

Evaluation the Impact of Human Interaction/Debate on Online News to Improve User Interfaces for Debate Applications

Abdulrahman Alqahtani^{1,2}, Marius Silaghi²

(1Department of Computer Science, Najran University, Najran, Saudi Arabia)

(2Department of Computer Science, Florida Institute of Technology, Melbourne, FL, USA)

ABSTRACT

The average of many people trust online comments for any news as much as personal recommendations [1], [2]. In this paper, we analyzed the impact of the online news's comments to evaluating the threading models of electronic debates by using online surveys. In this paper, based on the results of our online survey of 500 participants, we evaluated whether forums with comments concerning online news are appropriate for the study of debates. In particular, we have to verify whether the nature of discussions around news is argumentative and whether the participating people expect to engage in multiple rounds of arguments. We presented DirectDemocracyP2P application as a user interface for decentralized debates. In this paper, we evaluated and analyzed the comments that were collected from online surveys to improve the DirectDemocracyP2P applications. Also we have to verify whether the actual comments commonly submitted around news do go beyond the simple advertisement of one own's merchandise and attacks of competitors, into fair reviews of news features and quality.

Keywords–User interfaces for Online News, Evaluation of the Impact of Threading Models on Online News, Methodology, and Results

I. INTRODUCTION

Some of people who read the online news would like to read and/or write comment on the article that has been read. We live in a society where almost everyone uses social media applications on devices such as smart phones or tablets. These applications include Facebook, Twitter, YouTube, Gmail, Google Plus, Yahoo messenger, WhatsApp etc. [3], [4]. Most of the recent studies focus on the impact of online comments on open news [5]. This paper focuses on analyzing and comparing threading models for representing knowledge stemming from debates. Peer to Peer starts systems for distributing, editing, and contorting the online news items [5]. Direct DemocracyP2P (DDP2P) provides a system, with multiple graphical interfaces, for news that related to motion in a given organization [6]. In this paper, we present DDP2P application for debate threading. Online networks, including Facebook and MySpace, provide an easy way to create circles of friends. These applications enable the people to connect with the worldwide community. However, they have not yet been studied from the perspective of how they improve group decision making [1], [2]. Large amounts of data and spam that can occur in decentralized social networks raise a challenge for debate user interfaces. The problem that emerges is to scientifically decide whether the debate user interface mechanisms proposed in debate systems do really offer the enhancements for which they were proposed.

II. Background

Recently, we can describe the evolution of a news as open news or open publishing e.g. Bruns. Open news platforms can used in many interest fields and topic from politics to entertainment e.g. Slashdot [5].

Defining the Debate: In certain decision making fora (such as parliaments, or electronically in DirectDemocracyP2P), a debate focuses on a clear motion (i.e., proposal of a decision) that is relevant to a given organization. Users can vote on it with justifications.

We differentiate between debates and brainstorming sessions, namely where a question and its possible answers were not yet crystallized. The regular discussions commonly available with blogs and electronic news are classified as brainstorming sessions, while discussions associated with common polls, news reviews and petition drives platforms are classified as debates.

III. User Interfaces for Online News

The output of debates consists not only in the decision to be done, but also in a better understanding of the issue by the involved participants. Besides improving their understanding of the discussed matter, involved participants can also improve their understanding of each other's point of view. The obtained classification of the goals of a debate is as follows:

a) making a decision (count supports vs count oppositions)

b) improving the understanding

– understanding the abstract matter

– understanding the participants’ point of view

The goal of our research is to identify ways of measuring how well a debate platform helps in achieving the goals mentioned at point (b), namely to improve understanding. Understanding is improved as the user gets acquaintance with the relevant justification provided by other participants. An essential ingredient comes from the correct evaluation of the importance of a justification as yielded by the number of participants supporting it.

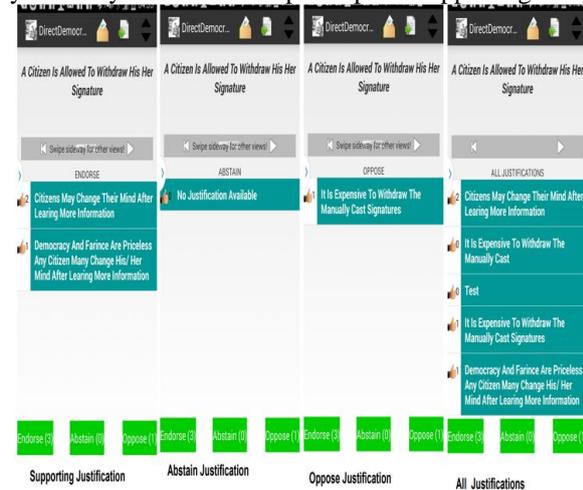


Figure 1 User Interfaces for DirectDemocracyP2P Android application

Another important factor in catalyzing the understanding of a justification is the intensity with which each participant supports that justification. In electronic debates, users can support somebody else’s justification as an alternative to providing his/her own justification. Justifications with large support can be favored by viewers, as they may better represent the opinion of the group. A further mechanism to help users locate relevant justifications is based on threading. Namely, new justifications can point to old justifications that they claim to refute or enhance. Thereby people visualizing old justifications are notified of the presence of the refutation and enhancement claims. In a DDP2P application, all debates and arguments with news have to be related to a motion in a given organization. The user can vote on any with only one justification and he/she can post news linked to motions or justifications. Here we introduce an Android application for DDP2P according to motions and justifications that are in the exit organization.

