**WEB APPENDIX**

Automated Text Analysis for Consumer Research

ASHLEE HUMPHREYS

REBECCA JEN-HUI WANG

This supplement presents a sample text analysis of online word of mouth for a newly-introduced electronic device and uses this example to explore and test potentially interesting theories regarding expertise and word of mouth communication in product reviews. It is intended as an illustration of top-down, dictionary-based methods and bottom-up classification methods according to the stages discussed in “Automated Text Analysis for Consumer Research.”

Automated text analysis is appropriate for tracking systematic trends in language over time and making comparisons between groups of texts. To illustrate both a top-down and a bottom-up approach to text analysis, this appendix will present a short study of consumer response to the product launch of an mp3 player/wireless device, the Apple iTouch. We’ve selected this case because it can be used to illustrate both comparison between groups and change over time and because it’s relatively agnostic regarding theoretical framework. One could study word-of-mouth communication from a psychological, sociological, anthropological, or marketing strategy point of view (c.f. Godes and Mayzlin 2004; Kozinets et al 2010; Phelps et al. 2004; Winer 2009). We also map each step in this example to the six stages illustrated in the main paper.

***Stage 1: Research Question***

This study proposes a specific, strategic research question: After a product launch, do experts respond differently from non-experts? Further, how does word-of-mouth response change in expert versus non-expert groups as the product diffuses? Word of mouth from experts can be particularly influential in product adoption so it’s important to know how their views may change over time and in- comparison with non-expert groups. The context chosen for this study, the launch of the Apple iTouch, is a good case to study because both the product category and the criteria for evaluating the product were ambiguous at the time of launch.

***Stage 2: Data Collection***

*Data.* Data was collected from two websites, Amazon.com and CNET.com. Consumer comments from Amazon were used to reflect a non-expert or mixed consumer response, while user comments from CNET were used to measure expert response. Amazon is a website that sells everything from books to toys and has a broad audience. CNET, on the other hand, is a website dedicated exclusively to technology and is likely to have posters with greater expertise. Archival data also suggests that there are differences amongst visitors to the two sites.[[1]](#footnote-1)

Data was scraped from the internet, stored in a spreadsheet, and segmented by post. The comment date, poster name, rating, location of the poster, and the text of the comment itself were all stored as separate variables. Two levels of analysis were chosen. The most basic level of analysis is at the comment level. Each comment was coded for its content so that correlations between the content of that post and the date, poster experience and location could be assessed. The second level of analysis is the group-level, between Amazon and CNET. Comparisons can thus be made between expert and non-expert groups based on the assumption that Amazon posters are non-experts or a mix of experts and non-experts, while dedicated members of the CNET community have more expertise. Lastly, because the time variable exists in the dataset, it will also be possible to periodize the data. This may be relevant in assessing the effects of different product launches (e.g. 1st vs. 2nd generation iPods) on the textual content of posts. 204 posts were collected from Amazon and 269 posts were collected from CNET, yielding a sample size high enough to make statistical comparisons between groups.

After a file structure was created, the data was cleaned by running a spell-check on all entries. Slang words (e.g. “kinda”) were replaced with their proper counterparts. Text was scanned for problematic words. For example, “touch” appeared with greater frequency than usual because it was used to refer to the product, not to the sense. For that reason, “touch” was replaced with a non-codable character like “TTT” so that it would not be counted in the haptic category used in the standard dictionary.

***Stage 3: Construct Definition***

Work in information processing suggests that experts process information differently from novices (Alba and Hutchinson 1987). In general, experts view products more cognitively, evaluating product attributes over benefits or uses (Maheswaran and Sternthal 1990; Maheswaran, Sternthal, and Gurhan 1996; Sujan 1985). Whereas novices use only stereotypical information, experts use both attribute information and stereotypical cues (Maheswaran 1994). Experts are able to assimilate categorical ambiguity, which means one would expect for them to adjust to an ambiguous product more quickly than non-experts (Meyers-Levy and Tybout 1989). They also tend to approach judgment in an abstract, higher level construal than non-experts (Hong and Sternthal 2010).

From previous research, several working hypotheses can be developed. The strategic comparison we wish to make is about how experts vs. non-experts evaluate the product and whether or not this changes over time. First, one might expect that experts would use more cognitive language and that they would more critically evaluate the device.

H1: Experts will use more cognitive language than novices.

Secondly, one would also expect that experts would attend to features of the device, but non-experts would attend more to uses of the device (Maheswaran et al. 1996). Note that this is based on the necessary assumption that users discuss or verbally elaborate on what draws their mental attention, which is reasonable according to previous research (Carley 1997).

H2: Experts will discuss features more than non-experts.

H3: Non-experts will discuss benefits and uses more than experts.

Thirdly, over time, one might predict that experts would be able to assimilate ambiguous product attributes while non-experts would not. Because experts can more easily process ambiguous category information and because they have a higher construal level, one would predict that they would like this ambiguous product more than novices and would learn to assimilate the ambiguous information. For example, in this case the capacity of the device makes it hard to categorize (cell phone vs. mp3 player). One would expect that experts would more quickly understand this ambiguity and that over time their elaboration on this feature would decrease.

