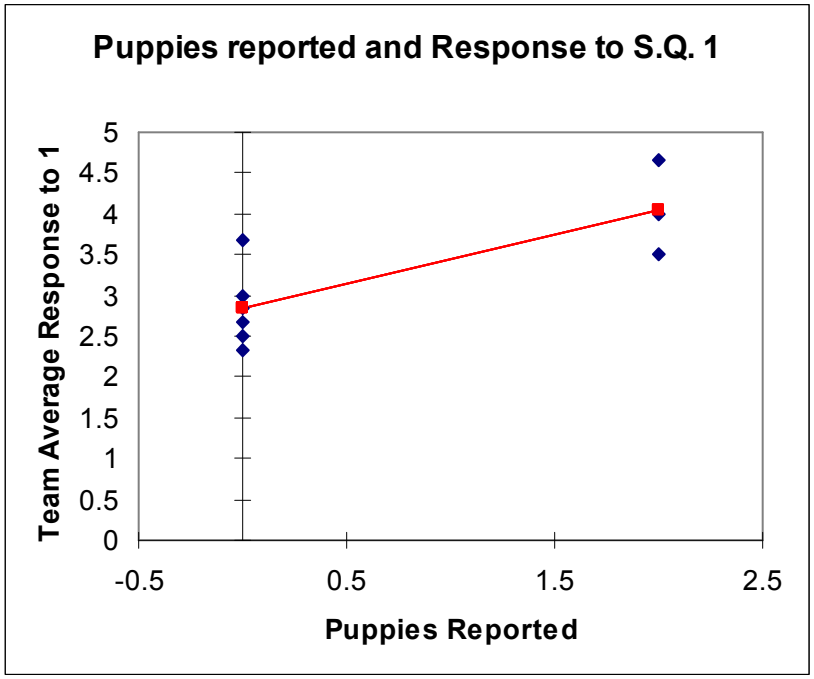

Figure 24

Puppies and S.Q. 1 Avg.	
R^2	0.6245
Slope	0.0625
P value	0.0113

Figure 25

	Photographers	Total	%
Men	4	6	66.667
Women	3	17	17.647

Figure 26

Discussion

Similar Sites Analysis

Figure 1 shows the average values of continuous numerical traits (damage, wear and vandalism), regressed against the same traits of another zone analyzed by the same team, in the same area. As we can see, the r^2 is not perfect, but is very strong and highly significant ($P<.05$). This is precisely what we would expect in any situation; similar sites yield similar results. Figure 3 shows a regression of similar sites, this time those conducted by different teams. Again, similar sites show similar results. In figure 5, we see different sites conducted by the same teams. The results are not significantly similar ($P\sim.2$). Compare this to figure 7, different sites by different teams, which are again not significantly similar. From this we can see that that the difference between sites is more significant than the similarity of teams; we can thus conclude that there is no inherent bias. The small scale may account for this; in a ten-point scale, for example, we might see that certain people are more likely to produce certain favoured numbers. That being the case, the use of a four point scale was ideal for finding the 'actual' data of any given region, and the instructions and training provided resulted in data ideal for accurate analysis.

Survey Analysis

Figure 9 shows the correlation between the number of zones an individual surveyed, versus their assessment of the statement, "I liked this project." Although the r^2 value and slope are weak, the significance of the data is quite strong ($P<.01$). This could be explained in a few ways- more enthusiastic students wanted to complete more zones; alternately, the more time a student spent in the field, getting exercise and petting puppies (see Figures 9 and 10), the happier they were by the end of the survey period. The combination of the two might represent an enjoyment feedback cycle- enthusiasm leads to surveying, which leads to fun and more enthusiasm.

