Intercoder Reliability Report

OpenFile.ca 2011

OpenFile is an online news source that reports local news based upon community suggestions. While the website provides news coverage for 7 Canadian cities, only the Toronto coverage was considered for this project. All Toronto news items posted on the OpenFile website between January 1, 2011, and August 31, 2011, were coded by two coders. Each month was assigned an edition number. For example, news items from January were part of edition 401 and news items from February were considered part of edition 402.

The two coders received approximately 10 hours of training prior to the test. Both had already done coding work for the Local News Research Project so they were familiar with the data entry and coding processes as well as coding principles. To prepare for the test, the coders each coded the same group of 63 news items from edition 401 (January, 2011). As they worked through the training items, they compared their coding decisions and discussed discrepancies.

Two randomly selected editions were coded as test issues. The test editions were 402 (February, 2011) and 405 (June, 2011). 89 news items were coded for February and 52 news items were coded for June. The test sample represents approximately 39% of all news items coded for OpenFile, based on a total of 360 local news items from the entire study period. This exceeds the minimum of 10% required for test samples (Riffe, Aust & Lacy, 1993). Intercoder reliability was measured by testing all variables within the test issues for agreement between coders.

Prior to running the intercoder reliability tests, the two coders verified the accuracy of their own data entries. Each coder randomly selected 10 news items that they had coded. The data for all elements of the selected news items was then verified to ensure the coder had not made a mistake in entering the data. If any errors or discrepancies were found within the 10 selected items, the coder completed checking that series of 10 and then selected another 10 items and verified the data entries. Each coder repeated the process until a set of 10 randomly selected items were found to be without errors. This process helped minimize errors by coders due to data entry and provided coders with an opportunity to review their decision making.

Intercoder reliability was measured for interval variables using Lin's Concordance, with the aid of PRAM (Program for Reliability Assessment with Multiple Coders) software. Lin's Concordance measures the correlation between coder responses, and takes into account systematic coding errors (coder bias), the possible range of responses, the magnitude of difference between coders' responses, and the agreement expected by chance. We considered reliability to be acceptable at or above .700 using Lin's Concordance.

Intercoder reliability was measured for nominal variables using Cohen's Kappa, with the aid of PRAM (Program for Reliability Assessment with Multiple Coders) software. Cohen's Kappa is a relatively conservative index that measures the extent to which coders make identical coding decisions, and takes into account the agreement expected by chance. We considered reliability to be acceptable at or above .700 using Cohen's Kappa (Lombard, Snyder-Duch & Bracken, 2002).

Lin's Concordance coefficient was above .700 for all item count variables.

Variable	**Lin's Concordance Test Result**
Number of Links	0.99
Number of Spatial References	0.986

Cohen's Kappa coefficient was at or above .700 for all variables, with the exception of Race/Ethnicity 4. The Cohen's Kappa coefficient for Race/Ethnicity 4 was .498. The low level of agreement for Race/Ethnicity 4 was due to the small sample size. Only 2 news items contained entries for this variable. We decided that the Race/Ethnicity 4 variable could still be considered valid because the variable did not meet the minimum sample size of 20 required for statistical testing (Baxter, 2000; Preacher & MacCallum, 2002). Also, the other three race/ethnicity variables were coded using the same methodology and the test threshold in all of those cases was achieved.

Variable	**Cohen's Kappa Test Result**
Item Form	1.00
Interactive	1.00
Photo Status	0.98
General Photo Content	0.941
Race/Ethnicity 1	0.882
Race/Ethnicity 2	0.739
Race/Ethnicity 3	0.831
Race/Ethnicity 4	0.498
Religion	1.00
General Subject	0.738
Subject Detail	0.735

REFERENCES

Baxter, M. J., Beardah, C. C. & Westwood, S. 2000. Sample Size and Related Issues in the Analysis of Lead Isotope Data. *Journal of Archaeological Science*, 27, 973–980.

Lombard, M., Snyder-Duch, J. & Bracken, C. C. 2002. Content analysis in mass communication: Assessment and reporting of intercoder reliability. Human Communication Research 28: 587-604.

Preacher, K. J. & MacCallum, R.C. 2002. Exploratory Factor Analysis in Behavior Genetics Research: Factor Recovery with Small Sample Sizes. *Behavior Genetics*, 32(2), 153-161.

Riffe, D., Aust, C. F. & Lacy, S. 1993. The effectiveness of random, consecutive day, and constructed week sampling. Journalism Quarterly 70: 133-139.