Rule of Thumb – Effect of Social Button Icons on Interaction

Adrian Holzer
EPFL
1015 Lausanne, Switzerland
adrian.holzer@epfl.ch

Andrii Vozniuk
EPFL
1015 Lausanne, Switzerland
andrii.vozniuk@epfl.ch

Samuel Bendahan
University of Lausanne
1015 Lausanne, Switzerland
samuel.bendahan@unil.ch

Denis Gillet
EPFL
1015 Lausanne, Switzerland
denis.gillet@epfl.ch

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org. MobileHCI '16 Adjunct, September 06 - 09, 2016, Florence, Italy Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM 978-1-4503-4413-5/16/09$15.00 DOI:
http://dx.doi.org/10.1145/2957265.2961842

Abstract
Social buttons are now widespread in social media apps. They are used to assign weight to user content and trigger user engagement. They come in different shapes (e.g., thumb in Facebook, arrows in Reddit or StackOverflow, plus one in Google+) but very little is known about the influence of the shape on user behaviour. This paper, addresses this issue by presenting results of a controlled randomized experiment with 173 users. The results suggest that *thumbs up / thumbs down* icons are significantly more engaging than the *plus one / minus one* icons. At the same time the result shows that type of the icon used has no significant influence on the direction of the vote.

Author Keywords
Social Media; HCI; Social buttons; Interaction design

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction styles

Introduction
February 9, 2009 – two years after the first discussions about potentially adding some kind of *props* feature to its interface to allow users to engage with each other's content, Facebook introduces its *like* button with its

iconic thumb up image [23].In the early days of the Web, user engagement was measured through hit counters [8] and the first social counters were introduced by bookmarking sites such as Delicious and content aggregation sites such as Digg or Reddit [13]. Nowadays in social media, voting on user generated content has become a standard feature but it is implemented in different ways across systems. For instance, YouTube uses a *thumb up* and a *thumb down* buttons for voting on videos. While Facebook has only the *thumb up* icon at the moment of writing, there are predictions that it is going to introduce a *thumb down* option as well [3]. Sites such as Stackoverflow and Reddit use *arrows* to allow users to rate up or down messages. Google introduced the *plus one* (+1) button in March 2011 for the same purpose [21] .

The social buttons have become so mainstream that commentators now talk about the *like economy*, which transforms user engagement into number of button clicks, which can in turn be traded and monetized [12]. Experts assign a monetary value to every single like click and some go as far as estimating the value of USD 214.81 for each Facebook like click on the page of non-profits seeking to attract donation [5]. With such numbers, these buttons have become central in a social media application's interaction design, yet there is a lack of studies investigating how social button design can influence user engagement [1]. In this exploratory paper, we specifically want to address this issue by investigating the following research question:

RQ: Do social button icons influence user interaction?

In order to answer this question we performed a real-world randomized controlled experiment with 173 users of SpeakUp, a social media app for classroom interaction. With SpeakUp, users can post anonymous messages and upvote or downvote them using social buttons. We performed an experiment over five 90-minute sessions, where we randomly assigned one of three social button interfaces to users (thumbs, arrows, plus one), analyzed users behavior and surveyed their attitudes.

This paper is structured as follows. We first discuss related work. Second, we present the SpeakUp app, which will serve as test case for the experiment in this paper. Third, we present our research methods. Fourth, we describe our interaction and survey analyses. Finally, we conclude by highlighting findings, discussing limitations and pointing to future work.

Related work

Interface designers can be viewed as choice architects [24], as described by Thaler and Sunstein in their seminal book *Nudge* [24]. Indeed designer can influence the choices people make by performing interface changes. For instance, making a choice a default option (opt out vs opt in) can significantly nudge user behavior towards making that choice [24]. As such nudging can be seen as being an element in the toolkit of persuasive computing [11].

