

A Sentiment and Linguistic Analysis of Online Reviews of Online Universities

Alan Peslak
arp14@psu.edu
Information Sciences and Technology
Penn State University
Dunmore, PA 18512 USA

Abstract

The growth and proliferation of online higher education have been staggering. This comes despite an overall decline in university enrollments. But the quality of online courses has been questioned by many researchers, suggesting it may be less than a traditional face to face experience. Many researchers have explored this area but our research study reviews online universities from an impartial non-profit organization and analyzes both the content and the sentiment in these reviews. This analysis provides a unique perspective on how online universities and their programs are viewed by their students. Implications of this analysis are examined.

Keywords: online education, reviews, sentiment analysis, linguistic analysis, online universities, online reviews

1. INTRODUCTION

The growth of online studies in higher education has been phenomenal. The number of students taking at least one online university course has grown from 1.2 million in 2002 to 5.8 million in 2014 (WICHE, 2016). According to the 2016 Online Report Card that Babson Survey Research Group publishes yearly, nearly 1/3 (32%) of all higher education students are enrolled in at least one online course. Of this 32%, 15% are fully online and 17% take a mix of traditional delivery with some online courses. (Seaman, Allen and Seaman, 2016). The concentration of distance education enrollments is noteworthy. "Almost half of distance education students are concentrated in just five percent of institutions, while the top 47 institutions (just 1.0% of the total) enroll 22.4% (1,421,703) of all distance students." (Seaman et al. 2016). The growth in online education has been non-stop with increased enrollments for the fourteenth straight year. This growth in online students came to a large extent

at the expense of face-to-face instruction. Overall, students in higher education declined by a total of 800,000 students between 2012 and 2016. This means that with the growth in online education total face-to-face students are down more than one million students over the past 4 years. (Seaman, Allen, and Seaman, 2016). Given this tremendous growth in online higher education, it is essential that the quality of online higher education be thoroughly examined. Many studies have been prepared that begin to address this issue. Specifically, our research questions are what are the current views of students of online higher education. We see student opinions as an approach to understanding and estimating the current quality of online higher education. Our method is to analyze the current status via a review of social media, specifically by analyzing sentiment and linguistics of online reviews of online higher education institutions. These reviews are the opinions of students who have taken courses in these universities. They will only serve as a surrogate to overall quality but student

opinions and reviews of course quality have shown to correlate strongly with other means of evaluations (Marsh, 1987).

2. LITERATURE REVIEW

Many researchers have explored the attitudes of students regarding the quality and effectiveness of their online university classes. Tallent-Runnels (2006) reviewed all research literature in 2006 and found many problems and inconsistencies in online learning. Braun (2008) performed a much-cited study of 90 grad students at one Western US university. He found that 90% of students found online courses were viewed as more challenging than traditional face to face courses and that only 8% would not want to take an online course again. Margaryan, A., Bianco, M., & Littlejohn, A. (2015) found that many Massive Open Online Courses (MOOCs) had poor instructional design.

There is also research that suggests that online education has improved over time. Allen and Seaman (2016) found 57% of students rated their online graduate courses as good as or better than face to face in 2003. This percentage increased over the years with 77% of respondents rating online as good or better. This is strong improvement over the decade and comes along with the growth and proliferation of online higher education.

Our research analyzes online reviews of online universities. The study of user reviews has a robust track in current research. Much work has been accomplished in analyzing product reviews on sites such as Amazon (Mudambi, & Schuff, 2010) Dellarocas, Zhang, & Awad (2007) and Duan, Gu, & Whinston (2008) analyzed movie sales and popularity based on online movie reviews and found a positive correlation. Ye, Law, & Gu (2009) studied the impact of online reviews and their effect on hotel room bookings. Likewise, Clemons, Gao, & Hitt (2006) studied how online reviews affect craft beer differentiation. Hu, Bose, Koh, & Liu (2012) began work on analyzing the sentiment of online reviews. Their work centered on use of sentiment analysis to determine bias and manipulation of reviews. Bowen, Chingos, Lack, & Nygren (2014) concluded that if online courses are well designed that "have the potential" to achieve "equivalent educational outcomes"

There has also been preliminary work done on what has come to be known as E-WOW or electronic word of mouth. Filieri, & McLea (2014) analyzed social media and reviews of travel sites to determine relevant variables.

As to validity of ratings, Marsh (1987) generally found that student ratings of courses generally were "valid against a variety of indicators of effective teaching". Therefore, we believe we can use online students reviews as a relatively accurate reflection of learning effectiveness.