a) **Motions:** A motion is a proposal related to a statute, constitutional amendment, or discussion

issue that is raised for the vote of a committee or constituency. In some organizations (e.g., US towns), the motions can be submitted to the town council only by certain members of the council. In other organizations (e.g., Swiss towns) a motion can be submitted by any group of citizens that gathers signatures [6].

The mechanism of disseminating motions can be used to help the community converge towards enhanced versions of a motion. Discovery of better versions of a motion can be boosted by an appropriate threading mechanism, with each new motion referring back to previous motions on which it claims to improve. These references create a thread that can be traversed by a user, or can be used by automatic reasoning tools helping users in locating promising motions.

b) **Justifications:** One of the benefits of gathering votes for a motion is that constituents can get an understanding concerning the positions of their peers, and therefore better grasp the implications of a given motion on their organization and on peer members.

Namely, if a majority of peer members disagree with a motion that the user has earlier believed to be good, he may reconsider his position on the motion. The peers could have potentially discovered problems with that motion, problems communicated via justifications that can make the constituent withdraw his/her support. Withdrawing support for an unpopular motion will save the time of the other constituents who will be less tempted to spend time reading it, and this will help the organization to save the resources needed to move on the proposal and organize an official ballot [6], [7]. Common alternatives when voting on a motion are Support, Oppose, and Abstain. However, each submitted motion can be customized to allow for any set of possible reactions as appreciated by the author of the motion. Poor choices are supposed to be correctable by enhancements. As previously explained, the understanding of the opinion of one’s peers can be further improved by enabling the submitter of votes

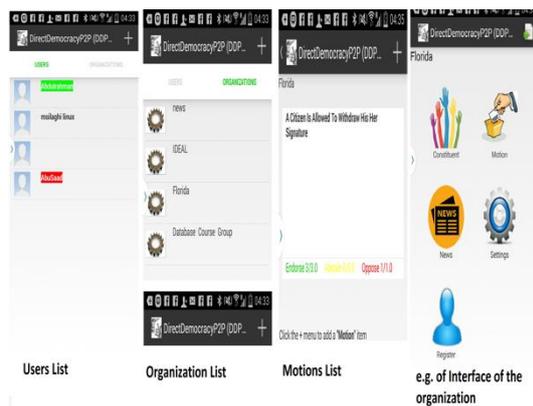


Figure 2 User Interfaces for DirectDemocracyP2P Android application2

to associate a justification of their support or opposition to the motion. Threading and thumbs (technical term that everyone knows) in fora are used for training an automatic moderator [7], [8].

c) **Problem:** The problem that emerges is to scientifically quantify how some debate mechanism supports decision making [7], [9], [10].

IV. METHODOLOGY

There are several types of debates, function of the ad- dressed topic:

- **Online Products:** Online products sellers asking customers to leave a review/comment on their site exist.
- **Online News:** Readers of the online news posting comments on a news article.
- **Religion:** People sharing information about their religious beliefs.
- **Science:** People raising concerns related to the significance and correctness of scientific issues.
- **Politics:** People sharing information about their political beliefs.
- **Sports:** Comments and arguments around news concerning sports.

V. Evaluation of the Impact of Threading Models on Online News

We use surveys to extract the properties of each type of debates, and to see the differences in their rules as expected and deemed appropriate by users. We started an online study from August to September 2015 in which we presented a survey to participants and asked them to answer its questions using the Survey Monkey platform. In our study we use the technique of submitting the questions in an online survey to collect the data from online users. We designed and distributed questionnaires in our survey in a couple of languages (English and Arabic). B. Study Questions Study questions contained three groups:

- **Participation Agreement:** The first question in our survey is a participation agreement. Participation is voluntary.
- **General Information:** We collected general information like gender, age range, secondary language, and level of education to evaluate our survey population.
- **Understanding Questions:** Participants answered a chain of multiple choice questions to determine the factors that attract the users while reading or taping reviews (comments/threads) for any online news.

C. Goals of Our Survey

The purpose of our surveys is to:

- Evaluate how comments for online news, differ from other debates.
- Gather suggestion of how to improve user interfaces for debate applications.

VI. RESULTS AND DISCUSSION

A. Participation Agreement (Institutional Review Board (IRB))

The first question in our survey asked the user to accept a participation agreement.

We asked: Do you agree to the above terms? By clicking Yes, you consent that you are willing to answer the questions in this survey.

Key Finding:

99 percent accepted to answer our survey.
1 percent rejected to participate in our survey.

