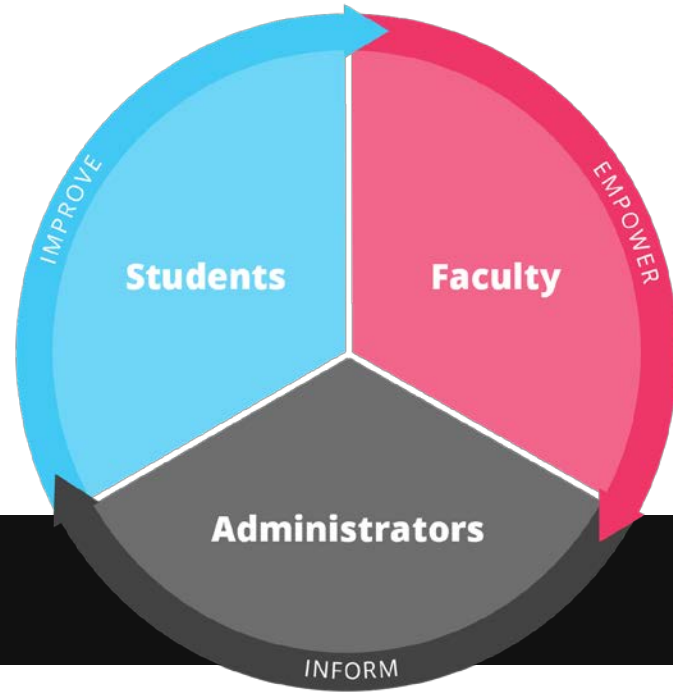


# The Role of Instructor and Peer Feedback in Improving the Cognitive, Interpersonal, and Intrapersonal Competencies of Student Writers in STEM Courses\*



 **MY REVIEWERS**

Joe Moxley, Norbert Elliot, Alex Rudniy, and Val Ross,  
IWAC, June 23, 2016

\*This research is supported by the National Science Foundation under Award #154423



1. Demonstrate ways the assessment community can use big data, real-time assessment tools to create valid measures of writing development
2. Provide quantitative evidence regarding the effects of particular commenting and scoring patterns on student
3. Inform STEM faculty regarding the efficacy of particular high impact practices, especially peer review
4. Provide a domain map to help us better understand non-cognitive competencies and student success in the STEM curriculum
5. Provide the evidence necessary to build interactive assessment loops and algorithms to provide more helpful feedback and assessments

# My Reviewers: What Is It?

A comprehensive suite of tools, *My Reviewers* is:

an e-learning environment

a document markup tool that facilitates peer review and team projects

an e-portfolio tool

an assessment tool

a publication platform for e-texts

a research project for universities to examine student success, pedagogy, the development of writing competencies, and more

ENC 1101 Project 1 Early

Student 3, Test



Student

Teacher

Demo 123

## Demo Paper

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse iaculis libero justo, non tincidunt dui hendrerit nec. Sed id orci convallis, commodo justo ac, faucibus ligula. Donec maximus risus non diam scelerisque, eget sagittis neque vestibulum. Phasellus risus nisi, suscipit nec odio non, ornare vehicula diam. Donec elementum sapien et nisl mattis porttitor. Nulla vel metus nec est porttitor faucibus eget finibus arcu. Maecenas sodales lacus orci, a facilisis diam sagittis non. Integer dapibus viverra ligula, fermentum vulputate libero. Nulla ut odio eu tortor varius fringilla in a nibh. Proin vel lacus convallis, eleifend elit eget, feugiat diam. Integer sit amet gravida urna. Proin lacinia quam id eros eleifend malesuada. Phasellus pharetra ultrices nunc imperdiet pellentesque. Vivamus eu ante ut mauris feugiat ullamcorper non posuere quam.

Duis ornare facilisis nisi. Nulla at imperdiet ipsum. Proin quis varius justo. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Vestibulum auctor augue ligula, id consectetur nisl auctor non. Maecenas ac arcu ac neque placerat eleifend id id erat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Etiam tempor suscipit massa at ultricies. Sed laoreet justo lacinia, sodales odio nec, pretium lacus.

Integer placerat libero quis varius maximus. Sed ipsum massa, consequat vitae efficitur non, feugiat at risus. Pellentesque sodales rhoncus lacus. Integer posuere dui eleifend velit pulvinar luctus. Donec condimentum tincidunt neque a volutpat. Fusce eget metus congue, cursus justo non, tincidunt risus. Maecenas suscipit lectus ligula, ac eleifend sem iaculis eu. Aliquam sit amet uma semper, consequat tortor a, porta urna. Praesent ornare lorem vitae ligula malesuada, sit amet

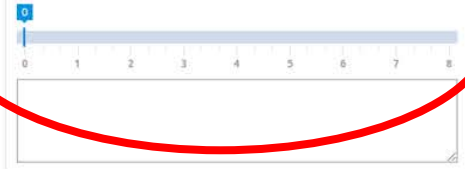
Rubric Comments

Submit

## ANALYSIS - 30%



## EVIDENCE - 30%



## FORMAT/ORGANIZATION - 20%



## STYLE - 20%



ENC 1101 Project 1 Intermediate

Anonymous PR

B I U

DEL

sds

Student

Teacher

Demo 123

Demo Paper

Highlight

Your thesis statement could be clearer. I feel like you do a good job defending your points in the body of your paper, but that your thesis does not encompass all of your points. I think that you should re-read your paper so that you can formulate a thesis that matches your final argument.

>Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse iaculis libero justo, non tincidunt dui hendrerit nec. Sed id orci convallis, commodo justo ac faucibus ligula. Donec maximus risus non diam scelerisque, eget sagittis neque. **Active vs. passive voice** ~~risus nisi, suscipit nec odio non, ornare vehicula diam.~~ Donec elementum sapien et nisl mattis portitor. Nulla vel metus nec est porttitor faucibus eget finibus arcu. Maecenas sodales lacus orci, a facilisis diam sagittis non. Integer dapibus viverra ligula, fermentum vulputate libero. Nulla ut odio eu tortor varius fringilla in a nibh. Proin vel lacus convallis, eleifend elit eget, feugiat diam. Integer sit amet gravida urna. Proin lacinia quam id eros eleifend malesuada. Phasellus pharetra ultrices nunc imperdiet pellentesque. **Vivamus eu ante ut mauris feugiat ullamcorper non posuere quam.**

Duis ornare facilisis nisi. Nulla at imperdiet ipsum. Proin quis varius justo. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Vestibulum auctor augue ligula, id consectetur nisl auctor non. Maecenas ac arcu ac neque placerat eleifend id id erat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Etiam tempor suscipit massa at ultricies. Sed laoreet justo lacinia, ~~sodales odio nec, pretium lacus.~~

Integer placerat libero quis varius maximus. Sed ipsum massa, consequat vitae efficitur non, feugiat at risus. Pellentesque sodales rhoncus lacus. Integer posuere dui eleifend velit vulvar luctus. Donec condimentum tincidunt neque a volutpat. Fusce eget metus congue, cursus justo non, tincidunt risus. Maecenas suscipit lectus ligula, ac eleifend sem iaculis eu. Aliquam sit amet urna semper, consequat titor a porta urna. Praesent ornare lorem vitae ligula malesuada, sit amet

Rubric

Comments

Submit

## ANALYSIS - 30%



You meet most of the assignment requirements, however you needed one more section on what potential counterarguments could be made against your claim. Also, your research question is too broad/narrow, needs development, and lacks focus. You

## EVIDENCE - 30%



Most of your sources are appropriate/credible for scholarly research. You just need to check the reliability of the one I noted in your paper. Overall though, well done!

## FORMAT/ORGANIZATION - 20%



Your paper is correctly formatted.

## STYLE - 20%



## E-text

Community Comments

 Network, Collaborate,  
Compose

Rhetoric Matters

A Rhetorical Approach

Publish, Don't Perish

The Winning Brief

## Help

## Student:

Peer Review Written By	Reviewer Rubric Comments	Reviewer In-Text Comments
Reviewer 1 <a href="#">View This Peer Review</a>	<p><b>Evidence:</b> Most of your sources are appropriate/credible for scholarly research. You just need to check the reliability of the one I noted in your paper. Overall though, well done!</p> <p><b>Style:</b> You have a few style issues which I noted in your paper. The largest issue that I saw was with your transitions between paragraphs. I placed a Community Comment on the topic so you can practice this. You are definitely on</p>	<p>(1) Nice work here!</p> <p>(2) Your thesis statement could be clearer. I feel like you do a good job defending your points in the body of your paper, but that your thesis does not encompass all of your points. I think that you should re-read your paper so that you can formulate a thesis that matches your final argument.</p>
Reviewer 2 <a href="#">View This Peer Review</a>	<p><b>Evidence:</b> Good job! Your evidence all comes from credible sources. You might spread out the discussion of the sources a bit in your next draft, but overall this is good work.</p> <p><b>Style:</b> You have a few style errors, which I noted in the body of your paper. Specifically, I would work most on varying your sentence syntax and transitioning between your points for the next draft.</p>	<p>(1) I feel like you can expand on this intro and split it into two paragraphs.</p> <p>(2) Your thesis is on the right track, but needs work. Your thesis is very broad and general, but your actual argument is quite specific.</p> <p>(3) I like your use of sources. However you seem to discuss them mostly in one paragraph. I might break that up a bit.</p>
Reviewer 3 <a href="#">View This Peer Review</a>	<p><b>Analysis:</b> This is a very good start. It seems like you've put a lot of thought into your topic and arguments in the body of your paper. However, it seems like you spent less time on your intro. You mention a few ideas in the intro that never gets developed later in the paper. I would focus on rewriting your intro for the next draft.</p> <p><b>Evidence:</b> You've found some really great sources! However, it</p>	<p>(1) I think that you may want to switch this paragraph and the one before it. This paragraph seems to relate more to your point earlier in the paper.</p> <p>(2) You may want to rework your thesis</p> <p>(3) I feel like you need to better transition into this next paragraph.</p>

## Features

Learn More



### COMMUNITY COMMENTS

Search through our library of comments to instantly give content-rich feedback.



### CUSTOMIZABLE RUBRICS

Build and manage rubrics to address unique contexts, genres, and disciplines.



### TEAM PROJECTS

Save time by managing and grading from one simple interface



### PORTFOLIO ASSESSMENT

Assign multiple reviewers, facilitate grade norming, and run powerful reports.



### WRITING ANALYTICS

Access an extensive corpus for data analysis and insight on student success.



### PEER REVIEW

Use our powerful interface to improve peer review processes.

# My Reviewers @ USF

From the Fall 2009 to the Spring of 2016, students have completed 253,148 peer reviews and instructors have completed 174,366 reviews

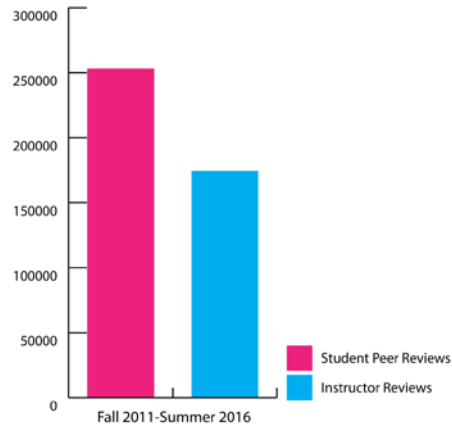


Figure x. Total number of completed reviews. Users have completed 253,148 Student Peer Reviews and 174,366 Instructor Reviews.