3. METHODOLOGY

Based on past studies, we have selected student online opinions as the basis for determining perceived relative views of online universities. These student opinions will then serve as a proxy for perceived quality of the online university program. We thus searched for a reliable review source available for study.

All reviews have been obtained from the website Guide to Online Schools. Data were gathered in February of 2018. <https://www.guidetoonlineschools.com/online-reviews> Permission has been obtained to use this data. The parent organization which solicits and hosts these reviews is the Washington based SR Education Group. Their overall mission is

"Our products are designed to help prospective students find a college suited to their individual needs, whether that means low tuition costs, high satisfaction reported by recent graduates, or degrees that lead to career advancement. We feature schools with great student reviews and strong success metrics, and provide unbiased, comprehensive information."

Each review was analyzed for content and sentiment.

Sentiment analysis (Polarity) was obtained via Meaning Cloud for Excel and summary results for each were analyzed via SPSS (N=negative, P=Positive, etc.)

The following glossary from Meaning Cloud (2018)describes their coding:

"Agreement This field marks the agreement between the sentiments detected in the text, the sentence or the segment it refers to. It has two possible values:

- AGREEMENT: the different elements have the same polarity.
- DISAGREEMENT: there is disagreement between the different elements' polarity.

Subjectivity This field marks the subjectivity of the text. It has two possible values:

- OBJECTIVE:** the text does not have any subjectivity marks.
- SUBJECTIVE:** the text has subjective marks.

Confidence This field represents the confidence associated with the sentiment analysis performed on the text. Its value is an integer number in the 0-100 range.

Irony This field indicates the irony of the text. It has two possible values:

- **NONIRONIC:** the text does not have ironic marks
- **IRONIC:** the text has ironic marks

Detection of irony

identifies comments in which what is expressed is the opposite of what is said.

Graduated polarity

distinguishes very positive and very negative opinions, as well as the absence of sentiment.

Agreement and disagreement

identifies opposing opinions and contradictory or ambiguous messages."

In addition, LIWC (Linguistic Analysis and Word Count) from Pennebaker was used to analyze linguistic meaning embedded in the reviews. The use of LIWC is well established in the literature. Robinson, Navea, and Ickes (2013) used LIWC analysis of students written self-introductions to grades that students achieved. Cordova, Cunningham, Carlson, and Andrkowski (2001) used LIWC to analyze how individuals adjusted to having breast cancer. There are many more examples of the use of LIWC used for scholarly research.

LIWC (Linguistic and Word Count) software (Pennebaker, Booth, Boyd, and Francis, 2015) is one of the most accepted and popular linguistic analysis tool. "The way that the **L**inguistic **I**nquiry and **W**ord **C**ount (LIWC) program works is fairly simple. Basically, it reads a given text and counts the percentage of words that reflect different emotions, thinking styles, social concerns, and even parts of speech. Because LIWC was developed by researchers with interests in social, clinical, health, and cognitive psychology, the language categories were created to capture people's social and psychological states. The text analysis module then compares each word in the text against a user-defined dictionary. As

described below, the dictionary identifies which words are associated with which psychologically-relevant categories.

After the processing module has read and accounted for all words in a given text, it calculates the percentage of total words that match each of the dictionary categories. For example, if LIWC analyzed a single speech that was 2,000 words and compared them to the built-in LIWC2015 dictionary, it might find that there were 150 pronouns and 84 positive emotion words used. It would convert these numbers to percentages, 7.5% pronouns and 4.2% positive emotion words." (Pennebaker Conglomerates, 2015).

LIWC was used in our study to enhance understanding of online review content as well as to enhance our findings.

4. RESULTS

The results of our sentiment and linguistic analyses are presented in the following tables.

		Frequency	Percent
Valid	N+	9	.7
	N	176	13.6
	NEU	141	10.9
	NONE	6	.5
	P	835	64.5
	P+	127	9.8
	Total	1294	100.0

Table 1 Meaning Cloud Polarity

Overall results of all reviews and the Polarity of each review are shown in table 1. The scale runs from N+ (very negative), N (Negative), NEU(Neutral), None, P (Positive) and P+ (Very Positive) and reflects the overall sentiment of the review. Specific dictionaries and other analytical techniques are used by Meaning Cloud to determine whether a particular review expresses an overall good (positive) or bad (negative) expression. According to Meaning Cloud "Our Sentiment Analysis API performs a detailed, multilingual sentiment analysis on information from different sources."

The text provided is analyzed to determine if it expresses a positive, neutral or negative sentiment (or if it is impossible to detect). In order to do so, the individual phrases are identified and the relationship between them is evaluated, which results in a global polarity value of the text as a whole.

