A Multi-Aspect Evaluation Framework for Comments on the Social Web

Theodore Patkos
FORTH-ICS, Greece
patkos@ics.forth.gr

Antonis Bikakis
University College London, UK
a.bikakis@ucl.ac.uk

Giorgos Flouris
FORTH-ICS, Greece
fgeo@ics.forth.gr

Abstract

Users' reviews, comments and votes on the Social Web form the modern version of *word-of-mouth* communication, which has a huge impact on people's habits and businesses. Nonetheless, there are only few attempts to formally model and analyze them using Computational Models of Argument, which achieved a first significant step in bringing these two fields closer. In this paper, we attempt their further integration by formalizing standard features of the Social Web, such as commentary and social voting, and by proposing methods for the evaluation of the comments' quality and acceptance.

Introduction

While there is a lot of research on arguments within the context of Computational Argumentation, the types of arguments that populate the Social Web have not been formally studied yet. These arguments usually have the form of comments, opinions or reviews, and are the main ingredients of online discussion forums, social networks, online rating and review sites, debate portals and other online communities - the electronic version of *word-of-mouth* communication. Since the emergence of the Social Web, their impact ranges from health-related (Chou et al. 2012), to buying (Cheung and Thadani 2012), travelling (Ye et al. 2011) and voting habits (Bond et al. 2012), but also to the marketability of products and businesses (Luca 2011).

In this paper, we are interested in investigating comments in online debates from the scope of two questions:
Q1. *How closely do the participants of an online debate share the opinion expressed by a given comment?*
Q2. *How helpful do the participants find a comment, and what exactly contributes to its helpfulness?*

Answering the first question is an attempt to value how universally acceptable the position reflected by the comment is, while the aim of the second is to measure and analyze the comment's usefulness. Traditional argumentation frameworks focus primarily on Q1, but most comment-enabled websites (e.g., eBay, Amazon and IMDb), also consider Q2 by ranking comments based on a voting mechanism. Regarding Q2, some recent empirical studies attempt to gain a deeper understanding of what makes online comments useful (Schindler and Bickart 2012; Willemsen et al. 2011), by identifying different content-related characteristics such as relevance, informativeness, clarity and conformity.

Consider the following example:

Example 1. In the imaginary SuggestYourWine forum, comment a_1 expresses the opinion that a particular Cabernet Sauvignon 1992 red Bordeaux wine is an excellent choice for stew dishes. The author of comment a_2 supports this position, by informing that she recently tasted it and noticed how nicely it paired with the Irish stew she cooked. Comment a_3, written by what seems to be an expert in wines, further supports a_2 explaining that the full body of this type of grapes is a perfect match to dishes rich in fat and that the 1990s were golden years for Bordeaux wines. Another person attacks a_1 with opinion a_4, which states that consuming wine is a dangerous habit and should be taken with care. □

Although wine preference is largely a subjective matter, one can expect that comments a_1, a_2, a_3 of Example 1 may eventually enjoy wide acceptance, as they refer to commonly held opinions about Bordeaux wines. Still, a_3 should stand in front of the other two in terms of quality or completeness, as it seems to express an expert and well-explained opinion. As for a_4, although being true in principle, it does not seem relevant to the discussion and its attack should not significantly reduce the acceptance of a_1. Existing argumentation frameworks for the Social Web, e.g. (Leite and Martins 2011; Evripidou and Toni 2014; Eilmez, Martins, and Leite 2014; Baroni et al. 2015), do not distinguish between the acceptance and the quality of arguments. They blend together the combined strength of attacking and supporting arguments with a fuzzy aggregation of votes, even though each of these features can carry different semantics that can lead to a more accurate valuation of arguments. Moreover, they do not isolate irrelevant arguments (trolls) in an intuitive way.

In this paper, we formalize a framework that is flexible enough to model diverse features of comment-enabled sites, providing the machinery for extending them with new ones, if needed; we describe a simple mechanism that exploits users' feedback, in order to distinguish between the score assigned to a comment for valuating the *position that it expresses* and that for valuating *how this position is presented*; and we suggest a set of properties that guarantee an intuitive behavior for comment-enabled sites.