# Chemistry Courses @ USF

We began our partnership with the USF Chemistry department in the Spring 2016 term. The courses that use *My Reviewers* include:

[CHM 3941](#) (Peer Leading)

CHM 4411 (Physical Chem)

CHM 2045 (Gen Chem 1)

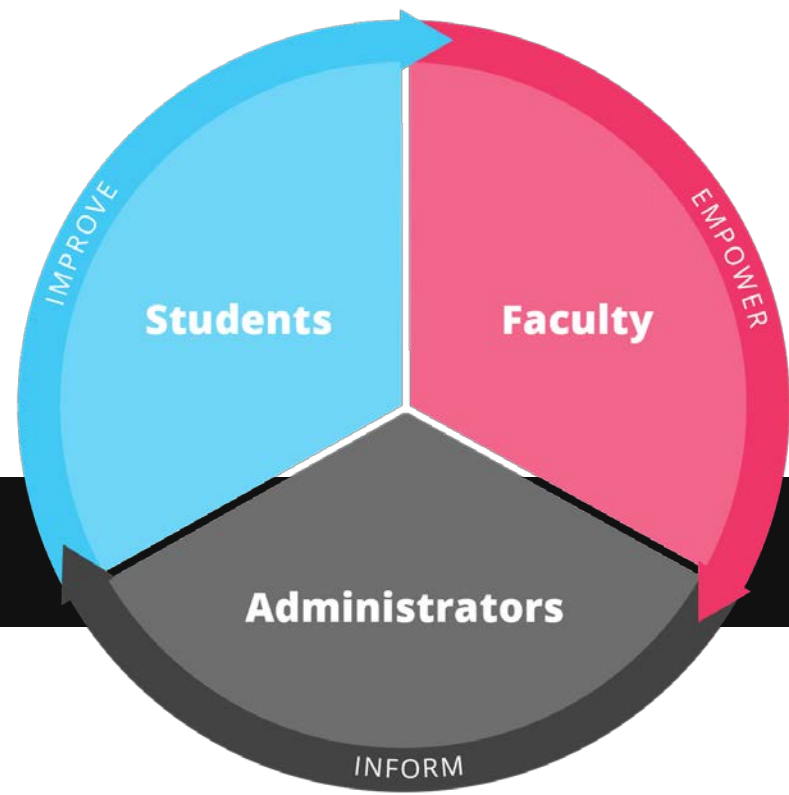
CHM 2046 (Gen Chem 2).

Courses use *My Reviewers* for peer reviews and final grading of lab and research reports

N = 2,027 students and 6,517 reviews



# The Role of Instructor and Peer Feedback in Improving the Cognitive, Interpersonal, and Intrapersonal Competencies of Student Writers in STEM Courses

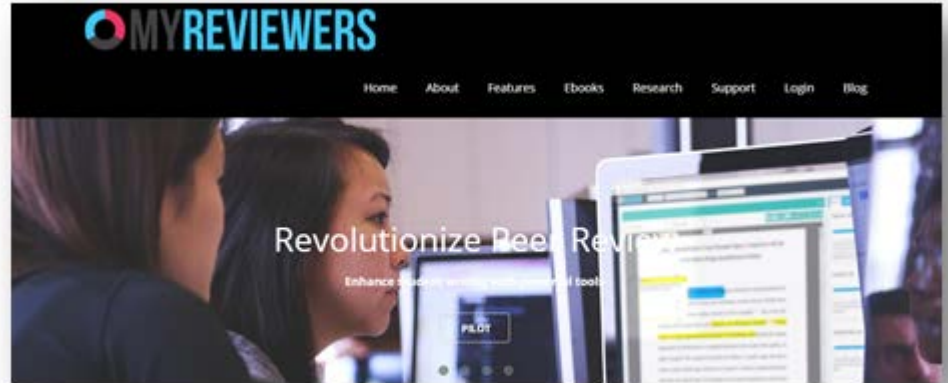


 **MY REVIEWERS**

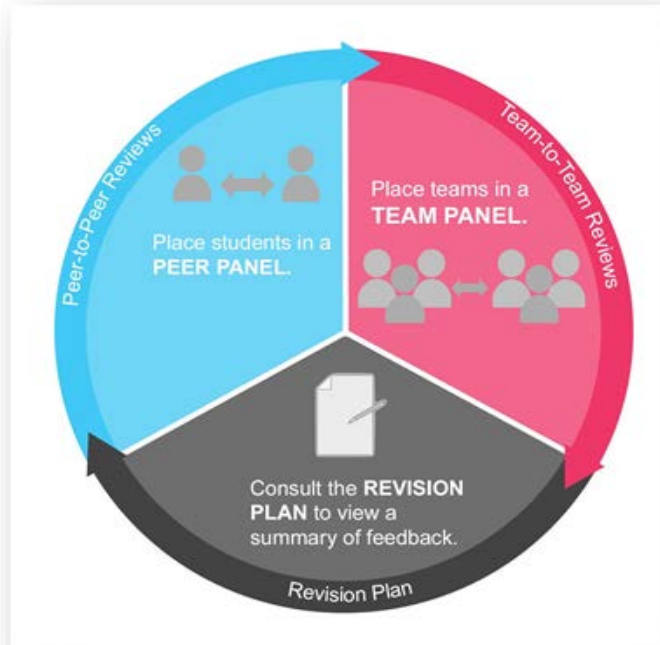
Norbert Elliot  
Program Evaluator for Award 1544239  
International Writing Across the  
Curriculum Conference  
June 23, 2016

# Outline

- Domain Specific Construct Modeling
- Mapping the Writing Construct
- Research Planning
- Sampling Plan
- Early Research Example
- Future Research
- Imaging the Future

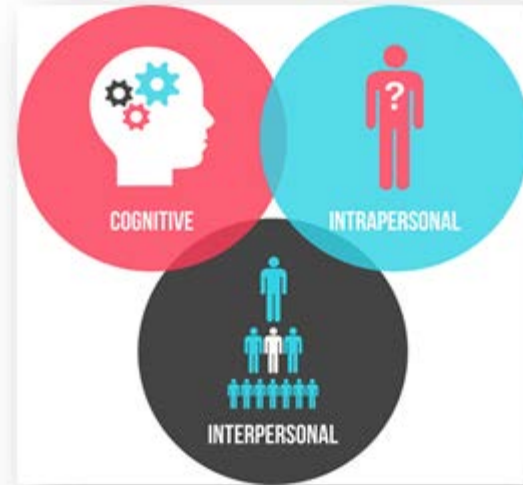
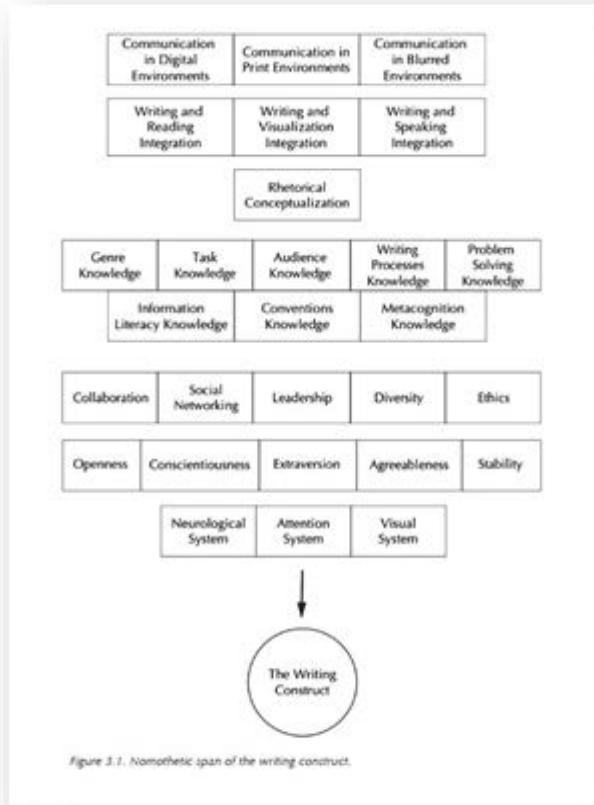


# Precision: Domain Specific Construct Modeling



Naturalistic Observation Emphasizing Sociocognitive and Sociocultural Construct Modeling  
Moss, P. A., Pullin, D. C., Gee, J. P., Haertel, E. H. & Young, L. J. (Eds.). (2008). *Assessment, equity, and opportunity to learn*. Cambridge, UK: Cambridge University Press.

# Target: Mapping the Writing Construct



National Research Council of the National Academies. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington D.C.: National Academic Press.

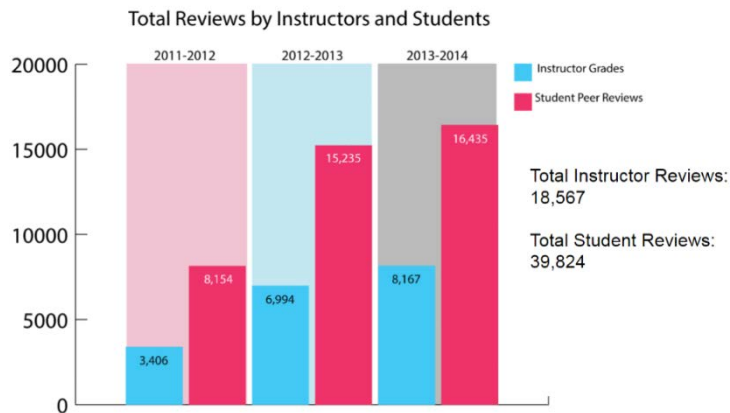
# Planning: Design for Assessment Approach to Research



White, E. M., Elliot, N., & Peckham, I. (2015). *Very like a whale: The assessment of writing programs*. Logan, UT: Utah State University.



# Sampling Plan: Massive Data Analysis:



- Basic Statistics
- Generalized  $N$ -Body Problems
- Graph-Theoretic Computations
- Linear Algebraic Computations
- Optimizations
- Integration
- Alignment Problems

National Research Council (2013). *Frontiers in massive data analysis*. Washington, D.C.: The National Academies Press.

# Early Research: N-Gram Analysis

## WS-2: Writing Analytics, Data Mining, and Writing Studies

*Val Ross,  
University of Pennsylvania  
Alex Rudniy,  
Fairleigh Dickinson University  
Joe Moxley,  
University of South Florida  
David Eubanks,  
Furman University*



Dataset	Comments	Comments
Dataset Trait 1. Focus	1,516	1,859
Dataset Trait 2. Evidence	2,976	3,809
Dataset Trait 3. Organization	1,219	1,682
Dataset Trait 4. Style	1,252	1,870
Dataset Trait 5. Format	2,549	4,084



N-gram analysis lead:  
Alex Rudniy  
arudniy@fdu.edu



# Research Questions and Sampling Plan

1. How can n-gram analysis be used to examine concept proliferation of course terms students should know?
2. How can n-gram analysis be used to examine concept proliferation of assessment traits used to assess student work?
3. What type of n-gram analysis is best suited to examine concept proliferation?

