Technical Supplement:
The Arts and Individual Well-Being in Canada

Connections between Cultural Activities and Health, Volunteering, Satisfaction with Life, and Other Social Indicators in 2010

Statistical Insights on the Arts

INTRODUCTION

The goal of this study is to examine predictors of three indicators of well-being: self-reported health; volunteering; and self-reported satisfaction with life. Data drawn from Statistics Canada's Social Survey of 2010 are used to determine which demographic factors and cultural activities can predict very good or excellent health, volunteering, and very strong satisfaction with life.

In addition to a number of demographic factors, six cultural activities were also selected for this analysis: 1) art gallery attendance; 2) theatre attendance; 3) popular music attendance; 4) classical music attendance; 5) cultural festival attendance; and 6) book reading. Statistics Canada asked Canadians 15 years of age or older about their participation in these (and other) cultural activities during the 12 months prior to the survey. Phrased in the language of the 2010 General Social Survey questionnaire, the percentage of Canadians 15 or older participating in each of these cultural activities was:

- 36% visiting "a public art gallery or art museum (including attendance at special art exhibits)".
- 44% attending "a theatrical performance such as a drama, musical theatre, dinner theatre, comedy".
- 13% attending "a symphonic or classical music performance".
- 39% attending "a popular music performance such as pop/rock, jazz, blues, folk, country and western".
- 37% attending "a cultural or artistic festival (such as film, fringe, dance, jazz, folk, rock, buskers or comedy)".
- 76% reading at least one book.

This technical supplement discusses variable selection and coding as well as the results of three logistic regression models. Logistic regression analysis was used to determine the odds ratio (or likelihood) of someone with the characteristic of an independent variable (e.g., someone between 35 and 44 years of age) having very good or excellent health, volunteered in the past year, or very strong satisfaction with life, compared with others (e.g., those between 15 and 24 years of age), holding other factors constant.

The main report for this project, available at www.hillstrategies.com, provides a brief summary of the results of the regression models. The report also provides exploratory data concerning the connections between cultural attendance by Canadians and their personal well-being. The exploratory data highlights the potential relationships between 18 arts, culture and heritage activities and eight social indicators:

1. Self-reported health
2. Self-reported mental health
3. Volunteering
4. Feeling trapped in a daily routine
5. Stress level
6. Knowledge of neighbours
7. Doing a favour for a neighbour
8. Self-reported satisfaction with life

<div align="center">

METHODS

</div>

Logistic Regression

Logistic regression was used to predict the probability of attendance at an arts activity due to other characteristics under investigation (the demographic variables and other cultural activities). More specifically, binary logistic regression was used, as the dependent variables in this study are dichotomous (i.e., yes/no variables). This method measured the statistical significance of the independent (predictor) variables, the multicollinearity of the independent variables, and the effect of the predictor variables on the dependent variable.

Binary (or binomial) logistic regression is a form of analysis where logistic regression equations are solved iteratively, a contrast from other forms of regression analysis where a mathematical equation is solved explicitly. In binary logistic regression a trial equation is fitted and adjusted in order to improve the fit:

$$\text{logit}[p] = \log[p/1-p] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_i x_i \qquad \text{[Equation 1]}$$

where p is the probability of the event occurring, α is the constant of the equation and β_i are the coefficients of the predictor values. The user determines when the iterations stop. That is, either an explicit number of iterations are performed or a cut-value is selected so that the iterations stop when the improvement from step to the next is less than the cut-value.

Once the equation is fitted, odds ratios can be determined. The odds ratio is a prediction about how likely the respondent was to have participated in the arts activity in question. This is calculated by raising the base of the natural logarithm to the βth power, where β is the coefficient of the predictor variable from the trial equation. That is:

$$\text{ODDS} = e^{\beta} = p/1-p \qquad \text{[Equation 2]}$$

where β is the slope of the logistic regression equation for a specific dependent variable. Once the odds of an event occurring have been calculated, they can be converted to probabilities:

$$p = \text{ODDS}/(1 + \text{ODDS}) \qquad \text{[Equation 3]}$$

Probabilities can be more useful when describing the effects of the predictor variables on the dependent variable.