Analysis: Most of the participants accepted to answer the questions in our survey (497 of 502) as shown in Figure 3. This was an easy question because this question confirmed the participation in our survey.

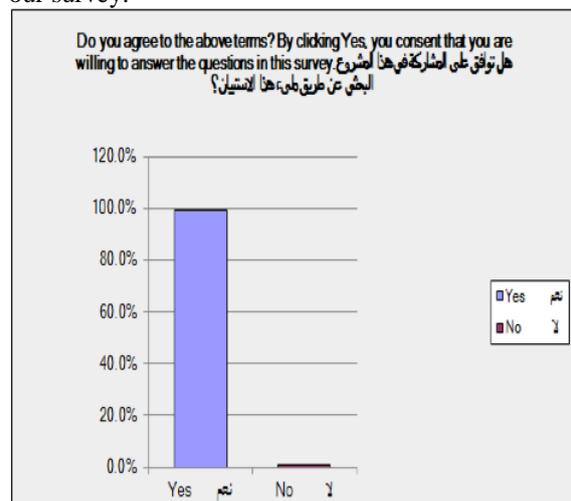


Figure 3 Institutional Review Board (IRB)

B. General Information:

There were four questions about general information. Every question had to fulfill the purpose of the study.

1) Gender: This question aimed to know who was more active in the debate, men or women. This question will help to structure debate user interfaces (DDP2P applications) according to the interaction of humans.

We asked: What is your gender?

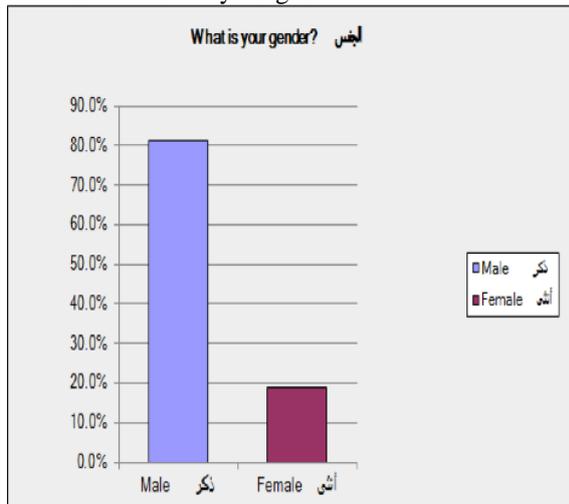


Figure 4 Gender percent of participants in our survey

Key Finding:

81.3 percent of participants were male.
 18.7 percent of participants were female.

Analysis: Bother gender were willing to debate in online news, according to this survey. This question will lead to focusing on their interesting more, by using ads, news, topics, etc., in a debate user interface (DDP2P applications), in order to attract them into successful debates.

2) Age Range: the age range question targeted the age range of participants who are willing to debate.
 We asked: What is your age range?

Key Finding:

4.8 percent of participants were less than 20.
 46.3 percent of participants were between 20 and 30.
 48.9 percent of participants were over than 30.

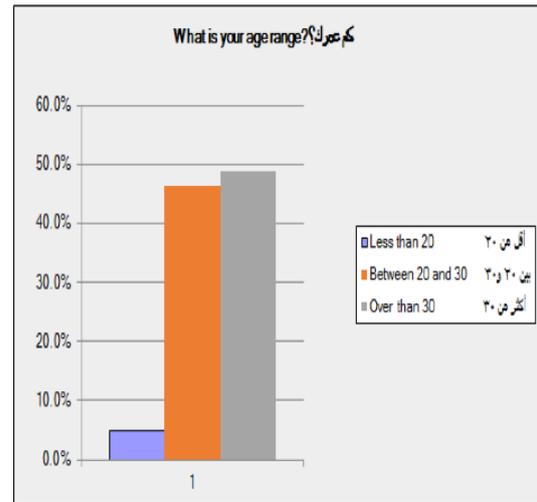


Figure 5 Age Range of participants in our survey

Analysis: The greatest age range of participants, who were willing to debate in this study, was older than 30 (465 of 924), then between 20 and 30 (225 of 460) as shown in Figure 5. This question gave us the age range of participants whom we should focus on when we improve the user interface of DDP2P applications.

3) Secondary Language: The secondary language question aimed to discover which languages are the most popular in the debate.

We asked: What is your secondary language, if any?

Key Finding:

80.4 percent of participants whose second language was English.
 0.9 percent of participants whose second language was Chinese.
 0.4 percent of participants whose second language was French.
 2.0 percent of participants whose second language was Spanish.
 75.0 percent of participants whose second language was Other.

Analysis: We found English was the most popular language in our study (370 of 460) as shown in Figure 6. From this question, in DDP2P applications, we will suggest using English as a formal language to communicate between users. Also, we will put English as the default user interface for DDP2P applications.