Dataset	Instructor Comments	Peer Comments
Dataset Trait 1. Focus	1,516	1,859
Dataset Trait 2. Evidence	2,976	3,809
Dataset Trait 3. Organization	1,219	1,682
Dataset Trait 4. Style	1,252	1,870
Dataset Trait 5. Format	2,549	4,084

Study 1: N-gram analysis of course terms  
Study 2: N-gram analysis of assessment terms

# Early Research: Study 1 (Course Terms)

Context: English Composition II			
Topics	Purpose	Genre	Terms Students Should Know
Project 1: Analyzing Visual Rhetoric	"In Project One, you will learn how to identify one stakeholder's argument and analyze that stakeholder's use of visual and rhetorical strategies."	Source-based essay: identify one stakeholder's argument and analyze that stakeholder's use of visual and rhetorical strategies.	stakeholder, rhetorical appeals, ethos, pathos, logos, Kairos, visual rhetoric, visual fallacies
Project 2: Finding Common Ground	"In Project Two, you will learn how to present an unbiased analysis of two arguments created by stakeholders with seemingly incompatible goals about an issue or topic and create a feasible, objective compromise that would benefit both stakeholders."	Source-based essay: analyze two stakeholders with seemingly incompatible goals regarding the same issue or topic; identify common ground between stakeholders.	compromise, empathy, negotiation, Rogerian argument
Project 3: Composing Multimodal Assignments	"Project 3 brings all you have done full circle. You will use your understanding of the rhetorical situation to decide how to craft the most effective means of engaging your audience and empowering the audience to take the action you recommend."	Multimedia Argument Website: produce a complementary argument using the digital medium of a website to address these aims: educate an audience of non-engaged stakeholders about the issue or topic, engage the audience by convincing them that they should care about this issue or topic, and empower the audience to take action in some way.  Formal Essay: produce a complimentary essay that addresses the website aims, Presentation: present their multimodal remediation (or a portion of it) for an audience of their peers. Individual instructors will dictate the specific requirements of these presentations.	multimodality, remediation, non-engaged stakeholder

*My Reviewers* allows free response textual comments and designation of numeric score on a 4-point scale 5 rubric traits: focus, evidence, organization, style, and format.

# Study 1 Results

Instructor

Student



Evidence	Instructor	Student
stakeholder	(761)	(740)
rhetorical	(1011)	(502)
ethos	(0)	(0)
pathos	(470)	(0)
logos	(508)	(0)
Kairos	(0)	(0)
visual	(659)	(0)
fallacies	(477)	(0)
compromise	(633)	(436)
empathy	(0)	(0)
negotiation	(0)	(0)
Rogerian	(0)	(0)
argument	(927)	(998)
multimodality	(0)	(0)
remediation	(0)	(0)
non-engaged	(0)	(0)

**Course Terms:** Patterns of congruence, disjuncture, and absence:

- *Congruence:* Regarding the trait of evidence, stakeholder, rhetorical, compromise, and argument are used in both sets of comments.
- *Disjuncture:* Regarding the trait of evidence, the term rhetorical is used twice more by instructors than by students; while instructors use the term visual, students do not use that term.
- *Absence:* Notable absence of key terms by both groups: ethos, pathos, logos, Kairos, fallacies, empathy, negotiation, Rogerian, multimodality, remediation, and non-engaged.

# Early Research: Study 2 (Assessment Terms)

	Trait 1: Focus	Trait 2: Evidence	Trait 3: Organization	Trait 4: Style	Trait 5: Format
Terms in Rubric	critical thinking, thesis, ideas, analysis, assignment requirements	critical thinking, credible sources and supporting details, synthesis, visuals, personal experience, anecdotes, writer's idea, source's ideas	critical thinking, introduction, topic sentences, segues, transitions, conclusion	critical thinking, grammar, punctuation, point of view, syntax, diction, word choice, vocabulary	documentation style, MLA, APA, formatting, in-text citations, annotated bibliographies, works cited, document design

# Study 2 Results

Instructor



Evidence	sources (1946) evidence (1817) source (1742) use (1600) sure (1562)	make sure (221) use evidence (211) good use (185) final draft (175) relevant stuff (174)	smart relevant stuff (174) support paper s (110) sources establish credibility (106) introductions sources establish (104) article relevant research (89)	introductions sources establish credibility (104) article relevant research published (89) biochemist Jane Chen discusses (89) credible magazine biochemist Jane Chen (89) Jane Chen discusses significance (89)	article relevant research published credible (89) biochemist Jane Chen discusses significance (89) credible magazine biochemist Jane Chen (89) magazine biochemist Jane Chen discusses (89) published credible magazine biochemist jane (89)
----------	---	--	---	--	--

Student



Evidence	sources (2458) evidence (2312) paper (2309) good (2044) used (1932)	text citations (365) credible sources (219) make sure (202) sources used (160) throughout paper (155)	works cited page (92) use text citations (47) good use evidence (45) good use sources (37) just make sure (36)	fair selection credible sources (21) credible sources supporting details (17) selection credible sources supporting (10) ideas source s ideas (9) good use text citations (8)	selection credible sources supporting details (10) fair selection credible sources supporting (7) relationship thesis primary secondary sources (7) across backed paper just make (5) also really good quoted gave (5)
----------	---	---	--	---	--

**Assessment Terms:** Patterns of congruence, disjuncture, and absence:

- *Congruence:* Unigram and bigram analysis for instructor and students are largely congruent.
- *Disjuncture:* Regarding evidence, trigram analysis reveals some disjuncture. Instructors note that sources establish credibility; students, in contrast, note the presence and features of the works cited page—a format substitution for the complexities of establishing claims.
- *Absence:* Absent are references to traits such as synthesis, personal experiences, anecdotes, segues, diction, and document design.

# NSF Research (Award #1544239): DFA Approach

## Concurrent Study 1: Deployment: Tools and Resources in STEM Courses

- ❖ To support the claim that *MyR* was deployed across all institutions in a ways leading to student and instructor motivation

## Concurrent Study 2: Analysis: Coding the Corpus

- ❖ To support the claim that coding categories will allow identification and mapping of the writing construct in its three domains

## Concurrent Study 3: Variable Mapping: Construct Modeling

- ❖ To support the claim that the construct model can disaggregated by student groups in order to structure opportunity to learn

## Concurrent Study 4: Foundations: Fairness, Validity, and Reliability

- ❖ To support the claim that foundational measurement principles can be used to analyze information across all groups in terms of gender, gender identification, race, ethnicity, and socioeconomic status

## Core Study 1: The Scoring Study

- ❖ To support the claim that an empirical research core can be established

## Core Study 2: Data Mining the Corpus

- ❖ To support the claim that digitally-based analytics allows systems such as *MyR* to transform course management systems into instructional and assessment environments

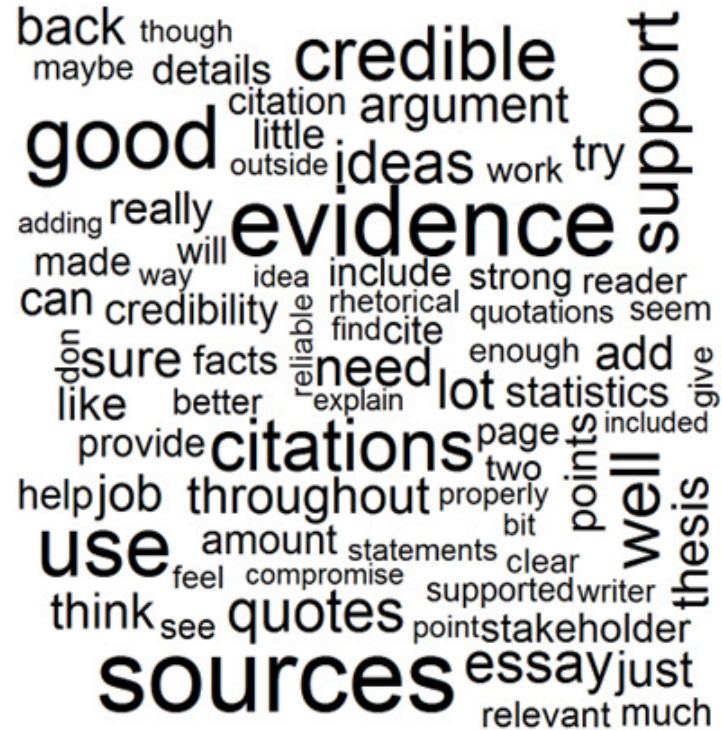
# Imagine: Visual Analytics and Actionable Information



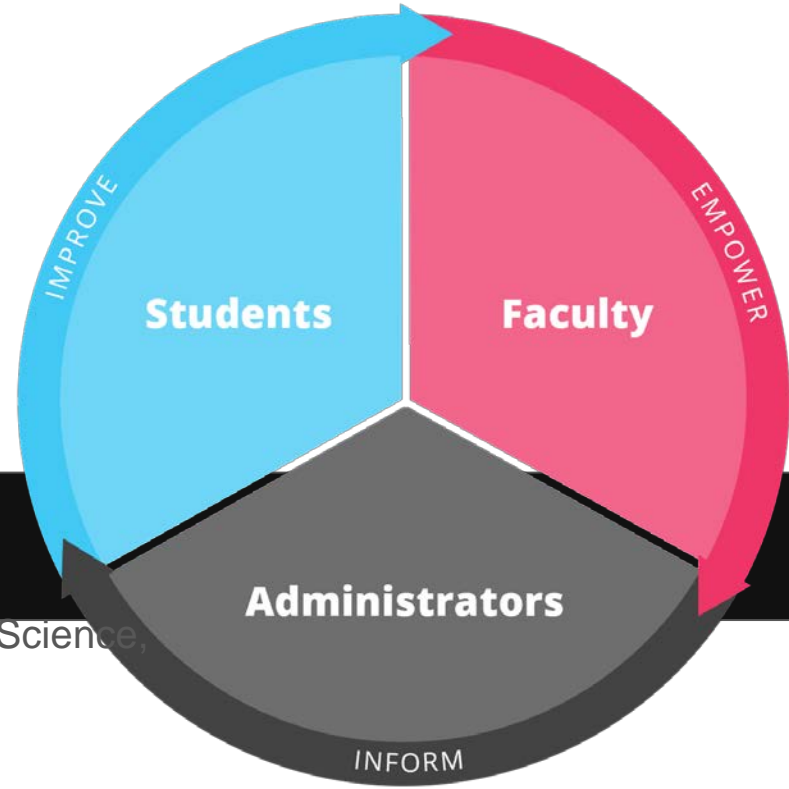
The R Project for Statistical Computing

R, RStudio, and the TM package:

- Word cloud of the 100 most frequent words by students responding to the trait of evidence



# N-gram Study IWAC, 2016



## **MY REVIEWERS**

Alex Rudniy, Assistant Professor of Computer Science,  
FDU

NSF Award 1544239



# Purpose of the Study

Explore the use of n-gram analysis

Analyze instructor and student comments elicited within *My Reviewers*, a web-based learning environment.

Study instructor and student use of concepts

Prepare a base for future analysis

# What is N-Gram?

N-gram is a sequence of  $n$  items as they appear in text

Letters, words, phonemes, part-of-speech tags or other elements.

$N$  is the number of items in a sequence.