Nudging user behaviour through user interface changes and interaction has been the hallmark of recommender systems [7] where user interactions are used to compute similarity and recommend relevant products (e.g., [4, 16]). Nudging behaviour by adding of removing social feedback features has also come under scrutiny lately (e.g.,[6, 9, 17]). With these features, the community can help its members understand its norms [17] by rewarding good behavior

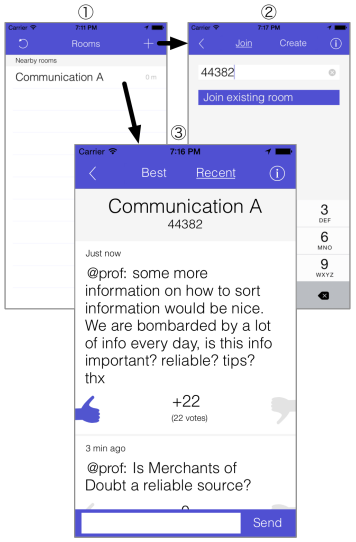

Figure 1: Screenshots of the SpeakUp application (with the thumbs interface).

through positive feedback (e.g. scores) and punish bad behaviour. Note that designing adequate feedback mechanisms is not trivial since they can lead to higher (e.g., [6]) but also lower (e.g., [9]) engagement.

Whereas recommender systems and social feedback systems use social interactions to create nudges, in this paper, we investigate how the design of social interaction buttons can by itself nudge behavior. Previous work in this field is still very limited as recently observed by Alonso and Kandylas [1]. In their descriptive work, they surveyed social button placement on web pages. They survey 10000 URLs and look at them over a period of 6 months, tracking their use of social buttons and the number of clicks on them. They find that Facebook trumps Google, Twitter and LinkedIn in both in the number of pages that contain the buttons, but also in the mean number of clicks per page. This result can potentially be explained by the number of users of the platforms. A surprising result shows that the Twitter *tweet* button, which is less present than the Google *g+* button, has similar or higher clicks per page number. Icon design could potentially explain the difference in behaviour, however it is impossible to reach such a conclusion without a controlled experiment, which we provide in this paper.

The authors of [27] present some design principles for the use of social buttons for their OpenChoice prototype to encourage facilitated participation among community users for voting, rating or discussing. For instance, easy to read, dynamically updated interaction, rating information visible to the whole community, simple rating scale (e.g., up/down, 5 stars), and clear task flow. In our work we followed these principles in the design of our social rating system.

Finally, the authors of [26] investigate whether dynamic voting, i.e., showing the results of the casted votes showed before voting, influences voting behavior. In their experiment they used both a graphical representation with a numerical representation of the numbers of likes and dislikes and also a condition with only numerical representations. They find that graphical representation of prior votes strongly influences user votes, whereas numerical representations alone do not seem to be noticed. In our work, we show a numerical representation of the prior votes, thus we might expect that prior votes would not strongly affect behavior.

In this paper we address an open research avenue pointed out by [6], namely how simple interface changes, which can be as minute as changing the shade of a color [20], can affect interaction. We chose design alternatives for the up and downvote icons of SpeakUp among ones employed by popular social media sites. In Dreamgrows top 10 social media list [10] there are four types of social interaction buttons: thumbs, arrows, hearts, +1. We used thumbs, arrows and +1, since they could all be used for up and downvoting. We measure the effect of interface change on user activity (i.e., posting [19]. We hypothesize that the design change will influence the number of votes, but not the number of posts.