H4: Experts will talk about ambiguous attributes (e.g. capacity) less over time, while non-experts will continue to discuss ambiguous attributes.

Lastly, previous research suggests that these differences in focus, experts on features and non-experts on benefits, would differentially influence product ratings. That is, ratings for non-experts will depend on evaluation of benefits such as entertainment, but expert ratings would be influenced more by features.

H5: Ratings will be driven by benefits for non-experts.

H6: Ratings will be driven by features by experts.

These are only a few of the many potential hypotheses that could be explored in an analysis of online word of mouth communication. One could equally explore the cultural framing of new technologies (Giesler 2008) or the co-production of brand communications by seeding product reviews with bloggers (Kozinets et al 2010). The question posed here—do experts respond differently to new products than non-experts over time?—is meant to be illustrative of what can be done with automated text analysis rather than a rigorous test of the psychological properties of expertise.

In this illustrative example, the key constructs in examining H1 through H6 are known: expert and non-experts, cognitive expressions, affect, product features and benefits. We therefore proceed with a top-down approach. Operationalization for some of the constructs—cognitive and affective language—is available through a standardized measure (LIWC; Pennebaker et al 2009), and we can therefore use a standardized dictionary for their operationalization. However, some constructs such a features and benefits are context specific, and a custom dictionary will be necessary for operationalization. In addition, there may be other characteristics that distinguish experts from non-experts. We will therefore also perform a bottom-up approach of classification.

***Stage 4: Operationalization***

For this analysis, the standard LIWC dictionary developed by Pennebaker et al (2007) was used in addition to a custom dictionary. Table W1 presents the categories used from both the standardized and the custom dictionaries. The standard dictionary includes categories for personal pronouns such as “I”, parts of speech such as adjectives, psychometrically pre-tested categories such as positive and negative emotion, and content-related categories such as leisure, family, and friend related language.

------- Insert Table W1 about here -------

A custom dictionary was also developed to identify categories specific to the product word-of-mouth data analyzed here. Ten comments from each website were selected and open coded, with the researcher blind to the site from which they came. Then, ten more comments from each website were selected and codes were added until saturation was reached (Weber 2005). In all, the subsample required to develop the custom dictionary was 60 comments, 30 from each website, about 11% of all comments. 14 categories were created, each containing six words on average.

The qualitative analysis of comments revealed posters tended to talk about the product in terms of features or aesthetics. Dictionary categories were therefore created for words associated with features (e.g. GPS, camera, hard drive, battery) and for aesthetics (e.g. sharp, clean, sexy, sleek). Posters also had recurring concerns about the capacity of the device, the cost of the product and reported problems they experienced using the product. Categories were created for each of those concerns. Because there might be some researcher-driven interest in product uses and because posters frequently mentioned entertainment and work-related uses, categories were created for each type of use. Categories of ‘big’ vs. ‘small’ were included because previous theorization in sociology has suggested that the success of the iPod comes from its offerings of excess—large screen, excess capacity, etc. (Sennett 2006). Two categories were created to count when competitive products were mentioned, either within the Apple brand or outside of it.

The dictionary categories were validated by three coders who suggested words for inclusion and exclusion. Percent agreements between coders on each dictionary category can be found in table W1. Average agreement was 90%. Text files were run through the LIWC program, first using the standard dictionary, then using the custom dictionary. A spreadsheet was created from three sets of data: 1) the comment data collected directly from the website (e.g. date of post, rating of product), 2) the computer output from the standard dictionary, and 3) the output from the custom dictionary.

*Validation.* Once rough findings were gleaned, the coding was validated. Twenty instances from each category were pulled from the dataset and categorized. “Hits” and “false hits” were then calculated. This yielded an average hit rate of 85% and a “false hit” rate of 15%. The least accurate category was aesthetics, with a hit rate of 70% and a false hit rate of 30%. The most accurate category was “small,” which had a hit rate of 95% and a false hit rate of 5%.

***Stage 5: Interpretation and Analysis***

Overall, the findings indicate that there are systematic differences between the way experts and non-experts interpret the new device. As with most textual data, there are many potential variables and measures of interest. The standard LIWC dictionary contains 61 categories, and in the dataset studied here, 28 of these categories were significantly different amongst text from the three websites. We will report some of the most notable differences, including those needed to test the hypotheses.

*Comparison between groups.* First, we assessed differences amongst the two groups of comments. This was done by comparing differences in the percent of words coded in each category between groups using the Mann-Whitney test due to the skewed distribution of the data. Tables W2 and W3 show the differences by category. With the standard dictionary, several important differences between the word of mouth of non-experts and experts can be discerned.