In addition to the local and global polarity, the API uses advanced natural language processing techniques to detect the polarity associated with both the entities and the concepts of the text. It also allows users to detect the polarity of entities and concepts they define themselves, which makes this tool applicable to any kind of scenario."

The table 1 results indicate that 74% of the posted reviews were positive or very positive. In addition, 11% were neutral. The overall sentiment analysis of the 1295 results (6 showed no sentiment) clearly suggests that online universities are currently viewed favorably by an overwhelming majority. Table 2 however, does suggest that this may not be unqualified. The reviews often have mixed emotions with 800 showing some disagreement, whereas only 494 showed agreement. Different elements of the texts have different sentiment or polarity.

		Frequency	Percent
Valid	AGREEMENT	494	38.2
	DISAGREEMENT	800	61.8
	Total	1294	100.0

Table 2 Meaning Cloud Agreement/Disagreement

Irony was not apparent in the online university reviews with only 30 of the 1294 reviews expressing ironic marks as shown in Table 3.

		Frequency	Percent
Valid	IRONIC	30	2.3
	NONIRONIC	1264	97.7
	Total	1294	100.0

Table 3 Meaning Cloud Irony Measure

As expected with reviews, the vast majority of the reviews (85%) are of a subjective nature (Table 4), expressing individual opinions rather than objective facts. Chi Square analysis confirms the significance of this finding at $p < .001$. (Table 5 and 6)

		Frequency	Percent
OBJECTIVE	204	15.8	
SUBJECTIVE	1090	84.2	
Total	1294	100.0	

Table 4 Meaning Cloud Subjectivity

	OBJECTIVE	SUBJECTIVE	Total
Count No Resp.	7	0	0
Expected Count	.0	1.1	5.9
% within Agreement	100.0%	0.0%	0.0%
Count	0	110	384
Expected Count	2.7	77.5	413.9
AGREE % within Agreement	0.0%	22.3%	77.7%
Count	0	94	706
Expected Count	4.3	125.4	670.3
DIS % within Agreement	0.0%	11.8%	88.3%
Total Count	7	204	1090
Expected Count	7.0	204.0	1090.0
% within Agreement	0.5%	15.7%	83.8%

Table 5 Subjectivity/Expected Levels

	Asymptotic Significance (2-sided)
Pearson Chi-Square	.000
Likelihood Ratio	.000
N of Valid Cases	

Table 6 Chi Square Significance Variation

The next analysis performed was to determine whether all online universities were viewed the same or whether there were differences in polarity based on the university. A summary by University was performed and via SPSS ANOVA. We found that mean scores were significantly different among the 340 universities at $p < .003$.

There is significant difference between universities with regard to polarity. Overall averages by school show that there were 341 schools represented and the overall mean was 3.8, nearly a full 4.0 positive (table 8). No schools had an average score of N+ or very negative.

	N	Min	Max	Mean
Mean	341	2.00	5.00	3.8143
Valid N (listwise)	341			

Table 8 Central Tendency of Polarity

A word count analysis was also performed and relevant value-laden and associated words were collected and shown in table 9. As is apparent, most words are positive with great, well, good, helpful and best all with over 150 mentions in the content. This supports the preliminary findings of generally positive emotions associated with online universities today.

Order	Unfiltered word count	Occurrences	%
56	great	474	0.273
88	well	280	0.161
92	recommend	269	0.155
98	good	250	0.144
111	helpful	210	0.121
149	best	159	0.092
153	however	151	0.087
166	easy	139	0.08
171	nothing	136	0.078
172	better	136	0.078
176	support	134	0.077
187	care	126	0.073
195	hard	121	0.07
203	different	113	0.065
205	challenging	111	0.064

Table 9 Value Laden Words and Count

In general, online university programs are viewed favorably with 74% of the 1294 reviews from 340 online universities.

Linguistic Results

Linguistic analysis was performed using LIWC software and is shown in table 10.

The linguistic analysis yielded some interesting results about the type and style of reviews for online universities.

According to Pennebaker, Booth, Boyd, and Francis (2015) "Analytical thinking -- a high

number reflects formal, logical, and hierarchical thinking; lower numbers reflect more informal, personal, here-and-now, and narrative thinking." The analytic measure had a mean for all 1294 reviews of 61.4908 which suggests a mid-high range of logical thought versus informal thought. Standards for different types of communication are shown in table 11 for comparison. This table shows scores for each analytic measure for common forms of communication such as NY Times, expressive writing, etc. As the table shows, blog posts have a much lower rating than these reviews, at 49.89 and natural speech is only 18.43. This suggests that significant thought was incorporated into the review and their reliability can be considered high. These do not appear to be random thoughts, but well thought-out opinions.