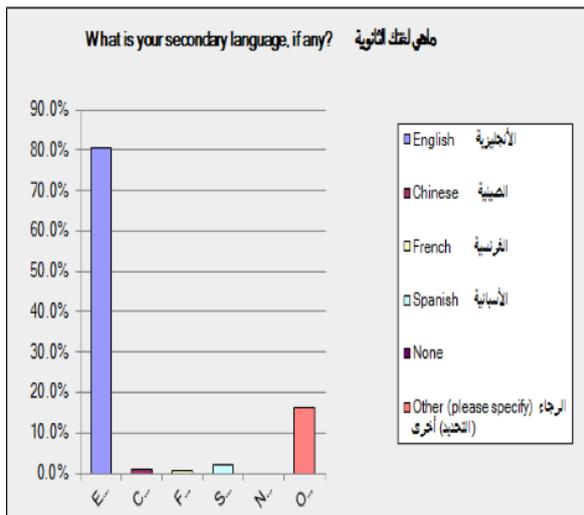


Figure 6 Secondary Language of participants in our survey

4) **Level of the Education:** The level of education question referred to the impact of level of education on the debate.

We asked: What is your education level?

Key Finding:

- 15.2 percent of participants have a High School degree
- 48.3 percent of participants have a Bachelor’s degree
- 26.7 percent of participants have a Master’s degree
- 9.8 percent of participants have a Ph.D degree

Analysis: Most participants have a Bachelor’s degree (222 of 460) as shown in Figure 7. This question showed us that most participants could be familiar with any updates or developments for improving the user interface of DDP2P applications because the majority of participants had a Bachelor’s degree.

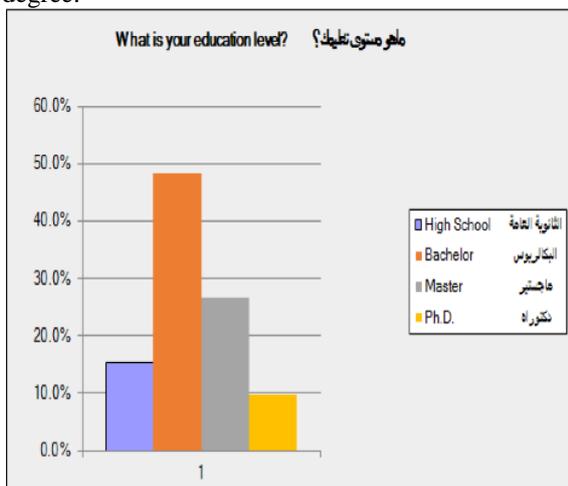


Figure 7 Level of Education of participants in our survey

C. Survey Validity

We had a question to test the validity of our online survey. The validity question depended on asking questions which measured what we were supposed to be measuring.

We asked: How likely would you be to read the comments/threads of a news article after read article online?

Key Finding:

- 19.8 percent of participants usually read the comments/threads of a news article after read article online.
- 73.5 percent of participants sometimes read the comments/threads of a news article after read article online.
- 6.7 percent of participants never read the comments/threads of a news article after read article online.

Analysis: The result of this question referred to whether the majority of participants would read a news comments "Usually" (91 of 460) or "Sometimes" (338 of 460) as shown in Figure 8. Whoever answered "Never" for this question could not continue to the next series of questions because the remaining questions focused on actual readers of news comments.

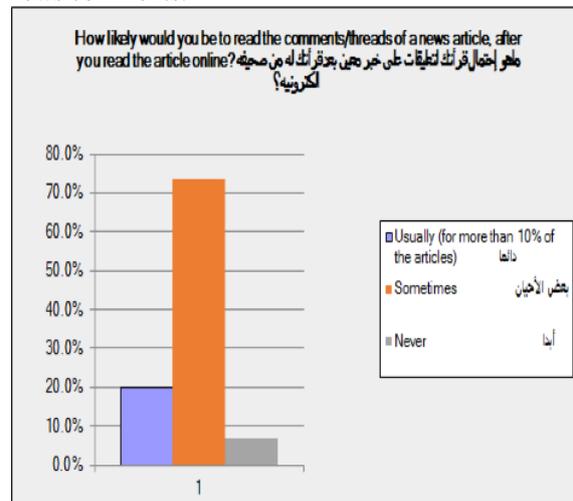


Figure 8 The Validity of our Survey

D. Threads Questions:

We have several questions which focused on comments of online news. Our samples were the participants who read comments. They were supposed to answer a chain of multiple choice questions to determine the factors that attract users while reading or taping comments/threads for any online news. The results of those questions will help us to improve our DirectDemocracyP2P applications.

1) Trusting the Justifications:

Forty-three percent of participants were trusted to read a brief comment (388 of 720) as shown in Figure 9.

We asked: When you read comments on ant online news, what types of comments/threads do you trust the most?

Key Finding:

43.5 percent of participants were trusted to read a brief comment.
 22.3 percent of participants were trusted to read a long comment.
 34.2 percent of participants were not likely to trust any online comments.