------- Insert Tables W2 and W3 about here -------

First, experts use more cognitive words (Mcog|CNET=16.57, Mcog|Amazon=15.64, Mann-Whitney U = 30,562, z=2.12 p<.05) than non-experts, but they also use more affective (both positive and negative) language (Maffect|CNET=7.3 vs Maffect|Amazon=6.53, U = 30, 581, z = 2.14, p<.05) as well. The finding that experts evaluate the product cognitively is congruent with previous research (Maheswaran et al. 1996), and the highly affective tone indicates that they are likely more involved in product evaluation (Kelting and Duhacheck 2009). However, CNET posters use more negation (Mneg|CNET=2.47, Mneg|Amazon=1.74, U = 34,487, z = 4.81, p<.001). Together with the presence of cognitive language, this indicates that they may be doing more critical evaluation. The first hypothesis was therefore supported.

Secondly, non-experts focus on distal rather than proximate uses, while experts focus on device-related issues like features. Non-experts on Amazon use more distal social, time-, family-related language (e.g. Msocial|Amazon=5.55 vs Mscoial|NET=4.23, U = 22,259.5, z = -3.52, p< .001 and Mtime|Amazon=5.65, Mtime|CNET=3.89, U = 18,527 z = -6.01, p< .001). Experts on CNET, on the other hand, focus on features (Mfeatures|CNET=.61 vs. Mfeatures|Amazon=.41, U = 30,012.5, z = 2.10, p< .05) and capacity (Mconnect|CNET=1.08 vs. Mconnect|Amazon=.756, U = 35,819, z = 6.14, p< .001), but also on aesthetics (Maesth|CNET = .833 vs Maesth|Amazon=.168, U = 33,518, z = 5.02, p< .001). Experts discussed aesthetics about 8 times more than the mixed group on Amazon. These differences indicate that, in general, experts focus on the device itself while non-experts focus on uses. This lends convergent evidence to support to H2 and H3.

One other finding not specified by the hypotheses is notable. Non-experts use more past-oriented language (Mpast|Amazon=3.58 vs Mpast|CNET=2.13, U = 21, 289, z = ‒4.20, p< .001), while expert posters use more future-oriented language (Mfuture|CNET=1.01, Mfuture|Amazon=.76, U = 31,446, z = 2.83, p< .01). This suggests that experts might frame the innovation in the future while non-experts focus on the past. Recent research suggests experts and novices differ in temporal construal (Hong and Sternthal 2010). Experts focus on the far future while novices focus on the near future. The results here provide convergent evidence that supports previous research and suggests a further hypothesis—that novices focus on past-related information—for future experimental research.

In an extended analysis, adding a third group could help the researcher draw more rigorous conclusions through techniques of analytic induction (Mahoney 2003; Mill 1843). That is, if an alternative explanation is possible, the researcher could include a comparison set to rule out the alternative explanation. For example, one might propose that the difference in “cost” discourse is because Amazon.com users make less money than CNET users, on average, and are therefore more concerned about price. One could then include an expert website where the users are known to have a lower income than the posters on Amazon to address this explanation. If the same results are found, this would rule out the alternative hypothesis.

*Trends Over Time.* Because the product studied here is an innovation, the change of comments over time as the product diffuses is of interest. Time was analyzed first as a continuous variable in a correlation analysis and then as a discrete variable in ordinary least squares regression analyses, where the release of the 1st and 2nd generation of iTouch marked each period.

A correlation analysis was used to analyze time as a continuous variable (Table W4). We find that affect increases over time in the expert group, which indicates that group becomes more involved (r(affect, Date|CNET) =.144, p<.01). Experts become less concerned with capacity (r(capacity, Date|CNET) = -.203 p<.01) while Amazon users don't change in their concern for capacity. This indicates that experts learn something about the product category: the limited capacity was initially a shock to reviewers, as it was unorthodox for an mp3 player. But, over time, experts learned that this new category segment—mp3 wireless devices—did not offer as much memory. This supports Hypothesis 4.

------- Insert Figure W1 about here -------

Besides the correlation analysis, we also did ordinary least square linear regression analyses to analyze whether reviewers’ expressions changed over time (Table W5). We created a binary variable, which is set to “1” if the review is posted after the 2nd generation of iTouch is released, and “0” if the review is for the 1st generation of iTouch. To account for asymmetry in their distributions due to non-normality, we log-transformed the term frequency measurements of affect and capacity, our variables of interest. The results from the OLS analyses are congruent with the correlation analysis. We observe that in general expert reviewers discussed capacity more than non-experts (, *p* < 0.001). However, as predicted by Hypothesis 4, such discussions decreased after the release of the second-generation iPod (, *p* < 0.001).

------- Insert Table W5 about here -------

Affect also changes differentially in each group (Figure W2). The OLS analysis (Table W5) shows that in the first time-period, affective language is roughly equivalent, but experts on CNET use more affective language in the second time-period than they do in the first time-period (, *p* < 0.05). In short, site and period have a positive interactive effect on affective expressions. These are just two examples of how automated content analysis can be used to assess changes in word of mouth communication.