A single word is a unigram (1-gram)

Two words—bigram (2-gram)

Three words—trigram (3-gram)

Four words—four-gram (4-gram)

Five words— five-gram (5-gram)

# Software Tools



**Package 'tm'**

July 3, 2015

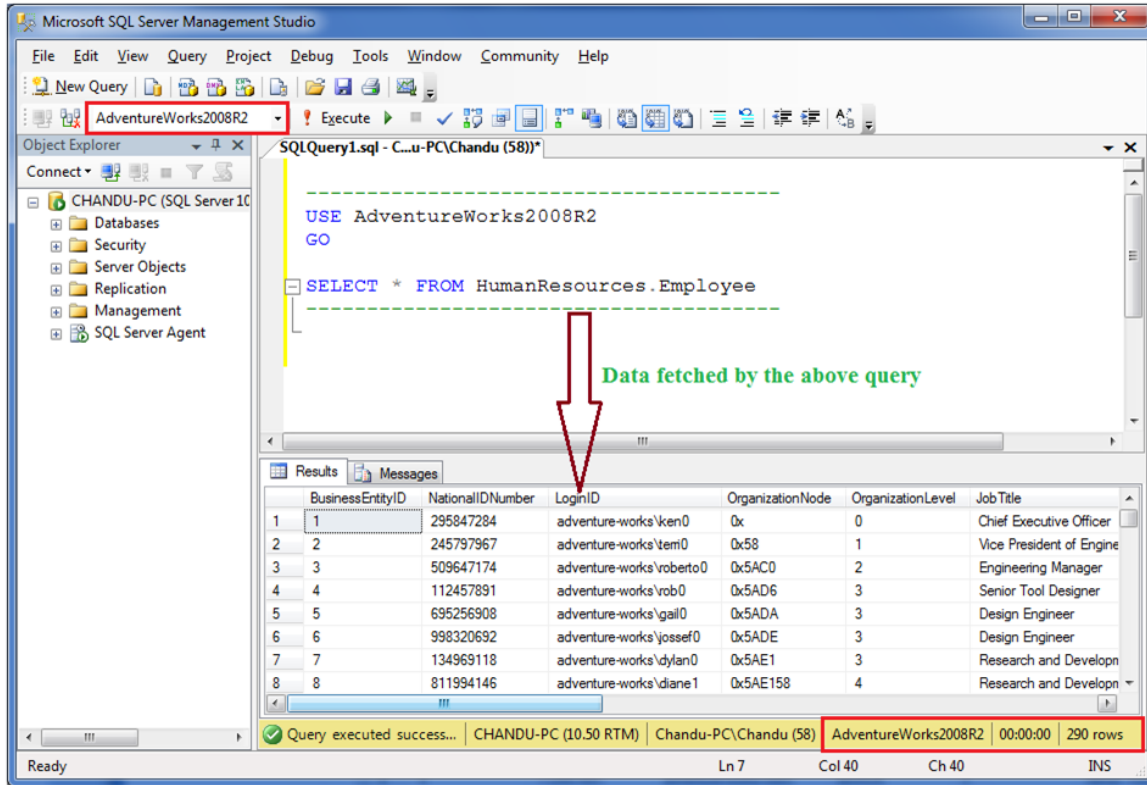

DirSource . . . . .	Source . . . . .
Docs . . . . .	stemCompletion . . . . .
findAssocs . . . . .	stemDocument . . . . .
findFreqTerms . . . . .	stopwords . . . . .
foreign . . . . .	stripWhitespace . . . . .
getTokenizers . . . . .	TermDocumentMatrix . . . . .
getTransformations . . . . .	termFreq . . . . .
inspect . . . . .	TextDocument . . . . .
meta . . . . .	tm_combine . . . . .
PCorpus . . . . .	tm_filter . . . . .
PlainTextDocument . . . . .	tm_map . . . . .
plot . . . . .	tm_reduce . . . . .
readDOC . . . . .	tm_term_score . . . . .
Reader . . . . .	tokenizer . . . . .
readPDF . . . . .	URISource . . . . .
readPlain . . . . .	VCorpus . . . . .
readRCV1 . . . . .	VectorSource . . . . .
readReut21578XML . . . . .	weightBin . . . . .
readTabular . . . . .	WeightFunction . . . . .
readTagged . . . . .	weightSMART . . . . .
readXML . . . . .	weightTf . . . . .
removeNumbers . . . . .	weightTfIdf . . . . .
removePunctuation . . . . .	writeCorpus . . . . .
removeSparseTerms . . . . .	
removeWords . . . . .	

# SQL Server

- is a Microsoft product to manage and store data.
- is a relational database management system (RDMS).
- uses Structured Query Language (SQL)

## Available Editions:

Enterprise  
Business Intelligence  
Standard  
Web  
Developer (free)  
Express (free)



The screenshot displays the Microsoft SQL Server Enterprise Manager interface. The Object Explorer on the left shows the server instance 'CHANDU-PC (SQL Server 10.50.1713.1)' with various folders like Databases, Security, and Server Objects. The main window shows a query window with the following SQL code:

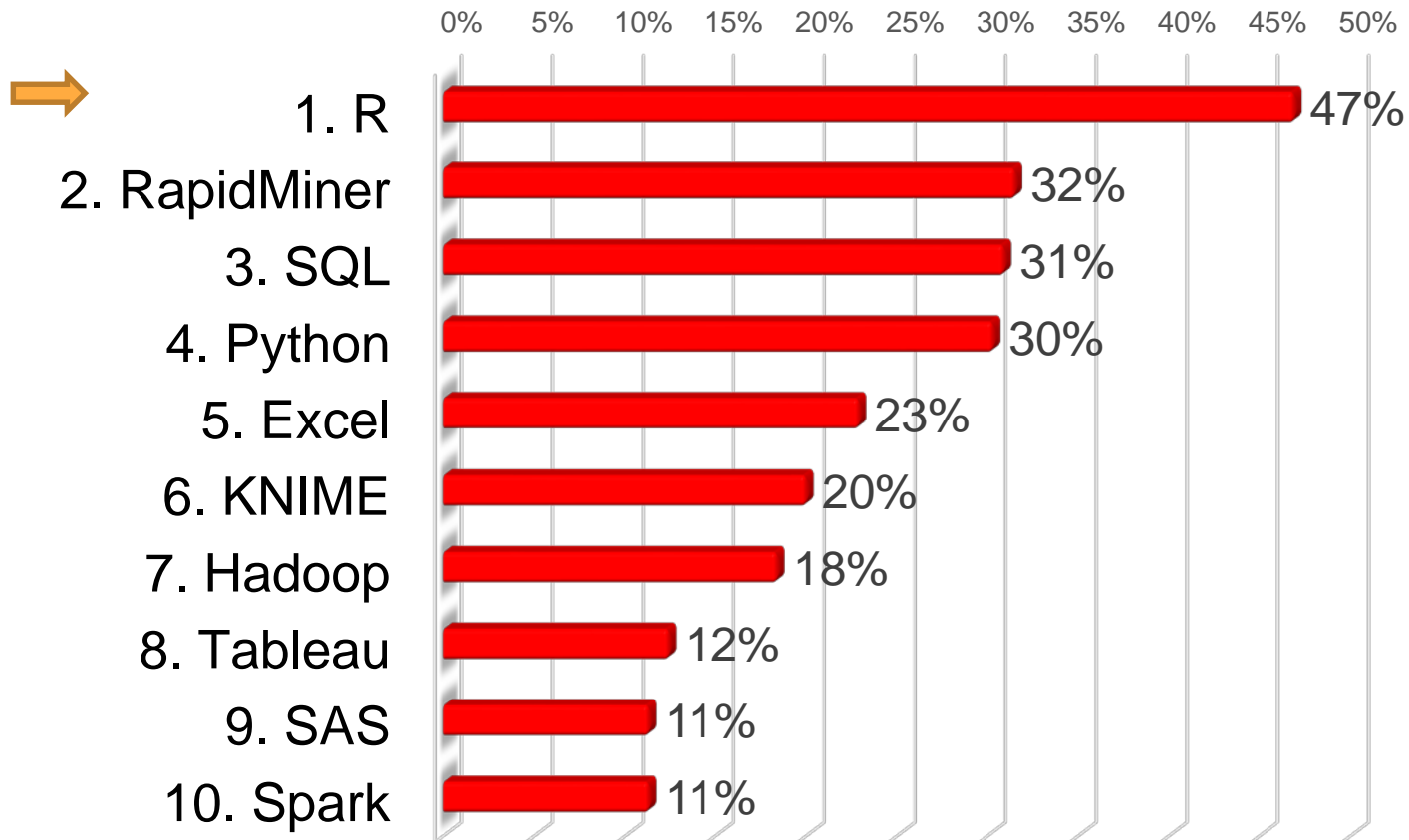
```
USE AdventureWorks2008R2
GO
SELECT * FROM HumanResources.Employee
```

A red arrow points from the text 'Data fetched by the above query' to the 'Results' pane, which displays the following table:

	BusinessEntityID	NationalIDNumber	LoginID	OrganizationNode	OrganizationLevel	JobTitle
1	1	295847284	adventure-works\ken0	0x	0	Chief Executive Officer
2	2	245797967	adventure-works\tem0	0x58	1	Vice President of Engine
3	3	509647174	adventure-works\roberto0	0x5AC0	2	Engineering Manager
4	4	112457891	adventure-works\rob0	0x5AD6	3	Senior Tool Designer
5	5	695256908	adventure-works\gail0	0x5ADA	3	Design Engineer
6	6	998320692	adventure-works\jossef0	0x5ADE	3	Design Engineer
7	7	134969118	adventure-works\dylan0	0x5AE1	3	Research and Developm
8	8	811994146	adventure-works\diane1	0x5AE158	4	Research and Developm

The status bar at the bottom indicates 'Query executed success...' and '290 rows'.

# Top 10 Analytics & Data Science Software, 2015



# R

**Creator:** Ross Ihaka and Robert Gentleman, University of Auckland, New Zealand and R Foundation

**Year Released:** 1995

**R** is an implementation of the **S** programming language by Bell Labs

**The design and evolution** are controlled by the R-core group and R foundation

**R is written** in C, Fortran and R.

**R** has been **used** in academia and finding its way to industry.

# What is R?

Freely available **language** and **environment** for statistical computing and graphics

R provides a wide variety of statistical and graphical techniques:

linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc.

Consists of a **language plus** a **run-time environment** with:

Graphics

A debugger

Access to functions stored in packages

Currently, the **CRAN package repository** features 7,802 available packages (<https://cran.r-project.org/>).

And the ability to run programs stored in script files.

# Top 10 Most Downloaded R Packages for Machine Learning, January-May 2015

1. **E1071**. Latent class analysis, short-time Fourier transform, fuzzy clustering, support vector machines, shortest path computation, bagged clustering, naïve Bayes classifier, etc. (142,479 downloads)
2. **RPart**. Recursive Partitioning and Regression Trees. (135,390 downloads)
3. **Igraph**. A collection of network analysis tools. (122,930 downloads)
4. **Nnet**. Feed-forward Neural Networks and Multinomial Log-Linear Models. (108,298 downloads)
5. **RandomForest**. Breiman and Cutler's random forests for classification and regression. (105,375 downloads)
6. **Caret**. Classification and REgression Training of predictive models. (87,151 downloads)
7. **Kernlab**. Kernel-based Machine Learning Lab. (62,064 downloads)
8. **Glmnet**. Lasso and elastic-net regularized generalized linear models. (56,948 downloads)
9. **ROCR**. Visualizing the performance of scoring classifiers. (51,323 downloads)
10. **Gbm**. Generalized Boosted Regression Models. (44,760 downloads)



# RStudio Interface

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R Markdown code for a document titled "Untitled". The code includes a title, author, date, and output format, followed by introductory text about R Markdown.
- Environment Pane:** Shows the "Global Environment" which is currently empty.
- Console:** Displays the output of the R Markdown rendering process, showing progress bars and labels for different chunks of text.
- Package Manager:** Lists installed and available packages from the System Library.

```
1 ---
2 title: "Untitled"
3 author: "My Name"
4 date: "Monday, May 25, 2015"
5 output: html_document
6 ---
7
8 This is an R Markdown document.
9 Markdown is a simple formatting syntax
10 for authoring HTML, PDF, and MS word
documents. For more details on using R
Markdown see <http://rmarkdown.rstudio
.com>.
9
10 when you click the **knit** button a
document will be generated that
includes both content as well as the
```