Predictive analytic software (SPSS) was used to perform the iterations necessary to derive the logistic regression equation. The weighted dataset for the 2010 General Social Survey was used.[1]

Once the logistic regression equation was fitted, the software's output presented the coefficients of the predictor variables and also included a 95% confidence interval used for error analysis.

Variable Selection

The goal of this study is to examine predictors of three indicators of well-being: self-reported health; volunteering (in the past year); and self-reported satisfaction with life. A model was constructed for each dependent variable using the forward stepwise conditional logistic regression procedure.

---

[1] The weights were rescaled to have the average weight equal to 1. This was done to provide more meaningful calculations of variances. See Statistics Canada, *General Social Survey, Cycle 24: Time-Stress and Well-Being, Public Use Microdata File Documentation and User's Guide*, Catalogue no. 12M0024X, page 19.

The statistical models include demographic variables that exploratory statistics and similar models from other jurisdictions indicated might be significant. This was done in order to create the best possible models using only demographic factors.[2]

At this point, the six arts and culture activities were added separately to the models. The goal was to find whether each of the arts and culture activities has some explanatory value in the models above and beyond demographic information. The goal was not to find the cultural activities that created the best possible overall models of the three indicators of well-being.

As is the case with all secondary research, there was no opportunity to include customized questions in the survey. For example, the General Social Survey did not ask whether respondents smoke or consume alcohol. If these were available, they might have been a useful predictor of health.

While the statistical models might provide evidence of a connection between cultural activities and well-being, it is very difficult to provide irrefutable evidence of a cause and effect relationship between the variables in a statistical model in the absence of an experiment to directly measure the impacts of culture on personal well-being.

Table 1 lists all of the independent variables that were entered into the models.

| Table 1: Independent variables in the logistic regression models ||
Demographic factors	Cultural activities
Level of education	Art gallery visits
Household income	Theatre attendance
Age group	Classical music attendance
Sex	Pop music attendance
Children at home	Cultural festival attendance
Region (BC, Prairies, Ontario, Quebec, Atlantic)	Read a book
Urban / rural	
Household language	
Aboriginal person	
Visible minority person	
Immigrant	
Physical activity in the past week	

---

[2] Model accuracy was measured by the statistical significance of the model, in particular, the -2 log likelihood. The steps of adding and removing different combinations of variables was done until the -2 log likelihood was as small as possible, while other statistical tests showed that the data fit the model well and that the predictor variables did not exhibit multicollinearity.

Variable Coding

A dichotomous structure was created for each of the dependent variables (i.e., the three indicators of well-being). For self-reported health, ratings of "excellent" or "very good" were separated and compared with ratings of "good", "fair" or "poor". Volunteer activity in the past year is a yes/no variable. Self-reported satisfaction with life was separated into very strong ratings (i.e., ratings of 8-10 on a scale from 1-10) and lower ratings (1-7).

Many of the variables in the model have two possible response options, typically "yes" and "no". For these dichotomous variables, "no" was coded "0" and "yes" was coded "1".

Other categorical variables have multiple response options. For the respondent's household language, the response options were grouped by Statistics Canada as "English only", "French only", and "Another language only or multiple languages". The coding of these responses in the statistical model was: English only = 0; French only = 1; Another language only or multiple languages = 2.

In cases where the questions included the response options "don't know" and "not stated", these responses were removed from the regression analysis. When these results are deleted for all variables, the revised dataset includes 7,330 complete records rather than the 7,500 original records.

RESULTS

Data Interpretation

The results obtained from the SPSS output are provided in Appendix 1 for three "core" models (i.e., models with demographic variables only).

Information about the statistical importance of six cultural activities is provided below the core models in Appendix 1. The six cultural activities, which were entered separately to the core demographic models, include art gallery attendance, theatre attendance, classical music attendance, popular music attendance, cultural festival attendance, and book reading.

Some of the coefficients for the core demographic factors changed slightly when a statistically significant cultural activity was entered into the model. These changes were very small and are not provided here.

The odds ratio – denoted as exp(B) – for each independent variable is a positive real number. If the odds ratio is greater than one, then the independent variable predicts a greater likelihood of having very good health, volunteered in the past year, or very strong satisfaction with life, compared with people in the reference class for that variable. Conversely, if the odds ratio is less than one, there is a lower likelihood.