------- Insert Figure 2 about here -------

*Regression with ratings*. Now that relationships between semantic elements in the text have been discerned, their relationship to other, non-semantic variables is of interest. For example, what factors impact ratings for experts vs. non-experts? To test the impact of discourse on rating, an OLS regression was run with rating as the dependent variable and the discursive categories as the independent variables. Several discursive variables were significant predictors of ratings overall (FAmazon = 2.55, p<.05; FCNET = 2.30, p<.05). Results are shown in table W6. These reveal that the ratings of non-experts were influenced by entertainment and features, while the ratings of experts were affected by connectablity and by the (negative) evaluation of the features. This provides support for H5 and H6. However, they also indicate a more complicated relationship. Features are correlated with both expert and non-expert ratings. However, for non-experts, features are positively correlated with ratings while for experts, they are negatively correlated. Problems and cost, although much discussed in the posts, appeared to have little effect on ratings. The unimportance of cost may be explained by the fact that the ratings data is non-behavioral, that is, most posters had already purchased the device.

------- Insert Table W6 about here -------

*Classification*. We also took a classification approach to discover which words were unique to experts versus non-experts. We first processed all of the Amazon and CNET reviews, tokenizing them into individual words, which were then stemmed and lemmatized so that the words with the same semantic meaning but different morphologies could be considered to be the same word, e.g., “applications” and “application.” For each unique stemmed and lemmatized word, we then counted the word occurrences in each review. For words that occur in both Amazon and CNET reviews, the top five content words that non-experts use more than experts are *email*, *download*, *purchase*, *day*, and *app*. Experts, on the other hand, use the words *video*, *screen*, *touch*, *application*, and *ipod* more than experts, and they are more likely to include external websites that refer readers to other reviews or blogposts than non-experts. Non-experts use words associated with uses such as “email” and “download” while experts use feature based words like “video,” “screen,” and “application.” These descriptive findings add convergent support to the claim that non-experts focus on benefits while experts focus on features (H2 and H3).

After processing the review files and storing them into a database management system, we created a term-frequency matrix to include the counts of first-person words (*I*, *I'm*, *we*, *we're*, *me*, *us*, *my*, *our*, *mine*, *ours*, *myself*, *ourselves*), second-person words (*you*, *you're*, *yours*, *yourself*, *yourselves*), third-person words (*he*, *she*, *it*, *he's*, *she's*, *it's*, *they*, *him*, *her*, *them*, *his*, *her*, *its*, *their*, *hers*, *theirs*, *himself*, *herself*, *itself*, *themselves*), demonstratives (*this*, *that*, *those*, *these*), and numbers. We then ran a classification analysis using binary logistic regression with these count-based variables, and the estimation results can be found in Table W7. Other alternative models including individual words with the most occurrences are also estimated, but word-based variables do not indicate statistical significance, while the pronoun, demonstrative, and number variables remain significant. We find that experts are more likely to use numbers than non-experts (, *p* < 0.01). They are also more likely to use third-person pronouns (, *p* < 0.0001) and demonstratives (, *p* < 0.05). However, they are less likely to express with first-person pronouns (, *p* < 0.0001) and second-person pronouns (, *p* < 0.01). The concordance statistic is 0.879 and Nagelkerke R-squared is 0.525. Taken together, these patterns may indicate that experts evaluate the product objectively as a third party rather than personally, as do non-experts.

------- Insert Table W7 about here -------

We also ran a classification analysis using a tree regression, which yielded the classification tree as shown in Figure W3. The tree supports the findings derived from the logistic regression, suggesting that experts tend to use fewer first- and second-person pronouns but more numbers and demonstratives and supporting the claim that experts take a cognitive approach and focus on the device rather than product usage.

***Stage 6: Validation***

The previous analyses revelaed there were systematic differences in the amount of word use between experts and non-experts. To assess construct validity, we used a triangulation approach to explore the relationships between the concepts through a correlation analysis of word association within comment (table W4). This means that we are looking for how the dictionary categories occur together within one post. To assess construct validity of affect, we included another operationalization of affect, star rating, in the correlational analysis. We calculated Pearson correlations for all categories in the set and compared them with cosine similarities. Both tables produced directionally similar results, and here we report Pearson correlations, as it accounts for both presence and absence of collocation. First, a few expected correlations between categories were checked. For both sites, positive emotion is correlated with rating (r(posem, rating) =.335, p<.01), as one would expect. Negative emotion is negatively correlated with positive emotion (r(negemo,posemo) =‒.348, p<.01). More can be learned, however, by comparing word association in expert versus non-expert groups.

------- Insert Table W4 about here -------

In general, non-experts use positive language alongside distal uses for the iPod such as work and family (r(work,posem|Amazon)=.243, p<.01 and r(family, posemo|Amazon)=.190, p<.01). For the non-experts, negative emotion is correlated with problems, as one would expect (r(problems,negem|Amazon) =.470). For experts, positive emotion occurs alongside aesthetics (r(aesth,posem|CNET)=.409, p<.01). For experts, there is also a positive correlation between Apple and love (r(Apple, love|CNET) =.203, p<.01) that doesn't exist for non-experts. These correlations indicate that aesthetics are viewed positively by experts and that they are involved with not only the device but the brand as well. Cosine similarities produce directionally similar results.

