RR-98-15

RESEARCH REPORT

COMPUTER ANALYSIS OF ESSAY CONTENT FOR AUTOMATED SCORE PREDICTION

A PROTOTYPE AUTOMATED SCORING SYSTEM FOR GMAT ANALYTICAL WRITING ASSESSMENT ESSAYS

Jill Burstein, Lisa Braden-Harder,
Martin Chodorow, Shuyi Hua,
Bruce Kaplan, Karen Kukich,
Chi Lu, James Nolan,
Don Rock, and Susanne Wolff

ETS Educational Testing Service
Princeton, New Jersey

March 1998

Computer Analysis of Essay Content for Automated Score Prediction

A Prototype Automated Scoring System for GMAT Analytical Writing Assessment Essays

Jill Burstein, Lisa Braden-Harder, Martin Chodorow, Shuyi Hua, Bruce Kaplan, Karen Kukich, Chi Lu, James Nolan, Don Rock, and Susanne Wolff[1]

[1] Authors are listed in alphabetical order. Lisa Braden-Harder is an independent consultant. Martin Chodorow is a consultant affiliated with Hunter College of the City University of New York. James Nolan is a consultant affiliated with Siena College in Loudenville, NY.

Abstract

This report discusses the development and evaluation of a research prototype system designed to automatically score essay responses to the GMAT Analytical Writing Assessments: (a). *Analysis of an Argument* (Argument essays) and (b). *Analysis of an Issue* (Issue essays) item types. The system, Electronic Essay Rater (*e-rater*), was designed to automatically analyze several features of an essay and score the essay based on the features of writing as specified in holistic rubrics. *E-rater* uses a hybrid feature methodology. It incorporates several variables that are derived statistically, extracted through NLP techniques, or achieved by simple "counting" procedures. The version of the *e-rater* described in this report uses five sets of critical feature variables to build the final linear regression model used for predicting scores. The same set of critical variables was used to fit models for the issue and argument training essays and the following results were achieved. For the set of 275 cross-validation data, exact or adjacent agreement with human rater scores reached 95%. For the 282 cross-validation issue essays exact or adjacent agreement with human rater scores achieved 93%. The rich feature variables used as score predictors in *e-rater* could potentially be used to generate explanation of score predictions, and diagnostic and instructional information.

Table of Contents

1. Introduction

This report discusses the development and evaluation of a prototype system designed to automatically score essay responses to the GMAT Analytical Writing Assessments: (a). *Analysis of an Argument* (Argument essays) and (b). *Analysis of an Issue* (Issue essays) item types.[2] The system, Electronic Essay Rater (*e-rater*)[3], was designed to automatically analyze several features of an essay and score the essay based on the features of writing as specified at each score interval in the 6-point holistic rubric (i.e., scoring guide).[4]

Previous work in automated essay scoring, such as Page and Petersen (1995), reports that predicting essay scores using surface feature variables, e.g., the fourth root of the length of an essay, shows correlations as high as .78 between a single human rater score and machine-based scores for a set of PRAXIS essays. A study done at ETS by Kaplan, et al (forthcoming), using the same set of essays, shows that using grammar checker variables in addition to word counts based on essay length yields up to 99% agreement between machine-based scores that match human rater scores within 1 point on a 6-point holistic rubric. Kaplan, et al's results using grammar checker variables have added value, since grammar checker variables may have substantive information about writing competency that might reflect rubric criteria such as, *essay is free from errors in mechanics, usage and sentence structure.*

This study was conducted to explore the extent to which essays could be automatically rated using features that reflected the 6-point holistic rubrics used by human raters to assign scores to essay. There were two primary goals underlying the development of this prototype. First, the system developed needed to be completely automated, so that it could be quickly moved into an operationally-ready mode. Secondly, it was highly desirable that the system use rubric-based features to evaluate the essay, such as rhetorical structure, topic and grammatical features. It is important that computer-based score predictions are supported by information in the rubrics that is used by human raters (see Burstein, et al (1997)). In addition, the rich feature information could potentially be used for explanation of scores, diagnostics and instructional information.

Literature in the field of discourse analysis points out that lexical (word) and structural (syntactic) features of discourse can be identified (Mann and Thompson (1988)) and represented in a machine, for computer-based analysis (Cohen (1984), Hovy, et al (1992), Hirschberg and Litman (1993), and Vander Linden and Martin (1995)). At a glance, lexical and structural features of rhetorical language did not appear in the set of responses from the issue essay item type. Rhetorical language information was obvious in the argument essays. Therefore, an information extraction program designed to identify rhetorical structure features in essays was based solely on the training response set from

[2] See Appendices A1 and A2 for test items used for this study.
[3] See Appendix D for a flowchart of *e-rater*.
[4] In the 6-point scoring guides, "1" was the lowest (worst) score and "6" was the highest (best) score.

the argument essay item type. Topic-based and syntactic structure features were identified using NLP techniques for issue and argument training responses sets. In addition, surface feature variables, such as number of words in the essay, were identified for issue and argument response sets. Rhetorical structure, topic, syntactic structure, and surface feature variables were all included in the linear regression analyses used for score prediction. Although the feature variables for rhetorical structure were specified based on the argument essays, variables from the set of rhetorical structures were also excellent predictors for responses in the issue essay item type. The same sets of critical features were used to fit linear regression analysis models to predict scores for responses to the argument and issue items. There were 403 training responses for the issue essays and 383 training responses for the argument essays. Using the same set of critical variables to fit models for the issue and argument training essays, the following results were achieved. For the set of 285 cross-validation data, exact or adjacent agreement with human rater scores reached 95%. For the 292 cross-validation issue essays exact or adjacent agreement with human rater scores achieved 93%.[5,6]

2. Methodology

E-rater uses a hybrid feature methodology. It incorporates several variables that are derived statistically, extracted through NLP techniques, or achieved by simple "counting" procedures. The version of the *e-rater* described in this report uses five sets of critical feature variables to build the final linear regression model used for predicting scores. These feature variable sets will henceforth be referred to as *predictor variables*. All predictor variables and counts of predictor variables were generated automatically by several independent computer programs. For both essay types, all relevant information about predictor variables was introduced into a stepwise linear regression in order to evaluate which variables accounted for most of the variation between essays at different score intervals. Predictor variables included in *e-rater*'s final score prediction model for argument and issue essays were: (a) surface feature variables, (b) structural features, (c) rhetorical structure analyses, (d) content vector analyses, and (e) content vector analyses by argument (henceforth, argument vector analyses). A conceptual rationale and a description of how each predictor variable was generated is described below. Also the specific features from each predictor variable set used in the final linear regression analysis model is given for the training response sets of issue and argument essays.

[5] For each essay type, human rater scores were provided for 60 essay responses at the beginning of the study. For the remaining responses in the cross-validation sets, human rater scores were provided only after machine scores predictions had been completed.

[6] *IntelliMetric* developed by TRIA Systems, Inc. in Princeton, NJ yields comparable results on the same data set. However, the specific variables used by IntelliMetric for score predictions are not available. (see Elliot, et al (1997)).

2.1 Surface Features

Breland, et al (1995) showed correlations of up to .74 between the number of words written for the College Board's English Composition Tests essay and the essay score based on a 6-point holistic rubric. Studies by Page and Petersen (1995) and Kaplan, et al (forthcoming) support Breland, et al's observation. Page and Petersen show correlations of .78 between a human rater score and a machine score based on the fourth root of the number of words in an essay. Using the same essay sample from Page and Petersen, Kaplan, et al (forthcoming) pointed out that the fourth root of the number of words was a strong score predictor variable. Using only the fourth root of the length of an essay as a variable, agreement between the human rater score and a machine score was 99% on exact or adjacent scores. In this study, correlations between a human rater score and a machine predicted score the number of words in an essay were as high as .80. Simple counts of number of words in an essay, and transformations of number of words were calculated to find surface feature variables that could be used in the score prediction model. The surface feature variables used in *e-rater* to predict argument and issue essays are shown in Table 1.

Table 1: Surface Feature Variables Used in e-rater to Predict Essay Scores

Argument Essays	• Square Root of the Total Number of Words (NN2)
Issue Essays	• Total Number of Words (NP1) • Total Number of Words to the Fifth Power (NP5) • Average Number of Sentences/Paragraph

Number of words in an essay (or some transformation thereof, such as the fourth root of the number of words) has been consistently shown to be a strong predictor of essay score. It should be kept in mind, that though it can be speculated that the more words a person writes might be related to an individual's verbal fluency, this has not been empirically established. In addition, the number of words written in an essay does not reflect any criterion in the holistic rubrics, and is therefore not a substantive predictor. Also, essay length is clearly an easily coachable variable. Number of words in an essay can be used as a variable to predict essay scores, but it should be used in combination with more substantive variables, less coachable variables. It is important to use additional, more substantive variables if richer information about scoring decisions is to be provided.

2.2 Structural Features

The scoring guides for both argument and issue essays indicate that one feature used to rate an essay is "syntactic variety." Syntactic structures in essays can be identified using NLP techniques. In this study, all sentences in the essays were parsed using the Microsoft Natural Language Processing tool (MSNLP). MSNLP takes a sentence string as input and returns a syntactically analyzed version of a sentence, as illustrated below in Figure 1. Examination of syntactic structures in essay responses yields information about the "syntactic variety" in an essay. For example, information about what types of clauses or verb types can reveal information about "syntactic variety."

Young people often express discontent with taxation levels to support the aged.

```
DECL1   NP1     AJP1    ADJ1*   "Young"
                NOUN1*  "people"
        AVP1    ADV1*   "often"
                VERB1*  "express"
        NP2     NOUN2*  "discontent"
        PP1     PP2     PREP1*  "with"
                NP3     NOUN3*  "taxation"
                        NOUN4*  "levels"
        INFCL1  INFTO1  PREP2*  "to"
                        VERB2*  "support"
                NP4     DETP1   ADJ2*   "the"
                                NOUN5*  "aged"
                CHAR1   "."
```

Figure 1: MSNLP's Syntactic Parse for a Sentence from an Issue Essay

A program was run on all the essays that counted the number of complement clauses, subordinate clauses, infinitive clauses, relative clauses and the subjunctive modal auxiliary verbs, *would*, *could*, *should*, *might* and *may,* for each sentence in an essay. The linear regression analysis selected the variables in Table 2 as predictors in the final score prediction model.

Table 2: Grammatical Structural Variables Used in e-rater to Predict Essay Scores

Argument Essays	• Total Number of Modal Auxiliary Verbs • Ratio of Complement Clauses Per Sentence
Issue Essays	• Total Number of Infinitive Clauses • Total Number of Modal Auxiliary Verbs/Paragraph

2.3 Rhetorical Structure Analysis

In both the argument and issue essays, the scoring guides indicate that an essay will receive a score based on the examinee's demonstration of a well-developed essay. For the argument essay, the scoring guide states specifically that a "6" essay "develops ideas cogently, organizes them logically, and connects them with clear transitions." The correlate to this for the issue essay would appear to be that a "6" essay "...develops a position on the issue with insightful reasons..." and that the essay "is clearly well-organized." Nolan (1997) points out that language in holistic scoring guides, such as, in this case, "cogent", "logical," "insightful," and "well-organized" have "fuzzy" meaning, since they are based on imprecise observation. Nolan uses methods of "fuzzy logic" to automatically assign these kinds of "fuzzy" classifications to essays. In this study, we try to identify organization of an essay through automated analysis of the rhetorical (or argument) structure of the essay.

The linguistic literature about rhetorical structure (Cohen (1984) Hovy, et al (1992), Hirschberg and Litman (1993), and Vander Linden and Martin (1995)) points out that rhetorical (or discourse) structure can be characterized by words, terms and syntactic structures. For instance, words and terms that provide "clues" about where a new argument starts, or how it is being developed are discussed in the literature (not surprisingly) as "clue words". We adapted the conceptual framework of conjunctive relations from Quirk, et al (1985) in which terms, such as, "In summary" and "In conclusion," which we consider to be clue words, are classified as conjuncts used for summarizing. Clue words such as "perhaps," and "possibly" are considered to be "Belief" words used by the writer to express a belief in developing an argument in the essay. Words like "this" and "these" may often be used to flag that the writer has not changed topics (Sidner (1986)). It was also observed that in certain discourse contexts that structures such as infinitive clauses (INFCL) mark the beginning of a new argument, e.g., *To experimentally support their argument Big Boards* (INFCL) *would have to do two things.*

An automated **argument partitioning and annotation program (APA)** was implemented. APA output a file for each essay that was partitioned into *argument units.*

In addition, APA output a second file in which each sentence in the essay was annotated with a word, term or structure classification that denoted argument structure. A specialized dictionary (lexicon) is used by APA to identify relevant clue words and terms.[7] APA is rule-based, and selects the dictionary-based clue words, terms, and structures based on these rules. The rules that APA uses to partition and annotate arguments denote the syntactic contexts in which clue word and terms, or specific structures denote new arguments and argument development. APA uses parsed essays to identify syntactic structures in essays.[8] Descriptions of the rules used by APA appear in Appendix B2. Figure 2 illustrates original essay text and the text output by APA with corresponding argument partitioning annotations.[9]

[7] The lexicon used *e-rater* is in Appendix B1.

[8] Parsed essays have been processed through MSNLP. Each sentence in the essay has a syntactic analysis.

[9] Wording in argument unit annotations has been revised for comprehensibility.

Figure 2: Example of Automated Argument Partitions and Annotations

Essay Text	Argument Partitioned/Annotated Essay Text
"*...Another problem with the argument is found in the evidence used to support it. Big Boards takes responsibility for increasing River City's awareness of the marathon runner over the three-month period by putting her names on billboards, but they also stated that there was "extensive national publicity" of the woman during that time. The 30% increase in awareness of this women could have been a result of the national publicity. Big Boards did not necessarily have anything to do with River City learning the woman's name - they may have learned it from TV or magazines or newspapers....*"	*Another problem with the argument is found in the evidence used to support it.* >Start Argument at *Parallel Word*: Another *Big Boards takes responsibility for increasing River City's awareness of the marathon runner over the three-month period by putting her names on billboards, but they also stated that there was "extensive national publicity" of the woman during that time.* > Argument Development at *Contrast Word*: but > Argument Development at *Complement Clause:* that... *The 30% increase in awareness of this women could have been a result of the national publicity.* > Speculation at *Auxiliary Word*: could *Big Boards did not necessarily have anything to do with River City learning the woman's name - they may have learned it from TV or magazines or newspapers* > Speculation at *Auxiliary Word*: may

Table 3 shows the rhetorical variables were used by *e-rater* for predicting scores.

Table 3: Rhetorical Structure Variable Used to Predict Scores for Issue and Argument Essays

Argument Essays	• Total Occurrences of Independent Arguments in the Final Paragraph of the Essay • Total Occurrences of Subjunctive Modal Verbs in the Final Paragraph of the Essay • Total Occurrences of Parallel Words that Begin an Independent Argument • Total Occurrences of Argument Development Using Belief Words
Issue Essays	• Total Occurrences of Arguments Starting With a Summary Phrase • Total Occurrences of Arguments Starting With an Infinitive Clause • Total Occurrences of Arguments Starting With a Subordinate Clause • Total Occurrences of Argument Development Using an Infinitive Clause • Total Occurrences of Argument Development Using a Belief Word • Total Number of Independent Arguments in the Body of the Essay, Excluding The First And Final Arguments/Paragraph

2.4 Content Vector Analysis

The scoring rubric suggests that certain ideas are expected in an essay by stating that the essay "*effectively supports the main points of the critique*" for argument essays and *explores ideas and develops a position on the issue with insightful reasons and/or persuasive examples* for the issue essays. Content vector analysis (CV) is a statistical weighting technique used to identify relationships between words and documents. With regard to the approximate specifications in the rubric about essay content, CV can used identify language (or, content words) in essays that appear to contribute to essay score.

Assigning one of six scores to a GMAT test essay is a standard type of classification problem. Statistical approaches to classification define each class (score) by the distribution of characteristics found in labeled training examples. Then, each test item is analyzed, and its distribution is compared to that of the known classes. The class which best matches the test item is selected.