Multi-Aspect Comment Evaluation

Our framework generalizes previous approaches in two ways. First, given an argument set $\mathcal{A}$, it assigns two different scores to characterize the strength of an argument $a \in \mathcal{A}$: the *quality score* $\mathcal{QUA} : \mathcal{A} \rightarrow \mathbb{I}$ and the *acceptance score* $\mathcal{ACC} : \mathcal{A} \rightarrow \mathbb{I}$, with $\mathbb{I} = [0, 1]$.

Second, it enables the definition of diverse criteria or *aspects* to calculate such scores, denoted as $\mathcal{D}_{aspect}$. Depending on the domain of interest, different aspects can be defined, such as how relevant an argument is to the topic of a discussion, how reliable, well-justified or subjective an argument is considered, whether an argument can be characterized as an "expert opinion", and so on. Each of these aspects may influence the quality and acceptance score of a target argument in different ways. In order to calculate scores related to an aspect, one may decide to blend different features, such as positive votes, negative votes and/or supporting and attacking replies (i.e., other arguments).

Definition 1. *An aspect $\mathcal{D}_x$ corresponding to an argument set $\mathcal{A}$ is a quadruple $\langle \mathcal{R}_x^{supp}, \mathcal{R}_x^{att}, V_x^+, V_x^- \rangle$, where $\mathcal{R}_x^{supp} \subseteq \mathcal{A} \times \mathcal{A}$, $\mathcal{R}_x^{att} \subseteq \mathcal{A} \times \mathcal{A}$ are binary acyclic support/attack relations (respectively) on $\mathcal{A}$, and $V_x^+ : \mathcal{A} \rightarrow \mathbb{N}^0$ and $V_x^- : \mathcal{A} \rightarrow \mathbb{N}^0$ are total functions mapping each argument to a number of positive/negative votes (respectively).*

As arguments in online debates are added in chronological order, $\mathcal{R}_x^{supp}, \mathcal{R}_x^{att}$ are acyclic. The goal is to evaluate the strength of arguments considering one or more aspects.

Definition 2. *An mDiCE (multi-Dimensional Comment Evaluation) framework is an (N+1)-tuple $\langle \mathcal{A}, \mathcal{D}_{d1}, \ldots, \mathcal{D}_{dN} \rangle$, where $\mathcal{A}$ is a finite set of arguments and $\mathcal{D}_{d1}, \ldots, \mathcal{D}_{dN}$ are aspects (dimensions), under which an argument is evaluated.*

Using Definitions 1 and 2, we can formalize the forum of Example 1 as an mDiCE framework $\langle \mathcal{A}, \mathcal{D}_{crt}, \mathcal{D}_{inf}, \mathcal{D}_{rlv} \rangle$ where $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ and $\mathcal{D}_{crt}$ refers to correctness, $\mathcal{D}_{inf}$ to informativeness, and $\mathcal{D}_{rlv}$ to relevance. $\mathcal{R}_{crt}^{supp}$ (the support relation with respect to correctness) contains $\{a_2, a_1\}$ and $\{a_3, a_2\}$, while $\mathcal{R}_{crt}^{att}$ contains $\{a_4, a_1\}$. V_{inf}^+ is expected to assign a bigger value to a_3 compared to all other arguments assuming that participants will find it more informative, while $V_{rlv}^-(a_4)$ will probably be big assuming that many participants will find a_4 irrelevant.

Not all aspects are appropriate to any domain. For example, how recent a comment is may not be important when discussing about a music band or a movie, but when it comes to rating a product, the effect of outdated comments may need to be neutralized. By making explicit which aspect is being evaluated by users when placing votes or supporting/attacking arguments, we enable a much more accurate evaluation, avoiding the correlation of unrelated aspects.

The Blank Argument Metaphor

Another way in which our framework extends previous ones is by introducing the following intuition: if votes on some argument a denote answers to an - explicit or implicit - aspect-related question, e.g., "is this a helpful argument?", they themselves express an opinion that can be represented as a supporting argument to the target argument with a measurable strength. Since this new argument has no actual content, rather it shares the same content with the target argument, we name it *blank argument* of a on aspect x and denote it as $\mathring{a}_x$. Note that an attack to a is also an attack to $\mathring{a}_x$, since they both share the same content and rationale.