Secondly, features are interpreted differentially between the two groups. Novices interpret some features using standards of other categories (like an mp3 player), while experts are more willing to judge them relative to the standards for a new category. For example, from the correlation between small and capacity among the non-expert group (r(capacity,small|Amazon) = .144, p<.01), one can conclude that posters feel the capacity is too small. No such correlation exists for experts. This could be because the iTouch is a product without a known category. Experts can interpret size for this ambiguous product, but novices are uncertain about what capacity is appropriate for the device. These are just a few of the findings that can be gleaned using a correlation table. A full spatial analysis might compare the network of meanings in the Amazon group to the network of meanings in the CNET group.

For the binary logistic classification, *k-*fold cross validation was performed, and per convention, we set *k* = 10. The resulting comparisons between predicted values based on our model and the real values show that overall, the model is 80.13% accurate (95% accuracy confidence interval = [0.7624, 0.8363]). Table W8 shows the confusion matrix.

***Summary and Conclusion***

In sum, the automated text analysis presented here shows that that experts evaluate new products in a systematically different way from non-experts. Using comparison between groups, we show that experts evaluate products by focusing on features while non-experts focus on the uses and benefits of the devices. Using correlation analysis, we find that experts associate aesthetics with positive emotion while non-experts associate positive emotion with uses of the device and negative emotion with problems. Further, the correlation analysis provides some validation for the method of automated content analysis by demonstrating the correlation between positive emotion and ratings, a variable used in previous studies of online word of mouth communication (Godes and Mayzlin 2004, 2009). We find that, over time, experts focus less on problematic features like capacity and speak more affectively about the product. A regression analysis of the elements of discourse on ratings demonstrates that ratings for experts are driven by features, while ratings by non-experts are better predicted by both features and the amount of talk about entertainment, a benefit. Note that, like field research, these findings make sense in convergence with previous findings from experimental data and provide ecological validity to previous findings obtained in laboratory settings. These are not meant to be a rigorous test of expertise, but rather an illustration of the way in which text analysis can provide convergent evidence that is meaningful to consumer researchers.

Data Collection

Data was collected by the first author with the help of a research assistant from Amazon.com and CNET.com from September 5, 2007 to November 62009. Keyword search for “iPod Touch” was used to gather all customer reviews available for the product at the time of analysis. Reviews for multiple versions of the device (1st and 2nd generation) were included and segmented in the analysis according to release date. The iPod Touch 1st generation was released on September 5, 2007, and the 2nd generation was released on September 9, 2008.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **abbv** | **Words** | **No. of Words** | **Alpha\*** |
| social processes | social | Mate, talk, they, child | 455 | 97% |
| affective processes | affect | Happy, cried, abandon | 915 | 97% |
| positive emotion | posemo | Love, nice, sweet | 406 | 97% |
| negative emotion | negemo | Hurt, ugly, nasty | 499 | 97% |
| cognitive processes | cogmech | cause, know, ought | 730 | 97% |
| past tense | past | Went, ran, had | 145 | 94% |
| present tense | present | Is, does, hear | 169 | 91% |
| future tense | future | Will, gonna | 48 | 75% |
| discrepancy | discrep | should, would, could | 76 | 80% |
| exclusive | excl | But, without, exclude | 17 | 67% |
| perceptual processes | percept | Observing, heard, feeling | 273 | 96% |
| relativity | relativ | Area, bend, exit, stop | 638 | 98% |
| space | space | Down, in, thin | 220 | 96% |
| time | time | End, until, season | 239 | 94% |
| work | work | Job, majors, xerox | 327 | 91% |
| aesthetics | aesth | sleek, cool, shiny, perfect | 9 | 83% |
| capacity | cap | capacity, space, storage | 7 | 93% |
| cost | cost | price, cost, dollars | 6 | 100% |
| big | big | large, huge, full | 5 | 83% |
| problems | prob | bugs, crash, freeze | 7 | 100% |
| competitors | comp | Zune, Microsoft, Archos | 4 | 67% |
| Apple | Apple | nano, iPod, iPhone | 4 | 100% |
| entertainment | ent | music, video, fun | 9 | 85% |
| job | job | work, commute, conference | 9 | 100% |
| connectability | connect | wifi, internet, web | 9 | 95% |
| features | feat | GPS, camera, battery | 5 | 87% |
| love | love | amazing, best, love | 7 | 100% |
| small | small | empty, small, tiny | 4 | 100% |
| expertise | expert | jailbreak, jailbroke, keynote | 4 | 67% |
| \*Alpha is the percent agreement of three coders on dictionary words in the category. | | | | |
|  | |
|  | | | | |

Table W1: Standard and Custom Dictionaries

|  |  |  |
| --- | --- | --- |
|  | **Amazon** | **CNET** |
| WC | 160.99 | 149.11 |
| social\*\*\* | 5.55 | 4.23 |
| affect† | 6.53 | 7.20 |
| posemo | 5.50 | 5.94 |
| negemo | 1.10 | 1.31 |
| cogmech\* | 15.64 | 16.57 |
| past\*\*\* | 3.58 | 2.13 |
| present | 8.91 | 9.22 |
| future\* | 0.76 | 1.01 |
| certain | 1.66 | 1.87 |
| excl\*\* | 2.68 | 3.20 |
| percept\*\*\* | 3.34 | 4.86 |
| relativ\*\*\* | 11.26 | 9.53 |
| space\* | 4.06 | 4.64 |
| time\*\*\* | 5.65 | 3.89 |
| work | 2.08 | 1.92 |
| achieve | 2.24 | 2.58 |
| leisure† | 3.28 | 3.80 |