For text, the characteristics may be physical (the number or length) of words, sentences, paragraphs, or documents), lexical (the particular words that occur), syntactic (the form, complexity, or variety of constructions), rhetorical (the number or type of arguments), logical (the propositional structure of the sentences), or a combination of these.

Standard content vector (CV) analysis characterizes each text document (essay) at the lexical (word) level. The document is transformed into a list of word-frequency pairs, where frequency is simply the number of times that the word appeared in the document. This list constitutes a vector which represents the lexical content of the document. Morphological analysis can optionally be used to combine the counts of inflectionally related forms so that "walks," "walked," and "walking" all contribute to the frequency of their stem, "walk." In this way, a degree of generalization is realized across morphological variants. To represent a whole class of documents, such as a score level for a set of essays, the documents in the class are concatenated and a single vector is generated to represent the class.

CV analysis actually refines this basic approach by assigning a weight to each word in the vector based on the word's salience. Salience is determined by the relative frequency of the word in the document (or class) and by the inverse of its frequency over all documents. For example, "the" may be very frequent in a given document, but its salience will be low because it appears in all documents. If the word "pterodactyl" appears even a few times in a document, it will likely have high salience because there are relatively few documents that contain this word.

A test item is compared to a class by computing a cosine correlation between their weighted vectors. The larger the value of the correlation, the closer the test item is to the class. The class which is closest to the test item is selected. These steps are summarized below.

- Vector construction for each document (or class):
    - Extract words from document (or combined documents)
    - Apply morphological analysis (optional)
    - Construct frequency vector
    - Assign weights to words to form weighted vector

- Testing:
    - Compute cosine correlation between test vector
    - and the vector of each class
    - Select class with highest correlation

As discussed in section 2.5, CV analysis can also be applied to units of text smaller than essays. For example, it can be used to evaluate the rhetorical arguments within an essay. In this case, each argument is treated like a mini-document and is compared to the classes independently of the other arguments. The result of testing is a vector of classes (scores), one for each argument in the essay.

A CV analysis computer program was written to automatically predict scores for both argument and issue essays. In the final prediction model, the scores assigned by the CV analysis are used as a predictor variable for the set of argument essays. [10]

2.5 Argument-Content Vector Analysis

An important goal of this study was to be able to predict essay scores based on 'what the writer was saying.' CV, as it is used above, identifies word associations over the essay as a whole. It looks at words randomly in the essay. Though this tells us something about possible essay content, it is important to capture words in a more structured way, so that topic may be identified using closely clustered word groupings.

The rubric specifies that relevant essay content (or, relevant words used in an essay) should be well-organized and should address relevant content. Therefore, a revised version of the content vector analysis program was implemented and run on the "argument partitioned" training essays for argument and issue essays. The purpose was to identify the associations between 'the words in each argument of an essay' and 'the essay score assigned by the human rater.' By considering vocabulary information as it is distributed in each argument of the essay, the system is closer to extracting information about topic -- or relevant essay content. For each essay, a score was generated for each argument in an essay. The vector of argument scores, followed by the mode of the argument scores is illustrated in Figure 3. The mode of the argument score vector was used as a predictor variable for the argument and issue essays in the final score prediction model.

Figure 3: Argument Score Vectors from Argument-Content Vector Analysis

Candidate Identification	Argument Score Vector	Mode
55555	3 5 5 5 4	5

[10] The 100 most heavily weighted words in each of the six scores for the GMAT argument essays are given in Appendix B1 - B12.

3. Results

As is discussed in previous sections, the same set of critical variables was used to fit a model on an independent set of training data for the issue and argument essays used in this study. Table 4 indicates the summary set of predictor variables used for the score prediction models for the issue and argument essays.

Table 4: Summary Table of Variables Used for Score Prediction for Issue and Argument Essays

	Issue Essay	Argument Essay
Surface Feature Variables	• Total Number of Words in the Essay • Number of Words in the Essays to the Fifth Power • Average Number of Sentences/Paragraph	• Square Root of the Total Number of Words in the Essay
Grammatical Structure Variables	• Total Number of Infinitive Clauses • Total Number of Subjunctive Modal Verbs/Paragraph	• Total Occurrences of Subjunctive Modal Verbs • Ratio of Complement Clauses/Sentence
Rhetorical Information Variables	• Total Occurrences of Arguments Starting With a Summary Phrase • Total Occurrences of Arguments Starting With an Infinitive Clause • Total Occurrences of Arguments Starting With a Subordinate Clause • Total Occurrences of Argument Development Using an Infinitive Clause • Total Occurrences of Argument Development Using a Belief Word • Total Number of Independent Arguments in the Body of the Essay (Excluding the first and final arguments)	• Total Occurrences of Independent Arguments in the Final Paragraph of the Essay • Total Occurrences of Subjunctive Modal Verbs in the Final Paragraph of the Essay • Total Occurrences of Parallel Words that Begin an Independent Argument • Total Occurrences of Argument Development Using Belief Words
Content Vector Analysis Variable	NO	YES
MODE (of Vector of Argument Scores)	YES	YES

Table 5 below shows the Multiple Correlation-Squared (R^2), and the percentage of agreement between two human rater scores and the machine score for the training data and the cross-validation data sets.

Table 5: Summary Results Table for Agreement Between Human Rater (HR) Scores and Machine Scores for Issue and Argument Data[11]

	Training Data		Scored Cross-Validation Data		Unscored Cross-Validation Data	
	Issue Essay (n=403)	Argument Essay (n=383)	Issue Essay (n=60)	Argument Essay (n=60)	Issue Essay (n=232)	Argument Essay (n=225)
R^2	.8646	.8175	N/A	N/A	N/A	N/A
% Agreement: HR1 and Machine	95.8	91.4	88.3	88.3	93.0	92.4
% Agreement: HR2 and Machine	94.5	91.6	93.3	95.0	92.2	94.6

Tables 6 and 7 below show the percentage agreement between human raters and *e-rater* for individual scores intervals.

Table 6: Percentage Agreement Between Human Raters and Machine Scores by Score Interval for Issue Essay Cross-Validation Sets

Score	% Agreement HR1 - M (n=60)	% Agreement HR2 - M (n=60)	% Agreement HR1 - M (n = 232)	% Agreement HR2 - M (n = 232)
1	77.8	100.0	100.0	100.0
2	77.8	80.0	98.6	92.6
3	88.9	99.0	86.0	84.7
4	90.0	88.2	90.0	95.1
5	92.2	99.0	93.0	97.7
6	100.0	100.0	93.0	92.2

[11] Human rater cores were provided for "scored cross-validation" at the beginning of the study. For the data set referred to as "unscored cross-validation data," human rater scores were provided only after machine generated score predictions were available.

Table 7: Percentage Agreement Between Human Raters and Machine Scores by Score Interval for Argument Essay Cross-Validation Sets

Score	% Agreement HR1 - M (n=60)	% Agreement HR2 - M (n=60)	% Agreement HR1 - M (n = 225)	% Agreement HR2 - M (n = 225)
1	76.6	75.0	100.0	86.9
2	100.0	100.0	89.4	91.6
3	75.0	91.6	86.6	95.4
4	83.3	100.0	100.0	91.8
5	100.0	100.0	97.2	97.5
6	100.0	100.0	92.4	100.0

3.1 Interrater Reliabilities

The following table reports summary statistics that show the percentage agreement, reliabilities (Fisher (1970)) and correlations between human rater scores, and between human rater and machine scores for the combined cross-validation data sets for issue and argument essays. The summary statistics for interrater and human rater-machine reliabilities report on exact and adjacent scores between human raters and between human raters and the machine predicted scores. These statistics suggest that agreement, reliabilities and correlations between two human raters, or between a human rater and the machine are comparable.

Table 8: Summary Statistics of Interrater Reliabilities, and Human Rater (HR) and Machine (M) Interreliabilities, Interrater Correlations and Human-Machine Correlations

	Issue Essay (n=292)			Argument Essay (n=285)		
	HR1-HR2	HR1-M	HR2-M	HR1-HR2	HR1-M	HR2-M
% Agreement (Exact Score)	56.1	40.7	47.2	56.4	43.8	48.0
% Agreement (Adjacent Score)	92.8	92.1	92.4	94.3	91.5	94.7
Reliability	0.860	0.808	0.823	0.879	0.822	0.860
Correlations	0.864	0.879	0.880	0.853	0.848	0.857

Table 9 shows the relative contribution of individual features sets as predictor variables. R^2 values in the table below show the amount of variation that can be explained by each feature set if it is used as the only predictor variable as opposed to using all the predictor variables for scoring, as is done in *e-rater*.

Table 9: R^2 for individual feature sets and the "model" for issue and argument prompts

Feature or Feature Set	R^2 - Issue Item	R^2 - Argument Item
Model (all predictor variables)	.86	.82
Model - Length (all predictor variables excluding Length)	.85	.81
Rhetorical	.80	.73
Length	.73	.68
CVA	.64	.63
CV	.60	.62
Structure	.48	.61

4. Neural Network Approaches

During the course of this study we experimented with using neural networks to predict essay scores. Our preliminary successes using in artificial neural networks (ANN) for essay score prediction warrant further examination of ANN approaches for essay scoring.

ANN's can be described as an approximate computer representation of neuro-physiological structures found in the human brain. The principle unit of an ANN is a neuron which is also the basic building block of the human brain. ANN systems try to model features resembling how the neuron operates in the human brain. Attributes of ANN's are that they can (a) handle noisy or incomplete data inputs and (b) draw inferences and make classifications based on non-overt relationships in data input. A natural application for an ANN is to use the argument score vectors in Figure 3 to predict essay scores. On the surface, the patterns of scores in the argument score vectors do not reveal any overt patterns from which one would be able to predict an essay score. For experimentation purposes, a simple neural network architecture was built to predict scores for the training and 'scored' cross validation data sets for the argument and issue essays.

A three-layer backpropogation network was used (Hecht-Nielsen (1989)). The input consisted of 7 input neurons -- one for each of the first 7 argument scores in the argument score vectors (for each candidate). Only the first 7 argument scores were used, since during training it was revealed that argument scores beyond the seventh in the vector did not seem to make a difference for score prediction. If an argument vector had fewer than 7 argument scores, the remaining places in the vector were zero-filled. The hidden layer of the network had 24 neurons which was the minimum necessary. The output layer had 1 neuron. This simple architecture is used in over 85% of working neural net applications.

The neural net was trained using the training data sets for argument essays (n=383) and issue essays (n=403). The neural net weights used in training were saved for predicting scores on both cross-validation sets for the argument and issue essays and a second cross-validation set for the argument essays and the issue essays. Accuracy results for scores predicted by the neural net are in Table 10. Final ANN predicted scores were a direct result of the neural net, and were independent from the linear regression analyses models. These results are promising and suggest that using a neural network architecture to predict scores based on the argument score vectors. Neural network approaches should be explored further using the score predictor variables used in *e-rater* to train a neural network to generate score predictions for argument and issue essays.

Table 10: ANN Score Predictions

Cross-Validation Data	ANN Score Prediction Agreement With Human Rater 1 <= 1 point (Correct Classification)	ANN Score Prediction Agreement With Human Rater 2 <= 1 point (Correct Classification)
Argument Essays (n=60)	90.6	90.6
Issue Essays (n=60)	91.7	91.7
Argument Essays (n=225)	82.2	83.1
Issue Essays (n=232)	86.2	83.6

5. Discussion and Conclusions

The development of *e-rater* positively demonstrates an automated capability for scoring general essay item types. The results of the study show that a *e-rater* can predict exact or adjacent scores with as high as 95% agreement with human rater scores. In addition, it is indicated that score reliability statistics are comparable between two human raters, and between a single human rater and a machine score. High reliabilities such as these suggest that machine score predictions might be used for quality reassurance purposes amongst raters. One of the primary benefits of *e-rater* is that it is designed to consider features of essays writing specified by holistic rubric that are used by human raters. As does a human rater, *e-rater* considers features, such as rhetorical language and structure, syntactic structure, and relevant topic content for score prediction. A second benefit of *e-rater* is that the variety of sophisticated variables that *e-rater* uses for score prediction would make it far less coachable than a system that uses only surface feature variables for score prediction. Another highlight of this system is that the rhetorical, structural, and topic features used for scoring can also be used for the automated generation of explanation of score predictions and diagnostic and instructional commentary. In some preliminary conversations with colleagues, some possibilities for diagnostics have been suggested, such as the generation of outlines of examinee essays based on rhetorical features, diagnostic comments about sentence complexity or vocabulary content in

essays. Further research needs to be done in order to appropriately design an extension to *e-rater* for generation of explanatory outputs that would be meaningful to test-takers.

Though *e-rater* has not been evaluated on additional argument and issue essay prompts, the method generalized across item types to two different stimuli. These results would suggest that the method is generalizable across prompts. To be certain, however, the system must be tested on responses from additional prompts of both argument and issue essay item types. It is important to bear in mind, however, that the set of GMAT essays used in this study were from a paper-and-pencil administration. The handwritten essays were manually transcribed and typed into electronic files. Future GMAT argument and issue essays will come from computer-delivered administrations. Essays that are a product of computer-delivered administrations may introduce new types of error, as well as additional factors that are currently unforeseeable.

Overall, *e-rater* shows extremely positive implications for computer-based writing tests and research in this area should continue in order to thoroughly explore all of the potential. Alternative approaches, such as neural networks, not currently used in *e-rater*, need to also be explored.

References

Breland, H. M., Bonner, M. W., & Kubota, M. Y. (1996). Factors in Performance on Brief, Impromptu Essay Examinations (College Board Rep. No. 95-4 and ETS Research Report 95-41). New York, NY: College Entrance Examination Board.

Burstein, J., Wolff, S., Lu, C., & Kaplan, R. (1997). An Automatic Scoring System for Advanced Placement Biology Essays. Proceedings of the Fifth Conference on Applied Natural Language Processing (pp. 174-181) . Washington D.C.: Association for Computational Linguistics.

Cohen, R. (1984). A computational theory of the function of clue words in argument understanding. Proceedings of 1984 International Computational Linguistics Conference (pp. 251-255). California.

Elliot, S., Burnham, W., Chernoff, M., & Kern, K. (1997). Computerized Scoring of Open-Ended Bar Examination Questions. Paper presented at the Annual Conference of the American Educational Research Association, Chicago, IL.

Fisher, R.A.(1970) Statistical Methods for Research Workers(14th edition) New York:Hafner Press.

Hirschberg, J., & Litman, D. (1993). Empirical Studies on the Disambiguation of Cue Phrases. Computational Linguistics 19 *(3),* 501-530.

Hecht-Nielsen, R. (1989). NEUROCOMPUTING. Reading, MA: Addison-Wesley.

Hovy, E., Lavid, J., Maier, E., Mittal, V., & Paris, C. (1992). Employing Knowledge Resources in a New Text Planner Architecture. In Dale, Hovy, Rosner & Stoch (Eds), Aspects of Automated NL Generation, (pp. 57-72). Springer-Verlag Lecture Notes in AI no.587.

Kaplan, R., Wolff, S., Burstein, J., Rock, D., Kaplan, B., & Lu, C. (in press). Scoring Essays Automatically Using Essay Features (Graduate Record Exam Monograph). Princeton, NJ: Educational Testing Service.

Mann, W. C., & Thompson, S.A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. Text 8(3), 243-281.

Nolan, J. (in press). The Architecture of a Hybrid Knowledge-Based System for Evaluating Writing Samples. Niku-Lari, A. (Ed.) Expert Systems Applications and

Artificial Intelligence Technology Transfer Series, EXPERSYS-97. Gournay S/M, France: IITT International.

Quirk, R., Greenbaum, S., Leech, G., & Svartik, J. (1985). A Comprehensive Grammar of the English Language. New York, NY: Longman.

Page, E. B., & Petersen, N. (1995). The computer moves into essay grading: updating the ancient test. Phi Delta Kappan, March , 561-565.

Sidner, C. (1986). Focusing in the Comprehension of Definite Anaphora. In B. Grosz, K. Sparck Jones, & B. L. Webber (Eds.), Readings in Natural Language Processing (pp. 363-394). Los Altos, CA: Morgan Kaufmann Publishers.