The same is not true, however, of resources spent (figure 14); there is no correlation between enjoyment and the use of resources toward the project. This may be an example of a poorly worded question, or one that does not take in all possible effects (students possessing a metro card will probably not consider that an expenditure, while those using tickets will). There is some indication that students who disliked the project would rather have done an equivalent essay, but the correlation is not strong (figure 15). One must ask, what would they rather do? In the comments section, many mentioned the timing of the work as the greatest problem, rather than any aspect of the volume or type of project. The problem isn't time restraints, though- there is almost no correlation between liking the project and responses to "We had enough time to collect data" (figure 16). Rather, the end of the undergraduate semester is packed with conferences, papers, and work study hours; getting three people to coordinate their schedules is cited as a major problem for certain groups. And neither is the enjoyment linked to an individual's satisfaction with their own project (figure 18). As a matter of fact, overall project satisfaction is not significantly correlated with any of the other scaled questions (other tables); this points to a simple fact statistics cannot capture. Every person is going to enjoy a project for different reasons, and will invest energy, like their project, and so on for those same reasons! Some people would have preferred to do the project in February, some want to move the class to fall; some wanted to work alone, others cited the group work as their favourite part of the project. Some individuals loved it and added no comments, some took the time to detail problems with the project's method, and others actually used "hate" as a descriptor. This great variety of responses could have created a portfolio effect, meaning the individual variances add up to a "normal," more accurate whole- students with different ideas of fun (say, taking pictures of stop signs) compensated for team members who would prefer to spend a day reading articles. There are some limits to this diversity- all students are in either the Arts or Arts and Science faculties, and list Anthropology as one of their majors. This weeds out quite a bit of diversity already. But as we can see from the variety of responses to all survey questions, quite a bit of variation remains in the sample.

Division of labour shows an interesting quirk. The male students (a minority in the class) represent over half of the photographers (4/7). A full 66% of male students conducted the photography for their team, while a measly 17% of women were photographers (figure 26). This is probably due to the myth of better female handwriting and map-reading skills (these were the other two general jobs assigned in a three person team, along with 'combination' jobs).

There are a few small points which may indicate an error on the part of the surveyor, by asking non-specific questions (or at any rate receiving non-specific answers). Figure 22 shows the relatedness of

a yes/no question, "Did this project teach us surveying methods?," and Scaled Question 2 ("This was a valuable project."). These, at least, should be correlated, a main purpose of the project being an education in field surveying. Yet there was no relationship between the two. Additionally, scaled questions 9 and 10 ("My field team divided up work fairly." and, "My field team was able to agree on things and solve most problems without friction.") are shown in figure 23. One certainly expects a degree of cooperation between these statements- teams that can agree can divide work fairly- and though there is some correlation it is not significant (P~.1) and the resultant slope is negative, meaning the better a team agreed the less fairly the work was divided. It is also possible that since the questions were in order, students felt the need to respond differently to them (Bertrand, 67).

Having concluded that bias does not play a significant part in the data collected, and that since opinion is a personal factor, how then do we explain the results of figure 11? Although the result is not technically significant, there is some correlation between a team's average enjoyment and the number of zones they completed. The small sample size comes into play, with only nine teams to analyze. Regressing the data for number of puppies teams reported encountering and average response to scaled question one actually produces significant results (P<.05); this is probably not a causal relationship, and so we can see this is probably not a sufficient sample size to undertake serious calculations.

The potential for additional research in this area is tremendous. Simply by repeating the Stop project (using the same field methods), and combining the results with an identical survey, the sample size might be large enough to yield significant results. Also of interest would be a zone analyzed separately by 2 different teams- any differences would be due to interpersonal variation, rather than variation of location. Hopefully, future generations of McGill archaeology will get to participate in this project and just maybe, develop further those problems we have been unable to surmount.

References

Bertrand, Marianne and Mullainathan, Sendhil. *Do People Mean What They Say? Implications for Subjective Survey Data.* The American Economic Review. 91 (2). May 2001. pp 67-72

Appendix

Table 1.0- Same People, Similar Places

	16	19
Damage	0.297872	0.595238
Wear	0.228916	0.571429
Van Fr	0.385542	0.380952
Vand B	0.148936	0.333333

	27	31
Damage	0.454545	0.458333
Wear	0.854545	0.891667
Van Fr	0.254545	0.125
Vand B	0.018182	0.033333

Table 1.1- Different people, Similar Places

	10	14
Damage	0.461538	0.608247
Wear	0.333333	0.43299
Vand Fr	0.217949	0.206186
Vand B	0.141026	0.086957