Console Output:

```
.../R Code/Fin Aid - 14 - Naive Bayes.Rmd
| 44%
label: unnamed-chunk-2
| .....
| 56%
ordinary text without R code
| .....
| ..... | 67%
label: unnamed-chunk-3
| .....
| ..... | 78%
ordinary text without R code
```

Name	Description	Versi...
<input type="checkbox"/> arules	Mining Association Rules and Frequent Itemsets	1.1-6
<input type="checkbox"/> arulesViz	Visualizing Association Rules and Frequent Itemsets	1.0-0
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-16
<input type="checkbox"/> caTools	Tools: moving window statistics, GIF, Base64, ROC AUC, etc.	1.17.1
<input type="checkbox"/> class	Functions for Classification	7.3-12
<input type="checkbox"/> cluster	Cluster Analysis Extended Rousseeuw et al.	2.0.1
<input type="checkbox"/> codetools	Code Analysis Tools for R	0.2-11
<input type="checkbox"/> coin	Conditional Inference Procedures	1.0-24

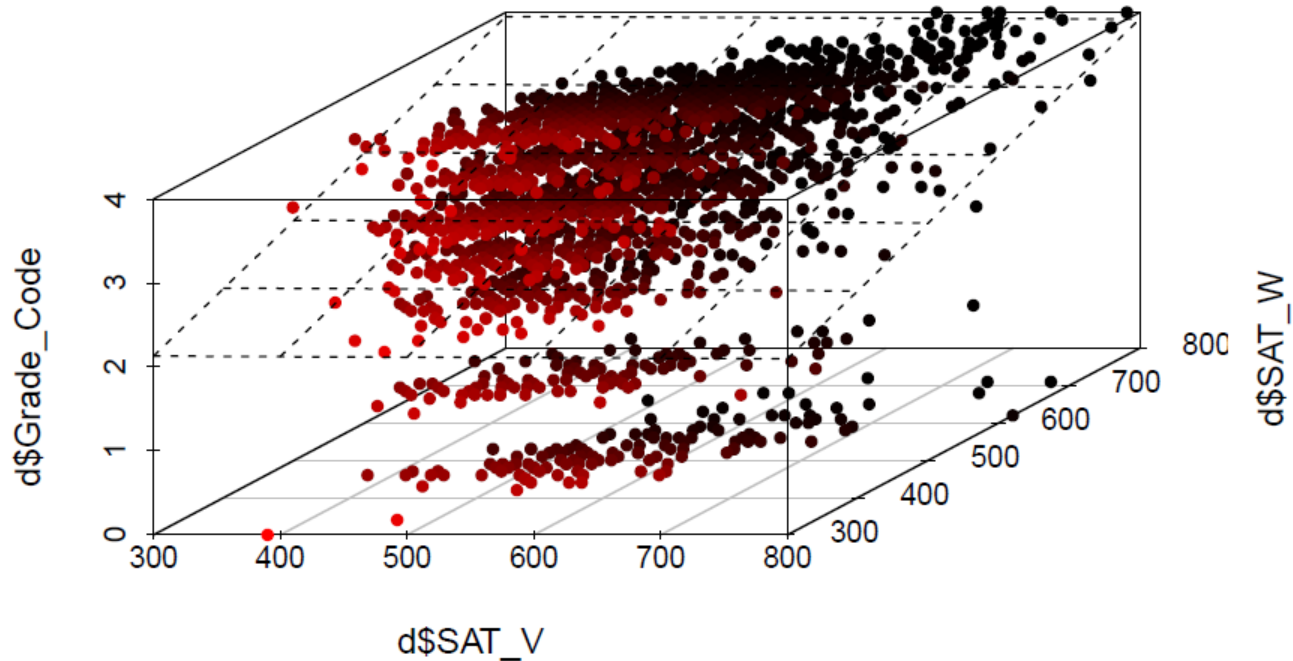
# R vs. SPSS vs. Excel

R	SPSS	Excel
<ul style="list-style-type: none"><li>• Freeware</li><li>• Flexible</li><li>• A lot of online help</li><li>• Powerful graphics</li><li>• Data-oriented programming language</li><li>• Statistics, data mining, and advanced machine learning</li><li>• Growing popularity and</li></ul>	<ul style="list-style-type: none"><li>• Expensive</li><li>• Point-and-click interface</li><li>• Does not require programming (though possible)</li><li>• Visualization, plotting, and statistics</li><li>• Popular in social sciences</li></ul>	<ul style="list-style-type: none"><li>• Data entry</li><li>• Data analysis and exploration</li><li>• Quick and easy data visualization</li><li>• Basic statistical analysis</li><li>• Widely known tool</li></ul>

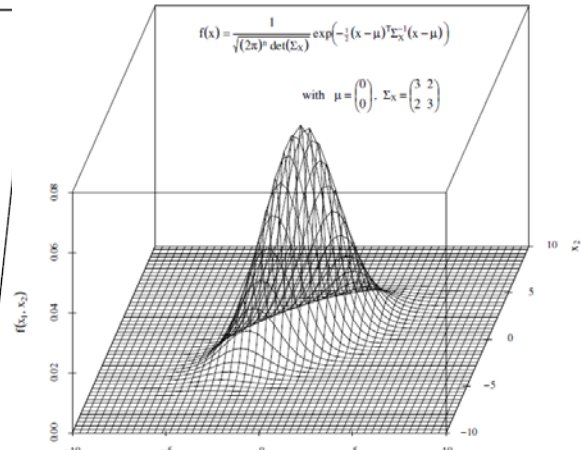
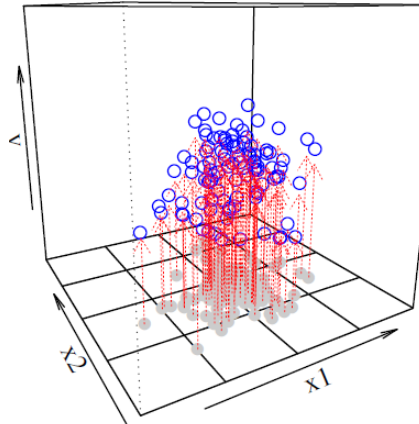
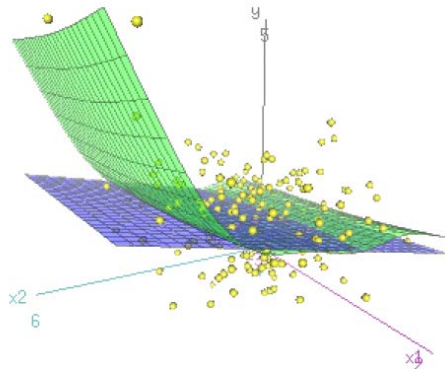
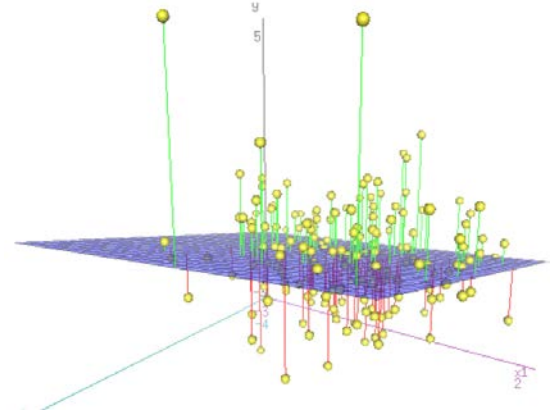
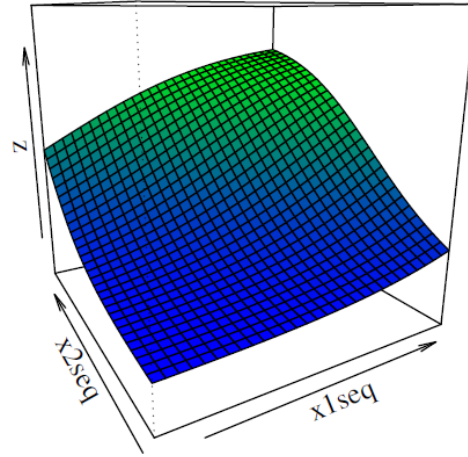
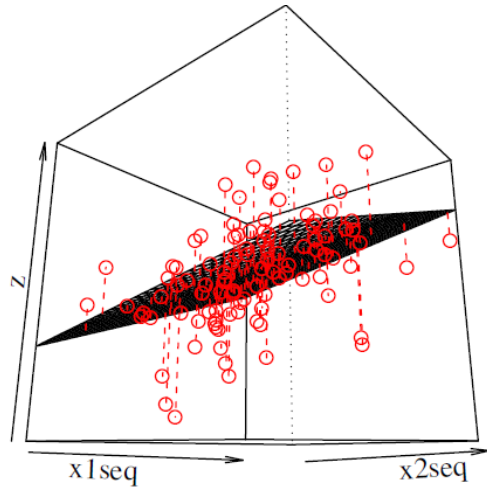
# R Graphics Example