"Emotional tone -- a high number is associated with a more positive, upbeat style; a low number reveals greater anxiety, sadness, or hostility. A number around 50 suggests either a lack of emotionality or different levels of ambivalence." (Pennebaker, Booth, Boyd, and Francis, 2015). The mean tone of the reviews also was fairly high at 67.3723. This supports the positive polarity found via Meaning Cloud since a higher number reflects an upbeat versus hostile tone. Surprisingly, the tone is about equal to the New York Times at 68.17 and much more upbeat than blogs at 54.5.

"Clout -- a high number suggests that the author is speaking from the perspective of high expertise and is confident; low Clout numbers suggest a more tentative, humble, even anxious style." (Pennebaker, Booth, Boyd, and Francis, 2015). The mean clout variable expresses the degree of confidence expressed in the review. This was fairly low at 44.6430 and about on par with blogs. We see this as confirming the Meaning Cloud finding of a high level of Disagreement in the reviews. Reviewers seem to have presented both sides of the issue in their reviews, both positive and some negative and this reflects in a humbler clout factor.

"Authentic -- higher numbers are associated with a more honest, personal, and disclosing text; lower numbers suggest a more guarded, distanced form of discourse". (Pennebaker, Booth, Boyd, and Francis, 2015). The authenticity or honest communication aspect of the reviews shows a mean of 50.3574, nearly exactly in the middle of the honest to guarded spectrum. Whereas the NY Times has a rating of high honesty at 24.84 since they present mostly facts,

the reviews reflect an even mix between honest facts and opinion.

Further analysis of the reviews for online universities was performed with LIWC and included WPS or words per sentence, Sixltr or number of words of six or more letters, and Dic or Dictionary words to determine complexity and level of writing. Function words are non-content words. Function words act as connectors between meaningful content words. Pronouns such as (I, me, you) suggest the written words have a more personal comment. These measures in table 10 compared to table 11 show mixed levels of complexity. WPS and Dictionary words match about at Blog level but Sixltr words are about at NY Times levels. Function words are also at blog level and pronouns match up to Twitter, well above NY Times levels but below Blogs and expressive writing. Overall, the reviews are fairly complex, fairly personal, and contain relatively high levels of non-content words.

	Mean	Std. Deviation
Analytic	61.4908	24.20618
Tone	67.3723	32.18595
Clout	44.643	25.16279
Authentic	50.3574	29.64372
WPS	17.3064	7.37118
Sixltr	22.7386	7.18486
Dic	89.4139	7.84134
Function	52.8078	7.19261
Pronoun	13.6895	4.88103

Table 10 LIWC results of Reviews

Category	Blogs	Expressive Writing	Natural Speech	NY Times	Twitter
Analytic	49.9	44.9	18.4	92.6	61.9
Clout	47.9	37.0	56.3	68.2	63.0
Authentic	60.9	76.0	61.3	24.8	50.4
Tone	54.5	38.6	79.3	43.6	72.2
Words/sentence*	18.4	18.4	-	21.9	12.1

Words> 6 letters	14.4	13.6	10.4	23.6	15.3
Dictionary words	85.8	91.9	91.6	74.6	82.6
Total function words	53.1	58.3	56.9	42.4	46.1
Total pronouns	16.2	18.0	20.9	7.4	13.6

Table 11 LIWC standards (Pennebaker, 2015)

A final correlation analysis was performed to determine what linguistic aspects of the review correlated with polarity (positive or negative evaluation).

Category	Correlation	Sig. (2-tailed)
Polar	1	
WC	-.154**	0
Analytic	.220**	0
Clout	-0.004	0.899
Authentic	-0.054	0.05
Tone	.622**	0
Sixltr	.232**	0
Dic	-0.049	0.078
Function	-.202**	0
Pronoun	-.149**	0

Table 12 LIWC and Polarity Correlation and Significance

The results of this analysis are shown in table 12. Many linguistic measures have significant correlation with degree of polarity. Higher word count is negatively correlated which suggests that more words in the review, the more it tends to be negative. Conversely shorter reviews tend to be positive. The analytic measure correlates positively with polarity and suggests more thoughtful and detailed reviews tend to be positive. Negative reviews provide less analytical support. Tone has high and positive correlation with polarity as anticipated since positive tone would suggest positive comments. This analytic also supports our findings from the Meaning Cloud results. Six letter or longer words positively correlate with polarity with positive reviews having more complexity. Both function words and pronouns are negatively correlated with positive polarity suggesting again that more meaningful words are included in thoughtful positive reviews.

Clout and the number of dictionary words do not have significant correlation.

