Introduction

As our world becomes increasingly digital, it is no surprise that the internet has become an increasingly popular place for consumers to share their reviews about products. Over the past decade, numerous online forums for eliciting and aggregating customer evaluations have sprouted – Yelp.com and Epinions.com, for example, are entirely dedicated to organising user reviews and recommendations of restaurants, doctors, cars, and a wide range of other products and services. These types of website contain reviews of virtually any product or service that is available in the marketplace.

At the same time, online stores have become a significant part of today’s commerce. Perhaps one of the best examples of how online stores have become integral to how people shop is that of online application (app) stores; these app stores have become a huge venue for the sale of software applications for devices such as smartphones and tablets, as well as for more traditional computers. As a result, app stores have become a huge source of business not only for traditional computer firms (such as Apple), but also for new start-up firms (such as Zynga, the firm that created Angry Birds). For example, Apple collects 30% commission on all paid apps downloaded from its app store; to date, over one billion apps have been downloaded from Apple’s app store. In 2010, Apple’s commission on paid apps generated $1.8 billion in revenues (Kent 2010); in 2011 this figure had jumped to $5.0 billion, according to Apple, Inc. (2012). Clearly, there is an enormous market for online app stores; further, online reviews available at app stores are very important sources of information for both buyers and sellers.

This paper is exploratory in nature, and serves to fill a gap in the existing literature about the ways in which researchers and marketers can explore online customer reviews. To accomplish this, reviewer input on the top four iPhone games available in Apple’s app store as of March 2011 was evaluated using a content analysis package known as Leximancer. Responses from users who rated games as five-star were compared to responses from users who rated them as one-star.
Literature review

When faced with the multitude of products and services available for sale, consumers seek tools to help them decide whether to make a purchase or not. One of the most common tools is the consumer recommendation: consumer reviews, suggestions or references can be the driving force behind decisions ranging from which hotel to stay in to which smartphone application to download (Fagerstrom & Ghinea 2011).

In general, consumer recommendations come in the form of word-of-mouth (WOM), or what other consumers say about a product or service. Such recommendations are known to be particularly influential in purchasing decisions as they are generally perceived as more trustworthy than recommendations from firms or advertisers (Arndt 1967). As a result, WOM can have an enormous impact on purchasing decisions: consumer WOM has been touted as the single most powerful factor in predicting the long-term success of experience goods (DeVany & Walls 1996). Recently, Twitter comments have been cited as some of the most influential determinants of brand images (Macale 2011). Indeed, WOM can make or break a consumer’s decision to buy.

Hennig-Thurau et al. (2004, p. 39), cited in Bronner and de Hoog (2010), present the following definition of eWOM: ‘any positive or negative statement made by potential, actual or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet’. The authors suggest that this definition might require some additional consideration and offer the following as an enhanced definition: ‘any statement – positive, negative or neutral – made by potential, current or former stakeholders about a product, service, company or person, which is made available to a multitude of people, organisations or institutions, via a digitally networked platform’. Such a definition incorporates current and future developmental communication activity more comprehensively.

Empirically, the emergence of online reviews is a significant step in WOM research. Dellarocas et al. (2007) note how the practice of reviewing products online has greatly increased potential for empirical understanding of WOM marketing: while articulated reviews vanish shortly after they are spoken, and are therefore tremendously difficult to capture and analyse, online reviews persist long after they are posted. Breazeale (2009) states that ‘digital platforms are changing our very understanding and nature of the significance of eWOM’, and sums it nicely thus: ‘WOM no longer vanishes instantly and it is not necessarily spontaneous. There may also be a reduced perception that the communicator is in fact non-commercial or even an actual consumer at all.’ As a result, online reviews leave an easily accessible and long-lasting record of opinions; moreover, anyone is free to tap into these records, although care should be taken to validate the source of the eWOM. Further, Zhu and Zhang (2010) indicate that online reviews are a strong reflection of overall WOM, and can therefore be used as a proxy for traditional WOM.