Example 2. Consider the graph shown on Figure 1(a) presenting a debate involving three arguments, where argument a_2 supports a_1 and argument a_3 attacks it. Each argument is annotated with the number of positive and negative votes it received (shown next to the boxes) and with two values denoting its quality (left box, also depicted by the portion of the painted area of each circle) and acceptance scores (right box, also depicted by the size of the circle). For simplicity, we assume a single aspect in this example. In order to calculate the quality and acceptance scores for a_1, we introduce the blank argument $\mathring{a}_1$ of a_1 (Figure 1(b)). Notice how $\mathring{a}_1$ is attacked by a_3 resulting in a weak support to a_1. □

In accordance with Leite and Martins' notion of *social support* (2011), where votes denote the support of the audience, our approach relies on the intuition that negative votes act as a means to weaken the support towards a and not as a way to strengthen the attack[1]. This is ascribed to the fact that positive votes denote a more self-explanatory response to the aspect-related question than negative ones. Positive votes have a very clear semantics, signifying congruence with the comment in terms of content, justification and stance towards the topic of the discussion. Arguably, the ideal comment has only positive votes and no supporting arguments, as the latter should ideally be asserted only in order to add material or explain better the opinion stated.

Negative votes, on the other hand, are more ambiguous. It is not clear if the person submitting a negative vote disagrees with the position stated; or if she finds it poorly explained to stand as an acceptable position; or if she just feels uncertain whether the comment qualifies for the aspect it is being asked for. Indeed, some rating sites try to interpret the meaning of negative responses using follow-up clarification questions (e.g., "offensive?", "off-topic?", etc.).

As a result, in our framework the strength of a blank argument is associated with the degree with which people have identified themselves with the target argument. And this strength, as already explained, is affected by the combined strength of arguments attacking a, as formalized next.

Definition 3. *Let $\mathcal{F}$be an mDiCE framework and $\mathcal{D}_x = \langle \mathcal{R}_x^{supp}, \mathcal{R}_x^{att}, V_x^+, V_x^- \rangle$ be an aspect of $\mathcal{F}$. For each argument $a \in \mathcal{A}$, we define $\mathring{a}_x$ a new argument related to $\mathcal{D}_x$, called the blank argument of a on x, such that*

- *$V_x^+(\mathring{a}_x) = V_x^+(a)$, $V_x^-(\mathring{a}_x) = V_x^-(a)$,*
- *$(\mathring{a}_x, a) \in \mathcal{R}_x^{supp}$, and*
- *for all $(a_i, a) \in \mathcal{R}_x^{att}$ it also holds that $(a_i, \mathring{a}_x) \in \mathcal{R}_x^{att}$.*

[1] We assume a positive stance when expressing aspect-related questions, i.e., positive votes weight in favor of the target argument. With proper reformulations, one can also support questions of the form "Is this argument out-of-date?", where positive votes characterize a negative stance towards the argument.

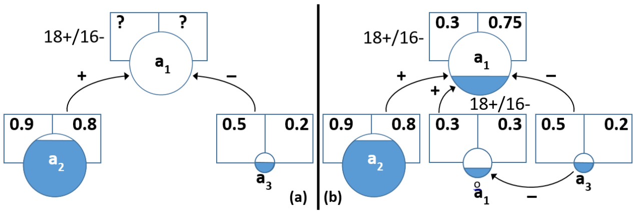

Figure 1: (a) A typical comment exchange graph, (b) The same graph extended with additional mDiCE features.

For notational convenience, we use $\mathring{\mathcal{A}}$ to refer to the set of blank arguments of an mDiCE framework $\mathcal{F}$ and $\widetilde{\mathcal{A}}$ for the set of user-generated arguments ($\mathcal{A} = \mathring{\mathcal{A}} \cup \widetilde{\mathcal{A}}$). Moreover, given an aspect $\mathcal{D}_x = \langle \mathcal{R}_x^{supp}, \mathcal{R}_x^{att}, V_x^+, V_x^- \rangle$, we define the set of direct supporters of an argument $a \in \mathcal{A}$ as $\mathcal{R}_x^+(a) = \{a_i : (a_i, a) \in \mathcal{R}_x^{supp}\}$. Similarly, the set of direct attackers of a is defined as $\mathcal{R}_x^-(a) = \{a_i : (a_i, a) \in \mathcal{R}_x^{att}\}$.