Vander Linden, K., & Martin, J.H. (1995). Expressing Rhetorical Relations in Instructional Text: A Case Study in Purpose Relation. Computational Linguistics 21 (1), 29-57.

Appendix A1: Analysis of An Argument Item

ANALYSIS OF AN ARGUMENT

Time - 30 minutes

Directions: In this section you will be asked to write a critique of the argument presented below. *You are not being asked to present your own views on the subject.*

Read the argument and the instructions that follow it, and then make any notes in your test booklet that will help you plan your response. Begin writing your response on the separate answer document. Make sure that you use the answer document that goes with this writing task.

The following is from a campaign by Big Boards, Inc., to convince companies in River City that their sales will increase if they use Big Boards billboards for advertising their locally manufactured products.

"The potential of Big Boards to increase sales of your products can be seen from an experiment we conducted last year. We increased public awareness of the name of the current national women's marathon champion by publishing her picture and her name on billboards in River City for a period of three months. Before this time, although the champion had just won her title and was receiving extensive national publicity, only five percent of 15,000 randomly surveyed residents of River City could correctly name the champion when shown her picture; after the three-month advertising experiment, 35 percent of respondents from a second survey could supply her name."

Discuss how well reasoned you find this argument. In your discussion be sure to analyze the line of reasoning and the use of evidence in the argument. For example, you may need to consider what questionable assumptions underlie the thinking and what alternative explanations or counterexamples might weaken the conclusion. You can also discuss what sort of evidence would strengthen or refute the argument, what changes in the argument would make it more sound and persuasive, and what, if anything, would help you better evaluate its conclusion.

NOTES

Use the space below or on the facing page to plan your response. Any writing on these pages will not be evaluated.

STOP

IF YOU FINISH BEFORE TIME IS CALLED, YOU MAY CHECK YOUR WORK ON THIS SECTION ONLY. DO NOT TURN TO ANY OTHER SECTION IN THE TEST.

Appendix A2: Analysis of An Issue Item

ANALYSIS OF AN ISSUE

Time - 30 minutes

Directions: In this section you will need to analyze the issue presented below and explain your views on it. The question has no "correct" answer. Instead, you should consider various perspectives as you develop your own position on the issue.

Read the statement and the instructions that follow it, and then make any notes in your test booklet that will help you plan your response. Begin writing your response on the separate answer document. Make sure that you use the answer document that goes with this writing task.

"Everywhere, it seems, there are clear and positive signs that people are becoming more respectful of one another's differences."

In your opinion, how accurate is the view expressed above? Use reasons and/or examples from your own experience, observations, or reading to develop your position.

NOTES

Use the space below or on the facing page to plan your response. Any writing on these pages will not be evaluated.

STOP

IF YOU FINISH BEFORE TIME IS CALLED, YOU MAY CHECK YOUR WORK ON THIS SECTION ONLY. DO NOT TURN TO ANY OTHER SECTION IN THE TEST.

Appendix B1: Lexicon (CLUELEX) used by APP

argument content#ALTERNATIVE_:: or, either
argument development class#BELIEF_word:: certainly, clearly, obviously, plainly, possibly, perhaps, potentially, probably, fortunately, generally, maybe, presumably, unless, albeit, luckily, unluckily, normally, for_sure, apparently, herein, likely, surely, ideally, undoubtedly, naturally
argument development class#BELIEF_phrase:: for_certain, for_sure, of_course, to_some_extent, above_all, if_only, in_order_to, in_order_for, so_that, so_as_to
argument initialization class#CLAIM_N:: argument, arguments, assumption, assumptions, claim, claims, issue, issues, evidence, evidences, idea, ideas, flaw, flaws, study, studies, point, points, position, positions, leap_of_logic, conclusion, conclusions, omission, generalization, indication, indications, deduction, passage, factor, factors, problem, problems, statement, statements, discussion, discussions, question, questions, example, examples, counterexample, counterexamples, reason, reasons
argument initialization class#CLAIM_phraseRO:: I, we
argument initialization class#CLAIM_THAT:: that
argument development class#CONTRAST_word:: otherwise, conversely, however, nonetheless, though, yet, meanwhile, while, but, instead, although, still, notwithstanding, anyway, unlike
argument development class#CONTRAST_phrase:: on_the_contrary, in_contrast, by_comparison, in_any_case, at_any_rate, in_spite_of, rather_than, on_the_other_hand, even_then, even_if, even_though, even_wordith, apart_from, instead_of
argument development class#DETAIL_word:: if, specifically, particularly, when, namely
argument development class#DETAIL_phrase:: for_example, for_instance, e.g, in_this_case, in_that_case, such_that, as_well_as, in_that, such_as, about_how, in_addition, in_addition_to
argument development class#DISBELIEF_word:: unfortunately
argument development class#EVIDENCE_word:: since, because, actually
argument development class#EVIDENCE_phrase:: in_fact, after_all, as_a_matter_of_fact, because_of
argument development class#INFERENCE_word:: accordingly, consequently, hence, thus, ultimately, so, thereby, then, therefore, following, after, afterward, afterwards
argument development class#INFERENCE_phrase:: as_a_consequence, as_a_result, if_so, if_not, as_such, according_to, in_turn, right_after
argument initialization class#PARALLEL_word:: firstly, essentially, additionally, first, second, another, third, secondly, thirdly, fourth, next, finally, final, last, lastly, moreover, too, also, likewise, similarly, initially, further, furthermore
argument initialization class#PARALLEL_phrase:: first_of_all, in_the_first_phraselace, for_one_thing, for_a_start, second_of_all, many_times, more_importantly, most_importantly

argument development class#REFORMULATION_word:: alternatively, alternately
argument development class#REFORMULATION_phrase:: that_is, in_other_words, i.e., briefly
argument development class#RHETORICAL_word:: ?, suppose, supposedly, supposing
argument development class#RHETORICAL_phrase:: what_if
argument initialization class#SUMMARY_word:: altogether, overall
argument initialization class#SUMMARY_phrase:: all_in_all, in_conclusion, in_sum, in_summary, in_summation, in_short, on_the_whole
arg_auxiliary_verb#SPECULATE_word:: might, may, should, could, would
argument initialization class#TRANSITION_phrase:: let_us

Appendix B2: Rules Used By APP

I. "AFTER" RULE

A. Extracts "after", "after", and "afterwards" if they occur sentence initially as conjunction.

II. "ALSO" RULE

A. Constrains argument extraction for "also", classified in the lexicon as arg-init#Parallel, and for additional adverbs classified as arg_dev#Belief such that all are extracted if they appear in sentence initial position or if they modify the main verb of the sentence (defined as the first verb that occurs in the second column of the parse tree).

III. LEXICALLY-BASED RULE FOR BEGINNING AN ARGUMENT

A. Constrains the extraction of nouns and pronouns classified as arg-init#CLAIM words in the lexicon to main clause subject NPs and in sentences beginning with "There" , to the position after a form of the verb "to be".

IV.

A. Controls the extraction and labeling of Nouns in arg_init position that are modified by "this" or "these" that are labeled arg_dev#SAME_TOPIC when they occur in the second or later sentence of a paragraph.

B. If "This", "These" or "It" occur as a pronoun in the first noun phrase of the parse tree of sentences that are not paragraph-initial, they are output with the label arg_dev#SAME_TOPIC. This label is generated dynamically. "This, "these" and "it" are not stored in the lexicon

V. "BUT" RULE

A. Extracts "but" if it is labeled as a conjunction.

VI. COMPLEMENT CLAUSE RULE

A. Extracts complement clauses introduced by "that" as well as complement clauses that do not begin with "that."

B. Labels complement clause as arg_init#CLAIM_THAT* when it is the first or only sentence of a paragraph, otherwise it is labeled as arg_dev#CLAIM_THAT*

C. Extracts the conjunction "that" if it occurs in a complement clause, or a complement clause not introduced by "that" under the following conditions:
   1. the complement clause is not embedded in another COMPCL or SUBCL
   2. the complement clause is not further embedded than the third column of the parse tree

VII. SUBORDINATE CLAUSE" RULE FOR BEGINNING AN ARGUMENT

A. If the very first sentence of a paragraph begins with a subordinate clause, extract the noun or pronoun from the main clause NP and consider it to be the beginning of a new argument. The noun or pronoun extracted is labeled arg_init#D-SPECIFIC if it is not listed in the lexicon.

VIII. "FIRST" RULE

A. Constrains words listed in lexicon that are classified as arg_init#Parallel words.
B. All words of this category in sentence initial position are extracted (cf. ALSO RULE).
C. If the word is not sentence-initial one of the following conditions must be satisfied.
   1. It must be in the first constituent of the parse tree,provided that the first constituent is not a subordinate clause and that it is not further embedded in the parse tree than the third column.
   2. It must be the first NP following a sentence-initial adverb.
   3. If the first constituent is the pronoun "I"followed by a verb , then the "FIRST" item must be immediately following the verb.

IX. "FURTHER" RULE

A. Extracts "further" "overall" or "altogether" if they occur sentence-initially and do not modify another constituent.

X. INFINITIVE CLAUSE RULE

A. Extracts an infinitival clause that is not further embedded than the third column of the parse tree and either follows or precedes the main verb of the sentence. The clause is not embedded in a subordinate clause or a complement clause. Infinitival clauses that are extracted are labeled as

arg_init#To-INFL if it is the first or only sentence of a paragraph, otherwise arg_dev#To_INFL.

XI. RULE FOR BEGINNING AN ARGUMENT AT A NEW PARAGRAPH

A. If a paragraph has no lexical or structural "argument initializations" then a label arg_init#NEW_PARAGRAPH is applied..

XII. "OR" RULE

A. Extracts the conjunctions "or" and "either" when they occur in the second column of the parse tree, and the node immediately following the conjunction is a verb phrase.

XIII. PARALLEL TERM RULE

A. Prevents the extraction of arg_init#Parallel lexical entries terms if they modify a verb or a noun at any level of embedding. (cf also FIRST.DOC)

XIV. "SHOULD" RULE

A. The words, *would*, *should*, *might, may*, and *could* are be picked up for each essay. They are classified as arg_aux#SPECULATE in the lexicon.

B. These words occur in parse trees in the structure

C. AUXP VERB "_____"

XV. "SO" RULE

A. Extracts the conjunction *so* if it occurs initially in a subordinate clause or if it is a sentence-initial adverb

XVI. "THEN" RULE

A. Extracts "then" if it occurs as an adverb or a conjunction that is not further embedded than the second column of the parse tree.

XVII. VERB+ING RULE

A. Extracts sentence-initial nouns and verbs ending in "-ing", as well as "-ing" verbs that immediately follow a prepositional phrase or an adverb that is in the second column of a parse tree. These extracted "-ing" words are labeled as arg_init#CLAIM_Ving if in the first or only sentence of a paragraph, and arg_dev#CLAIM_Ving otherwise.

B. If the base form of the verb is "do", then the label will be arg_dev#Inference.

XVIII. "WHEN" RULE

A. Extracts all occurrences of "when" in the following structure

1. **ABBCL* CONJUNCTION PHRASE* CONJUNCTION* "when"**

if this structure occurs no further embedded than the fourth column of the parse.

XIX. "WHILE" RULE

A. Extracts "while" under the following conditions.
1. It is the first constituent of a sentence
2. It is a conjunction in a subordinate clause that is not further embedded than the third column

Appendix C1: GRADE 6: TOP 100 WORDS BY WEIGHT - ARGUMENT ESSAYS

1	2	3	4	5	*6*	
0.000	0.038	0.063	0.187	0.412	0.690	RECOGNITION
0.047	0.114	0.107	0.195	0.160	0.317	LOCAL
0.000	0.007	0.071	0.098	0.178	0.315	CLAIM
0.229	0.163	0.207	0.207	0.295	0.304	SURVEY
0.070	0.118	0.134	0.304	0.198	0.294	CAMPAIGN
0.000	0.050	0.071	0.172	0.232	0.268	ASSUMPTION
0.000	0.057	0.045	0.109	0.141	0.243	BEEN
0.028	0.161	0.249	0.244	0.211	0.241	ARGUMENT
0.000	0.029	0.085	0.086	0.088	0.232	ATHLET
0.081	0.192	0.212	0.217	0.219	0.224	AWARENES
0.000	0.052	0.053	0.089	0.114	0.220	ASSUM
0.087	0.215	0.197	0.206	0.211	0.212	BILLBOARD
0.036	0.059	0.078	0.117	0.112	0.208	SAMPL
0.086	0.081	0.098	0.161	0.089	0.205	PUBLICITY
0.022	0.120	0.155	0.090	0.230	0.203	EVIDENC
0.000	0.027	0.081	0.096	0.126	0.198	DATA
0.000	0.070	0.138	0.076	0.095	0.197	SUPPORT
0.032	0.092	0.159	0.104	0.217	0.196	RESULT
0.180	0.138	0.129	0.163	0.138	0.184	NATIONAL
0.000	0.048	0.042	0.119	0.103	0.182	QUESTION
0.000	0.057	0.020	0.034	0.096	0.181	BOARD'
0.000	0.000	0.017	0.176	0.252	0.180	RUNNER
0.129	0.096	0.144	0.153	0.127	0.178	SECOND
0.090	0.142	0.123	0.141	0.171	0.174	SAL
0.089	0.174	0.236	0.256	0.202	0.171	INC
0.044	0.124	0.074	0.080	0.154	0.170	CONCLUSION
0.076	0.095	0.144	0.188	0.145	0.164	INFORMATION
0.041	0.000	0.019	0.147	0.092	0.163	EXPOSUR
0.110	0.096	0.079	0.072	0.119	0.163	POTENTIAL
0.000	0.061	0.072	0.084	0.094	0.162	ADDRES
0.000	0.023	0.000	0.035	0.116	0.161	STUDY
0.027	0.144	0.187	0.136	0.175	0.160	EXAMPL
0.157	0.092	0.056	0.084	0.146	0.158	THEN
0.102	0.206	0.195	0.158	0.137	0.158	PUBLIC
0.186	0.236	0.159	0.107	0.142	0.158	MARKET
0.000	0.054	0.069	0.134	0.202	0.155	PROVID
0.000	0.050	0.089	0.060	0.069	0.154	FAIL
0.000	0.031	0.027	0.061	0.062	0.153	CAUS
0.091	0.147	0.080	0.159	0.200	0.153	BILL
0.051	0.055	0.091	0.184	0.141	0.151	CONDUCT

0.152	0.132	0.258	0.279	0.205	0.150	EXPERIMENT
0.099	0.099	0.094	0.055	0.152	0.149	EVEN
0.089	0.102	0.118	0.148	0.120	0.149	FIRST
0.000	0.010	0.036	0.084	0.055	0.149	LEAD
0.130	0.125	0.190	0.093	0.118	0.148	SOME
0.122	0.148	0.151	0.143	0.141	0.148	OTHER
0.063	0.076	0.037	0.063	0.058	0.148	CASE
0.227	0.250	0.224	0.201	0.155	0.148	CHAMPION
0.061	0.179	0.149	0.110	0.154	0.147	WHO
0.000	0.034	0.050	0.017	0.052	0.146	WHIL
0.173	0.218	0.217	0.154	0.275	0.146	RIVER
0.138	0.081	0.132	0.107	0.114	0.144	SAME
0.000	0.054	0.088	0.054	0.104	0.144	DUE
0.000	0.060	0.038	0.058	0.079	0.143	FINALLY
0.237	0.167	0.113	0.143	0.207	0.142	BUSINES
0.026	0.076	0.074	0.187	0.191	0.142	DID
0.000	0.068	0.052	0.044	0.082	0.141	WHETHER
0.231	0.277	0.242	0.173	0.207	0.140	COMPANY
0.000	0.000	0.000	0.044	0.054	0.138	CORRELATION
0.000	0.010	0.026	0.167	0.186	0.137	POPULATION
0.121	0.106	0.101	0.049	0.168	0.136	MIGHT
0.028	0.098	0.121	0.125	0.191	0.136	PRESENT
0.000	0.067	0.112	0.151	0.125	0.136	FACTOR
0.000	0.047	0.131	0.081	0.178	0.136	NEED
0.161	0.165	0.158	0.138	0.136	0.135	MARATHON
0.395	0.240	0.273	0.146	0.184	0.135	PEOPL
0.032	0.078	0.100	0.104	0.126	0.134	PROV
0.000	0.014	0.000	0.021	0.064	0.133	STATISTICAL
0.000	0.000	0.032	0.064	0.084	0.132	FIGUR
0.000	0.022	0.010	0.033	0.117	0.132	INTO
0.000	0.017	0.085	0.111	0.174	0.131	FLAW
0.037	0.118	0.170	0.140	0.162	0.131	CONVINC
0.000	0.100	0.110	0.075	0.076	0.131	NUMBER
0.123	0.066	0.106	0.079	0.091	0.130	RECEIV
0.000	0.072	0.027	0.100	0.111	0.129	CONSUMER
0.137	0.089	0.079	0.094	0.209	0.129	RECOGNIZ
0.108	0.109	0.090	0.076	0.072	0.127	WELL
0.153	0.212	0.211	0.159	0.250	0.127	CITY
0.064	0.139	0.092	0.091	0.126	0.126	SPORT
0.143	0.085	0.143	0.092	0.124	0.125	LOCALLY
0.153	0.224	0.167	0.157	0.128	0.125	PRODUCT
0.000	0.000	0.018	0.070	0.174	0.125	NECESSARILY
0.078	0.147	0.211	0.152	0.144	0.125	MANUFACTUR
0.236	0.109	0.117	0.099	0.159	0.125	TIME
0.000	0.046	0.062	0.113	0.080	0.125	FACE