For dichotomous variables, the odds ratio indicates that someone who responded "yes" to the given question is exp(B) times as likely to have very good health (in the first example) than someone who responded "no" to the same question. For example, the first row of the cultural activities portion of Appendix Table 2 shows that art gallery visitors are 1.35 times as likely as non-visitors to have very good or excellent health.

All dichotomous variables that are not considered significant predictors were excluded from the tables in this technical supplement.

For categorical variables, if the respondent fits into one of the classes of the variable, then they are exp(B) times as likely to attend compared with someone in the reference class. The reference class is the first option listed and always has an empty entry for exp(B). For household language, the odds are shown relative to the reference category ("English only"). The SPSS output for satisfaction with life (Appendix Table 3, Exp(B) column) shows that respondents with French as their household language are 1.56 times as likely to report very strong satisfaction with life as respondents who speak English in their household. In other words, they are about 56% more likely to report very strong satisfaction with life. This result is statistically significant ("Sig" = 0.000, which is less than 0.05).

Care must be used when examining the results for the categorical variables, as not all classes are necessarily significant. Categorical variables with *at least one* class that is useful are included in the model. The SPSS output for satisfaction with life (Appendix Table 3) shows that the finding for respondents with another household language or multiple household languages is not statistically significant (Sig = 0.184, which is greater than 0.05). As such, the odds ratio in the Exp(B) column should be ignored.

For each of the classes that has a value for exp(B), there are columns entitled "Lower" and "Upper". These represent the 95% confidence interval (margin of error) for the odds ratio. If the 95% confidence interval contains the value "1", then it contains both a prediction of "yes" and "no" with regards to the dependent variable. This corresponds to a value larger than .05 in the "Sig" column, indicating that the independent variable is not an effective predictor in the model.

The independent variables were entered into the models with the coding of "0" as the reference level. The reference level is listed first in the following tables.

Appendix 1: SPSS Output for the Binary Logistic Regression Models of Three Dependent Variables (Indicators of Personal Well-being)

Appendix Table 1: Results of the model to predict very good or excellent health

Health model	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Education (no high school diploma)			29.515	4	0.000			
Education (high school diploma)	0.26	0.11	5.556	1	0.018	1.29	1.04	1.60
Education (some university or college)	0.26	0.11	5.823	1	0.016	1.29	1.05	1.59
Education (college diploma or certificate)	0.10	0.10	1.167	1	0.280	1.11	0.92	1.34
Education (university degree)	0.45	0.10	19.829	1	0.000	1.56	1.28	1.91
Household income under $20,000			110.712	6	0.000			
HH income $20,000 to $39,999	0.64	0.14	20.844	1	0.000	1.89	1.44	2.48
HH income $40,000 to $59,999	0.71	0.14	27.288	1	0.000	2.04	1.56	2.67
HH income $60,000 to $79,999	0.73	0.14	27.517	1	0.000	2.07	1.58	2.71
HH income $80,000 to $99,999	0.98	0.14	46.200	1	0.000	2.67	2.01	3.54
HH income $100,000 to $149,999	1.13	0.14	68.081	1	0.000	3.11	2.37	4.07
HH income $150,000 and over	1.31	0.15	81.705	1	0.000	3.72	2.80	4.94
Age 15-24			34.362	6	0.000			
Age 25-34	-0.06	0.11	0.355	1	0.551	0.94	0.76	1.16
Age 35-44	-0.04	0.11	0.109	1	0.741	0.96	0.78	1.19
Age 45-54	-0.37	0.10	12.532	1	0.000	0.69	0.56	0.85
Age 55-64	-0.23	0.11	4.550	1	0.033	0.79	0.64	0.98
Age 65-74	-0.06	0.13	0.190	1	0.663	0.95	0.73	1.22
Age 75 and over	-0.56	0.15	13.765	1	0.000	0.57	0.43	0.77
Region (Ontario)			28.396	4	0.000			
Region (BC)	0.26	0.09	8.416	1	0.004	1.29	1.09	1.54
Region (Prairies)	0.09	0.08	1.210	1	0.271	1.09	0.93	1.28
Region (Quebec)	0.39	0.08	25.158	1	0.000	1.48	1.27	1.72
Region (Atlantic)	0.12	0.11	1.191	1	0.275	1.13	0.91	1.40
Visible minority (no, yes)	-0.31	0.09	12.160	1	0.000	0.74	0.62	0.87
Physical activity (less than 3 days in past week)			123.500	2	0.000			
Physical activity (3-5 days in past week)	0.74	0.08	89.404	1	0.000	2.09	1.80	2.44
Physical activity (6-7 days in past week)	0.79	0.07	110.697	1	0.000	2.20	1.90	2.54
Constant	-1.48	0.17	78.454	1	0.000	0.23		
-2 Log likelihood = 7431.553								