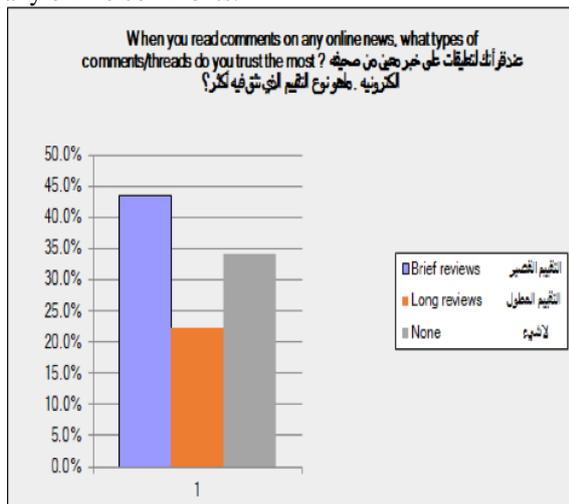


Figure 9: Trusting the Justifications

Analysis: Designing the brief comments by Limiting the length of the motion will help to attract users to debate according to the results of this question. Limitation of the length of the debate arguments will directly affect users' acquisition and, in turn, trusting the justifications about any given motion in the DDP2P applications.

2) Sorting the Important Justification:

Most users would read up to 10 reviews according to the results of this question (298 of 395) as shown in Figure 10.

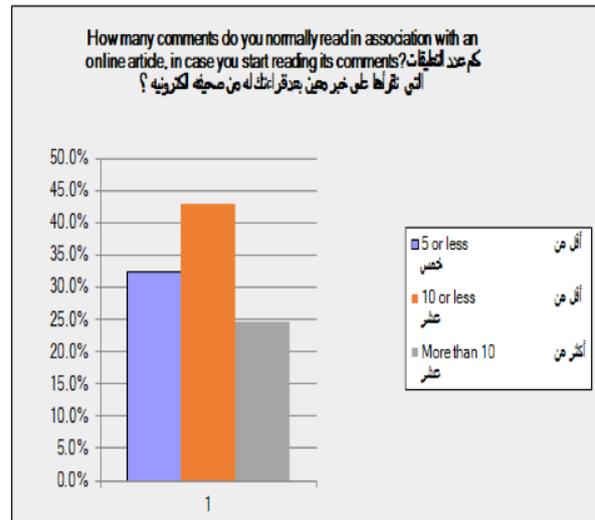


Figure 10: Sorting the Important Justification

We asked: How many comments do you normally read in association with an online article, in case you start reading its comments?

Key Finding:

32.4 percent of participants read 5 or less normally read in association with an online article, in case you start reading its comments.
 43.0 percent of participants read 10 or less normally read in association with an online article, in case you start reading its comments.
 24.8 percent of participants read more than 10 normally read in association with an online article, in case you start reading its comments.

Analysis: In DDP2P applications, sorting the important justifications among the top ten justifications (around a given motion) will give the user opportunity to read them.

3) Separating the Justification:

Sixty-nine percent of participants were likely to read any type of the arguments (Positive or Negative comments) for any debate (274 of 395) as shown in Figure 11.

We asked: When you read comments for some online news article, do you focus on comments that

Key Finding:

9.6 percent of participants were likely to read the comments that are agreed with the article.
 10.4 percent of participants were likely to read the comments that are disagreed with the article.
 6.3 percent of participants were likely to read the comments that are agreed with the user's opinion.
 69.4 percent of participants were likely to read the comments that are disagreed with the user's opinion.

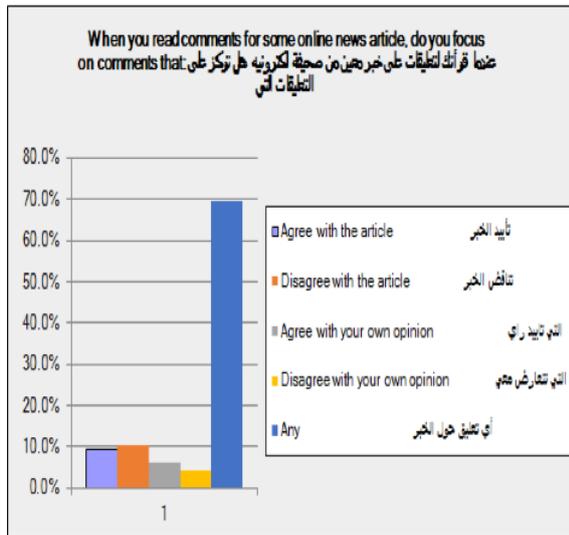


Figure 11: Separating the Justification

Analysis: In DDP2P applications, we have already separated the justification on a motion whether Support, Oppose, or Abstain.

4) Showing the number of Justifications and Witnesses:

Fifty-nine percent of participants disagreed with the statement, "Would a number of positive comments, the number of readers, or other rating criteria, be enough for you to trust a specific news from an online news?" (233 of 395) as shown in Figure 12.