The correlation between online ratings and online purchases has been well established. Senecal and Nantel (2004) revealed that subjects who sought product recommendations tended to follow these recommendations, and found that consumers who consulted online reviews selected recommended products twice as often as subjects who did not consult product recommendations. These results are consistent with numerous other studies in which online ratios have been significantly correlated with online purchases (Chevalier & Mayzlin 2006; Dellarocas et al. 2007; Tsang & Prendergast 2009). Breazeale (2009), in this journal, argued that WOM had changed the very ‘ways that marketers perceive and manage this important component of the marketing mix’. However, Breazeale (2009) concedes that ‘The possibility exists that, just as consumers ignore some traditional WOM, they may mentally compensate for the possibility that a firm is manipulating its own eWOM.’ Providing a healthy evaluation and a sceptical view is a fundamental condition in human assessment and communication, and
hence enables a considered decision.

Nearly a quarter of all internet users consult online reviews prior to purchasing a product or service (Zhu & Zhang 2010), and these users take reviews into account when making their purchasing decisions (Senecal & Nantel 2004; Chevalier & Mayzlin 2006; Dellarocas et al. 2007). Overall, the body of literature on online consumer reviews indicates that consumers are willing to seek online reviews, and to accept them as a valuable and credible source of information about product strengths and weaknesses. Leskovec et al. (2007), cited in Bronner and de Hoog (2010), conclude in their study that ‘eWOM is of great value in fortifying advertisements’ and this maybe highly applicable to a games-based market where players' attitudes can be switched from ‘bad’ to ‘good’, and vice versa.

The utility of online reviews extends beyond assisting consumers in their decision making: online product reviews also present an incredibly valuable source of information for marketers who wish to understand how consumers react to their products. Marketers can monitor online reviews in real time, and can quickly learn about gaps in product distribution or performance. As a result, online reviews create opportunities for marketers to engage in corrective measures in a timely fashion (Tsang & Prendergast 2009). Chen and Xie (2008) go so far as to suggest that online reviews are sufficiently powerful in helping consumers find products that could reasonably be incorporated into the marketing communications mix as ‘sales assistants’.

Ultimately, the body of research on consumer reviews reveals that the relationship between online reviews and online purchases is one that cannot be ignored by researchers, consumers or marketers. Gruen et al. (2006, p. 449), cited in Bronner and de Hoog (2010), write: ‘Similar to WOM, research has shown that eWOM may have higher credibility, empathy and relevance to customers than marketer-created sources of information on the Web.’ Hardey (2011) suggests that consumer-generated content (CGC) is still a difficult area to assess, and comments that ‘the range and scope of related research still feels foggy’. Paraphrasing from Hardey (2011) suggests that the tracking of consumer interest and deciding which channels of consumer content to measure, and thus demonstrating its relation to ROI, is complex and difficult at best, even given the post-Facebook era.

**Interpreting online reviews: challenges**

Online reviews present an undeniable wealth of information for consumers and marketers alike: consumers can use online reviews as a tool to assist decision making, while marketers can use them as a source of valuable feedback. However, extracting information from customer reviews is not without its challenges – in fact, both marketers and consumers face difficulty in extracting meaning from the massive amount of information available in online reviews. For example, online reviews can be based on ratings (i.e. a five-star rating), on a rating plus a comment, or on comments alone. Furthermore, even a five-star rating is not clearly defined (for instance, how do consumers define the threshold between stars?). Finally, a user may give a product a strong star rating but still comment on negative aspects of the product (or vice versa). In general, online reviews are unstructured and responses are unsystematic.

Despite the challenges that interpreting user reviews presents, written feedback from customers is one of the most important sources of information for consumers and marketers alike. Tsang and Prendergast (2009) revealed that consumer comments had a stronger impact in affecting purchase decisions and perceived trustworthiness than did star ratings. Similarly, Chevalier and Mayzlin (2006) indicated that, although consumers reviewed star ratings in their decision making, consumers read and applied information provided in written reviews.