\*\*\*p<.001

\*\* p<.01

\* p<.05

† p<.10

Table W2: Amazon vs. CNET Differences in Means, Standard Dictionary

|  |  |  |
| --- | --- | --- |
|  | **Amazon** | **CNET** |
| aesthetics\*\*\* | 0.168 | 0.833 |
| capacity\*\*\* | 0.538 | 1.408 |
| cost\* | 0.384 | 0.641 |
| big\*\* | 0.070 | 0.178 |
| problems† | 0.286 | 0.165 |
| competitors | 0.080 | 0.104 |
| Apple\* | 1.461 | 1.927 |
| entertainment\*\* | 1.377 | 1.838 |
| job† | 0.164 | 0.087 |
| connect\* | 0.756 | 1.075 |
| features† | 0.413 | 0.606 |
| love\*\*\* | 0.746 | 1.470 |
| small\* | 0.054 | 0.135 |
| expert\* | 0.009 | 0.028 |
|  |  |  |
| \*\*\*p<.001 |  |  |
| \*\* p<.01 |  |  |
| \* p<.05 |  |  |
| † p<.10 |  |  |
|  |  |  |

Table W3: Differences in Means, Custom Dictionary

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | | | | | | | | | | |
| Statistics=Pearson Correlation | | | | | | | | | | | | | | |
|  | site | Rating | Date | affect | posemo | negemo | aesth | capacity | ent | connect | feat | love | big | small |
| Rating | Amazon | 1 | .009 | .282\*\* | .387\*\* | -.200\*\* | .061 | .064 | .216\*\* | .002 | .128 | .273\*\* | .015 | -.024 |
| CNET | 1 | -.012 | .095 | .319\*\* | -.433\*\* | .024 | -.058 | .044 | .145\* | -.118 | .373\*\* | .091 | -.053 |
| Date | Amazon | .009 | 1 | -.087 | -.046 | -.118 | -.082 | .013 | .073 | .008 | -.040 | .022 | -.156\* | -.095 |
| CNET | -.012 | 1 | .144\* | .145\* | .011 | -.009 | -.203\*\* | .114 | .127\* | -.102 | -.006 | -.106 | -.001 |
| affect | Amazon | .282\*\* | -.087 | 1 | .910\*\* | .350\*\* | -.049 | -.098 | -.043 | -.187\*\* | .049 | .450\*\* | -.001 | -.036 |
| CNET | .095 | .144\* | 1 | .865\*\* | .263\*\* | .367\*\* | -.036 | .111 | .036 | .108 | .411\*\* | -.096 | .034 |
| posemo | Amazon | .387\*\* | -.046 | .910\*\* | 1 | -.056 | .005 | -.052 | .032 | -.164\* | .064 | .473\*\* | .006 | -.015 |
| CNET | .319\*\* | .145\* | .865\*\* | 1 | -.253\*\* | .409\*\* | -.019 | .156\* | .106 | .104 | .514\*\* | -.038 | -.056 |
| negemo | Amazon | -.200\*\* | -.118 | .350\*\* | -.056 | 1 | -.117 | -.140\* | -.194\*\* | -.104 | -.030 | -.013 | .026 | -.050 |
| CNET | -.433\*\* | .011 | .263\*\* | -.253\*\* | 1 | -.086 | -.026 | -.087 | -.139\* | .000 | -.205\*\* | -.119 | .167\*\* |
| aesth | Amazon | .061 | -.082 | -.049 | .005 | -.117 | 1 | .131 | -.019 | .016 | .005 | -.055 | .126 | .003 |
| CNET | .024 | -.009 | .367\*\* | .409\*\* | -.086 | 1 | -.025 | .040 | -.052 | .291\*\* | .015 | -.072 | -.053 |
| capacity | Amazon | .064 | .013 | -.098 | -.052 | -.140\* | .131 | 1 | .055 | .052 | -.044 | -.010 | -.046 | .144\* |
| CNET | -.058 | -.203\*\* | -.036 | -.019 | -.026 | -.025 | 1 | .079 | -.177\*\* | -.079 | -.048 | -.025 | .020 |
| ent | Amazon | .216\*\* | .073 | -.043 | .032 | -.194\*\* | -.019 | .055 | 1 | .139\* | -.022 | -.061 | .069 | .063 |
| CNET | .044 | .114 | .111 | .156\* | -.087 | .040 | .079 | 1 | .023 | -.141\* | .072 | .055 | -.012 |
| connect | Amazon | .002 | .008 | -.187\*\* | -.164\* | -.104 | .016 | .052 | .139\* | 1 | .007 | -.055 | -.077 | -.009 |
| CNET | .145\* | .127\* | .036 | .106 | -.139\* | -.052 | -.177\*\* | .023 | 1 | .008 | .139\* | .038 | -.056 |
| feat | Amazon | .128 | -.040 | .049 | .064 | -.030 | .005 | -.044 | -.022 | .007 | 1 | .000 | -.019 | -.024 |
| CNET | -.118 | -.102 | .108 | .104 | .000 | .291\*\* | -.079 | -.141\* | .008 | 1 | -.086 | -.045 | -.096 |
| love | Amazon | .273\*\* | .022 | .450\*\* | .473\*\* | -.013 | -.055 | -.010 | -.061 | -.055 | .000 | 1 | -.016 | -.048 |
| CNET | .373\*\* | -.006 | .411\*\* | .514\*\* | -.205\*\* | .015 | -.048 | .072 | .139\* | -.086 | 1 | .078 | .044 |
| big | Amazon | .015 | -.156\* | -.001 | .006 | .026 | .126 | -.046 | .069 | -.077 | -.019 | -.016 | 1 | .055 |
| CNET | .091 | -.106 | -.096 | -.038 | -.119 | -.072 | -.025 | .055 | .038 | -.045 | .078 | 1 | .059 |
| small | Amazon | -.024 | -.095 | -.036 | -.015 | -.050 | .003 | .144\* | .063 | -.009 | -.024 | -.048 | .055 | 1 |
| CNET | -.053 | -.001 | .034 | -.056 | .167\*\* | -.053 | .020 | -.012 | -.056 | -.096 | .044 | .059 | 1 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | | | | | | | | | | | |
| \*. Correlation is significant at the 0.05 level (2-tailed). | | | | | | | | | | | | | | |