Grade ~ SAT V + SAT W on full dataset

$r = .73, p < .01$



# More Charts in R



# Processing in R using TM package

Read a CSV file

Convert text to lower case

Remove

- Extra whitespace and non-printable characters

- Numbers

- Punctuation

Split text into n-grams

Build Term-Document Matrix

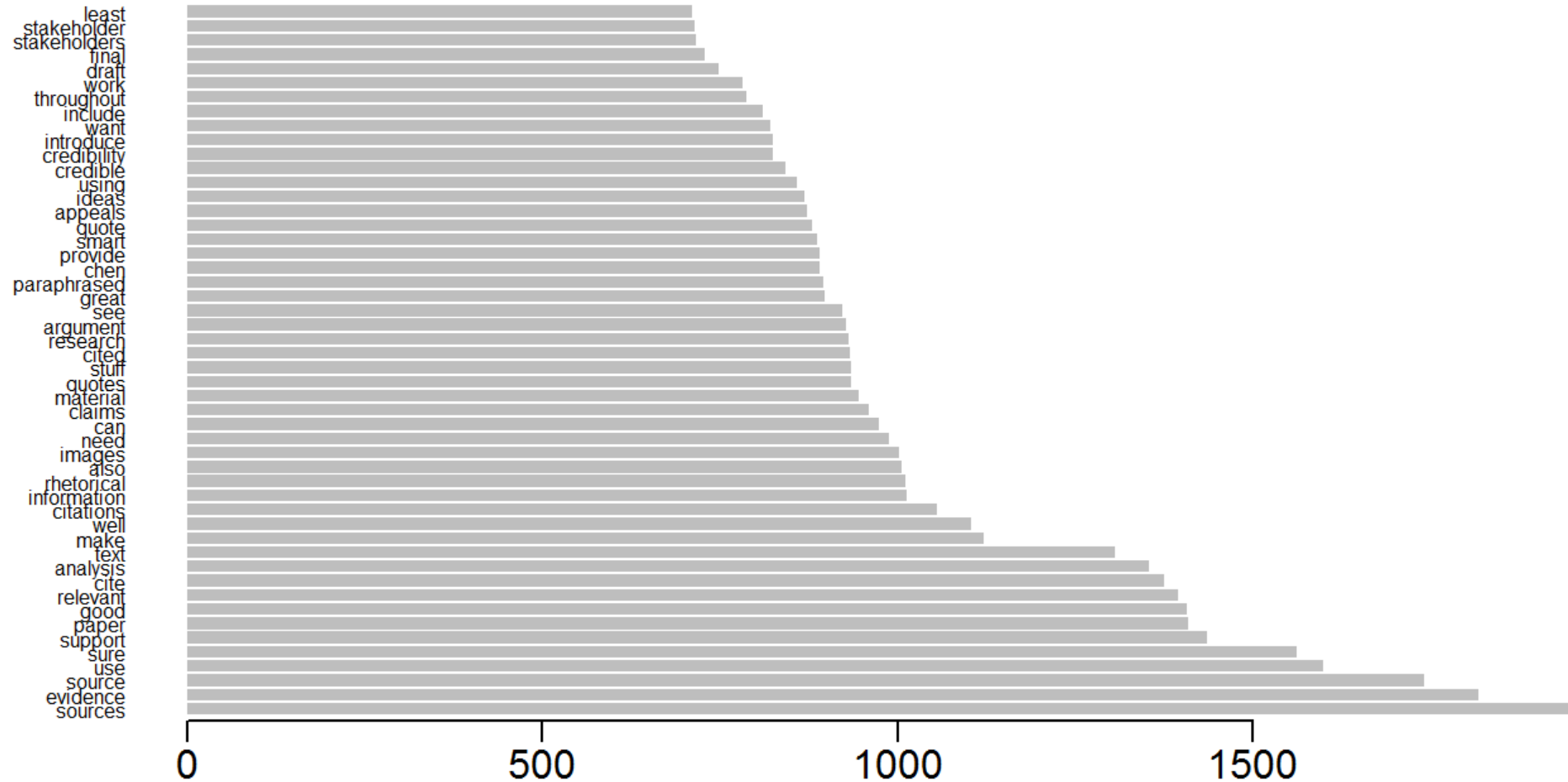
- N-grams are row headers



# Word Cloud of Most Frequent 1-grams

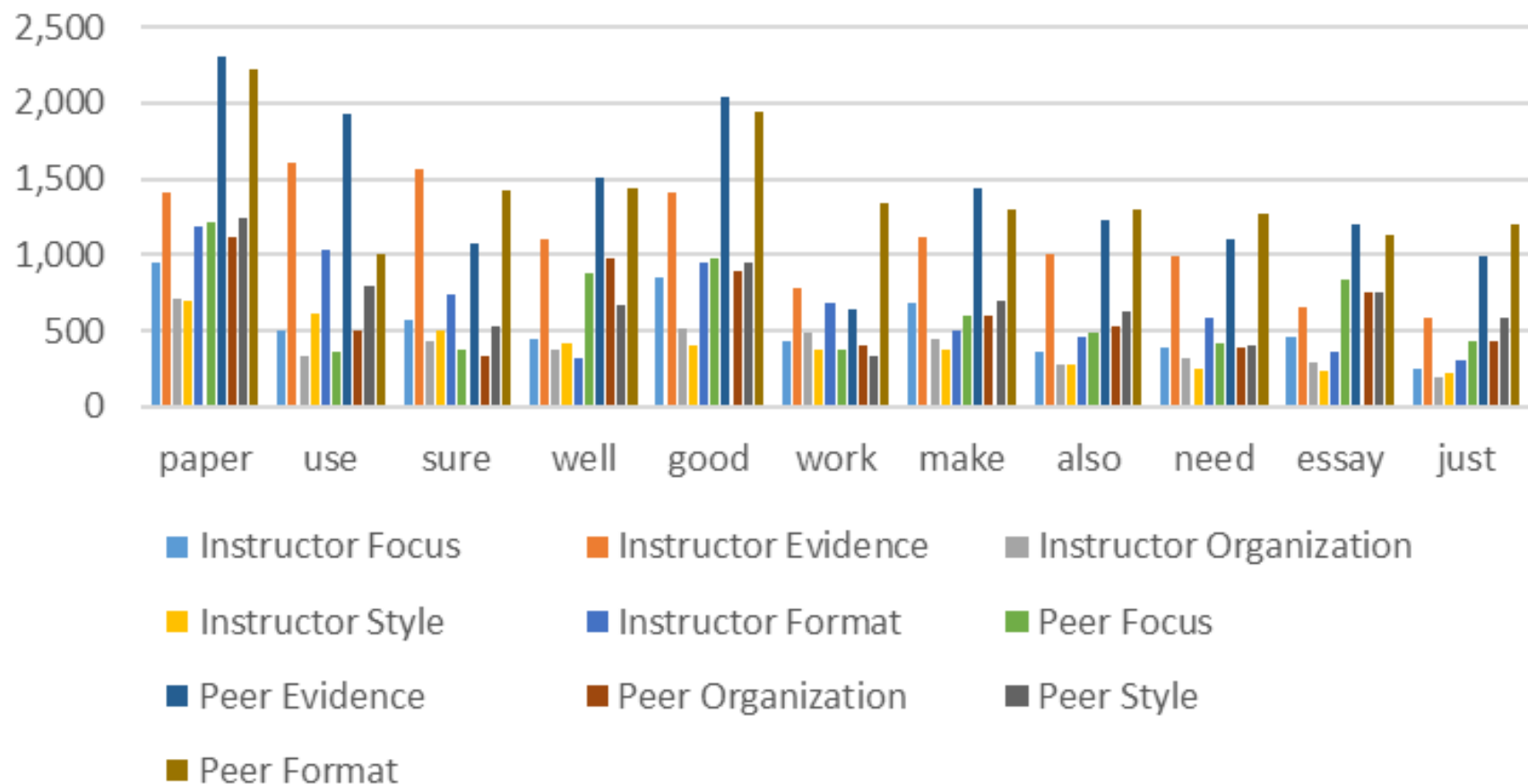


# Histogram of Most Frequent 1-grams

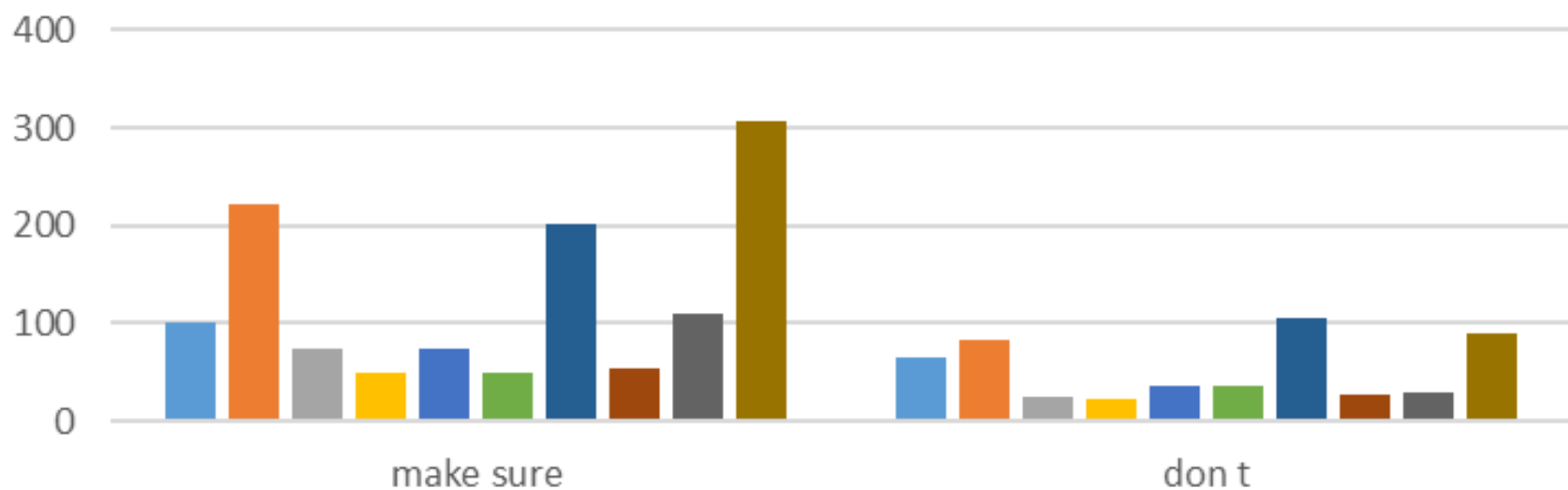




## Peer and Instructor Common Words



## Peer and Instructor Common Bigrams



Instructor Focus

Instructor Evidence

Instructor Organization

Instructor Style

Instructor Format

Peer Focus

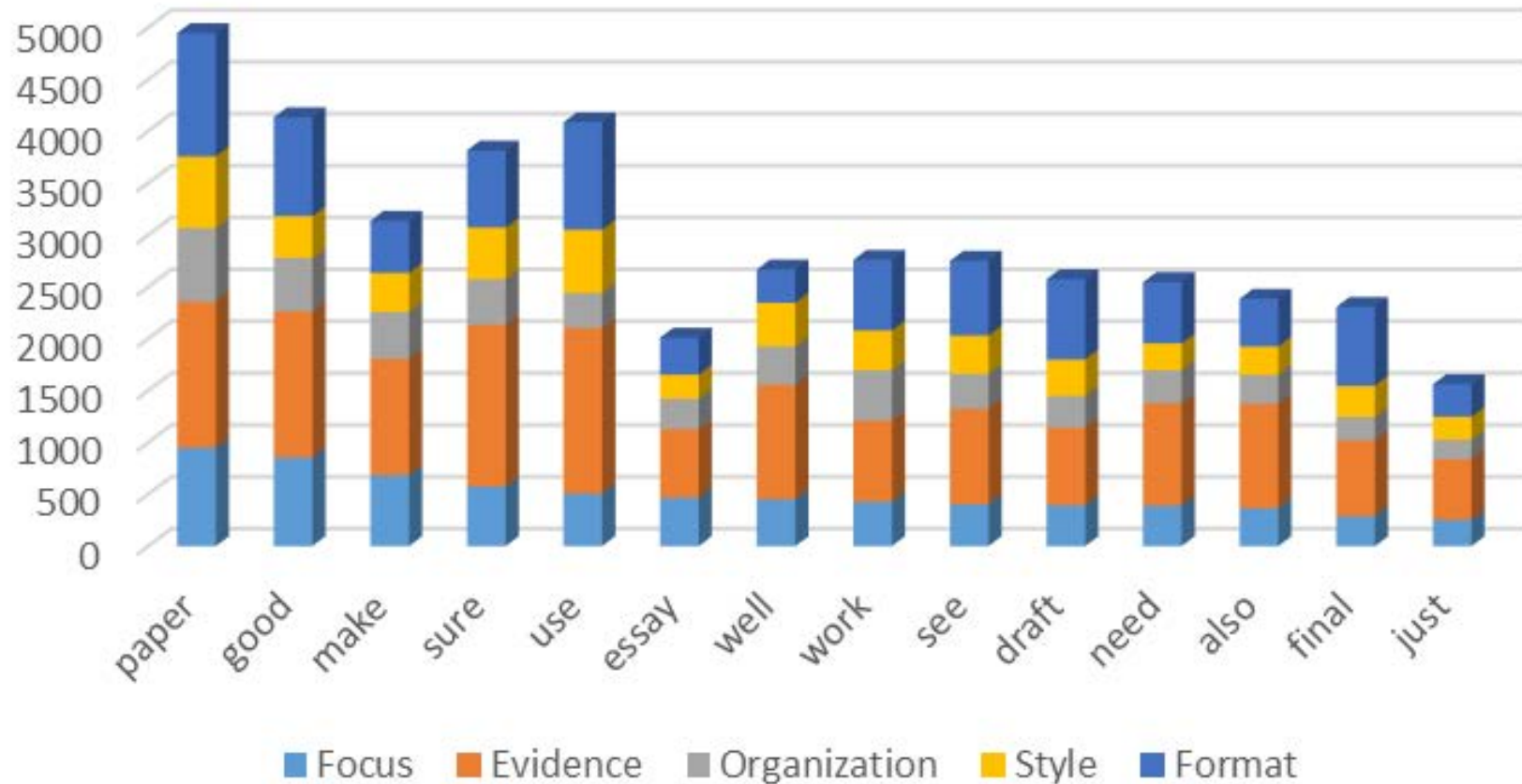
Peer Evidence

Peer Organization

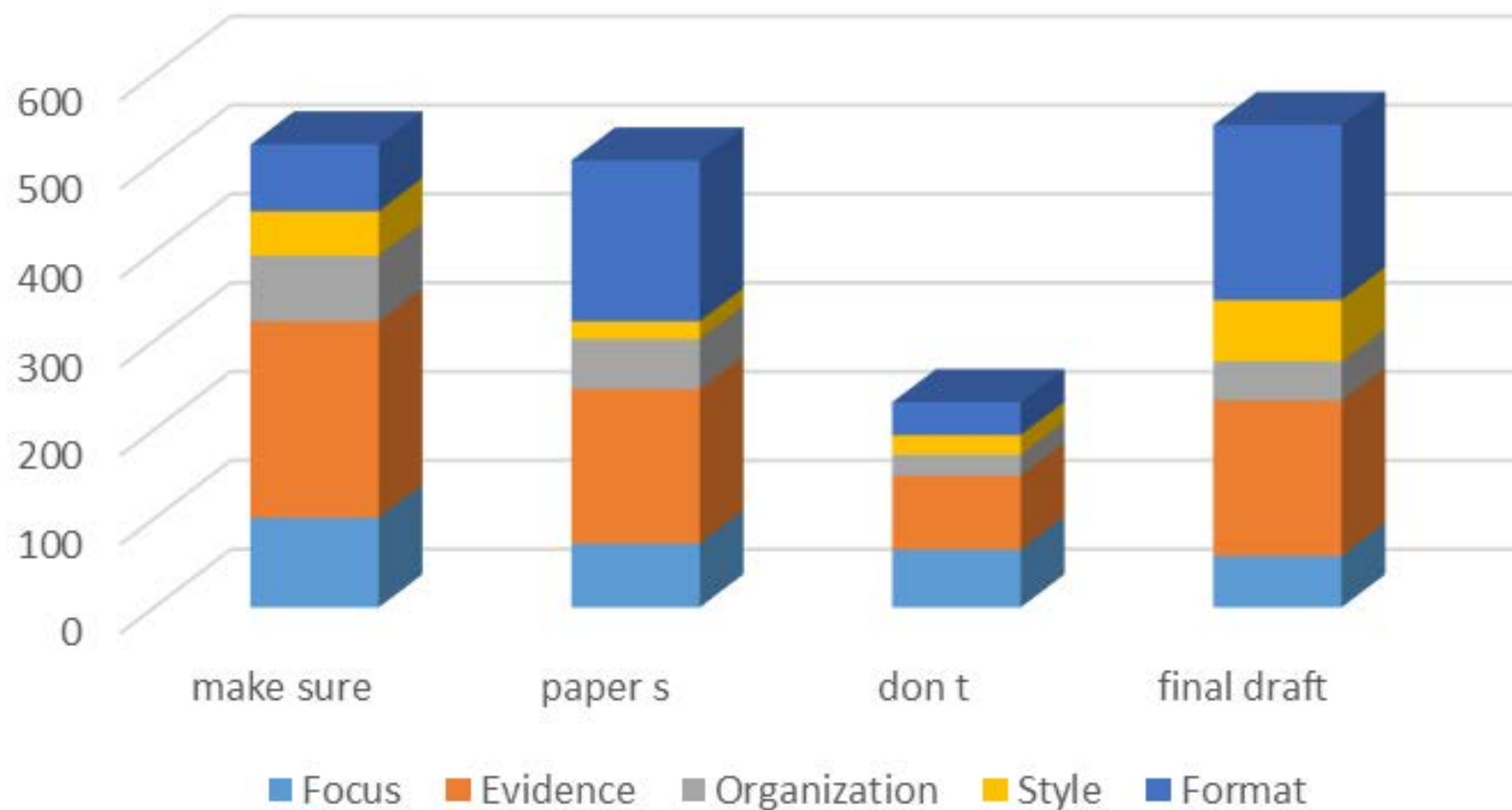
Peer Style

Peer Format

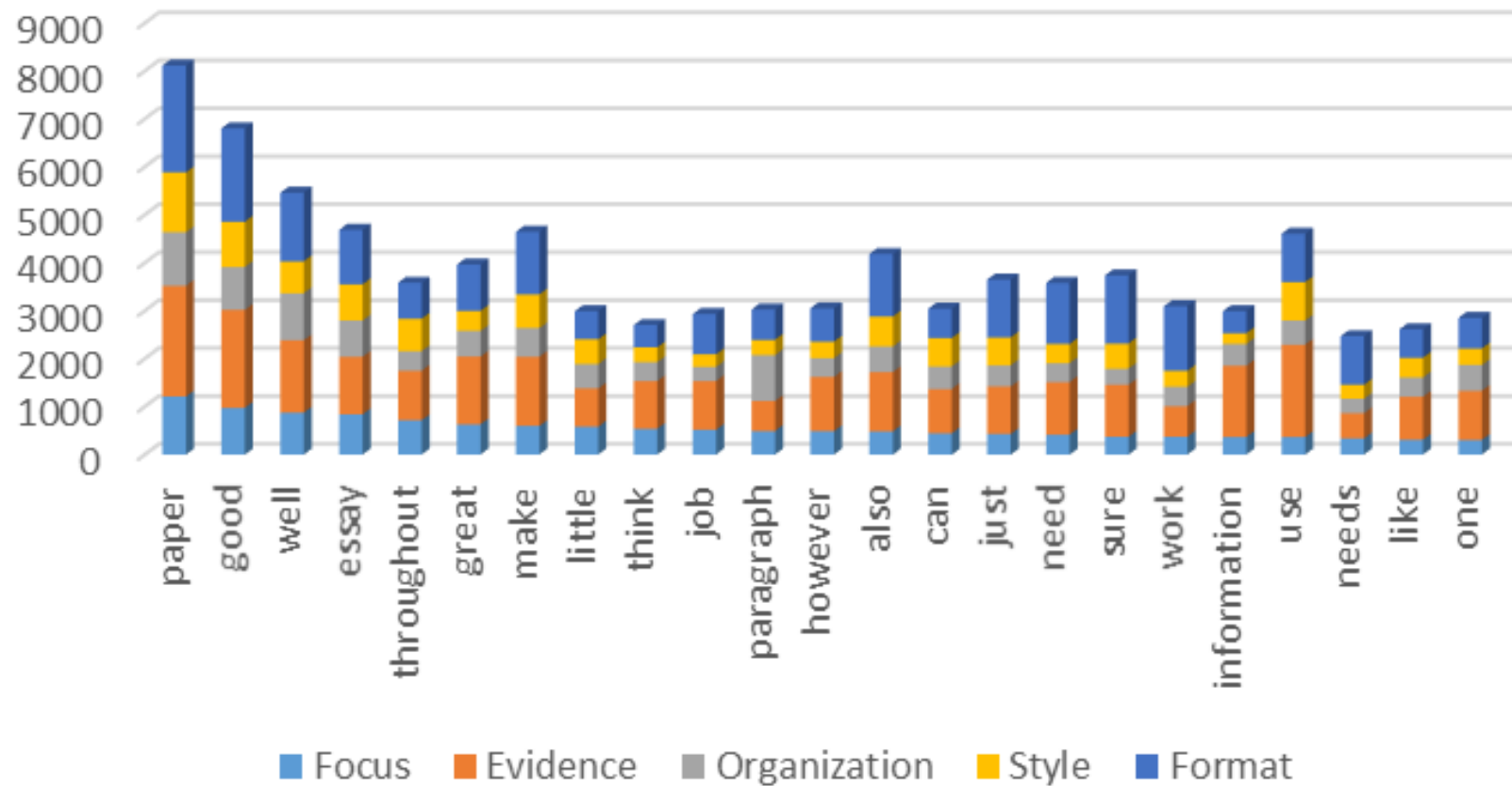