Interpretation of these reviews is both difficult and time consuming, and these challenges become even more daunting when
one considers the sheer volume of online consumer reviews: many products receive thousands of ratings and reviews. Indeed, the biggest challenges that consumers and marketers face when using online reviews is in synthesising the information into a useful message.

**Method**

Gebauer *et al.* (2008) note that interpreting customer reviews is typically based on automated or non-automated content analysis. Non-automated content analysis occurs when humans classify text according to a preset classification system, whereas a computer typically performs automated content analysis. Each method has benefits and limitations, and the decision as to which form of content analysis to employ depends on the goals of a particular project. Specifically, automated content analysis can lead to a greater reliability of the findings, as well as a greater ability to handle large volumes of text. However, as Gebauer *et al.* (2008) point out, automated content analysis is limited in its ability to reveal the communicative intent of word usage or symbolic meanings of words. Due to the enormous volume of text analysed in this paper, an automated approach to content analysis was deemed appropriate.

**Materials**

Leximancer, a data-mining software package, was chosen above others because it enables the researcher to navigate the complexity of text more flexibly and is more analytically comprehensive than its competitors. Kotsiantis (2007) states that ‘the automated categorisation of texts into predefined categories has witnessed a booming interest in the last ten years, due to the increased availability of documents in digital form and the ensuing need to organise them. In the research community the dominant approach to this problem [of organisation] is based on machine learning techniques’; according to Fabrizio (2002), ‘the most significant of which is data mining’.

In addition Leximancer is ideally suited to a natural language processing method. Warschauer and Healey (1998) define natural language processing as ‘the process of a computer extracting meaningful information from natural language input and/or producing natural language output’. Leximancer undertakes data-mining approaches but, coupled with natural language processing through concept development, achieves insight and discovery more readily than ‘categorisation’ learning approaches alone – it (Leximancer) enables the development of meaning.

It can easily handle large volumes of unstructured data text, it can identify ‘concepts’ within the text rather than merely keywords (which other, simpler analysis tools offer), and provides focus on developmental ‘discovery’ and not merely ‘data exploration’ (like other analysis tools). The ‘discovery’ element is crucial to ‘true meaning’; moreover, the utilisation of interconnectedness and co-occurrence within the software radically enhances the contextual understanding of the study. The result of this is to increase the overall integrity of the findings and their applicability in practical – both managerial and organisational – contexts. Leximancer automatically removes pronouns, conjunctions and two-letter words from the analysis, and gives users the option of merging or removing words as necessary. Ultimately, the program operates by discovering and organising the most common words contained in any body of text. Once Leximancer identifies a concept, words that are closely linked to that concept are developed; this process generates themes that surround particular groups of concepts. As a result, Leximancer allows common themes and concepts to be extracted and defined through how they are related to other words included in the text. The approach to this analysis is to state that individual words from the text will not be referred to as ‘themes’; moreover these will be referred to as ‘identifiers’. Themes will be displayed through the ‘theme maps’ produced by the software that includes a representation of multiple ‘words’ or ‘identifiers’.
For output, Leximancer generates a ‘concept map’ or a ‘theme map’ that visually depicts the main concepts and themes (and the links between these) for any piece of written work. Leximancer depicts the relative importance of concepts and themes through size, space and colour-coding; those themes that are highly important to the content analysed are depicted by large, brightly coloured circles. Additionally, Leximancer uses space to depict the relationships between concepts and themes; those items that exist in close proximity to one another are closely related, whereas those items that are separated by some distance are less closely related.

**Data collection and analysis**

The highest-grossing game apps as of 31 March 2011 were identified as *Angry Birds*, *Fruit Ninja*, *Tiny Wings* and *Cut the Rope*. These games were classified as the top-grossing games based on the number of times that they had been purchased; each game was priced at 99 cents.

Comments from consumers who rated one of the games either a one-star (the lowest rating) or a five-star (the highest rating) were each amalgamated into text-only files representing high and low ratings for each game. Additionally, separate files of all the comments for all the games from users who rated a game either one or five stars were created. Thus, a total of ten text files were created for analysis: four files of low-rating comments corresponding to individual games, four files of high-rating comments corresponding to individual games, one file representing all high-rating comments, and one file representing all low-rating comments. A Leximancer theme map was created for each of the text files.