The Set of mDiCE Aggregation Functions

To calculate the different strength scores, we define a set of aggregation functions, which are summarized in Table 1: the left column presents those that drive the process of calculating intermediate scores, whereas the right column involves the ones that can be used to rank arguments from the scope of the motivating questions considered.

Table 1: Overview of the mDiCE Functions

Internal Functions	External Functions
$s_x^{cng}, s_x^{dlg} : \mathcal{A} \to \mathbb{I}$	
$g_x^{cng}, g_x^{dlg} : \mathbb{I} \times \mathbb{I} \times \mathbb{I} \to \mathbb{I}$	$\mathcal{QUA}, \mathcal{ACC} : \mathcal{A} \to \mathbb{I}$
$s^{vot} : \mathbb{N}^0 \times \mathbb{N}^0 \to \mathbb{I}$	$g^{\mathcal{QUA}}, g^{\mathcal{ACC}} : \mathbb{I}^N \to \mathbb{I}$
$s^{set} : (\mathbb{N}^0)^{\mathbb{I}} \to \mathbb{I}$	

The two core functions that we use to characterize the strength of an argument, either blank or user-generated, given a certain aspect x, are the *congruence strength* and the *dialogue strength*, denoted as $s_x^{cng}()$ and $s_x^{dlg}()$, respectively. The former aims to reflect the degree of people's compliance with an argument along the given aspect. As explained in the previous section, the voting mechanism can be employed for this purpose, therefore this score will also characterize the supporting strength of the blank argument. The latter function aims to reflect the combined strength of supporting and attacking arguments that are attached to the target argument, i.e., the dialogue that it generated.

We first define the congruence strength as follows:

Definition 4. *Let $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}_{d1}, ..., \mathcal{D}_{dN} \rangle$ be an mDiCE framework and $\mathcal{D}_x = \langle \mathcal{R}_x^{supp}, \mathcal{R}_x^{att}, V_x^+, V_x^- \rangle$ be an aspect of $\mathcal{F}$. The congruence strength $s_x^{cng} : \mathcal{A} \to \mathbb{I}$ of an argument $a \in \mathcal{A}$ over aspect $\mathcal{D}_x$ is given by*

$$\begin{aligned} s_x^{cng}(a) = g_x^{cng}(&s^{vot}(V_x^+(a), V_x^-(a)), \\ &s^{set}(\{s_x^{dlg}(a_i) : a_i \in \mathcal{R}_x^+(a) \cap \widetilde{\mathcal{A}}\}), \\ &s^{set}(\{s_x^{dlg}(a_j) : a_j \in \mathcal{R}_x^-(a) \cap \widetilde{\mathcal{A}}\})) \end{aligned} \tag{1}$$

with

- *the generic score function $s^{vot} : \mathbb{N}^0 \times \mathbb{N}^0 \to \mathbb{I}$ valuating the strength of an argument considering its positive and negative votes;*
- *the generic score function $s^{set} : (\mathbb{N}^0)^{\mathbb{I}} \to \mathbb{I}$ valuating the combined dialogue strength of a set of arguments;*
- *the generic score function $g_x^{cng} : \mathbb{I} \times \mathbb{I} \times \mathbb{I} \to \mathbb{I}$ valuating the congruence score of an argument, considering the aggregation of the strength of the votes, the positive and the attacking arguments.*

That is, the congruence strength can be determined by aggregating the strength of votes, the strength of supporting arguments and that of attacking arguments. Typically, $g_x^{cng}(x_v, x_s, x_a)$ should lay more emphasis on x_v and x_a, as already described, increasing on x_v and decreasing on x_a; however, we keep the function generic to allow its instantiation to vary from system to system.