	1	2
Damage	0.39	0.462185
Wear	0.47	0.422811
Van Fr	0.37	0.495798
Vand B	0.69	0.672269

Table 1.2 – Same people, Different places

	7	9
Damage	0.859649	0.696296
Wear	0.54386	0.762963
Van Fr	0.22807	0.251852
Vand B	0.280702	0.022222

	3	13
Damage	0.854369	0.608247
Wear	0.271845	0.43299
Van Fr	0.786408	0.206186
Vand B	0.495146	0.086957

Table 1.3 – Different People, Different places

	1	8
Damage	0.39	0.575
Wear	0.47	0.108333
Van Fr	0.37	0.398438
Vand B	0.69	0.15873

	11	29
Damage	0.791209	0.37234
Wear	0.912088	1.053191
Van Fr	0.197802	0.287234
Vand B	0.076923	0.234043

Object 2.0- Email Survey

Name:
Program:
Concentrations:
Year:

What neighborhood do you live in?

What is your stop sign project? How much work (beyond surveying) do you expect to undertake?

Please list the zones you worked on, and a few words of description about each zone.

How was work divided in your survey group(s), and what role did you have?

In the past thirty days have you (a) walked (b) bus/metro-ed, (c) biked or (d) driven the most?

Since the beginning of the project, have you noticed a change in your attitude about stop signs? Please describe.

Do you feel this project was effective in teaching us surveying methods? Why or why not?

Please rate, on a scale of 1-5 (1- fully disagree, 5 fully agree), the following statements:
() I enjoyed this project.
() The stop project was a valuable experience.
() I put a lot of my own resources (time, energy, metro tickets) into this project.
() Given the choice I would rather have done an equivalent research essay.
() We had sufficient time to gather all the necessary data.
() I had a lot of choices in specific projects,
() and am pleased with the one I chose.
() This project took me places in Montreal I have never been to.
() My field team divided up work fairly.
() My field team could agree on most things and solve problems without friction.

Any additional comments to have on this project belong here. Things to improve the project. Things you loved and hated. The craziest dog you saw. Anything. Go nuts,

Table 2.1

DOL #	Change Attitude	Survey Y/N	1	2	3	4	5	6	7	8	9	10	Comments sc	Word count
1	0	+	3	3	4	5	1	5	2	2	5	3	0	0
0	0	0	2	2	4	2	1	3	1	5	5	5	-	87
2	0	-	2	1	4	5	1	2	4	4	5	5	-	152
1	0	0	3	2	5	1	2	4	4	2	3	5	-	66
3	0	+	3	2	5	4	3	4	4	5	4	4	0	30
1	0	0	2	3	4	3	4	2	3	3	5	5	0	32
1	-	-	2	1	4	3	1	3	4	5	5	5	-	44
1	+	+	5	5	5	1	5	3	5	5	5	4	+	57
4	-	+	4	3	5	2	4	2	3	1	4	5	-	198
4	-	+	4	4	5	1	5	4	5	5	5	5	0	0
4	-	0	4	4	5	2	3	5	4	4	5	5	0	50
3	0	+	2.5	3	2.5	1	5	5	3	1	5	5	-	65
2	0	+	5	5	5	2	1	1	4	5	5	5	+	65
3	-	+	3	4	4	4	3	5	4	5	4	5	0	27
3	+	+	5	5	5	5	3	5	5	5	5	5	+	208
2	0		3	4	4	4	3	5	4	5	5	5	-	82
1	0	+	2	4	4	1	3	3	5	5	4	5	0	0
3	0	+	2	1	5	2	4	5	3	5	5	5	-	134
1	-	-	3	1	1	5	1	5	3	5	5	5	-	76
2	0	0	3	4	5	1	4	5	5	5	3	5	+	262
3	+	+	3	3	5	3	5	5	4	5	5	5	-	68
2	0	+	5	5	5	5	4	5	5	5	5	5	+	113
1	0	+	3	4	4	3	2	4	4	5	5	5	+	57
3	0	+	5	5	5	3	2	3	4	5	5	5	+	29