Table W4: Correlation Table, Amazon vs. CNET

|  |  |  |  |
| --- | --- | --- | --- |
| Dependent Variable |  |  | |
| B | Std. Error |
| ln(Capacity) | (Intercept)\*\*\* | 0.275 | 0.058 |
| Is 2nd Gen | 0.024 | 0.081 |
| Is CNET\*\*\* | 0.407 | 0.069 |
| Is 2nd Gen × CNET\*\*\* | ‒0.546 | 0.158 |
| ln(Affect) | (Intercept)\*\*\* | 1.916 | 0.048 |
| Is 2nd Gen | ‒0.043 | 0.068 |
| Is CNET | 0.063 | 0.057 |
| Is 2nd Gen × CNET\* | 0.275 | 0.132 |

Table W5: OLS Regression Coefficient Estimates. Affect and Capacity by Time and Amazon vs. CNET.

+ p <.10, \* p < .05, \*\* p < .01, \*\*\* p < .001.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficients** | | | | | | |
| site | category | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| Amazon | (Constant) | 3.839 | .137 |  | 27.932 | .000 |
| aesthetics | .145 | .175 | .058 | .833 | .406 |
| capacity | .064 | .087 | .051 | .732 | .465 |
| problems | -.015 | .086 | -.012 | -.174 | .862 |
| **entertainment** | **.150** | **.047** | **.221** | **3.178** | **.002** |
| connect | -.035 | .073 | -.033 | -.476 | .635 |
| **features** | **.174** | **.088** | **.136** | **1.972** | **.050** |
| CNET | (Constant) | 3.799 | .144 |  | 26.373 | .000 |
| aesthetics | .031 | .031 | .062 | .978 | .329 |
| capacity | -.029 | .042 | -.043 | -.697 | .486 |
| problems | -.290 | .195 | -.091 | -1.484 | .139 |
| entertainment | .011 | .040 | .017 | .277 | .782 |
| **connect** | **.100** | **.049** | **.128** | **2.062** | **.040** |
| **features** | **-.126** | **.059** | **-.137** | **-2.138** | **.033** |

Table W6:

Regression Coefficients: Predictors of Product Rating for Experts vs. Non-Experts

|  |  |  |  |
| --- | --- | --- | --- |
| Dependent Variable  Is Expert | |  | |
| B | Std. Error |
|  | (Intercept)\*\* | 0.550 | .173 |
| Numbers\*\* | 0.224 | .072 |
| First-Person Pronouns\*\*\* | ‒0.607 | .066 |
| Second-Person Pronouns\*\* | ‒3.499 | 1.254 |
| Third-Person Pronouns\*\*\* | 0.282 | .056 |
| Demonstratives\* | 0.166 | .074 |

Table W7: Classification with Logistic Regression, Coefficient Estimates.

Likelihood-ratio test: *p* < 0.0001; C-statistic = 0.879 ; Pseudo Nagelkerke R-squared = 0.525.

+ p <.10, \* p < .05, \*\* p < .01, \*\*\* p < .001.

|  |  |  |  |
| --- | --- | --- | --- |
| Prediction | |  | |
| Expert | Not Expert |
|  | Expert | 237 | 62 |
| Not Expert | 32 | 142 |

Table W8: Confusion Matrix from 10-fold Cross Validation. Accuracy = 0.8013. *p*-Value [Accuracy > No Information Rate] = < 2e-16.