We asked: Are a number of positive comments, the number of readers, or other rating criteria, enough for you to trust a specific news from an online news?

Key Finding:

- 20.8 percent of participants answered (Yes).
- 20.3 percent of participants answered (Yes, if I do not have personal opinion).
- 59.0 percent of participants answered (No).

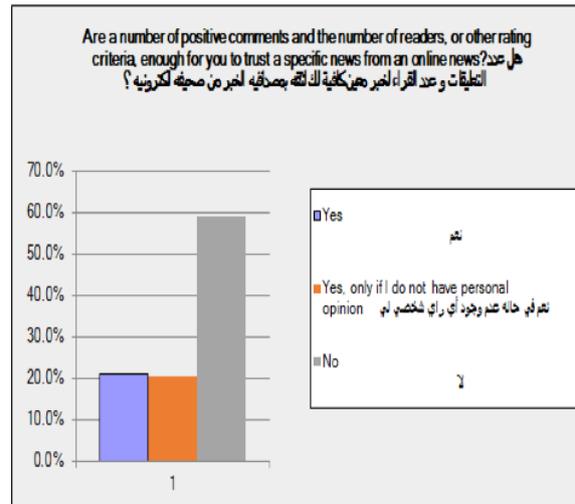


Figure 12: Showing a number of positive reviews (comments/threads) and the number of stars, or other rating criteria

Analysis: Showing a number of positive comments, the number of reader, or other rating criteria will attract users to read and write comments and make good arguments. In DDP2P, a number of justifications, the number of witnesses, or other rating criteria should be shown in the first page of the user interface for the motion.

4) Form for Attention-Grabbing-Words:

Sixty-two per- cent of participants would be attracted by any type of online news comments (Positive or Negative Words)(248 of 395) as shown in Figure 13.

We asked: What types of words attract you the most while reading comments for any online news?

Key Finding:

- 29.4 percent of participants were attracted by positive words of arguments.
- 7.8 percent of participants were attracted by negative words of arguments.
- 62.8 percent of participants were attracted by both sides of arguments (Positive or Negative words).

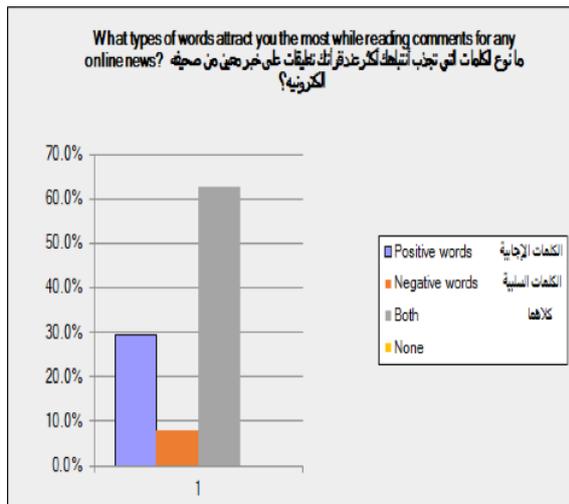


Figure 13 Types of words that attract users

Analysis: In DDP2P applications, we could design a form for attention-grabbing words which would attract users to become more involved in the debate.

5) Form for emphasizing words:

Forty-six percent of participants would expand words, when you type a comment in reviews as shown in Figure 14.

We asked: When you type a comment in reviews (com- ments/threads) for any online news, do you expand some words for emphasis? For example
 verrrrrrrrrrrry

Key Finding:

6.3 percent of participants were always likely to expand some words for emphasis when they typed a comment in online reviews.
 39.7 percent of participants were sometimes likely to expand some words for emphasis when they typed a comment in online reviews.
 53.9 percent of participants were never likely to expand some words for emphasis when they typed a comment in online reviews.

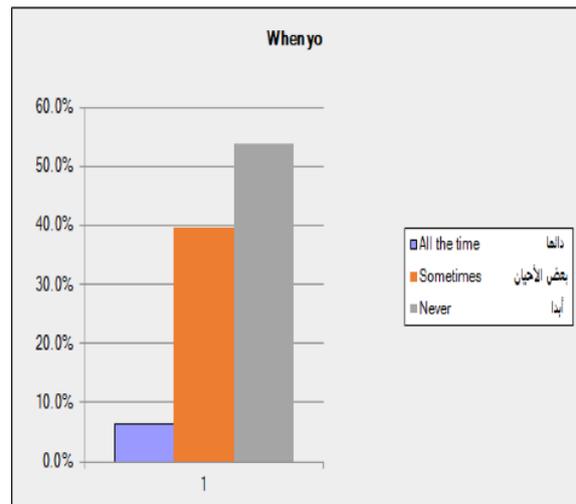


Figure 14 Expanding some words for emphasis

Analysis:

In DDP2P applications, we can design a form for emphasizing words which will attract users to become more involved in the debate.

6) Form for Translating Words of the User Region:

Sixty- three percent of participants would use argot language from their region (251 of 395) as shown in Figure 15.