Cultural activities added separately to demographic model

Health model	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Art gallery attendance (no, yes)	0.30	0.06	24.120	1	0.000	1.35	1.20	1.51
Theatre attendance (no, yes)	0.28	0.06	23.340	1	0.000	1.32	1.18	1.48
Classical music attendance (no, yes)	Classical music attendance was not retained as a significant factor.							
Pop music attendance (no, yes)	0.21	0.06	12.694	1	0.000	1.23	1.10	1.38
Festival attendance (no, yes)	0.13	0.06	4.960	1	0.026	1.14	1.02	1.28
Read at least 1 book (no, yes)	0.25	0.07	14.103	1	0.000	1.28	1.13	1.46

Appendix Table 2: Results of the model to predict having volunteered in the past year

Volunteering model	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Education (no high school diploma)			141.291	4	0.000			
Education (high school diploma)	-0.15	0.12	1.645	1	0.200	0.86	0.69	1.08
Education (some university or college)	0.36	0.11	10.890	1	0.001	1.44	1.16	1.79
Education (college diploma or certificate)	0.17	0.10	2.783	1	0.095	1.18	0.97	1.44
Education (university degree)	0.86	0.11	66.674	1	0.000	2.36	1.92	2.90
Household income under $20,000			35.773	6	0.000			
HH income $20,000 to $39,999	0.13	0.14	0.829	1	0.363	1.14	0.86	1.52
HH income $40,000 to $59,999	0.34	0.14	5.946	1	0.015	1.41	1.07	1.86
HH income $60,000 to $79,999	0.31	0.14	4.767	1	0.029	1.37	1.03	1.81
HH income $80,000 to $99,999	0.37	0.15	6.115	1	0.013	1.44	1.08	1.93
HH income $100,000 to $149,999	0.42	0.14	8.724	1	0.003	1.52	1.15	2.00
HH income $150,000 and over	0.72	0.15	23.632	1	0.000	2.05	1.54	2.75
Age 15-24			54.051	6	0.000			
Age 25-34	-0.71	0.12	36.502	1	0.000	0.49	0.39	0.62
Age 35-44	-0.34	0.12	8.079	1	0.004	0.71	0.56	0.90
Age 45-54	-0.35	0.11	9.657	1	0.002	0.70	0.56	0.88
Age 55-64	-0.41	0.11	12.959	1	0.000	0.66	0.53	0.83
Age 65-74	0.04	0.13	0.079	1	0.778	1.04	0.80	1.34
Age 75 and over	-0.29	0.15	3.638	1	0.056	0.75	0.55	1.01
Sex (female, male)	-0.30	0.06	26.944	1	0.000	0.74	0.66	0.83
Children at home (no, yes)	0.26	0.07	14.488	1	0.000	1.30	1.13	1.48
Region (Ontario)			112.786	4	0.000			
Region (BC)	-0.01	0.09	0.006	1	0.937	0.99	0.83	1.18
Region (Prairies)	0.22	0.08	7.575	1	0.006	1.25	1.07	1.47
Region (Quebec)	-0.70	0.08	76.228	1	0.000	0.50	0.42	0.58
Region (Atlantic)	-0.02	0.11	0.019	1	0.891	0.98	0.79	1.23
Urban or rural residence (urban, rural)	-0.46	0.08	37.227	1	0.000	0.63	0.54	0.73
Visible minority (no, yes)	-0.33	0.09	13.232	1	0.000	0.72	0.60	0.86
Constant	-0.25	0.17	2.110	1	0.146	0.78		
-2 Log likelihood = 7111.760								