0.148	0.229	0.159	0.135	0.095	0.123	ABOUT
0.244	0.150	0.129	0.116	0.086	0.123	AFTER
0.074	0.130	0.124	0.146	0.229	0.123	WHAT
0.000	0.036	0.000	0.000	0.014	0.122	AUTHOR'
0.040	0.075	0.076	0.040	0.107	0.121	METHOD
0.029	0.134	0.154	0.142	0.193	0.121	MEDIA
0.169	0.130	0.080	0.082	0.094	0.121	BEFOR
0.000	0.020	0.009	0.075	0.107	0.121	PERHAP
0.000	0.000	0.036	0.051	0.000	0.120	RAIS
0.000	0.012	0.042	0.036	0.064	0.119	APPEAR
0.000	0.017	0.059	0.025	0.026	0.119	AUTHOR
0.077	0.097	0.125	0.125	0.125	0.118	BOARD
0.026	0.110	0.085	0.134	0.163	0.118	USED
0.227	0.130	0.169	0.109	0.105	0.118	RESPONDENT
0.000	0.000	0.000	0.068	0.089	0.118	COVERAG

Appendix C2: GRADE 5: TOP 100 WORDS BY WEIGHT - ARGUMENT ESSAYS

1	2	3	4	*5*	6	
0.000	0.038	0.063	0.187	0.412	0.690	RECOGNITION
0.229	0.163	0.207	0.207	0.295	0.304	SURVEY
0.173	0.218	0.217	0.154	0.275	0.146	RIVER
0.000	0.000	0.017	0.176	0.252	0.180	RUNNER
0.153	0.212	0.211	0.159	0.250	0.127	CITY
0.000	0.050	0.071	0.172	0.232	0.268	ASSUMPTION
0.022	0.120	0.155	0.090	0.230	0.203	EVIDENC
0.074	0.130	0.124	0.146	0.229	0.123	WHAT
0.081	0.192	0.212	0.217	0.219	0.224	AWARENES
0.032	0.092	0.159	0.104	0.217	0.196	RESULT
0.087	0.215	0.197	0.206	0.211	0.212	BILLBOARD
0.028	0.161	0.249	0.244	0.211	0.241	ARGUMENT
0.137	0.089	0.079	0.094	0.209	0.129	RECOGNIZ
0.237	0.167	0.113	0.143	0.207	0.142	BUSINES
0.231	0.277	0.242	0.173	0.207	0.140	COMPANY
0.152	0.132	0.258	0.279	0.205	0.150	EXPERIMENT
0.000	0.054	0.069	0.134	0.202	0.155	PROVID
0.089	0.174	0.236	0.256	0.202	0.171	INC
0.091	0.147	0.080	0.159	0.200	0.153	BILL
0.070	0.118	0.134	0.304	0.198	0.294	CAMPAIGN
0.029	0.134	0.154	0.142	0.193	0.121	MEDIA
0.028	0.098	0.121	0.125	0.191	0.136	PRESENT
0.026	0.076	0.074	0.187	0.191	0.142	DID
0.000	0.047	0.031	0.105	0.188	0.093	MANUFACTURER
0.000	0.010	0.026	0.167	0.186	0.137	POPULATION
0.000	0.011	0.030	0.034	0.184	0.090	ABILITY
0.395	0.240	0.273	0.146	0.184	0.135	PEOPL
0.091	0.049	0.087	0.086	0.181	0.082	BETTER
0.104	0.140	0.124	0.142	0.180	0.107	REASON
0.000	0.047	0.131	0.081	0.178	0.136	NEED
0.000	0.007	0.071	0.098	0.178	0.315	CLAIM
0.027	0.144	0.187	0.136	0.175	0.160	EXAMPL
0.000	0.017	0.085	0.111	0.174	0.131	FLAW
0.000	0.000	0.018	0.070	0.174	0.125	NECESSARILY
0.090	0.142	0.123	0.141	0.171	0.174	SAL
0.409	0.269	0.227	0.110	0.168	0.094	ADVERTISEMEN
0.121	0.106	0.101	0.049	0.168	0.136	MIGHT
0.169	0.196	0.165	0.198	0.167	0.106	SHOULD

0.000	0.013	0.024	0.040	0.165	0.045	LINE
0.026	0.110	0.085	0.134	0.163	0.118	USED
0.000	0.107	0.069	0.094	0.163	0.057	MENTION
0.037	0.118	0.170	0.140	0.162	0.131	CONVINC
0.488	0.444	0.188	0.189	0.162	0.101	YOU
0.124	0.167	0.127	0.161	0.161	0.112	US
0.039	0.010	0.074	0.055	0.160	0.109	GROUP
0.049	0.104	0.093	0.093	0.160	0.111	HOWEVER
0.047	0.114	0.107	0.195	0.160	0.317	LOCAL
0.000	0.000	0.000	0.021	0.160	0.081	RESPONSIBL
0.236	0.109	0.117	0.099	0.159	0.125	TIME
0.120	0.097	0.048	0.048	0.157	0.031	SEEN
0.000	0.050	0.104	0.144	0.156	0.098	ANOTHER
0.227	0.250	0.224	0.201	0.155	0.148	CHAMPION
0.061	0.179	0.149	0.110	0.154	0.147	WHO
0.044	0.124	0.074	0.080	0.154	0.170	CONCLUSION
0.000	0.084	0.056	0.032	0.153	0.061	DONE
0.099	0.099	0.094	0.055	0.152	0.149	EVEN
0.351	0.207	0.156	0.085	0.149	0.092	RESIDENT
0.157	0.092	0.056	0.084	0.146	0.158	THEN
0.076	0.095	0.144	0.188	0.145	0.164	INFORMATION
0.090	0.073	0.102	0.145	0.144	0.116	THAN
0.078	0.147	0.211	0.152	0.144	0.125	MANUFACTUR
0.041	0.098	0.048	0.065	0.142	0.045	RELAT
0.030	0.202	0.086	0.085	0.142	0.084	EFFECTIV
0.186	0.236	0.159	0.107	0.142	0.158	MARKET
0.057	0.054	0.061	0.092	0.141	0.115	FACT
0.122	0.148	0.151	0.143	0.141	0.148	OTHER
0.000	0.057	0.045	0.109	0.141	0.243	BEEN
0.051	0.055	0.091	0.184	0.141	0.151	CONDUCT
0.034	0.028	0.057	0.090	0.141	0.104	GIVEN
0.182	0.191	0.217	0.087	0.140	0.102	BECAUS
0.180	0.138	0.129	0.163	0.138	0.184	NATIONAL
0.102	0.206	0.195	0.158	0.137	0.158	PUBLIC
0.161	0.165	0.158	0.138	0.136	0.135	MARATHON
0.000	0.071	0.063	0.054	0.133	0.085	ISSU
0.151	0.068	0.108	0.096	0.130	0.113	SUCH
0.000	0.147	0.080	0.042	0.130	0.085	STATEMENT
0.153	0.224	0.167	0.157	0.128	0.125	PRODUCT
0.129	0.096	0.144	0.153	0.127	0.178	SECOND
0.116	0.141	0.055	0.135	0.126	0.065	MUCH
0.000	0.027	0.081	0.096	0.126	0.198	DATA
0.032	0.078	0.100	0.104	0.126	0.134	PROV
0.064	0.139	0.092	0.091	0.126	0.126	SPORT
0.129	0.035	0.069	0.084	0.126	0.115	MAK

0.000	0.034	0.069	0.097	0.125	0.100	STRENGTHEN
0.096	0.026	0.076	0.161	0.125	0.118	STAT
0.077	0.097	0.125	0.125	0.125	0.118	BOARD
0.000	0.067	0.112	0.151	0.125	0.136	FACTOR
0.143	0.085	0.143	0.092	0.124	0.125	LOCALLY
0.000	0.043	0.115	0.057	0.124	0.081	AWAR
0.000	0.025	0.033	0.065	0.123	0.072	IDENTIFY
0.040	0.021	0.028	0.048	0.122	0.093	POSSIBL
0.553	0.171	0.208	0.086	0.122	0.076	PICTUR
0.062	0.184	0.067	0.088	0.122	0.045	TRY
0.000	0.053	0.066	0.080	0.122	0.080	SELL
0.117	0.095	0.161	0.142	0.121	0.102	BUY
0.089	0.102	0.118	0.148	0.120	0.149	FIRST
0.000	0.000	0.000	0.015	0.120	0.057	ATHELET
0.000	0.000	0.000	0.000	0.120	0.098	POLL
0.000	0.033	0.030	0.150	0.120	0.070	FORM
0.000	0.044	0.085	0.098	0.120	0.098	TWO

Appendix C3: GRADE 4: TOP 100 WORDS BY WEIGHT - ARGUMENT ESSAYS

1	2	3	*4*	5	6	
0.070	0.118	0.134	0.304	0.198	0.294	CAMPAIGN
0.152	0.132	0.258	0.279	0.205	0.150	EXPERIMENT
0.089	0.174	0.236	0.256	0.202	0.171	INC
0.028	0.161	0.249	0.244	0.211	0.241	ARGUMENT
0.508	0.253	0.242	0.241	0.052	0.046	PERCENT
0.081	0.192	0.212	0.217	0.219	0.224	AWARENES
0.229	0.163	0.207	0.207	0.295	0.304	SURVEY
0.087	0.215	0.197	0.206	0.211	0.212	BILLBOARD
0.051	0.000	0.134	0.205	0.021	0.017	30
0.227	0.250	0.224	0.201	0.155	0.148	CHAMPION
0.169	0.196	0.165	0.198	0.167	0.106	SHOULD
0.047	0.114	0.107	0.195	0.160	0.317	LOCAL
0.243	0.252	0.108	0.189	0.100	0.109	YOUR
0.488	0.444	0.188	0.189	0.162	0.101	YOU
0.076	0.095	0.144	0.188	0.145	0.164	INFORMATION
0.000	0.038	0.063	0.187	0.412	0.690	RECOGNITION
0.026	0.076	0.074	0.187	0.191	0.142	DID
0.051	0.055	0.091	0.184	0.141	0.151	CONDUCT
0.000	0.000	0.040	0.180	0.080	0.081	&
0.000	0.000	0.017	0.176	0.252	0.180	RUNNER
0.000	0.058	0.064	0.173	0.077	0.078	ARGUEMENT
0.231	0.277	0.242	0.173	0.207	0.140	COMPANY
0.000	0.050	0.071	0.172	0.232	0.268	ASSUMPTION
0.130	0.315	0.194	0.170	0.080	0.069	CUSTOMER
0.000	0.010	0.026	0.167	0.186	0.137	POPULATION
0.180	0.138	0.129	0.163	0.138	0.184	NATIONAL
0.086	0.081	0.098	0.161	0.089	0.205	PUBLICITY
0.096	0.026	0.076	0.161	0.125	0.118	STAT
0.124	0.167	0.127	0.161	0.161	0.112	US
0.091	0.147	0.080	0.159	0.200	0.153	BILL
0.153	0.212	0.211	0.159	0.250	0.127	CITY
0.102	0.206	0.195	0.158	0.137	0.158	PUBLIC
0.460	0.181	0.067	0.157	0.015	0.040	GET
0.153	0.224	0.167	0.157	0.128	0.125	PRODUCT
0.173	0.218	0.217	0.154	0.275	0.146	RIVER
0.129	0.096	0.144	0.153	0.127	0.178	SECOND
0.078	0.147	0.211	0.152	0.144	0.125	MANUFACTUR
0.000	0.067	0.112	0.151	0.125	0.136	FACTOR
0.000	0.033	0.030	0.150	0.120	0.070	FORM

0.177	0.135	0.106	0.149	0.116	0.079	PERSON
0.089	0.102	0.118	0.148	0.120	0.149	FIRST
0.041	0.000	0.019	0.147	0.092	0.163	EXPOSUR
0.074	0.130	0.124	0.146	0.229	0.123	WHAT
0.167	0.015	0.013	0.146	0.023	0.031	COMMUNITY
0.395	0.240	0.273	0.146	0.184	0.135	PEOPL
0.090	0.073	0.102	0.145	0.144	0.116	THAN
0.108	0.019	0.095	0.145	0.059	0.109	SUCCES
0.000	0.050	0.104	0.144	0.156	0.098	ANOTHER
0.237	0.167	0.113	0.143	0.207	0.142	BUSINES
0.122	0.148	0.151	0.143	0.141	0.148	OTHER
0.104	0.140	0.124	0.142	0.180	0.107	REASON
0.117	0.095	0.161	0.142	0.121	0.102	BUY
0.029	0.134	0.154	0.142	0.193	0.121	MEDIA
0.090	0.142	0.123	0.141	0.171	0.174	SAL
0.145	0.207	0.149	0.141	0.069	0.092	VERY
0.037	0.118	0.170	0.140	0.162	0.131	CONVINC
0.161	0.165	0.158	0.138	0.136	0.135	MARATHON
0.091	0.024	0.065	0.137	0.056	0.051	KNEW
0.027	0.144	0.187	0.136	0.175	0.160	EXAMPL
0.116	0.141	0.055	0.135	0.126	0.065	MUCH
0.148	0.229	0.159	0.135	0.095	0.123	ABOUT
0.000	0.054	0.069	0.134	0.202	0.155	PROVID
0.026	0.110	0.085	0.134	0.163	0.118	USED
0.097	0.078	0.092	0.130	0.093	0.069	OUT
0.234	0.189	0.148	0.129	0.107	0.113	ADVERTIS
0.159	0.192	0.123	0.128	0.033	0.031	REMEMBER
0.028	0.098	0.121	0.125	0.191	0.136	PRESENT
0.036	0.049	0.070	0.125	0.082	0.110	BE
0.077	0.097	0.125	0.125	0.125	0.118	BOARD
0.000	0.083	0.042	0.124	0.018	0.044	SOUND
0.097	0.061	0.069	0.123	0.066	0.101	WHER
0.000	0.071	0.072	0.121	0.054	0.072	SAY
0.071	0.190	0.186	0.121	0.087	0.075	COST
0.106	0.076	0.101	0.121	0.036	0.067	TITL
0.000	0.048	0.042	0.119	0.103	0.182	QUESTION
0.154	0.141	0.110	0.118	0.038	0.055	THEREFOR
0.036	0.059	0.078	0.117	0.112	0.208	SAMPL
0.244	0.150	0.129	0.116	0.086	0.123	AFTER
0.000	0.046	0.062	0.113	0.080	0.125	FACE
0.000	0.074	0.056	0.111	0.040	0.079	SINC
0.000	0.017	0.085	0.111	0.174	0.131	FLAW
0.068	0.000	0.065	0.110	0.112	0.111	ASK
0.409	0.269	0.227	0.110	0.168	0.094	ADVERTISEMEN
0.061	0.179	0.149	0.110	0.154	0.147	WHO