We asked: When you type a comment in a review (comments/threads) for any online news, do you use argot language from your region?

Key Finding:

11.9 percent of participants were always likely to use argot language from their region.
 51.6 percent of participants were sometimes likely to use argot language from their region.
 36.5 percent of participants were never likely to use argot language from their region.

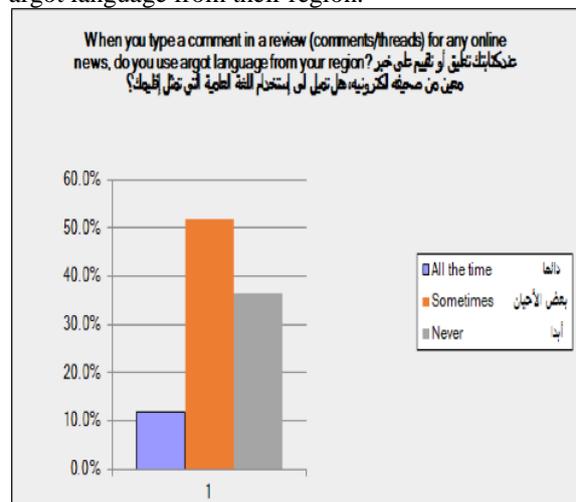


Figure 15: Using argot language from user region

Analysis: In DDP2P applications, we should design a form for translating words of the user's region to English, and give some space to clarify these words (enhancement).

7) Form for Supporting Translation:

Fifty-five percent of participants were never likely to use words from other languages (219 of 395) as shown in Figure 16.

We asked: When you type a comment in a review (comments/threads) for any online news, do you use some words from other languages?

Key Finding:

- 3.8 percent of participants were always likely to use words from other languages
- 40.8 percent of participants were sometimes likely to use words from other languages
- 55.9 percent of participants were never likely to use words from other languages

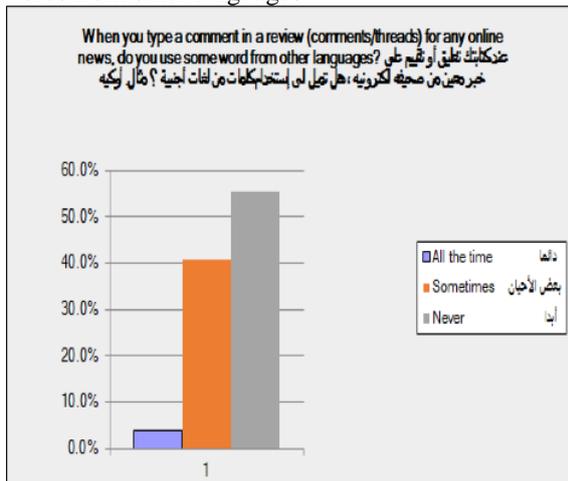


Figure 16: Using any word from other languages

Analysis: In DDP2P applications, we should design a form for supporting translation of words of the users' languages, and give the users space to explain these words (explanation).

8) Benefit of Study Threads:

Thirty-four percent of participants said that comment about any online news are argumentative reviews while thirty percent described online reviews as positive reviews. The rest of participants considered online reviews as negative reviews as shown in Figure 17.

We asked: From your perspective, how would you generally describe reviews (comments/threads) about any online news?

Key Finding:

34.4 percent of participants described online reviews

as argumentative comments

30.4 percent of participants described online reviews as positive comments

35.2 percent of participants described online reviews as negative comments

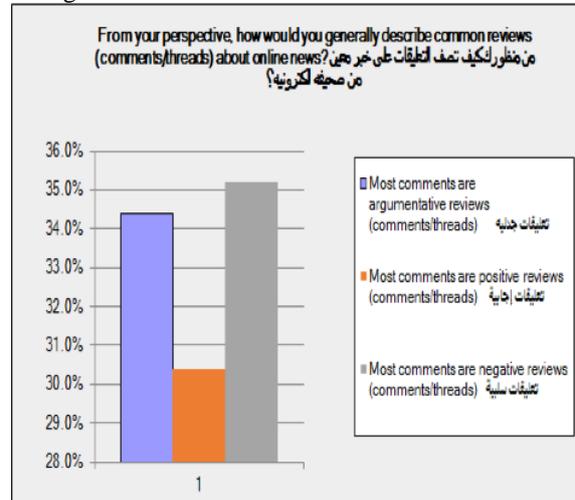


Figure 17: Benefit of Study Reviews/Threads

Analysis: The results of this question gave us the benefit of studying online reviews (comments/threads). There are a lot of users who trust online reviews, especially if they are serious and positive reviews.

10) Structured/Unstructured Platform for Threads:

Fifty-eight percent of participants were likely to prefer platforms for reviews (comments/threads) of the online news to be structured, which could be a specific question that the user should answer or comment on. Structured platforms helped extract a conclusion of arguments around the news as shown in Figure 18.

We asked: How do you prefer platforms for reviews (comments/threads) associated with with online to be structured platforms or unstructured platforms?