**Results and discussion**

Reviewers shared positive reviews far more often than negative reviews: the comments associated with five-star ratings totalled 3,825 pages and 759,854 words, whereas the comments associated with one-star ratings totalled only 636 pages and 128,006 words. Overall, this came to roughly 41,000 comments associated with five-star ratings and roughly 12,400 comments associated with one-star ratings.

The sizeable differences in the number of comments between high and low rater comments is consistent with previous research in which a positive bias has been found in consumer reviews. Chevalier and Mayzlin (2006), for example, explored the relationship between online book reviews and online book sales, and found that reviewers tended to share positive information more often than negative information. These researchers suggested that, because consumers make the decision to purchase products that they believe they will enjoy, it is not surprising that consumers of the product therefore have a positive bias towards the product.

When searching for an app, an average star rating is given. The researchers posit that the star rating is considered first by the customer, while the reviews offer an additional level of detail that better informs the purchaser. Anecdotally when one observes the shopping or information search behaviour of online consumers and when the authors consider their own behaviour, the ‘easy and quick’ option for assessment of products, services and value is something many of us prefer, for a number of reasons. Neuroscientists argue that the human brain is wired to look for patterns and easy recognition of any presented data in order to assess future events. Reilly (2002) states that, ‘Neuroscientists have repeatedly pointed out that pattern recognition is the key to understanding cognition in humans.’ He continues, ‘pattern recognition also forms the very basis of predicting future events’ and concludes: ‘evidence from numerous studies of human behaviour support[s] the hypothesis that humans perceive patterns from the external world in order to make predictions about the future and they do so even when the patterns are the product of random variability’. Thus the ‘star rating’, in this case for games, is an easy draw on
the eye and brain, enabling the ‘fast’ assessment and quick ‘hit’ of information needed so that decisions can be more acutely defined. Leximancer’s graphical outputs simplify the written feedback for easier interpretation by researchers and marketing practitioners in much the same way as the ‘star based rating’ does for the online consumer. The data become easier to decipher mentally, rather than reading through the text of customer reviews themselves, of which there are often a significant number.

**Analysis of negative comments**

A Leximancer theme map of all negative comments is presented in Figure 1. The strongest identifiers associated with low ratings of all four games were *game, money, fun* and *app*; however, the strongest identifiers with negative connotations were *boring, stupid, easy* and *crashes*. The main findings from the analysis of negative comments are discussed below.

![Figure 1 Leximancer theme map for all comments associated with one-star reviews of Angry Birds, Fruit Ninja, Tiny Wings and Cut the Rope](image)

- **Finding 1:** The main negative identifiers refer to perceptions of the game being boring, easy, or stupid.

  The presence of the identifiers *easy, boring, and stupid* in the comments made by one-star reviewers signals two important points about consumer perceptions of games available on Apple’s app store. First, these identifiers reveal that a game will not be given positive reviews unless it is perceived as exciting (i.e. the opposite of boring and stupid). Second, the identifiers reveal that a game must be challenging in order to give consumers a sense of satisfaction (a ‘theme’). The relationship between *levels* and *easy*, for instance, reflects the fact that users are frustrated by the ease at which they are able to pass through various levels (another ‘theme’).

- **Finding 2:** Negative reviews are strongly linked to technological difficulties.

  The theme map illustrates that technological difficulties are a significant driver of negative consumer reviews: the identifiers of *crashes, update* and *load*, for example, indicate consumer dissatisfaction with the game’s loading time and with the frequencies with which games crash. The link between technological difficulties and negative reviews becomes even more
clearly illustrated in the negative reviews associated with specific games. In the low ratings theme map for *Fruit Ninja*, for example, the identifiers *can’t*, *fix* and *game* were closely linked, showing that users who rated *Fruit Ninja* low suffered from technological glitches that users were not able to address themselves (Figure 2). In the low ratings theme map for *Angry Birds*, the identifiers *version*, *Angry Birds* and *update* were closely linked to one another, illustrating how users attributed negative aspects of the game to the specific version that they were using (Figure 2).