0.000	0.057	0.045	0.109	0.141	0.243	BEEN
0.202	0.181	0.105	0.109	0.056	0.060	LIKE
0.227	0.130	0.169	0.109	0.105	0.118	RESPONDENT
0.186	0.183	0.158	0.108	0.085	0.109	THRE
0.138	0.081	0.132	0.107	0.114	0.144	SAME
0.000	0.044	0.059	0.107	0.067	0.069	WEAK
0.186	0.236	0.159	0.107	0.142	0.158	MARKET
0.000	0.094	0.028	0.105	0.048	0.013	ADVERTIZ
0.000	0.025	0.023	0.105	0.029	0.095	SOURC
0.087	0.085	0.172	0.105	0.089	0.084	DIFFERENT
0.139	0.154	0.129	0.105	0.100	0.113	JUST
0.000	0.047	0.031	0.105	0.188	0.093	MANUFACTURER
0.044	0.082	0.031	0.105	0.072	0.073	WOMAN
0.032	0.092	0.159	0.104	0.217	0.196	RESULT
0.032	0.078	0.100	0.104	0.126	0.134	PROV
0.137	0.055	0.107	0.104	0.078	0.058	NEWSPAPER

Appendix C4: GRADE 3: TOP 100 WORDS BY WEIGHT - ARGUMENT ESSAYS

1	2	*3*	4	5	6	
0.395	0.240	0.273	0.146	0.184	0.135	PEOPL
0.152	0.132	0.258	0.279	0.205	0.150	EXPERIMENT
0.028	0.161	0.249	0.244	0.211	0.241	ARGUMENT
0.508	0.253	0.242	0.241	0.052	0.046	PERCENT
0.231	0.277	0.242	0.173	0.207	0.140	COMPANY
0.089	0.174	0.236	0.256	0.202	0.171	INC
0.409	0.269	0.227	0.110	0.168	0.094	ADVERTISEMEN
0.227	0.250	0.224	0.201	0.155	0.148	CHAMPION
0.182	0.191	0.217	0.087	0.140	0.102	BECAUS
0.173	0.218	0.217	0.154	0.275	0.146	RIVER
0.081	0.192	0.212	0.217	0.219	0.224	AWARENES
0.078	0.147	0.211	0.152	0.144	0.125	MANUFACTUR
0.153	0.212	0.211	0.159	0.250	0.127	CITY
0.553	0.171	0.208	0.086	0.122	0.076	PICTUR
0.229	0.163	0.207	0.207	0.295	0.304	SURVEY
0.045	0.012	0.205	0.082	0.047	0.071	AD
0.368	0.129	0.202	0.030	0.038	0.012	PUBLISH
0.087	0.215	0.197	0.206	0.211	0.212	BILLBOARD
0.102	0.206	0.195	0.158	0.137	0.158	PUBLIC
0.130	0.315	0.194	0.170	0.080	0.069	CUSTOMER
0.171	0.138	0.190	0.011	0.012	0.006	FAMOU
0.130	0.125	0.190	0.093	0.118	0.148	SOME
0.328	0.203	0.188	0.033	0.020	0.040	THINK
0.488	0.444	0.188	0.189	0.162	0.101	YOU
0.102	0.064	0.187	0.062	0.056	0.076	GO
0.027	0.144	0.187	0.136	0.175	0.160	EXAMPL
0.071	0.190	0.186	0.121	0.087	0.075	COST
0.087	0.085	0.172	0.105	0.089	0.084	DIFFERENT
0.037	0.118	0.170	0.140	0.162	0.131	CONVINC
0.227	0.130	0.169	0.109	0.105	0.118	RESPONDENT
0.153	0.224	0.167	0.157	0.128	0.125	PRODUCT
0.082	0.110	0.166	0.041	0.051	0.051	NEW
0.109	0.078	0.165	0.059	0.030	0.069	GIVE
0.169	0.196	0.165	0.198	0.167	0.106	SHOULD
0.237	0.151	0.162	0.089	0.049	0.026	CURRENT
0.117	0.095	0.161	0.142	0.121	0.102	BUY
0.000	0.070	0.160	0.060	0.038	0.062	ABOV
0.037	0.130	0.160	0.037	0.084	0.054	TV
0.186	0.236	0.159	0.107	0.142	0.158	MARKET

0.032	0.092	0.159	0.104	0.217	0.196	RESULT
0.148	0.229	0.159	0.135	0.095	0.123	ABOUT
0.186	0.183	0.158	0.108	0.085	0.109	THRE
0.161	0.165	0.158	0.138	0.136	0.135	MARATHON
0.207	0.167	0.157	0.101	0.096	0.114	MONTH
0.351	0.207	0.156	0.085	0.149	0.092	RESIDENT
0.022	0.120	0.155	0.090	0.230	0.203	EVIDENC
0.029	0.134	0.154	0.142	0.193	0.121	MEDIA
0.250	0.190	0.154	0.096	0.086	0.065	WOMEN'
0.135	0.181	0.153	0.068	0.062	0.076	WORK
0.122	0.148	0.151	0.143	0.141	0.148	OTHER
0.196	0.228	0.150	0.101	0.098	0.080	SO
0.061	0.179	0.149	0.110	0.154	0.147	WHO
0.145	0.207	0.149	0.141	0.069	0.092	VERY
0.219	0.098	0.148	0.051	0.060	0.069	INTEREST
0.234	0.189	0.148	0.129	0.107	0.113	ADVERTIS
0.129	0.096	0.144	0.153	0.127	0.178	SECOND
0.076	0.095	0.144	0.188	0.145	0.164	INFORMATION
0.143	0.085	0.143	0.092	0.124	0.125	LOCALLY
0.130	0.091	0.143	0.084	0.096	0.099	MOST
0.091	0.049	0.140	0.036	0.084	0.030	TOO
0.126	0.068	0.140	0.059	0.017	0.047	LOT
0.206	0.120	0.139	0.069	0.078	0.069	TH
0.000	0.070	0.138	0.076	0.095	0.197	SUPPORT
0.038	0.071	0.136	0.069	0.086	0.064	BECOM
0.444	0.327	0.134	0.094	0.115	0.086	GOOD
0.051	0.000	0.134	0.205	0.021	0.017	30
0.070	0.118	0.134	0.304	0.198	0.294	CAMPAIGN
0.000	0.116	0.133	0.095	0.053	0.058	DOESN'T
0.335	0.170	0.133	0.075	0.023	0.017	35
0.138	0.081	0.132	0.107	0.114	0.144	SAME
0.194	0.119	0.132	0.056	0.085	0.105	IMPORTANT
0.000	0.047	0.131	0.081	0.178	0.136	NEED
0.244	0.150	0.129	0.116	0.086	0.123	AFTER
0.180	0.138	0.129	0.163	0.138	0.184	NATIONAL
0.139	0.154	0.129	0.105	0.100	0.113	JUST
0.142	0.183	0.129	0.074	0.082	0.102	000
0.124	0.167	0.127	0.161	0.161	0.112	US
0.077	0.097	0.125	0.125	0.125	0.118	BOARD
0.038	0.071	0.125	0.045	0.070	0.076	COMPAR
0.000	0.250	0.125	0.023	0.000	0.013	SOCIETY
0.074	0.130	0.124	0.146	0.229	0.123	WHAT
0.104	0.140	0.124	0.142	0.180	0.107	REASON
0.195	0.105	0.124	0.085	0.080	0.106	POINT
0.325	0.184	0.124	0.085	0.027	0.044	DON'T

0.090	0.142	0.123	0.141	0.171	0.174	SAL
0.159	0.192	0.123	0.128	0.033	0.031	REMEMBER
0.028	0.098	0.121	0.125	0.191	0.136	PRESENT
0.184	0.086	0.120	0.009	0.094	0.021	IT'S
0.031	0.127	0.120	0.070	0.116	0.091	THOS
0.089	0.102	0.118	0.148	0.120	0.149	FIRST
0.000	0.000	0.117	0.000	0.029	0.008	SPECIFY
0.236	0.109	0.117	0.099	0.159	0.125	TIME
0.000	0.043	0.115	0.057	0.124	0.081	AWAR
0.030	0.120	0.114	0.066	0.104	0.067	SHOW
0.113	0.091	0.114	0.062	0.093	0.082	PERIOD
0.000	0.070	0.114	0.053	0.090	0.088	SERVIC
0.237	0.167	0.113	0.143	0.207	0.142	BUSINES
0.034	0.063	0.113	0.102	0.104	0.079	ORDER
0.040	0.085	0.113	0.048	0.033	0.075	CHAMPION'
0.000	0.067	0.112	0.151	0.125	0.136	FACTOR

Appendix C5: GRADE 2: TOP 100 WORDS BY WEIGHT - ARGUMENT ESSAYS

1	*2*	3	4	5	6	
0.488	0.444	0.188	0.189	0.162	0.101	YOU
0.291	0.360	0.090	0.041	0.048	0.078	WAY
0.444	0.327	0.134	0.094	0.115	0.086	GOOD
0.130	0.315	0.194	0.170	0.080	0.069	CUSTOMER
0.231	0.277	0.242	0.173	0.207	0.140	COMPANY
0.409	0.269	0.227	0.110	0.168	0.094	ADVERTISEMEN
0.508	0.253	0.242	0.241	0.052	0.046	PERCENT
0.243	0.252	0.108	0.189	0.100	0.109	YOUR
0.000	0.250	0.125	0.023	0.000	0.013	SOCIETY
0.227	0.250	0.224	0.201	0.155	0.148	CHAMPION
0.395	0.240	0.273	0.146	0.184	0.135	PEOPL
0.186	0.236	0.159	0.107	0.142	0.158	MARKET
0.148	0.229	0.159	0.135	0.095	0.123	ABOUT
0.196	0.228	0.150	0.101	0.098	0.080	SO
0.153	0.224	0.167	0.157	0.128	0.125	PRODUCT
0.173	0.218	0.217	0.154	0.275	0.146	RIVER
0.087	0.215	0.197	0.206	0.211	0.212	BILLBOARD
0.153	0.212	0.211	0.159	0.250	0.127	CITY
0.000	0.208	0.017	0.014	0.000	0.024	PUBLICATION
0.351	0.207	0.156	0.085	0.149	0.092	RESIDENT
0.145	0.207	0.149	0.141	0.069	0.092	VERY
0.102	0.206	0.195	0.158	0.137	0.158	PUBLIC
0.328	0.203	0.188	0.033	0.020	0.040	THINK
0.030	0.202	0.086	0.085	0.142	0.084	EFFECTIV
0.169	0.196	0.165	0.198	0.167	0.106	SHOULD
0.000	0.194	0.091	0.038	0.117	0.072	PROMOT
0.120	0.194	0.048	0.057	0.058	0.090	MY
0.081	0.192	0.212	0.217	0.219	0.224	AWARENES
0.159	0.192	0.123	0.128	0.033	0.031	REMEMBER
0.182	0.191	0.217	0.087	0.140	0.102	BECAUS
0.071	0.190	0.186	0.121	0.087	0.075	COST
0.250	0.190	0.154	0.096	0.086	0.065	WOMEN'
0.234	0.189	0.148	0.129	0.107	0.113	ADVERTIS
0.144	0.186	0.110	0.058	0.065	0.087	15
0.062	0.184	0.067	0.088	0.122	0.045	TRY
0.325	0.184	0.124	0.085	0.027	0.044	DON'T
0.186	0.183	0.158	0.108	0.085	0.109	THRE
0.142	0.183	0.129	0.074	0.082	0.102	000
0.135	0.181	0.153	0.068	0.062	0.076	WORK
0.460	0.181	0.067	0.157	0.015	0.040	GET

0.084	0.181	0.080	0.076	0.043	0.019	SALE
0.202	0.181	0.105	0.109	0.056	0.060	LIKE
0.056	0.180	0.027	0.034	0.011	0.031	ATTRACT
0.000	0.180	0.071	0.090	0.054	0.058	KNOWN
0.222	0.179	0.106	0.095	0.103	0.056	THEM
0.061	0.179	0.149	0.110	0.154	0.147	WHO
0.089	0.174	0.236	0.256	0.202	0.171	INC
0.000	0.172	0.000	0.000	0.016	0.000	BIGBOARD
0.553	0.171	0.208	0.086	0.122	0.076	PICTUR
0.335	0.170	0.133	0.075	0.023	0.017	35
0.237	0.167	0.113	0.143	0.207	0.142	BUSINES
0.124	0.167	0.127	0.161	0.161	0.112	US
0.207	0.167	0.157	0.101	0.096	0.114	MONTH
0.265	0.166	0.074	0.018	0.027	0.030	KIND
0.161	0.165	0.158	0.138	0.136	0.135	MARATHON
0.229	0.163	0.207	0.207	0.295	0.304	SURVEY
0.050	0.162	0.024	0.010	0.041	0.090	OUR
0.028	0.161	0.249	0.244	0.211	0.241	ARGUMENT
0.352	0.158	0.084	0.077	0.084	0.069	SEE
0.000	0.156	0.055	0.012	0.012	0.026	ALWAY
0.036	0.155	0.069	0.051	0.045	0.089	5%
0.139	0.154	0.129	0.105	0.100	0.113	JUST
0.066	0.151	0.111	0.087	0.075	0.056	ENOUGH
0.237	0.151	0.162	0.089	0.049	0.026	CURRENT
0.244	0.150	0.129	0.116	0.086	0.123	AFTER
0.000	0.148	0.047	0.055	0.056	0.092	PROBLEM
0.050	0.148	0.048	0.020	0.072	0.017	AROUND
0.069	0.148	0.057	0.090	0.106	0.085	HELP
0.122	0.148	0.151	0.143	0.141	0.148	OTHER
0.055	0.147	0.026	0.011	0.022	0.018	EASILY
0.000	0.147	0.052	0.011	0.056	0.012	RIGHT
0.000	0.147	0.080	0.042	0.130	0.085	STATEMENT
0.091	0.147	0.080	0.159	0.200	0.153	BILL
0.078	0.147	0.211	0.152	0.144	0.125	MANUFACTUR
0.000	0.146	0.000	0.014	0.000	0.023	FIELD
0.000	0.146	0.071	0.030	0.041	0.017	DEPEND
0.039	0.146	0.055	0.070	0.032	0.104	COMPANY'
0.465	0.145	0.105	0.099	0.096	0.088	WHEN
0.027	0.144	0.187	0.136	0.175	0.160	EXAMPL
0.090	0.142	0.123	0.141	0.171	0.174	SAL
0.203	0.142	0.068	0.033	0.058	0.041	SUPPLY
0.000	0.142	0.050	0.021	0.011	0.030	HAND
0.154	0.141	0.110	0.118	0.038	0.055	THEREFOR
0.116	0.141	0.055	0.135	0.126	0.065	MUCH
0.104	0.140	0.124	0.142	0.180	0.107	REASON

0.064	0.139	0.092	0.091	0.126	0.126	SPORT
0.000	0.139	0.037	0.042	0.096	0.012	FAMILIAR
0.171	0.138	0.190	0.011	0.012	0.006	FAMOU
0.180	0.138	0.129	0.163	0.138	0.184	NATIONAL
0.177	0.135	0.106	0.149	0.116	0.079	PERSON
0.029	0.134	0.154	0.142	0.193	0.121	MEDIA
0.152	0.132	0.258	0.279	0.205	0.150	EXPERIMENT
0.037	0.130	0.160	0.037	0.084	0.054	TV
0.227	0.130	0.169	0.109	0.105	0.118	RESPONDENT
0.169	0.130	0.080	0.082	0.094	0.121	BEFOR
0.074	0.130	0.124	0.146	0.229	0.123	WHAT
0.131	0.129	0.031	0.088	0.081	0.029	MAYB
0.431	0.129	0.023	0.068	0.020	0.043	READ
0.368	0.129	0.202	0.030	0.038	0.012	PUBLISH
0.000	0.128	0.047	0.064	0.082	0.058	ABLE