## Instructor Common Words



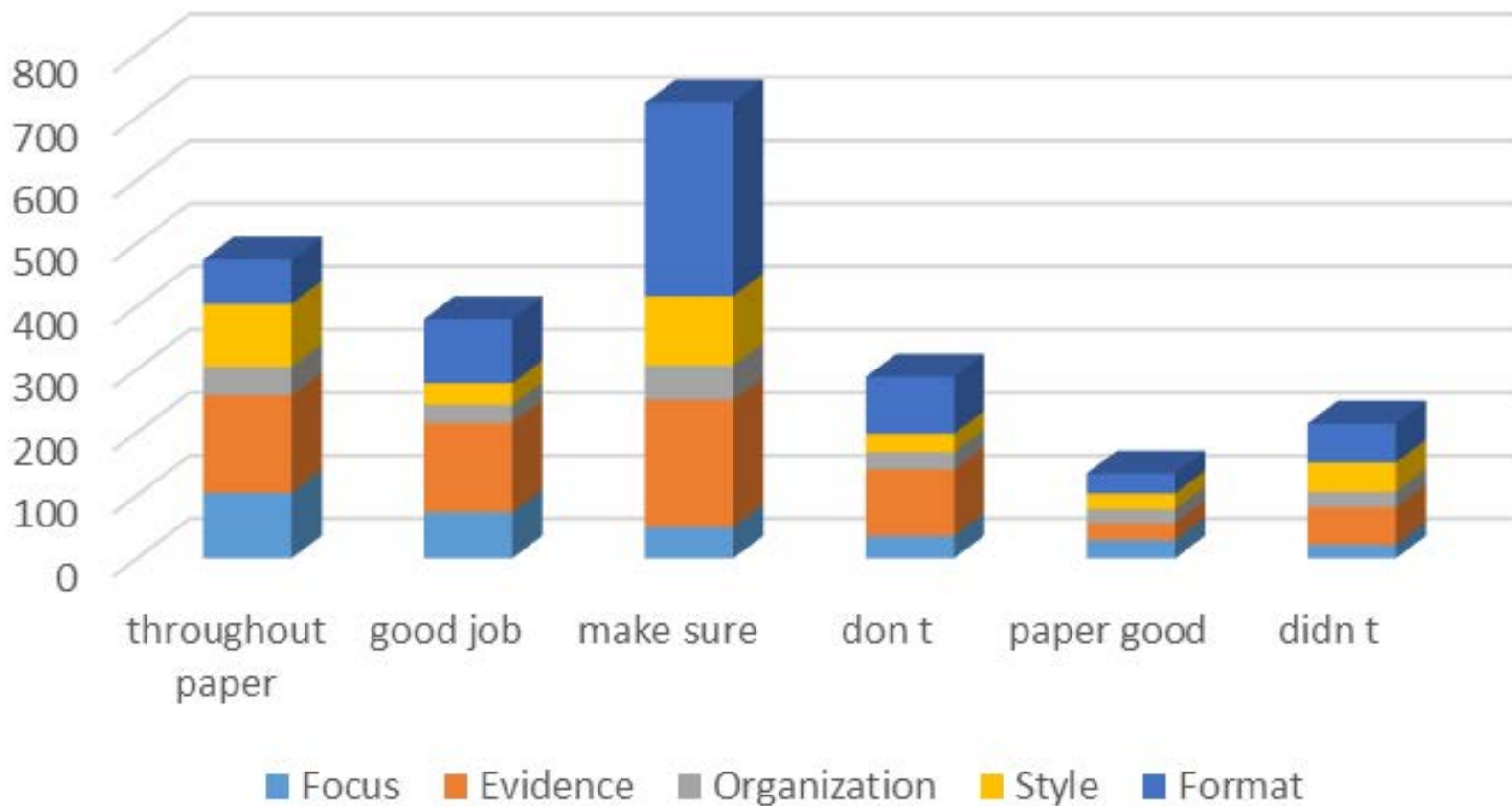
## Instructor Common Bigrams



## Peer Common Words



## Peer Common Bigrams



# Comparison of Peer Comments & Scores, Lower and Upper Quartile

Valerie Ross, Mark Liberman, Lan Ngo, Rodger LeGrand,  
University of Pennsylvania, June 2016

# How do peer comments correlate with peer scores?

Peer feedback is a common practice in writing instruction

Much attention has been paid to the kinds of comments and grades given by teachers (and tutors) to writing

Less attention has been focused on the content of peer assessment





# Findings

- ✦ Students in lower quartile appear to receive more direct instruction, more negative terms of evaluation, and more words in general from their peers.
- ✦ Students in upper quartile appear to receive more descriptive/indirect feedback, more positive terms of evaluation, and fewer words in general from their peers.