5. CONCLUSION

The results of the over 1290 reviews show that overall sentiment is generally positive. 74% of reviews are positive or very positive. Only 14% are negative or very negative clearly the results indicate that online universities have achieved an overall positive response from its students and other related stakeholders who have opted to complete a review. Since we are using student opinions as a proxy for quality, this suggests that overall quality of online universities is good. But not all online universities are viewed as good as others. The overall mean of our sample was 3.81 on a scale of 1 to 5. But there were statistically significant differences among the 340 universities at $p < .003$. The range was from negative to very positive. Not all schools are viewed positively. We recommend that students review ratings for the specific online university they are exploring to assure they are getting a school with positive ratings. For the schools themselves, they are recommended to review their ratings and address shortcomings that have been expressed.

Even though the reviews are generally positive, there is some level of disagreement found in the sentiment analysis. Different elements of a review have different levels of polarity. For schools this proposes that even though their review may be generally positive there still may be areas that can be improved in their online program.

The online reviews were overall found to be nonironic but also subjective. Subjectivity is expected since these are opinions students have of the program. Also, the nonironic tone portrays honest and straightforward opinions. We can thus rely on the sentiment classification.

Our review of the most common words used in reviews reinforces the findings that online universities are viewed favorably. The most six most frequent words used value laden words were positive.

Linguistic analysis also reveals strong positive tone and generally complex, authentic, and analytic levels especially in the positive reviews. This analysis serves as a check on the sentiment analysis. In all cases the findings of the linguistic analysis are consistent with the sentiment analysis.

Finally, though we believe these findings represent accurate opinions on online universities, the relationship between online ratings and actual product or service quality has not been fully studied. This is a limitation of the research. Further research needs to be performed to map online program outcomes with perceived quality opinions. This is an area that merits significant further research.

6. REFERENCES

- Allen, I. E., & Seaman, J. (2016). Online Report Card: Tracking Online Education in the United States. *Babson Survey Research Group*.
- Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management, 33*(1), 94-111.
- Braun, T. (2008). Making a choice: The perceptions and attitudes of online graduate students. *Journal of Technology and Teacher Education, 16*(1), 63.
- Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of management information systems, 23*(2), 149-171.
- Cordova, M. J., Cunningham, L. L., Carlson, C. R., & Andrykowski, M. A. (2001). Social constraints, cognitive processing, and adjustment to breast cancer. *Journal of consulting and clinical psychology, 69*(4), 706.
- Cui, G., Lui, H. K., & Guo, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce, 17*(1), 39-58.
- Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing, 21*(4), 23-45.
- Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter?—An empirical investigation of panel data. *Decision support systems, 45*(4), 1007-1016.
- Filieri, R., & McLeay, F. (2014). E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of

- information from online reviews. *Journal of Travel Research*, 53(1), 44-57.
- Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, 52(3), 674-684.
- King, Robert Allen, Pradeep Racherla, and Victoria D. Bush. "What we know and don't know about online word-of-mouth: A review and synthesis of the literature." *Journal of Interactive Marketing* 28.3 (2014): 167-183.
- Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education*, 80, 77-83.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International journal of educational research*, 11(3), 253-388.
- Meaning Cloud Sentiment Analysis (2018) <https://www.meaningcloud.com/products/sentiment-analysis>
- Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS quarterly*, 185-200.
- Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS quarterly*, 185-200.
- Pennebaker Conglomerates (2015) *LIWC How It Works*. <http://liwc.wpenigne.com/how-it-works/>
- Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015). *Linguistic Inquiry and Word Count: LIWC2015*. Austin, TX : Pennebaker Conglomerates (2015). www.LIWC.net
- Pennebaker, J., Boyd, R., Jordan, K., and Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX. University of Texas at Austin.
- Robinson, R. L., Navea, R., & Ickes, W. (2013). Predicting final course performance from students' written self-introductions: A LIWC analysis. *Journal of Language and Social Psychology*, 0261927X13476869.
- Seaman, I. Allen and J. Seaman (2016) Grade Increase Tracking Distance Education in The United States <http://onlinelearningsurvey.com/reports/gradeincrease.pdf>
- SR Education Group (2018). Mission. <https://www.sreducationgroup.org/#mission>
- Tallent-Runnels, Mary K., et al. "Teaching courses online: A review of the research." *Review of educational research* 76.1 (2006): 93-135.
- WICHE Cooperative for Educational Technologies(2016). WCET Distance Education Enrollment Report 2016 <https://wcet.wiche.edu/sites/default/files/WCETDistanceEducationEnrollmentReport2016.pdf>
- Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, 28(1), 180-182.