Appendix C6: GRADE 1: TOP 100 WORDS BY WEIGHT - ARGUMENT ESSAYS

1	2	3	4	5	6	
0.553	0.171	0.208	0.086	0.122	0.076	PICTUR
0.508	0.253	0.242	0.241	0.052	0.046	PERCENT
0.488	0.444	0.188	0.189	0.162	0.101	YOU
0.465	0.145	0.105	0.099	0.096	0.088	WHEN
0.460	0.181	0.067	0.157	0.015	0.040	GET
0.444	0.327	0.134	0.094	0.115	0.086	GOOD
0.431	0.129	0.023	0.068	0.020	0.043	READ
0.409	0.269	0.227	0.110	0.168	0.094	ADVERTISEMEN
0.395	0.240	0.273	0.146	0.184	0.135	PEOPL
0.368	0.129	0.202	0.030	0.038	0.012	PUBLISH
0.361	0.119	0.096	0.032	0.066	0.018	CORRECTLY
0.352	0.158	0.084	0.077	0.084	0.069	SEE
0.351	0.207	0.156	0.085	0.149	0.092	RESIDENT
0.335	0.170	0.133	0.075	0.023	0.017	35
0.328	0.203	0.188	0.033	0.020	0.040	THINK
0.326	0.125	0.067	0.037	0.067	0.021	IDEA
0.325	0.184	0.124	0.085	0.027	0.044	DON'T
0.291	0.360	0.090	0.041	0.048	0.078	WAY
0.290	0.067	0.040	0.042	0.077	0.051	WHY
0.281	0.000	0.033	0.000	0.029	0.016	UNDERSTAND
0.272	0.097	0.108	0.027	0.037	0.015	EVERY
0.271	0.018	0.065	0.014	0.000	0.008	SOMETIM
0.265	0.166	0.074	0.018	0.027	0.030	KIND
0.256	0.052	0.031	0.013	0.013	0.000	FUTUR
0.255	0.041	0.036	0.051	0.010	0.034	5
0.253	0.126	0.086	0.058	0.045	0.053	RANDOMLY
0.250	0.190	0.154	0.096	0.086	0.065	WOMEN'
0.246	0.121	0.108	0.074	0.034	0.028	FIVE
0.244	0.150	0.129	0.116	0.086	0.123	AFTER
0.243	0.252	0.108	0.189	0.100	0.109	YOUR
0.240	0.022	0.000	0.016	0.016	0.000	AWAY
0.237	0.167	0.113	0.143	0.207	0.142	BUSINES
0.237	0.151	0.162	0.089	0.049	0.026	CURRENT
0.236	0.109	0.117	0.099	0.159	0.125	TIME
0.234	0.189	0.148	0.129	0.107	0.113	ADVERTIS
0.233	0.075	0.055	0.094	0.048	0.052	QUALITY

0.232	0.047	0.014	0.070	0.036	0.007	EASY
0.231	0.277	0.242	0.173	0.207	0.140	COMPANY
0.229	0.163	0.207	0.207	0.295	0.304	SURVEY
0.227	0.250	0.224	0.201	0.155	0.148	CHAMPION
0.227	0.130	0.169	0.109	0.105	0.118	RESPONDENT
0.222	0.179	0.106	0.095	0.103	0.056	THEM
0.219	0.098	0.148	0.051	0.060	0.069	INTEREST
0.211	0.038	0.000	0.028	0.000	0.016	FAVOR
0.211	0.076	0.017	0.014	0.029	0.000	BILLBORD
0.207	0.022	0.040	0.050	0.034	0.102	RANDOM
0.207	0.167	0.157	0.101	0.096	0.114	MONTH
0.206	0.120	0.139	0.069	0.078	0.069	TH
0.203	0.142	0.068	0.033	0.058	0.041	SUPPLY
0.202	0.181	0.105	0.109	0.056	0.060	LIKE
0.198	0.096	0.095	0.080	0.057	0.036	PUT
0.197	0.035	0.000	0.080	0.014	0.007	GOT
0.196	0.228	0.150	0.101	0.098	0.080	SO
0.195	0.105	0.124	0.085	0.080	0.106	POINT
0.194	0.119	0.132	0.056	0.085	0.105	IMPORTANT
0.187	0.100	0.000	0.038	0.038	0.007	ADVERTISMENT
0.187	0.000	0.015	0.038	0.051	0.007	APPROACH
0.186	0.236	0.159	0.107	0.142	0.158	MARKET
0.186	0.183	0.158	0.108	0.085	0.109	THRE
0.184	0.086	0.120	0.009	0.094	0.021	IT'S
0.182	0.191	0.217	0.087	0.140	0.102	BECAUS
0.180	0.138	0.129	0.163	0.138	0.184	NATIONAL
0.177	0.135	0.106	0.149	0.116	0.079	PERSON
0.173	0.218	0.217	0.154	0.275	0.146	RIVER
0.171	0.138	0.190	0.011	0.012	0.006	FAMOU
0.171	0.046	0.000	0.046	0.012	0.038	RESPOND
0.169	0.196	0.165	0.198	0.167	0.106	SHOULD
0.169	0.130	0.080	0.082	0.094	0.121	BEFOR
0.167	0.015	0.013	0.146	0.023	0.031	COMMUNITY
0.164	0.044	0.013	0.022	0.034	0.037	GREAT
0.164	0.053	0.086	0.066	0.081	0.096	ALTHOUGH
0.161	0.165	0.158	0.138	0.136	0.135	MARATHON
0.160	0.000	0.000	0.016	0.000	0.009	SAFE
0.159	0.192	0.123	0.128	0.033	0.031	REMEMBER
0.157	0.092	0.056	0.084	0.146	0.158	THEN
0.156	0.059	0.074	0.081	0.038	0.098	EXTENSIV
0.155	0.084	0.049	0.021	0.021	0.041	MIND
0.155	0.084	0.074	0.031	0.053	0.017	PROMOTION
0.154	0.141	0.110	0.118	0.038	0.055	THEREFOR
0.153	0.224	0.167	0.157	0.128	0.125	PRODUCT
0.153	0.212	0.211	0.159	0.250	0.127	CITY

0.153	0.093	0.091	0.046	0.079	0.047	SHOWN
0.152	0.132	0.258	0.279	0.205	0.150	EXPERIMENT
0.152	0.020	0.000	0.000	0.016	0.017	PARTY
0.152	0.020	0.036	0.000	0.000	0.009	IMPORTANC
0.151	0.068	0.108	0.096	0.130	0.113	SUCH
0.150	0.027	0.024	0.020	0.041	0.056	BEGIN
0.148	0.229	0.159	0.135	0.095	0.123	ABOUT
0.146	0.039	0.000	0.029	0.000	0.000	ONE'
0.146	0.000	0.017	0.000	0.030	0.008	PUSH
0.145	0.207	0.149	0.141	0.069	0.092	VERY
0.144	0.107	0.060	0.058	0.097	0.073	AREA
0.144	0.090	0.057	0.048	0.010	0.043	REALLY
0.144	0.186	0.110	0.058	0.065	0.087	15
0.143	0.085	0.143	0.092	0.124	0.125	LOCALLY
0.142	0.038	0.068	0.076	0.107	0.026	EACH
0.142	0.183	0.129	0.074	0.082	0.102	000
0.140	0.000	0.067	0.014	0.000	0.008	SPECIAL
0.140	0.088	0.067	0.056	0.038	0.031	BEST
0.139	0.154	0.129	0.105	0.100	0.113	JUST

Appendix C7: GRADE 6: TOP 100 WORDS BY WEIGHT - ISSUE ESSAYS

1	2	3	4	5	*6*	
0.017	0.095	0.113	0.205	0.237	0.366	OUR
0.000	0.007	0.027	0.188	0.110	0.284	&
0.035	0.117	0.072	0.163	0.145	0.282	SOCIETY
0.022	0.035	0.079	0.171	0.285	0.268	GROUP
0.000	0.014	0.032	0.086	0.138	0.253	RELIGIOU
0.021	0.006	0.090	0.095	0.165	0.251	WOMEN
0.000	0.021	0.031	0.116	0.139	0.239	ETHNIC
0.013	0.039	0.054	0.072	0.155	0.228	THOS
0.000	0.006	0.028	0.069	0.124	0.228	INTOLERANC
0.000	0.028	0.056	0.038	0.124	0.223	POLITICAL
0.014	0.027	0.093	0.096	0.218	0.220	AMERICAN
0.000	0.000	0.042	0.041	0.121	0.219	TOLERANC
0.000	0.031	0.067	0.130	0.150	0.217	STAT
0.000	0.012	0.012	0.106	0.021	0.215	MINORITY
0.015	0.044	0.084	0.125	0.224	0.214	BLACK
0.028	0.041	0.113	0.127	0.132	0.211	ISSU
0.000	0.017	0.028	0.144	0.124	0.211	DIVERSITY
0.009	0.027	0.060	0.116	0.200	0.211	BEEN
0.000	0.000	0.033	0.048	0.100	0.208	RECENT
0.089	0.065	0.145	0.227	0.181	0.195	CULTUR
0.000	0.018	0.027	0.055	0.099	0.192	INTO
0.000	0.009	0.036	0.125	0.224	0.192	INCREAS
0.000	0.015	0.020	0.098	0.091	0.191	WHIL
0.039	0.005	0.032	0.053	0.123	0.191	CONTINU
0.054	0.070	0.098	0.126	0.154	0.190	EXAMPL
0.000	0.027	0.027	0.042	0.093	0.189	ACTION
0.000	0.000	0.090	0.079	0.198	0.189	RACIAL
0.000	0.004	0.009	0.070	0.229	0.188	CHURCH
0.000	0.005	0.048	0.065	0.092	0.187	NATION
0.067	0.090	0.112	0.166	0.136	0.186	WORLD
0.059	0.104	0.193	0.120	0.116	0.186	COUNTRY
0.000	0.000	0.022	0.056	0.090	0.184	OFTEN
0.023	0.056	0.059	0.132	0.141	0.183	SUCH
0.012	0.036	0.066	0.063	0.116	0.182	HOWEVER
0.040	0.086	0.111	0.072	0.189	0.182	STIL
0.000	0.010	0.058	0.086	0.146	0.181	COMMUNITY
0.014	0.065	0.073	0.095	0.072	0.180	FEEL
0.012	0.030	0.057	0.175	0.190	0.178	BE
0.000	0.021	0.050	0.189	0.202	0.177	INDIVIDUAL

0.000	0.020	0.055	0.151	0.117	0.177	BELIEF
0.000	0.019	0.053	0.085	0.146	0.177	OVER
0.000	0.010	0.046	0.131	0.083	0.177	CULTURAL
0.048	0.023	0.078	0.131	0.145	0.172	TODAY
0.019	0.020	0.055	0.090	0.123	0.170	LACK
0.000	0.000	0.012	0.047	0.099	0.169	MOVEMENT
0.036	0.048	0.048	0.141	0.175	0.168	CHILDREN
0.000	0.000	0.020	0.030	0.132	0.165	AFRICAN
0.059	0.107	0.110	0.130	0.178	0.164	RESPECT
0.020	0.027	0.033	0.061	0.089	0.164	GOVERNMENT
0.023	0.050	0.099	0.101	0.185	0.162	YEAR
0.000	0.018	0.032	0.118	0.151	0.160	TOWARD
0.038	0.016	0.026	0.070	0.091	0.160	GROW
0.000	0.000	0.000	0.000	0.055	0.160	AFFIRMATIV
0.000	0.069	0.078	0.184	0.146	0.160	ACCEPT
0.014	0.079	0.057	0.131	0.119	0.157	US
0.092	0.075	0.121	0.116	0.067	0.156	OPINION
0.027	0.011	0.105	0.125	0.139	0.156	RACE
0.046	0.053	0.109	0.073	0.128	0.156	VIEW
0.095	0.128	0.202	0.203	0.165	0.155	DIFFERENT
0.025	0.007	0.033	0.067	0.132	0.155	LAWS
0.048	0.043	0.051	0.154	0.140	0.155	SCHOOL
0.032	0.017	0.074	0.097	0.081	0.154	LESS
0.051	0.234	0.138	0.072	0.118	0.153	HE
0.044	0.056	0.060	0.110	0.146	0.153	EVEN
0.000	0.020	0.068	0.103	0.131	0.151	UNIT
0.030	0.096	0.069	0.095	0.106	0.151	THAN
0.091	0.061	0.100	0.090	0.129	0.150	SEEM
0.036	0.063	0.048	0.054	0.085	0.150	MUCH
0.013	0.060	0.091	0.067	0.177	0.149	RIGHT
0.016	0.038	0.081	0.152	0.204	0.148	RELIGION
0.017	0.028	0.056	0.042	0.148	0.148	ALTHOUGH
0.000	0.013	0.020	0.110	0.054	0.148	HATE
0.031	0.008	0.054	0.116	0.092	0.147	WHER
0.072	0.124	0.100	0.112	0.173	0.147	WHO
0.000	0.028	0.022	0.050	0.124	0.147	WAR
0.024	0.006	0.051	0.043	0.090	0.145	MEN
0.000	0.000	0.000	0.052	0.036	0.145	NATIONAL
0.000	0.030	0.035	0.124	0.088	0.145	EXIST
0.000	0.024	0.052	0.048	0.156	0.144	SEEN
0.025	0.054	0.109	0.160	0.115	0.143	STATEMENT
0.000	0.000	0.015	0.025	0.044	0.141	CITIZEN
0.000	0.010	0.026	0.093	0.091	0.141	BOTH
0.018	0.038	0.048	0.037	0.078	0.140	WELL
0.046	0.105	0.118	0.182	0.149	0.140	EACH

0.000	0.006	0.017	0.052	0.108	0.139	TOLERANT
0.026	0.007	0.000	0.032	0.041	0.139	ATTEMPT
0.016	0.043	0.065	0.054	0.097	0.138	POINT
0.125	0.111	0.061	0.095	0.115	0.137	POSITIV
0.071	0.065	0.100	0.104	0.068	0.137	UNDERSTAND
0.022	0.018	0.012	0.034	0.070	0.136	FEW
0.000	0.044	0.068	0.055	0.171	0.136	WHIT
0.023	0.019	0.044	0.043	0.089	0.134	CONFLICT
0.000	0.000	0.033	0.043	0.136	0.134	FORC
0.036	0.101	0.104	0.197	0.091	0.133	WORK
0.116	0.054	0.094	0.149	0.139	0.133	BELIEV
0.000	0.000	0.000	0.062	0.000	0.132	INCREASINGLY
0.000	0.040	0.079	0.094	0.093	0.131	AMERICA
0.000	0.013	0.038	0.014	0.090	0.130	YET
0.073	0.079	0.054	0.088	0.057	0.130	PROBLEM
0.065	0.116	0.091	0.059	0.082	0.130	OWN