![Figure 2](image1.png)

**Figure 2** Selection from theme maps of comments associated with one-star reviews of *Fruit Ninja* (left) and *Angry Birds* (right)

- Finding 3: Users who provide low ratings to apps urge other consumers not to spend their money on the game.

The presence of the identifier *money* illustrates that those users who rate a game only one-star did not believe that other users should spend their money on the game. Furthermore, that fact that the identifier of *money* appears in close proximity to the identifiers *boring* and *stupid* indicates that these game attributes are directly linked to low willingness to pay. The low ratings *Angry Birds* theme map further illustrates this point: the identifiers *money*, *don’t*, *paid* and *worst* were in close proximity to one another, indicating pleas from unhappy users asking other users to avoid wasting their money on the game (Figure 3).

![Figure 3](image2.png)

**Figure 3** Relationship among themes regarding paying for *Angry Birds* in comments made by one-star reviewers

- Finding 4: Negative reviews of an app are linked to the comments of ads that are linked to the game.

The identifier *ads*, which was directly linked to the identifiers *app* and *star*, suggests a direct link between the chosen star rating of the game and the presence of ads in that game. Of course, both the presence of ads in online applications and their potential to irritate internet users are well established: as internet users have become more in control of their internet use and
navigation, advertisers have been forced to find new ways to grab the attention of potential users (Chan et al. 2010). As a result, the uses of banner ads, pop-up ads and other forms of online advertising have become routine. However, these forms of advertising are well known to irritate internet users. Pop-up ads, for example, are highly effective in capturing the attention of online users (Beard 2001; Kamp 2001), however they are also considered to be the most irritating form of online advertising (Coursey 2001). As ads are already known to bother users, it is not surprising that the presence of ads that either flash on-screen during use of the game, or appear between levels of the game, may be a direct driver of low star ratings for games sold in Apple’s app store. Professionals involved in marketing games sold in app stores would be well advised to consider the advertisement structure associated with their games. The type of ads that lead directly to poor user reviews remains unclear – however, this is an area that could be explored in future research.

**Analysis of positive comments**

A Leximancer theme map of all comments associated with five-star reviews is presented in Figure 4, and main findings from this analysis are described below. The main identifiers associated with five-star reviews were game, app, addicting and fun.

![Figure 4 Leximancer theme map for all comments associated with five-star reviews of Angry Birds, Fruit Ninja, Tiny Wings and Cut the Rope](image)

- Finding 1: High star ratings are associated with user comments of fun, awesome, or addictive games.
- Finding 2: Overall, comments by users who give games high ratings do not urge other users to spend their money on the
Interestingly, the identifier *money* does not appear for the analysis of comments from high-raters. Furthermore, the main themes related to positive comments do not include recommendations to purchase the game (i.e. identifiers such as *buy* or *recommend* do not appear on the map). The fact that one-star comments tended to urge other consumers not to spend their money on a game, whereas five-star comments did not tend to urge other customers to purchase the game is somewhat surprising. However, this finding may shed some light on previous research: Chevalier and Mayzlin (2006) found that negative reviews were more powerful in decreasing book sales than positive reviews were in generating sales – it is possible that this finding is due to the content of the messages conveyed in negative and positive reviews. However, the question as to why users do not urge other people to purchase the game remains unanswered. Furthermore, it is not obvious whether a clear encouragement to purchase the game would result in a substantial increase in purchases above and beyond the most common form of positive reviews (i.e. ‘this game is awesome and fun’) present in current consumer reviews.

- Finding 3: Users who give games high ratings comment on the presence of multiple levels of the game.