Note that the domain of s^{set} is the set of multisets of numbers in $\mathbb{I}$. Moreover, although the congruence strength is defined for both blank and user-generated arguments, the valuation considers the dialogue strength of the underlying non-blank arguments only, avoiding redundancies. This dialogue strength is defined as follows:

Definition 5. *Let $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}_d, ..., \mathcal{D}_{dN} \rangle$ be an mDiCE framework and $\mathcal{D}_x = \langle \mathcal{R}_x^{supp}, \mathcal{R}_x^{att}, V_x^+, V_x^- \rangle$ be an aspect of $\mathcal{F}$. The dialogue strength $s_x^{dlg} : \mathcal{A} \to \mathbb{I}$ of an argument $a \in \mathcal{A}$ over aspect $\mathcal{D}_x$ is given by*

$$s_x^{dlg}(a) = \begin{cases} g_x^{dlg}(s^{vot}(V_x^+(a), V_x^-(a)), \\ \quad s^{set}(\{s_x^{dlg}(a_i) : a_i \in \mathcal{R}_x^+(a)\}), & \text{if } a \in \widetilde{\mathcal{A}} \\ \quad s^{set}(\{s_x^{dlg}(a_j) : a_j \in \mathcal{R}_x^-(a)\})) \\ \\ s_x^{cng}(a)\,, \text{ if } a \in \mathring{\mathcal{A}} \end{cases} \tag{2}$$

- *with the generic score function $g_x^{dlg} : \mathbb{I} \times \mathbb{I} \times \mathbb{I} \to \mathbb{I}$ valuating the dialogue strength of an argument for a given aspect, considering the aggregation of the strength of its votes, its supporting and its attacking arguments.*

The idea is that the dialogue strength of blank arguments coincides with their congruence strength; for the rest, we can consider the aggregation of all supports and attacks that have been placed. In contrast with g_x^{cng}, $g_x^{dlg}(x_v, x_s, x_a)$ should lay more emphasis on x_s and x_a, increasing on x_s and decreasing on x_a.

Finally, by considering the strength of different aspects defined within a particular mDiCE framework, the external scores of an argument can be determined:

Definition 6. *Let $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}_{d1}, ..., \mathcal{D}_{dtN} \rangle$ be an mDiCE framework. The quality and acceptance scores of an argument $a \in \mathcal{A}$ is given by the functions $\mathcal{QUA} : \mathcal{A} \to \mathbb{I}$ and $\mathcal{ACC} : \mathcal{A} \to \mathbb{I}$, respectively, which aggregate the strength of its aspects, such that*

$$\mathcal{QUA}(a) = g^{\mathcal{QUA}}(s_{d1}^{cng}(a), ..., s_{dN}^{cng}(a)) \tag{3}$$

$$\mathcal{ACC}(a) = g^{\mathcal{ACC}}(s_{d1}^{dlg}(a), ..., s_{dN}^{dlg}(a)) \tag{4}$$

with $g^{\mathcal{QUA}}, g^{\mathcal{ACC}} : \mathbb{I}^N \to \mathbb{I}$.

Appropriate instantiations of the aforementioned functions can be applied to comply with the demands and scope of different websites. In (Patkos, Bikakis, and Flouris 2016), the functions that are needed to compute the scores shown in Figure 1(b) are analyzed, along with the formal definition of a set of properties that should be satisfied, as explained next.

Desirable Properties

Our framework is generic enough to allow many different types of functions to be defined. However, there are certain useful properties for such functions, which would guarantee a “reasonable” behaviour for the task at hand, such as *monotonicity* and *smoothness* requirements. Monotonicity requirements constrain the relative effect of a new vote or argument, e.g., that the effect of a positive vote will always be non-negative. Smoothness requirements guarantee that “small” changes in some argument (e.g., a single new positive vote) cannot have “large” effects on the overall evaluation of arguments. Both properties are essential features for the adoption of a rating framework, as they rule out unreasonable effects that would cause users to lose their trust on the objectivity of the rating algorithms.

Discussion

The mDiCE framework can be used to model different types of review and debate web sites such as:

Single-aspect voting-based sites, where users can vote on the helpfulness of a comment, and reply to (but not explicitly support or dispute) other comments. Sites in this category, e.g. Amazon, IMDb, TripAdvisor, App Store and Google Play, can be modeled as mDiCE frameworks with a single “helpfulness” aspect: $\mathcal{D}_{hlp} = \langle \emptyset, \emptyset, V^+_{hlp}, V^-_{hlp} \rangle$.

Multiple-aspect voting-based sites, where users can vote on multiple aspects of a comment. In Slashdot, for example, users can vote other users’ posts with respect to their relevance, informativeness and overratedness. We can model Slashdot as an mDiCE framework with three aspects: $\mathcal{D}_{rlv} = \langle \emptyset, \emptyset, V^+_{rlv}, V^-_{rlv} \rangle$, $\mathcal{D}_{inf} = \langle \emptyset, \emptyset, V^+_{inf}, V^-_{inf} \rangle$ and $\mathcal{D}_{est} = \langle \emptyset, \emptyset, V^+_{est}, V^-_{est} \rangle$.