Appendix C8: GRADE 5: TOP 100 WORDS BY WEIGHT - ISSUE ESSAYS

1	2	3	4	*5*	6	
0.022	0.035	0.079	0.171	0.285	0.268	GROUP
0.017	0.095	0.113	0.205	0.237	0.366	OUR
0.000	0.004	0.009	0.070	0.229	0.188	CHURCH
0.000	0.009	0.036	0.125	0.224	0.192	INCREAS
0.015	0.044	0.084	0.125	0.224	0.214	BLACK
0.014	0.027	0.093	0.096	0.218	0.220	AMERICAN
0.016	0.038	0.081	0.152	0.204	0.148	RELIGION
0.000	0.021	0.050	0.189	0.202	0.177	INDIVIDUAL
0.009	0.027	0.060	0.116	0.200	0.211	BEEN
0.000	0.000	0.090	0.079	0.198	0.189	RACIAL
0.012	0.030	0.057	0.175	0.190	0.178	BE
0.040	0.086	0.111	0.072	0.189	0.182	STIL
0.023	0.050	0.099	0.101	0.185	0.162	YEAR
0.000	0.000	0.000	0.081	0.183	0.105	BURN
0.089	0.065	0.145	0.227	0.181	0.195	CULTUR
0.059	0.107	0.110	0.130	0.178	0.164	RESPECT
0.013	0.060	0.091	0.067	0.177	0.149	RIGHT
0.036	0.048	0.048	0.141	0.175	0.168	CHILDREN
0.072	0.124	0.100	0.112	0.173	0.147	WHO
0.000	0.007	0.014	0.016	0.171	0.116	PROGRES
0.000	0.044	0.068	0.055	0.171	0.136	WHIT
0.021	0.006	0.090	0.095	0.165	0.251	WOMEN
0.095	0.128	0.202	0.203	0.165	0.155	DIFFERENT
0.000	0.000	0.024	0.161	0.161	0.104	DISRESPECT
0.000	0.024	0.052	0.048	0.156	0.144	SEEN
0.013	0.039	0.054	0.072	0.155	0.228	THOS
0.092	0.177	0.189	0.142	0.154	0.120	MY
0.054	0.070	0.098	0.126	0.154	0.190	EXAMPL
0.000	0.018	0.032	0.118	0.151	0.160	TOWARD
0.000	0.031	0.067	0.130	0.150	0.217	STAT
0.046	0.105	0.118	0.182	0.149	0.140	EACH
0.050	0.041	0.027	0.066	0.148	0.113	UP
0.017	0.028	0.056	0.042	0.148	0.148	ALTHOUGH
0.064	0.106	0.096	0.149	0.148	0.128	ANOTHER
0.000	0.009	0.062	0.080	0.146	0.020	GANG
0.042	0.057	0.053	0.093	0.146	0.111	JUST
0.044	0.056	0.060	0.110	0.146	0.153	EVEN
0.019	0.035	0.060	0.090	0.146	0.121	AMONG
0.000	0.019	0.053	0.085	0.146	0.177	OVER

0.000	0.069	0.078	0.184	0.146	0.160	ACCEPT
0.000	0.010	0.058	0.086	0.146	0.181	COMMUNITY
0.035	0.117	0.072	0.163	0.145	0.282	SOCIETY
0.048	0.023	0.078	0.131	0.145	0.172	TODAY
0.107	0.107	0.126	0.146	0.144	0.125	OTHER
0.000	0.017	0.035	0.072	0.143	0.092	EVIDENC
0.023	0.056	0.059	0.132	0.141	0.183	SUCH
0.048	0.043	0.051	0.154	0.140	0.155	SCHOOL
0.116	0.054	0.094	0.149	0.139	0.133	BELIEV
0.000	0.021	0.031	0.116	0.139	0.239	ETHNIC
0.027	0.011	0.105	0.125	0.139	0.156	RACE
0.059	0.032	0.008	0.018	0.139	0.084	HANDICAP
0.033	0.036	0.059	0.096	0.138	0.075	TWO
0.000	0.014	0.032	0.086	0.138	0.253	RELIGIOU
0.000	0.000	0.033	0.043	0.136	0.134	FORC
0.067	0.090	0.112	0.166	0.136	0.186	WORLD
0.191	0.107	0.079	0.108	0.135	0.109	SIGN
0.000	0.000	0.031	0.035	0.135	0.058	HATR
0.067	0.036	0.042	0.109	0.134	0.091	DISCRIMINATI
0.020	0.148	0.093	0.147	0.134	0.124	STUDENT
0.000	0.013	0.020	0.023	0.133	0.076	CIVIL
0.145	0.089	0.100	0.100	0.133	0.118	WAY
0.000	0.000	0.020	0.030	0.132	0.165	AFRICAN
0.025	0.007	0.033	0.067	0.132	0.155	LAWS
0.028	0.041	0.113	0.127	0.132	0.211	ISSU
0.031	0.033	0.063	0.108	0.132	0.082	OUT
0.000	0.020	0.068	0.103	0.131	0.151	UNIT
0.000	0.009	0.046	0.104	0.130	0.119	AGAINST
0.091	0.061	0.100	0.090	0.129	0.150	SEEM
0.000	0.006	0.026	0.087	0.129	0.078	VIOLENC
0.055	0.088	0.067	0.174	0.129	0.125	LEARN
0.000	0.037	0.037	0.076	0.128	0.095	SHOW
0.046	0.053	0.109	0.073	0.128	0.156	VIEW
0.072	0.094	0.117	0.091	0.125	0.115	WHAT
0.000	0.006	0.028	0.069	0.124	0.228	INTOLERANC
0.000	0.007	0.027	0.045	0.124	0.080	PUBLIC
0.000	0.028	0.056	0.038	0.124	0.223	POLITICAL
0.000	0.028	0.022	0.050	0.124	0.147	WAR
0.000	0.017	0.028	0.144	0.124	0.211	DIVERSITY
0.000	0.000	0.024	0.045	0.123	0.055	RIOT
0.019	0.020	0.055	0.090	0.123	0.170	LACK
0.039	0.005	0.032	0.053	0.123	0.191	CONTINU
0.000	0.000	0.021	0.039	0.122	0.042	KING
0.000	0.000	0.037	0.067	0.122	0.068	UNIVERSITY
0.000	0.000	0.042	0.041	0.121	0.219	TOLERANC

0.017	0.052	0.033	0.116	0.121	0.096	TOGETHER
0.000	0.017	0.034	0.045	0.120	0.099	CLEARLY
0.107	0.071	0.103	0.108	0.120	0.124	SEE
0.000	0.000	0.009	0.000	0.120	0.042	INTERACTION
0.017	0.116	0.125	0.062	0.119	0.116	HIS
0.014	0.079	0.057	0.131	0.119	0.157	US
0.025	0.041	0.020	0.046	0.119	0.097	SEXUAL
0.051	0.234	0.138	0.072	0.118	0.153	HE
0.000	0.006	0.038	0.035	0.118	0.086	OCCUR
0.027	0.014	0.014	0.016	0.117	0.092	LAW
0.000	0.000	0.022	0.032	0.117	0.097	WORKPLAC
0.000	0.020	0.055	0.151	0.117	0.177	BELIEF
0.000	0.036	0.043	0.024	0.117	0.097	FIGHT
0.059	0.104	0.193	0.120	0.116	0.186	COUNTRY
0.000	0.040	0.020	0.045	0.116	0.055	ACT
0.000	0.010	0.021	0.076	0.116	0.099	RECENTLY

Appendix C9: GRADE 4: TOP 100 WORDS BY WEIGHT - ISSUE ESSAYS

1	2	3	*4*	5	6	
0.042	0.023	0.079	0.240	0.059	0.106	BUSINES
0.089	0.065	0.145	0.227	0.181	0.195	CULTUR
0.017	0.095	0.113	0.205	0.237	0.366	OUR
0.095	0.128	0.202	0.203	0.165	0.155	DIFFERENT
0.036	0.101	0.104	0.197	0.091	0.133	WORK
0.000	0.021	0.050	0.189	0.202	0.177	INDIVIDUAL
0.000	0.007	0.027	0.188	0.110	0.284	&
0.000	0.069	0.078	0.184	0.146	0.160	ACCEPT
0.046	0.105	0.118	0.182	0.149	0.140	EACH
0.012	0.030	0.057	0.175	0.190	0.178	BE
0.055	0.088	0.067	0.174	0.129	0.125	LEARN
0.022	0.035	0.079	0.171	0.285	0.268	GROUP
0.089	0.054	0.054	0.168	0.021	0.086	COMPANY
0.067	0.090	0.112	0.166	0.136	0.186	WORLD
0.132	0.024	0.060	0.165	0.074	0.096	PLAC
0.035	0.117	0.072	0.163	0.145	0.282	SOCIETY
0.000	0.000	0.024	0.161	0.161	0.104	DISRESPECT
0.025	0.054	0.109	0.160	0.115	0.143	STATEMENT
0.000	0.084	0.114	0.156	0.023	0.109	NEW
0.048	0.043	0.051	0.154	0.140	0.155	SCHOOL
0.016	0.038	0.081	0.152	0.204	0.148	RELIGION
0.000	0.020	0.055	0.151	0.117	0.177	BELIEF
0.064	0.106	0.096	0.149	0.148	0.128	ANOTHER
0.116	0.054	0.094	0.149	0.139	0.133	BELIEV
0.020	0.148	0.093	0.147	0.134	0.124	STUDENT
0.107	0.107	0.126	0.146	0.144	0.125	OTHER
0.000	0.017	0.028	0.144	0.124	0.211	DIVERSITY
0.000	0.038	0.086	0.144	0.094	0.061	DAY
0.092	0.177	0.189	0.142	0.154	0.120	MY
0.036	0.048	0.048	0.141	0.175	0.168	CHILDREN
0.106	0.156	0.098	0.137	0.104	0.121	SOME
0.402	0.425	0.331	0.137	0.053	0.058	YOU
0.044	0.030	0.018	0.133	0.076	0.098	CITY
0.000	0.016	0.031	0.133	0.097	0.121	THROUGH
0.023	0.056	0.059	0.132	0.141	0.183	SUCH
0.014	0.079	0.057	0.131	0.119	0.157	US
0.048	0.023	0.078	0.131	0.145	0.172	TODAY
0.000	0.010	0.046	0.131	0.083	0.177	CULTURAL
0.000	0.043	0.025	0.130	0.086	0.116	LEVEL

0.000	0.031	0.067	0.130	0.150	0.217	STAT
0.063	0.099	0.099	0.130	0.100	0.121	TIME
0.059	0.107	0.110	0.130	0.178	0.164	RESPECT
0.020	0.141	0.054	0.128	0.076	0.082	HUMAN
0.071	0.000	0.010	0.128	0.000	0.029	MARKET
0.028	0.041	0.113	0.127	0.132	0.211	ISSU
0.030	0.048	0.057	0.127	0.009	0.024	COMMUNICATIO
0.054	0.070	0.098	0.126	0.154	0.190	EXAMPL
0.027	0.011	0.105	0.125	0.139	0.156	RACE
0.015	0.044	0.084	0.125	0.224	0.214	BLACK
0.000	0.009	0.036	0.125	0.224	0.192	INCREAS
0.000	0.030	0.035	0.124	0.088	0.145	EXIST
0.061	0.093	0.057	0.123	0.058	0.110	FIRST
0.000	0.000	0.071	0.123	0.061	0.103	CRIM
0.059	0.125	0.122	0.122	0.090	0.104	VERY
0.125	0.111	0.105	0.121	0.097	0.121	ABOUT
0.059	0.104	0.193	0.120	0.116	0.186	COUNTRY
0.000	0.008	0.030	0.119	0.009	0.052	INVOLV
0.000	0.018	0.032	0.118	0.151	0.160	TOWARD
0.092	0.075	0.121	0.116	0.067	0.156	OPINION
0.000	0.034	0.054	0.116	0.057	0.111	PART
0.009	0.027	0.060	0.116	0.200	0.211	BEEN
0.017	0.052	0.033	0.116	0.121	0.096	TOGETHER
0.031	0.008	0.054	0.116	0.092	0.147	WHER
0.000	0.021	0.031	0.116	0.139	0.239	ETHNIC
0.000	0.027	0.043	0.113	0.100	0.092	ALLOW
0.000	0.032	0.025	0.113	0.089	0.129	GAY
0.017	0.041	0.041	0.113	0.097	0.118	AREA
0.101	0.109	0.110	0.112	0.032	0.055	REASON
0.072	0.124	0.100	0.112	0.173	0.147	WHO
0.039	0.146	0.073	0.111	0.061	0.051	HELP
0.044	0.056	0.060	0.110	0.146	0.153	EVEN
0.000	0.008	0.015	0.110	0.044	0.052	GLOBAL
0.000	0.013	0.020	0.110	0.054	0.148	HATE
0.068	0.012	0.025	0.110	0.072	0.069	FORM
0.067	0.036	0.042	0.109	0.134	0.091	DISCRIMINATI
0.191	0.107	0.079	0.108	0.135	0.109	SIGN
0.107	0.071	0.103	0.108	0.120	0.124	SEE
0.179	0.120	0.110	0.108	0.077	0.119	WHEN
0.031	0.033	0.063	0.108	0.132	0.082	OUT
0.051	0.037	0.046	0.108	0.086	0.090	LIV
0.218	0.165	0.109	0.107	0.098	0.112	THEM
0.000	0.000	0.040	0.107	0.028	0.054	DECISION
0.163	0.271	0.128	0.107	0.086	0.066	PERSON
0.000	0.045	0.073	0.107	0.099	0.089	BACKGROUND

0.000	0.029	0.052	0.107	0.034	0.067	ECONOMIC
0.091	0.070	0.102	0.106	0.086	0.117	MOST
0.160	0.121	0.156	0.106	0.035	0.062	GET
0.064	0.062	0.089	0.106	0.083	0.082	ANOTHER'
0.000	0.113	0.061	0.106	0.072	0.082	BETTER
0.017	0.118	0.150	0.106	0.058	0.096	LIVE
0.000	0.012	0.012	0.106	0.021	0.215	MINORITY
0.164	0.104	0.075	0.105	0.115	0.110	CLEAR
0.000	0.012	0.031	0.105	0.080	0.094	SEVERAL
0.000	0.008	0.015	0.104	0.027	0.064	ORGANIZATION
0.000	0.036	0.014	0.104	0.084	0.059	PEAC
0.000	0.009	0.046	0.104	0.130	0.119	AGAINST
0.071	0.065	0.100	0.104	0.068	0.137	UNDERSTAND
0.000	0.020	0.068	0.103	0.131	0.151	UNIT
0.046	0.125	0.063	0.103	0.078	0.082	LIFE
0.000	0.032	0.043	0.102	0.113	0.069	COLOR

Appendix C10: GRADE 3: TOP 100 WORDS BY WEIGHT - ISSUE ESSAYS

1	2	*3*	4	5	6	
0.402	0.425	0.331	0.137	0.053	0.058	YOU
0.095	0.128	0.202	0.203	0.165	0.155	DIFFERENT
0.059	0.104	0.193	0.120	0.116	0.186	COUNTRY
0.161	0.117	0.191	0.064	0.078	0.086	THINK
0.092	0.177	0.189	0.142	0.154	0.120	MY
0.020	0.085	0.165	0.054	0.056	0.072	FAMILY
0.160	0.121	0.156	0.106	0.035	0.062	GET
0.017	0.118	0.150	0.106	0.058	0.096	LIVE
0.089	0.065	0.145	0.227	0.181	0.195	CULTUR
0.200	0.162	0.139	0.043	0.036	0.006	YOUR
0.019	0.112	0.138	0.057	0.083	0.085	EDUCATION
0.051	0.234	0.138	0.072	0.118	0.153	HE
0.000	0.165	0.133	0.101	0.068	0.052	ME
0.163	0.271	0.128	0.107	0.086	0.066	PERSON
0.031	0.017	0.126	0.047	0.020	0.038	STYL
0.107	0.107	0.126	0.146	0.144	0.125	OTHER
0.017	0.116	0.125	0.062	0.119	0.116	HIS
0.051	0.087	0.123	0.082	0.027	0.076	TRY
0.059	0.125	0.122	0.122	0.090	0.104	VERY
0.000	0.019	0.122	0.079	0.045	0.097	LANGUAG
0.092	0.075	0.121	0.116	0.067	0.156	OPINION
0.000	0.093	0.119	0.037	0.101	0.055	CLAS
0.046	0.105	0.118	0.182	0.149	0.140	EACH
0.188	0.124	0.118	0.089	0.097	0.109	BECAUS
0.072	0.094	0.117	0.091	0.125	0.115	WHAT
0.150	0.081	0.117	0.091	0.080	0.105	EVERYWHER
0.091	0.150	0.117	0.067	0.080	0.089	SO
0.101	0.131	0.115	0.037	0.006	0.062	IMPORTANT
0.045	0.109	0.115	0.041	0.042	0.032	FRIEND
0.095	0.084	0.114	0.078	0.060	0.119	SAME
0.000	0.084	0.114	0.156	0.023	0.109	NEW
0.028	0.041	0.113	0.127	0.132	0.211	ISSU
0.017	0.095	0.113	0.205	0.237	0.366	OUR
0.062	0.109	0.113	0.061	0.073	0.120	SHOULD
0.092	0.068	0.112	0.090	0.029	0.070	PARENT
0.067	0.090	0.112	0.166	0.136	0.186	WORLD
0.040	0.086	0.111	0.072	0.189	0.182	STIL
0.059	0.107	0.110	0.130	0.178	0.164	RESPECT