Key Finding:

58.2 percent of participants preferred reviews (comments/threads) for online news to be structured platforms

41.8 percent of participants preferred reviews (comments/threads) for online news to be unstructured platforms.

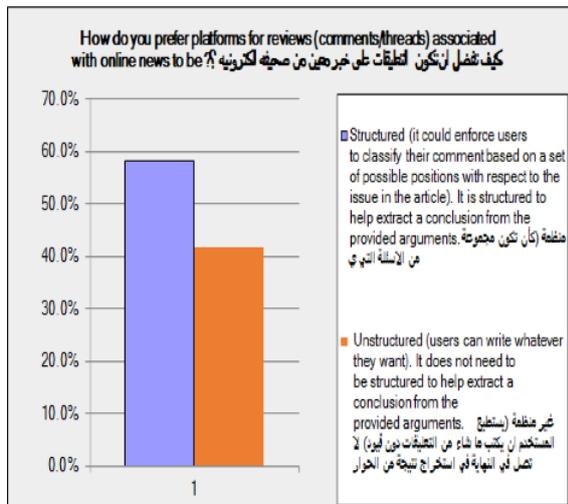


Figure 18: Structured/Unstructured Platform for Reviews/Threads

Analysis:

Forty-one percent of participants were likely to prefer re- views (comments/threads) for online news to be unstructured platforms. In DDP2P applications, we should have those two types of platforms. Unstructured platforms could be used for peers to join or create any organizations/motions, and structured platforms could be used for voting to post only one justification for any given motion, and whether they support it or are against it.

VII. ACKNOWLEDGMENT

We are sincerely grateful to Dr. Muzaffar Shaikh, Dr. John Lavelle, Dr. Khalid Abuhasel for their support and sharing their truthful and illuminating view on a number of issues related to the survey.

VIII. CONCLUSION

Here we have analyzed the impact of reviews (threads) for online news based on an online survey. We evaluated whether forums with comments concerning online news are appropriate for the study of debates on ideas. In particular we found that the discussions around news is argumentative and people expected to engage in multiple rounds of arguments. Also, we have gathered the following suggestions that may be used to improve user interfaces for debate applications:

- People believe that they “trust brief comments over long comments”. Most people believe that they “do not read more than 10 comments”. Most people believe that they “like to read both opposing arguments of a debate”, (despite the evidence suggesting that people end up reading only the news channels sharing their opinions).
- Most people believe that they are “satisfied by knowing the position of others without caring about their arguments”. Most people believe that they “are not biased by key- words in their

attention”.

- Half of the people believe that they “want to display their emotions in their comments”.
- Most people believe that they like to display “elements of their identity in their comments”.

REFERENCES

- [1.] M. S. Abdulrahman Alqahtani, “User interfaces for representing knowledge stemming from debates: Evaluating the impact of threading models (reviews) on online products,” 2015.
- [2.] M. S. Alqahtani, “Classification of debate threading models for representing decentralized debates,” 2015.
- [3.] S. Buchegger, D. Schiöberg, L.-H. Vu, and A. Datta, “Peerson: P2p social networking: early experiences and insights,” in Proceedings of the Second ACM EuroSys Workshop on Social Network Systems. ACM, 2009, pp. 46–52.
- [4.] H. R. Kim and P. K. Chan, “Learning implicit user interest hierarchy for context in personalization,” in Proceedings of the 8th international conference on Intelligent user interfaces. ACM, 2003, pp. 101–108.
- [5.] A. Bruns, “Stuff that matters: Slashdot and the emergence of open news,” 2003.
- [6.] M. C. Silaghi, K. Alhamed, O. Dhannoon, S. Qin, R. Vishen, R. Knowles, I. Hussien, Y. Yang, T. Matsui, M. Yokoo et al ., “Directdemocracyp2pâAT~decentralized deliberative petition drivesâAT~,” in Peer-to-Peer Computing (P2P), 2013 IEEE Thirteenth International Conference on. IEEE, 2013, pp. 1–2.
- [7.] K. Kattamuri, M. Silaghi, C. Kaner, R. Stansifer, and M. Zanker, “Supporting debates over citizen initiatives,” in Proceedings of the 2005 national conference on Digital government research. Digital Government Society of North America, 2005, pp. 279–280.
- [8.] A. Bondarenko, P.M. Dung, R.A. Kowalski, and F. Toni, “An abstract, argumentation-theoretic approach to default reasoning,” Artificial in-

- telligence, vol. 93, no. 1, pp. 63–101, 1997.
- [9.] C. Reed and G. Rowe, “Araucaria: Software for argument analysis, diagramming and representation,” *International Journal on Artificial Intelligence Tools*, vol. 13, no. 04, pp. 961–979, 2004.
- [10.] H. Mercier and D. Sperber, “Why do humans reason? Arguments for an argumentative theory,” *Behavioral and brain sciences*, vol. 34, no. 02, pp. 57–74, 2011.