# Writing Feedback

**Direct:** telling, suggesting, explaining, exemplifying  
(Mackiewicz 2015)

**Indirect:** open problem solving or discovery learning (e.g.,  
Kirschner, Sweller, & Clark, 2006).

**Direct:** delivers essential information but may dampen  
curiosity and motivation (GloggerFrey, Fleischer, Gruny,  
Kappich, & Renkl, 2015)

**Indirect:** lack of direct instruction may interfere with  
learning and transfer (GloggerFrey; Kirschner)

# Negative Feedback

- ✦ High self-efficacy learners view their performance optimistically, and therefore, may seek negative feedback to outperform on tasks (Hattie & Timperley, 2007).
- ✦ Negative feedback for low self-efficacy students may adversely impact their motivation and future performance (Brockner, Derr, & Laing, 1987; Hattie & Timperley, 2007; Moreland & Sweeney, 1984).
- ✦ Negative feedback from teachers or peers may be confusing and harmful to EFL students' confidence (Kaivanpanah, Alavi, and Sepehrinia (2015)]; these effects can be mitigated by presenting negative feedback in terms of guidance (Straub, 1997).

# Motivational Scaffolding

Direct encouragement appears to aid students with low self-efficacy but may not be helpful for high self-efficacy learners (Boyer et al, 2008).



# Positive Feedback

- \* **Feedback one of the strongest influences on learning and achievement** [meta-analysis, Hattie and Timperley (2007)]
- \* Positive feedback may increase a student's persistence. For **high self-efficacy students**, may teach coping skills for future negative (Deci, Koestner, & Ryan, 1999; Hattie & Timperley, 2007; Swann, Pelham, & Chidester; 1988).
- \* **However, low self-efficacy students** may react to positive feedback by avoiding tasks to limit the risk of receiving future negative feedback (Hattie & Timperley, 2007)



# Method:

## Weighted log-odds-ratio, informative Dirichlet prior method

Bottom quartile: 3046 reviews with scores between 2 and 3.3 out of 4

Top quartile: 3054 reviews with scores above 3.78.

Combined comments in bottom quartile: 1,022,709 words

Combined comments in the top quartile: 759,637 words.

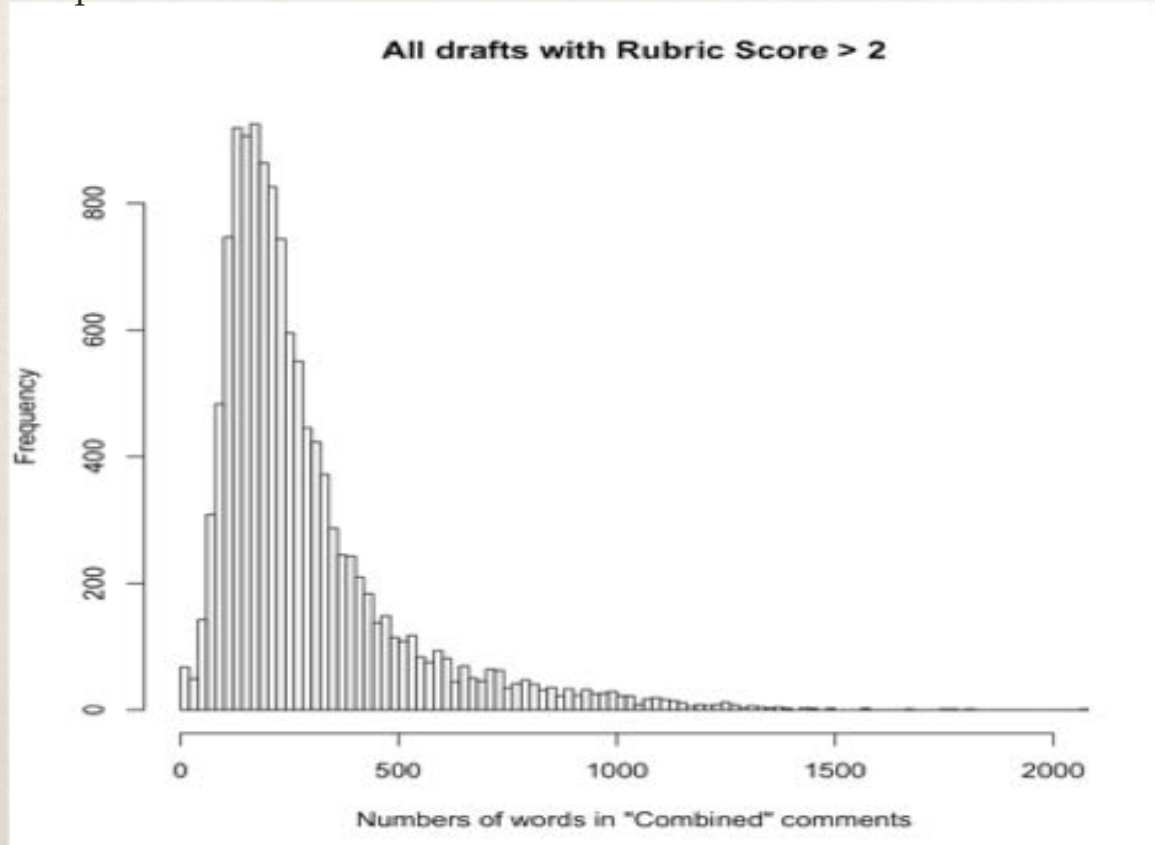
The word “should” occurs 3,780 times in the bottom-quartile comments, and 1,914 times in the top-quartile comments. Accounting for combined words, this tells us that the frequency of “should” is about 1.5 times greater in the bottom-quartile comments than in the top-quartile comments. But in this case, the overall frequency is high enough that we can be fairly confident that “should” will also be about 50% more frequent in the low-quartile comments in next semester’s sample – and “should” is common enough to be a useful indicator of overall review sentiment.

In order to evaluate the degree of association between individual words and score quartiles, we used the “algorithm from section 3.5.1” of Monroe et al. 2008. This method, originally developed for a study of political writing, starts with a simple ratio of estimated word frequencies in two collections of text.

# Data Set

- \* 1,183 undergraduate students (predominantly freshmen) drawn from Arts & Sciences, Wharton, Engineering and Nursing, who completed a writing seminar at the University of Pennsylvania in Spring 2016.
- \* Up to 5 drafts of a literature review
- \* Up to 6 peer reviews per draft, including rubric-guided scores and commentary
- \* Instructor commentary, feedback, and score

The bottom quartile has more words (per combined comment) than the top quartile: 336 v 249





The words most reliably associated with the bottom quartile include:

Word	Low Q Count	Low Q Freq Per Million	High Q Count	High Q Freq Per Million	Weighted Log Odds Ratio
be	10219	9992.09	5710	7516.75	7.261
sentence	9278	9071.98	5092	6703.2	7.221
more	7664	7493.82	3505	4614.05	10.11
paragraph	7001	6845.54	3443	4532.43	8.445
not	6424	6281.36	3516	4628.53	6.309
but	5123	5009.25	2742	3609.62	5.949
should	3780	3696.07	1914	2519.62	5.687
some	2984	2917.74	1529	2012.8	4.925
however	2701	2641.02	1255	1652.1	5.617
than	1938	1894.97	945	1244.02	4.536
seems	1719	1680.83	720	947.821	5.626
sure	1268	1239.84	549	722.714	4.78
rather	1052	1028.64	425	559.478	4.708
try	888	868.282	316	415.988	4.962
needs	793	775.392	253	333.054	5.439
media	731	714.768	208	273.815	5.235
pass	300	293.339	35	46.0746	5.514
chaplin	205	200.448	19	25.0119	4.765

The words most reliably associated with the top quartile include:

Word	Low Q Count	Low Q Freq Per Million	High Q Count	High Q Freq Per Million	Weighted Log Odds Ratio
the	71418	69832.2	56903	74908.1	-5.51
and	26778	26183.4	23808	31341.3	-8.526
is	20103	19656.6	16680	21957.9	-4.52
very	3391	3315.7	5269	6936.21	-14.319
well	3474	3396.86	4763	6270.1	-11.643
was	3185	3114.28	3842	5057.68	-8.508
good	3222	3150.46	3169	4171.73	-4.742
topic	2738	2677.2	2751	3621.47	-4.619
piece	2647	2588.22	2591	3410.84	-4.236
clear	2206	2157.02	2211	2910.6	-4.13
all	1773	1733.63	2083	2682.86	-5.637
job	1440	1408.03	1984	2611.77	-7.544
great	1149	1123.49	1811	2384.03	-8.504
really	1330	1300.47	1682	2214.22	-6.091
interesting	1447	1414.87	1653	2176.04	-4.985
easy	676	660.99	1096	1442.79	-6.835
strong	824	805.703	1091	1436.21	-5.281
read	906	885.882	1059	1394.09	-4.177
written	593	579.833	770	1013.64	-4.303
liked	130	127.113	324	426.52	-5.147
enjoyed	99	96.8017	257	338.319	-4.64
picasso	27	26.4005	112	147.439	-4.294
twins	1	0.977795	111	146.122	-5.658
identical	5	4.88898	76	100.048	-4.47



WORD	RATIO
unclear	2.004
incorrect	1.969
unnecessary	1.825
needs	1.729
clearer	1.688



WORD	RATIO
easy	2.939
great	2.857
very	2.816
nice	2.716
flows	2.553
logically	2.547
organized	2.500
job	2.497
well	2.485
supported	2.456
fits	2.419
strong	2.400
really	2.292
nicely	2.251

WORD	RATIO
convincing	2.211
presentation	2.155
persuasive	2.122
coherent	2.118
engaging	2.111
interesting	2.071
consistent	1.983
supports	1.949
clearly	1.932
helps	1.927
appropriate	1.925

## Questions:

How is peer review affecting students who struggle with writing?

How might we better prepare students to give and receive feedback?

Which peer feedback strategies appear to be most effective for students?

Are instructors demonstrating a similar feedback pattern?

# An Invitation: Join Us!

## Conference: 2017 Writing Analytics, Data Mining, and Student Success

Overview

Call for Participation

Organizers

Program

Writing Analytics, Data Mining and Student Success  
January 12-13, 2017, Tampa, FL, USA  
University of South Florida



Intended for Writing Program Directors, *Writing Analytics, Data Mining and Student Success* will explore innovations in writing analytics and data mining.