0.101	0.109	0.110	0.112	0.032	0.055	REASON
0.179	0.120	0.110	0.108	0.077	0.119	WHEN
0.025	0.054	0.109	0.160	0.115	0.143	STATEMENT
0.067	0.097	0.109	0.075	0.028	0.027	LOT
0.046	0.053	0.109	0.073	0.128	0.156	VIEW
0.218	0.165	0.109	0.107	0.098	0.112	THEM
0.027	0.011	0.105	0.125	0.139	0.156	RACE
0.000	0.000	0.105	0.011	0.067	0.007	HAIR
0.125	0.111	0.105	0.121	0.097	0.121	ABOUT
0.036	0.101	0.104	0.197	0.091	0.133	WORK
0.107	0.071	0.103	0.108	0.120	0.124	SEE
0.091	0.070	0.102	0.106	0.086	0.117	MOST
0.000	0.042	0.101	0.047	0.010	0.025	WORKER
0.000	0.059	0.101	0.056	0.010	0.038	CHINA
0.145	0.089	0.100	0.100	0.133	0.118	WAY
0.071	0.065	0.100	0.104	0.068	0.137	UNDERSTAND
0.185	0.060	0.100	0.073	0.058	0.068	EVERYON
0.046	0.050	0.100	0.035	0.058	0.076	TREAT
0.091	0.061	0.100	0.090	0.129	0.150	SEEM
0.072	0.124	0.100	0.112	0.173	0.147	WHO
0.019	0.037	0.099	0.070	0.067	0.079	AGO
0.063	0.099	0.099	0.130	0.100	0.121	TIME
0.023	0.050	0.099	0.101	0.185	0.162	YEAR
0.054	0.070	0.098	0.126	0.154	0.190	EXAMPL
0.106	0.156	0.098	0.137	0.104	0.121	SOME
0.062	0.120	0.097	0.097	0.063	0.101	LIKE
0.064	0.106	0.096	0.149	0.148	0.128	ANOTHER
0.000	0.032	0.096	0.053	0.046	0.006	CHINES
0.023	0.120	0.095	0.064	0.022	0.043	HIM
0.191	0.159	0.095	0.093	0.065	0.051	DON'T
0.211	0.209	0.095	0.102	0.067	0.066	GOOD
0.116	0.054	0.094	0.149	0.139	0.133	BELIEV
0.348	0.135	0.094	0.050	0.086	0.074	TH
0.000	0.000	0.093	0.021	0.022	0.035	UNIQU
0.014	0.027	0.093	0.096	0.218	0.220	AMERICAN
0.000	0.009	0.093	0.031	0.022	0.035	TRADITION
0.020	0.148	0.093	0.147	0.134	0.124	STUDENT
0.000	0.000	0.091	0.039	0.057	0.047	SHAR
0.013	0.060	0.091	0.067	0.177	0.149	RIGHT
0.065	0.116	0.091	0.059	0.082	0.130	OWN
0.021	0.006	0.090	0.095	0.165	0.251	WOMEN
0.000	0.000	0.090	0.079	0.198	0.189	RACIAL
0.064	0.062	0.089	0.106	0.083	0.082	ANOTHER'
0.044	0.065	0.089	0.086	0.014	0.067	FIND
0.000	0.000	0.088	0.009	0.065	0.066	INCIDENT

0.000	0.037	0.088	0.082	0.017	0.017	CARE
0.075	0.014	0.088	0.076	0.064	0.046	ENVIRONMENT
0.000	0.038	0.086	0.144	0.094	0.061	DAY
0.045	0.012	0.086	0.000	0.000	0.000	BUSY
0.134	0.090	0.086	0.061	0.079	0.062	EVERY
0.109	0.107	0.085	0.099	0.099	0.123	CHANG
0.034	0.079	0.084	0.078	0.087	0.102	EXPERIENC
0.015	0.044	0.084	0.125	0.224	0.214	BLACK
0.000	0.041	0.083	0.039	0.024	0.047	MATTER
0.051	0.021	0.083	0.054	0.088	0.068	MAN
0.023	0.044	0.082	0.064	0.089	0.033	IMPROV
0.000	0.032	0.082	0.050	0.067	0.043	WITHOUT
0.016	0.038	0.081	0.152	0.204	0.148	RELIGION
0.000	0.137	0.081	0.027	0.019	0.024	TEACHER
0.046	0.043	0.081	0.076	0.029	0.056	ABOV
0.000	0.012	0.081	0.090	0.051	0.121	AWAR
0.091	0.067	0.080	0.062	0.072	0.051	THOUGHT

Appendix C11: GRADE 2: TOP 100 WORDS BY WEIGHT - ISSUE ESSAYS

1	*2*	3	4	5	6	
0.402	0.425	0.331	0.137	0.053	0.058	YOU
0.163	0.271	0.128	0.107	0.086	0.066	PERSON
0.051	0.234	0.138	0.072	0.118	0.153	HE
0.211	0.209	0.095	0.102	0.067	0.066	GOOD
0.057	0.198	0.061	0.046	0.054	0.046	ALWAY
0.092	0.177	0.189	0.142	0.154	0.120	MY
0.158	0.171	0.017	0.019	0.000	0.000	EVERYBODY
0.218	0.165	0.109	0.107	0.098	0.112	THEM
0.000	0.165	0.133	0.101	0.068	0.052	ME
0.200	0.162	0.139	0.043	0.036	0.006	YOUR
0.191	0.159	0.095	0.093	0.065	0.051	DON'T
0.106	0.156	0.098	0.137	0.104	0.121	SOME
0.091	0.150	0.117	0.067	0.080	0.089	SO
0.020	0.148	0.093	0.147	0.134	0.124	STUDENT
0.039	0.146	0.073	0.111	0.061	0.051	HELP
0.075	0.141	0.034	0.090	0.024	0.041	SOMEON
0.020	0.141	0.054	0.128	0.076	0.082	HUMAN
0.000	0.137	0.081	0.027	0.019	0.024	TEACHER
0.348	0.135	0.094	0.050	0.086	0.074	TH
0.101	0.131	0.115	0.037	0.006	0.062	IMPORTANT
0.095	0.128	0.202	0.203	0.165	0.155	DIFFERENT
0.026	0.125	0.014	0.031	0.033	0.026	SOMETIM
0.046	0.125	0.063	0.103	0.078	0.082	LIFE
0.059	0.125	0.122	0.122	0.090	0.104	VERY
0.188	0.124	0.118	0.089	0.097	0.109	BECAUS
0.072	0.124	0.100	0.112	0.173	0.147	WHO
0.160	0.121	0.156	0.106	0.035	0.062	GET
0.034	0.121	0.028	0.010	0.033	0.014	PERSONALITY
0.179	0.120	0.110	0.108	0.077	0.119	WHEN
0.062	0.120	0.097	0.097	0.063	0.101	LIKE
0.023	0.120	0.095	0.064	0.022	0.043	HIM
0.017	0.118	0.150	0.106	0.058	0.096	LIVE
0.035	0.117	0.072	0.163	0.145	0.282	SOCIETY
0.161	0.117	0.191	0.064	0.078	0.086	THINK
0.017	0.116	0.125	0.062	0.119	0.116	HIS
0.065	0.116	0.091	0.059	0.082	0.130	OWN
0.018	0.114	0.040	0.050	0.110	0.101	SAY
0.021	0.114	0.028	0.096	0.053	0.116	IDEA
0.000	0.113	0.061	0.106	0.072	0.082	BETTER
0.000	0.113	0.078	0.047	0.033	0.016	DOESN'T

0.207	0.084	0.009	0.042	0.011	0.021	GOD
0.021	0.081	0.029	0.039	0.027	0.101	PRESENT
0.150	0.081	0.117	0.091	0.080	0.105	EVERYWHER
0.000	0.081	0.000	0.039	0.000	0.018	RESIDENT
0.030	0.081	0.000	0.009	0.028	0.067	DIDN'T
0.000	0.081	0.063	0.056	0.037	0.047	ESPECIALLY
0.165	0.080	0.009	0.060	0.000	0.020	PRODUCT
0.075	0.080	0.060	0.045	0.055	0.035	BEST
0.014	0.079	0.057	0.131	0.119	0.157	US
0.034	0.079	0.084	0.078	0.087	0.102	EXPERIENC
0.133	0.079	0.036	0.016	0.033	0.005	EVERYTH
0.073	0.079	0.054	0.088	0.057	0.130	PROBLEM
0.000	0.078	0.024	0.026	0.037	0.041	MANNER
0.000	0.078	0.052	0.029	0.031	0.013	FOOD
0.000	0.078	0.000	0.000	0.026	0.017	IDEAL

Appendix C12: GRADE 1: TOP 100 WORDS BY WEIGHT - ISSUE ESSAYS

1	2	3	4	5	6	
0.402	0.425	0.331	0.137	0.053	0.058	YOU
0.362	0.012	0.024	0.014	0.000	0.000	DIFFRENT
0.358	0.043	0.000	0.000	0.000	0.016	RULE
0.348	0.135	0.094	0.050	0.086	0.074	TH
0.241	0.009	0.019	0.031	0.022	0.007	DRIV
0.218	0.165	0.109	0.107	0.098	0.112	THEM
0.211	0.209	0.095	0.102	0.067	0.066	GOOD
0.209	0.039	0.051	0.069	0.079	0.068	FOLLOW
0.207	0.084	0.009	0.042	0.011	0.021	GOD
0.200	0.162	0.139	0.043	0.036	0.006	YOUR
0.191	0.159	0.095	0.093	0.065	0.051	DON'T
0.191	0.107	0.079	0.108	0.135	0.109	SIGN
0.188	0.124	0.118	0.089	0.097	0.109	BECAUS
0.186	0.042	0.059	0.009	0.020	0.025	PAY
0.185	0.060	0.100	0.073	0.058	0.068	EVERYON
0.179	0.120	0.110	0.108	0.077	0.119	WHEN
0.177	0.010	0.000	0.032	0.011	0.036	STOP
0.176	0.061	0.027	0.053	0.064	0.005	BAD
0.172	0.000	0.000	0.000	0.014	0.009	USEFUL
0.165	0.080	0.009	0.060	0.000	0.020	PRODUCT
0.164	0.104	0.075	0.105	0.115	0.110	CLEAR
0.163	0.271	0.128	0.107	0.086	0.066	PERSON
0.161	0.117	0.191	0.064	0.078	0.086	THINK
0.160	0.121	0.156	0.106	0.035	0.062	GET
0.158	0.171	0.017	0.019	0.000	0.000	EVERYBODY
0.150	0.081	0.117	0.091	0.080	0.105	EVERYWHER
0.145	0.089	0.100	0.100	0.133	0.118	WAY
0.136	0.000	0.000	0.014	0.000	0.018	SELL
0.136	0.012	0.012	0.000	0.029	0.000	HIGHWAY
0.134	0.090	0.086	0.061	0.079	0.062	EVERY
0.133	0.079	0.036	0.016	0.033	0.005	EVERYTH
0.132	0.024	0.060	0.165	0.074	0.096	PLAC
0.129	0.000	0.012	0.013	0.000	0.018	FUND
0.129	0.000	0.012	0.013	0.000	0.026	BANK
0.125	0.111	0.105	0.121	0.097	0.121	ABOUT
0.125	0.111	0.061	0.095	0.115	0.137	POSITIV
0.124	0.067	0.074	0.030	0.031	0.020	SOMETH
0.116	0.063	0.076	0.091	0.089	0.075	GO

0.116	0.054	0.094	0.149	0.139	0.133	BELIEV
0.112	0.020	0.000	0.011	0.024	0.023	NATURAL
0.109	0.107	0.085	0.099	0.099	0.123	CHANG
0.107	0.107	0.126	0.146	0.144	0.125	OTHER
0.107	0.071	0.103	0.108	0.120	0.124	SEE
0.106	0.156	0.098	0.137	0.104	0.121	SOME
0.101	0.055	0.009	0.010	0.043	0.021	BODY
0.101	0.047	0.027	0.038	0.040	0.051	JOB
0.101	0.131	0.115	0.037	0.006	0.062	IMPORTANT
0.101	0.109	0.110	0.112	0.032	0.055	REASON
0.099	0.044	0.009	0.010	0.021	0.027	NORMAL
0.095	0.084	0.114	0.078	0.060	0.119	SAME
0.095	0.089	0.045	0.050	0.037	0.063	HAND
0.095	0.128	0.202	0.203	0.165	0.155	DIFFERENT
0.093	0.059	0.017	0.038	0.010	0.025	CUSTOM
0.092	0.075	0.121	0.116	0.067	0.156	OPINION
0.092	0.177	0.189	0.142	0.154	0.120	MY
0.092	0.068	0.112	0.090	0.029	0.070	PARENT
0.091	0.037	0.018	0.034	0.100	0.102	CONSIDER
0.091	0.070	0.102	0.106	0.086	0.117	MOST
0.091	0.008	0.033	0.028	0.048	0.044	ENOUGH
0.091	0.061	0.100	0.090	0.129	0.150	SEEM
0.091	0.150	0.117	0.067	0.080	0.089	SO
0.091	0.067	0.080	0.062	0.072	0.051	THOUGHT
0.090	0.000	0.012	0.000	0.000	0.018	CHARACTERIZ
0.090	0.000	0.000	0.027	0.014	0.000	INTERVIEW
0.090	0.000	0.000	0.014	0.014	0.009	CHICAGO
0.089	0.054	0.054	0.168	0.021	0.086	COMPANY
0.089	0.065	0.145	0.227	0.181	0.195	CULTUR
0.088	0.064	0.048	0.018	0.046	0.024	MONEY
0.087	0.031	0.047	0.009	0.046	0.041	INFORMATION
0.086	0.000	0.012	0.013	0.000	0.026	LEADERSHIP
0.086	0.042	0.060	0.068	0.098	0.091	TRUE
0.086	0.099	0.064	0.059	0.048	0.039	GIVE
0.086	0.038	0.039	0.017	0.036	0.041	LISTEN
0.084	0.000	0.015	0.059	0.027	0.103	DIFFER
0.084	0.015	0.023	0.025	0.053	0.074	BUILD
0.084	0.045	0.023	0.042	0.035	0.074	SIDE
0.083	0.037	0.007	0.059	0.061	0.051	YOUNG
0.082	0.011	0.033	0.012	0.000	0.008	MISTAK
0.080	0.050	0.079	0.024	0.033	0.032	SAID
0.079	0.021	0.011	0.012	0.025	0.000	ROAD
0.079	0.000	0.065	0.000	0.000	0.032	PEACEFULLY
0.079	0.011	0.000	0.048	0.013	0.000	WALL
0.077	0.021	0.042	0.023	0.000	0.024	FASHION

0.077	0.014	0.077	0.039	0.033	0.053	START
0.077	0.028	0.035	0.062	0.057	0.042	HIGH
0.077	0.010	0.000	0.000	0.012	0.063	REGULATION
0.077	0.000	0.031	0.023	0.012	0.008	DECREAS
0.076	0.062	0.067	0.073	0.076	0.073	RESPECTFUL
0.075	0.080	0.060	0.045	0.055	0.035	BEST
0.075	0.141	0.034	0.090	0.024	0.041	SOMEON
0.075	0.020	0.000	0.000	0.012	0.053	POWER
0.075	0.014	0.088	0.076	0.064	0.046	ENVIRONMENT
0.073	0.079	0.054	0.088	0.057	0.130	PROBLEM
0.072	0.124	0.100	0.112	0.173	0.147	WHO
0.072	0.029	0.020	0.000	0.023	0.015	CAR
0.072	0.094	0.117	0.091	0.125	0.115	WHAT
0.071	0.029	0.057	0.011	0.011	0.000	NOBODY
0.071	0.029	0.000	0.053	0.022	0.014	SUCCESSFUL
0.071	0.065	0.100	0.104	0.068	0.137	UNDERSTAND
0.071	0.000	0.010	0.128	0.000	0.029	MARKET

Appendix D: System Flowchart