The presence of the identifier *levels* indicates that users enjoy completing different levels of a game. Given that satisfied users indicate identifiers such as *addictive*, and dissatisfied users indicate identifiers such as *boring* or *easy*, the identifier *level* may suggest that levels of increasing difficulty are the key to a top-rated game. That is, not only do multiple levels present a challenge for users (therefore obviating issues of boredom), but they also present opportunities for users to continue playing the game (i.e. avoid putting it down, or be addicted to playing it).

**Conclusion**

Although this study has investigated reviews of game apps specifically, it illustrates the utility of Leximancer to analyse large bodies of text; it is important to bear in mind that the methodology used in this research can be applied to the investigation of any form of written consumer review. An analysis of reviews associated with other categories of consumer goods or services may shed more light on patterns of behaviour in positive and negative reviewers. Moving forward, the vast quantity of online consumer reviews presents a tremendous number of research opportunities for exploring online WOM.

The current paper is consistent with past research in which a positive bias has been found in consumer reviews. Furthermore, the particular themes that one-star and five-star reviewers used to describe the games is incredibly telling, and provides managers, marketers and game developers with important information as to the qualities of game apps that users deem important.

Moving forward, this paper brings to light a number of research opportunities. For instance, the sharp contrast between perceptions of a game being addictive and those of a game being boring beg the question as to why these users perceive the same games to be drastically different. Using Leximancer, it is possible to explore this very question; links between concepts can be used to illustrate the finer details behind what consumers say. Furthermore, the fact that the only mention of money occurs when users urge other consumers not to purchase the game raises the question as to why the reverse does not occur.

In addition, further research might focus on the balance of the mix between ‘successful’ and ‘unsuccessful’ games (highest-grossing games indicates successful). The analysis of the negative feedback offered by online reviewers in terms of the difference between ‘dissenting’ feedback given about successful games versus ‘negative’ feedback given about unsuccessful games may generate knowledge in determining effects on online behaviour and consumer influence.
A final issue that this paper has not addressed is how app store consumers and managers can use the reviews that are available to them. For instance, it may be possible for managers to create a bank of buzz-words based on popular themes and concepts that are extracted from a Leximancer analysis. Such a word bank could be used as the basis for marketing campaigns that wish to capitalise on what consumers deem important or relevant.

References


**About the authors**

Karen Robson is a PhD candidate in Marketing at the Beedie School of Business, Simon Fraser University, Vancouver, Canada. She holds a Bachelor’s degree in psychology from Queens University, and Masters degrees in psychology and business administration from Simon Fraser University. A recipient of the Bombardier scholarship for PhD research, Karen has also published in journals such as *MIS Quarterly Executive, Journal of Public Affairs* and the *Journal of Marketing Education*.

Mana Farshid is a researcher in the Division of Industrial Marketing and eCommerce at Lulea University of Technology in Lulea, Sweden, where she also completed her PhD in Marketing at the end of 2012. Her research interests lie in the areas of social media and marketing strategy, and her work has been published in journals such as the *International Journal of Wine Business Research, South African Journal of Business Management* and the *Journal of Financial Services Marketing*.

John Bredican is a lecturer in Marketing, working with a number of European educational institutions and will complete his PhD in 2014 at KTH, The Royal Institute of Technology, Stockholm, Sweden. Building on his practical experience in broadcasting, retail and digital media, his research interests and educational practice include marketing and IS strategy, telecommunications, and mobile and digital media in the business to business area. His work has been published in the
Stephen Humphrey is Director of Training & Development at CatsEye Training and a lecturer in Marketing, working with a number of European universities. His working background and traditional practice has been in the field of engineering, telecommunications and entertainment. Receiving a Masters from the University of Greenwich, London, in Strategic Marketing, he now focuses his efforts as a PhD candidate with The Royal Institute of Technology, Stockholm, Sweden, on the integration and development of cross-discipline research activities. His research interests and educational practice include the fields of neuro-marketing, artificial intelligence and the functional interface between marketers, marketing, consumers and technology.

Address correspondence to: Karen Robson, Beedie School of Business, 8888 University Drive, Burnaby, B.C. Canada V5A 1S6.

Email: krobson@sfu.ca