Figure W1: Mean number of Capacity Words by Site and Time Period



Figure W2: Mean Number of Affect Words by Site and Time Period

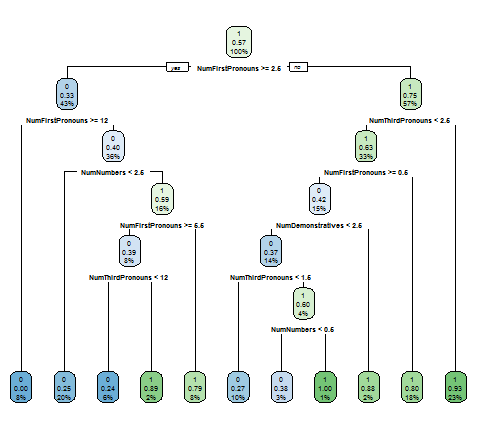


Figure W3: Expert Classification Tree. The Classification Tree (CART) analysis shows 10 nodes, each of which displays the sub-classification (0/1 as it is a binary split at each node), the probability of the sub-class, and the percentage of dataset used for the node. The branch goes to the left if the node condition is satisfied, and to the right if it is not.

References

Alba, Joseph W. and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (4), 411-54.

Carley, Kathleen (1997), "Network Text Analysis: The Network Position of Concepts," in *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*, ed. Carl W Roberts, Mahwah, NJ: Lawrence Erlbaum.

Giesler, Markus (2008), "Conflict and Compromise: Drama in Marketplace Evolution," *Journal of Consumer Research*, 34 (April) (6), 739-53.

Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545-60.

--- (2009), "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science*, 28 (4), 721-39.

Hong, Jiewen and Brian Sternthal (2010), "The Effects of Consumer Prior Knowledge and Processing Strategies on Judgments," *Journal of Marketing Research (JMR)*, 47 (2), 301-11.

Kozinets, Robert V., Kristine De Valck, Andrea C. Wojnicki, and Sarah JS Wilner, (2010), "Networked narratives: Understanding word-of-mouth marketing in online communities," *Journal of Marketing* 74, no. 2 (2010): 71-89.

Maheswaran, Durairaj (1994), "Country of Origin as a Stereotype: Effects of Consumer Expertise and Attribute Strength on Product Evaluations," *Journal of Consumer Research*, 21 (2), 354-65.

Maheswaran, Durairaj and Brian Sternthal (1990), "The Effects of Knowledge, Motivation, and Type of Message on Ad Processing and Product Judgments," *Journal of Consumer Research*, 17 (1), 66-73.

Maheswaran, Durairaj, Brian Sternthal, and Zeynep Gurhan (1996), "Acquisition and Impact of Consumer Expertise," *Journal of Consumer Psychology (Lawrence Erlbaum Associates)*, 5 (2), 115.

Mahoney, James (2003), "Strategies of Causal Assessment in Comparative Historical Analysis " in *Comparative Historical Analysis in the Social Sciences*, ed. James Mahoney and Dietrich Rueschemeyer, Cambridge, UK ; New York: Cambridge University Press, xix, 444.

Meyers-Levy, Joan and Alice M. Tybout (1989), "Schema Congruity as a Basis for Product Evaluation," *Journal of Consumer Research*, 16 (1), 39-54.

Mill, John Stuart (1843), *A System of Logic, Ratiocinative and Inductive : Being a Connected View of the Principles of Evidence, and Methods of Scientific Investigation*, London: J.W. Parker.

Phelps, Joseph E., Regina Lewis, Lynne Mobilio, David Perry, and Niranjan Raman (2004), "Viral Marketing or Electronic Word-of-Mouth Advertising: Examining Consumer Responses and Motivations to Pass Along Email," *Journal of Advertising Research*, 44 (4), 333-48.

Quantcast (2010a), ""Amazon Monthly Traffic (Estimated)"," [www.quantcast.com/amazon.com](http://www.quantcast.com/amazon.com).

--- (2010b), ""Cnet Monthly Traffic (Estimated)"," [www.quantcast.com/cnet.com](http://www.quantcast.com/cnet.com).

Sennett, Richard (2006), *The Culture of the New Capitalism*, New Haven: Yale University Press.

Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments," *Journal of Consumer Research*, 12 (1), 31-46.

Winer, R. S. (2009), "New Communications Approaches in Marketing: Issues and Research Directions," *Journal of Interactive Marketing*, 23 (2), 108-17.

1. According to Quantcast estimates Quantcast (2010b), ""Cnet Monthly Traffic (Estimated)"," www.quantcast.com/cnet.com., users to CNET.com are predominantly male and likely to visit websites like majorgeeks.com and read PC World. Amazon users, on the other hand, represent a broader demographic. They are more evenly divided between men and women (48/52), are more likely to have kids, and, visit websites like buy.com --- (2010a), ""Amazon Monthly Traffic (Estimated)"," www.quantcast.com/amazon.com. Data was collected in November of 2009. [↑](#footnote-ref-1)