Cultural activities added separately to demographic model

Volunteering model	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Art gallery attendance (no, yes)	0.64	0.06	105.290	1	0.000	1.89	1.67	2.13
Theatre attendance (no, yes)	0.83	0.06	186.264	1	0.000	2.29	2.03	2.58
Classical music attendance (no, yes)	0.62	0.09	52.587	1	0.000	1.86	1.58	2.21
Pop music attendance (no, yes)	0.50	0.06	67.280	1	0.000	1.64	1.46	1.85
Festival attendance (no, yes)	0.71	0.06	132.829	1	0.000	2.02	1.80	2.28
Read at least 1 book (no, yes)	0.55	0.07	57.948	1	0.000	1.74	1.51	2.01

Appendix Table 3: Results of the model to predict very strong satisfaction with life

Satisfaction with life model	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Household income under $20,000			79.294	6	0.000			
HH income $20,000 to $39,999	0.28	0.14	4.232	1	0.040	1.32	1.01	1.72
HH income $40,000 to $59,999	0.55	0.13	17.380	1	0.000	1.74	1.34	2.25
HH income $60,000 to $79,999	0.51	0.13	14.559	1	0.000	1.66	1.28	2.16
HH income $80,000 to $99,999	0.70	0.14	24.848	1	0.000	2.01	1.53	2.64
HH income $100,000 to $149,999	0.78	0.13	35.088	1	0.000	2.17	1.68	2.81
HH income $150,000 and over	1.02	0.14	53.989	1	0.000	2.78	2.12	3.66
Age 15-24			107.749	6	0.000			
Age 25-34	-0.18	0.11	2.677	1	0.102	0.84	0.68	1.04
Age 35-44	-0.18	0.11	2.525	1	0.112	0.84	0.67	1.04
Age 45-54	-0.25	0.11	5.605	1	0.018	0.78	0.63	0.96
Age 55-64	0.06	0.11	0.266	1	0.606	1.06	0.86	1.31
Age 65-74	0.78	0.14	33.177	1	0.000	2.19	1.68	2.86
Age 75 and over	0.88	0.16	30.686	1	0.000	2.41	1.77	3.30
Children at home (no, yes)	0.20	0.07	8.925	1	0.003	1.22	1.07	1.39
Urban or rural residence (urban, rural)	-0.26	0.07	12.610	1	0.000	0.77	0.66	0.89
Household language (English only)			38.982	2	0.000			
HH language (French only)	0.44	0.08	34.566	1	0.000	1.56	1.35	1.81
HH language (another language only or multiple languages)	-0.14	0.11	1.765	1	0.184	0.87	0.71	1.07
Visible minority (no, yes)	-0.20	0.10	4.348	1	0.037	0.82	0.68	0.99
Physical activity (less than 3 days in past week)			103.105	2	0.000			
Physical activity (3-5 days in past week)	0.62	0.08	62.857	1	0.000	1.85	1.59	2.16
Physical activity (6-7 days in past week)	0.74	0.07	97.901	1	0.000	2.09	1.81	2.42
Constant	-0.65	0.16	16.580	1	0.000	0.52		
-2 Log likelihood = 7271.447								

Cultural activities added separately to demographic model

Satisfaction with life model	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Art gallery attendance (no, yes)	Art gallery attendance was not retained as a significant factor.							
Theatre attendance (no, yes)	0.27	0.06	20.408	1	0.000	1.30	1.16	1.46
Classical music attendance (no, yes)	0.26	0.09	8.374	1	0.004	1.29	1.09	1.54
Pop music attendance (no, yes)	0.16	0.06	7.421	1	0.006	1.18	1.05	1.32
Festival attendance (no, yes)	0.22	0.06	13.205	1	0.000	1.25	1.11	1.40
Read at least 1 book (no, yes)	0.14	0.07	4.547	1	0.033	1.15	1.01	1.32

Bibliography

Logistic Regression, G. David Garson. 2006,
http://www2.chass.ncsu.edu/garson/PA765/logistic.htm

Logistic Regression, 2002, http://userwww.sfsu.edu/~efc/classes/biol710/logistic/logisticreg.htm

Logistic Regression, Gerard E. Dallal, Ph.D. 2001, http://www.jerrydallal.com/LHSP/logistic.htm

Statistics Canada, *General Social Survey, Cycle 24: Time-Stress and Well-Being, Public Use Microdata File Documentation and User's Guide*, Catalogue no. 12M0024X.