Debate-based sites, where users can explicitly support or dispute other users’ comments. In CreateDebate, users participate in debates by posting arguments or votes in favour or against other arguments, or by asking clarifications for existing arguments. We can identify two aspects in this case. The first one refers to the level of agreement with an argument’s content: $\mathcal{D}_{agr} = \langle \mathcal{R}^{supp}_{agr}, \mathcal{R}^{att}_{agr}, V^+_{agr}, V^-_{agr} \rangle$, where $\mathcal{R}^{supp}_{agr}$ and $\mathcal{R}^{att}_{agr}$ contain all pairs of arguments (b, a) such that b supports or disputes a, respectively, and V^+_{agr}, V^-_{agr} return the number of positive and negative votes. The second refers to an argument’s clarity: $\mathcal{D}_{clr} = \langle \emptyset, \mathcal{R}^{att}_{clr}, V^+_{clr}, \emptyset \rangle$, where $\mathcal{R}^{att}_{clr}$ contains all pairs of arguments (d, a) such that d requests clarification of a, and $V^+_{clr} = V^+_{agr}$, assuming that users who agree with an argument also find it clear.

Given the range of features of the mDiCE framework, it is obvious that the above sites can be extended by enabling users to vote or argue on multiple aspects of comments. Such extensions are consistent with the findings of several empirical studies (e.g. see Introduction) on the meaning of users’ votes on the helpfulness of online reviews and comments.

Summarizing, we proposed a formal multi-dimensional framework for evaluating online comments taking into account the responses and votes that they receive. Compared to previous efforts, we distinguish between the quality and the acceptance of a comment, and consider different assessment methods for comments. In the future, we plan to generalize the framework to support different kinds of user ratings, e.g., star-ratings instead of boolean votes. We will also study instantiations of the framework and evaluate them with real datasets, e.g., the Yelp and TripAdvisor datasets.

References

Baroni, P.; Romano, M.; Toni, F.; Aurisicchio, M.; and Bertanza, G. 2015. Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation* 6(1):24–49.

Bond, R. M.; Fariss, C. J.; Jones, J. J.; Kramer, A. D. I.; Marlow, C.; Settle, J. E.; and Fowler, J. H. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489(7415):295–298.

Cheung, C. M., and Thadani, D. R. 2012. The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems* 54(1):461 – 470.

Chou, W.-y. S.; Prestin, A.; Lyons, C.; and Wen, K.-y. 2012. Web 2.0 for health promotion: Reviewing the current evidence. *American Journal of Public Health* 103(1):e9–e18.

Evripidou, V., and Toni, F. 2014. Quaestio-it.com: a social intelligent debating platform. *Journal of Decision Systems* 23(3):333–349.

Eilmez, S.; Martins, J. a.; and Leite, J. a. 2014. Extending social abstract argumentation with votes on attacks. In Black, E.; Modgil, S.; and Oren, N., eds., *Theory and Applications of Formal Argumentation*, volume 8306 of *LNCS*. Springer Berlin Heidelberg. 16–31.

Leite, J., and Martins, J. 2011. Social abstract argumentation. IJCAI-11, 2287–2292.

Luca, M. 2011. Reviews, reputation, and revenue: The case of yelp.com. Technical Report 12-016, Harvard Bus. School.

Patkos, T.; Bikakis, A.; and Flouris, G. 2016. Rating comments on the social web using a multi-aspect evaluation framework. Technical Report TR-463, FORTH-ICS.

Schindler, R. M., and Bickart, B. 2012. Perceived helpfulness of online consumer reviews : the role of message content and style. *Journal of Consumer Behaviour* 11(3, (5/6)):234–243.

Willemsen, L. M.; Neijens, P. C.; Bronner, F.; and de Ridder, J. A. 2011. “Highly Recommended!” the content characteristics and perceived usefulness of online consumer reviews. *Journal of Computer-Mediated Communication* 17(1).

Ye, Q.; Law, R.; Gu, B.; and Chen, W. 2011. The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior* 27(2).