SpeakUp

SpeakUp is a mobile app that allows to create temporary chat rooms that users can join anonymously and where they can post messages and vote on them. It has been used in several classroom settings, has been shown to increase interaction, and is appreciated by both lecturers and students [15, 14]. SpeakUp

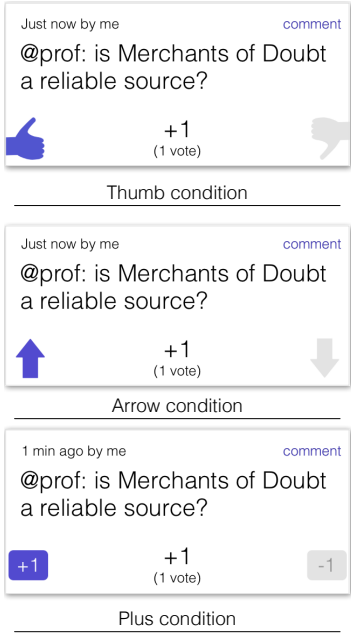

Figure 2: Screenshots of the three conditions. A third of the users where assigned thumbs, a third arrows and a third plus/minus icons on the social voting buttons. Icons are blue when pressed, gray when not.

scores 83 [15] on the System Usability Scale [2] (SUS) which indicates good to excellent usability. Figure 1 shows screenshots of SpeakUp. When users open the app, they see the nearby chat rooms (screen 1 in Figure 1). A user can simply click on it to enter, without the need to log in. If a user does not see the room in the nearby list, she can press the + button from the home screen to join the room with the room number (screen 2). In a room, users can anonymously post and vote on messages (screen 3). Each message can be rated up or down, which adds or removes a point to the score. Individual messages can be discussed by posting comments for a message.

Method

We set up an experimental study using randomly assigned conditions in real settings. The group of users were first year Bachelor engineering students following the Communication A course at our university. They were instructed to download and use SpeakUp during five lectures which lasted one hour and a half each. Upon opening the SpeakUp application the first time, the app automatically and randomly assigned one of the following three conditions: thumb, arrow, and plus one (+1). Thereafter, corresponding icons were displayed whenever the user opened the application as demonstrated on Figure 2. All other parts of the SpeakUp user interface were exactly the same across the three conditions.

Interaction analysis

The overall usage data over the 5 lectures shows that SpeakUp was used by a maximum of 174 users per session. Overall 924 messages were posted and 7934 votes were cast. The result of the automatic condition assignement was that there were 33% users in the thumb condition, 38% in the arrow condition and 29% in the plus one condition. The interface distribution was consistent with a random assignment. Overall, of all users it appears that almost 100% have at least once opened the app, 82% of them have at least voted once and 52% have at least created one message. Given that some commentators assume, as a rule of thumb, that 1% of social media users produce content, 10% interact with it and 89% views it [25], our setting can be seen as especially interaction friendly. In the analysis we focused only on the first level messages in the room, i.e., we excluded comments from the analysis, since we assumed that users would be more evenly exposed to the first level messages than to the comments which can only be reached by pressing an extra button on the interface. Additionally, we removed one outlier user who posted over 100 times the same message in the last session.

Table 1 shows descriptive data about the number of users per lecture each week (in each room), the number of messages created and number of votes per lecture. Note that if a user changed her vote on a message only the last selected option (vote or non-vote) was counted.

Table 1: Descriptive SpeakUp usage data.

	Week 1	Week 2	Week 3	Week 4	Week 5
Users	173	156	163	83	89
Messages	262	345	218	63	35
Votes	2687	3171	1669	223	184

To answer whether content rating icons influence user interaction, we assess the possible impact of the rating icons on (1) the number of votes, (2) the vote direction (positive or negative), and (3) the number of posts.

Table 2: Descriptives for each interface

Interface	Thumbs	Arrows	Plus
Obs. (N)	208	252	204
Interact. (DV)	5.82	5.17	4.70
Std Dev	3.00	3.04	3.13
Std Err Mean	0.21	0.19	0.22

Table 3: Regression table

	Interactions	
Thumbs	1.116**	(2.92)
Arrows	0.469	(1.27)
Constant (Plus)	4.706***	(16.07)
Observations	664 (282 clusters)	

t statistics in parentheses
*p<0.05, **p<0.01, ****p<0.001

Do content rating icons influence the number of votes? In order to analyze the influence of the interfaces on the voting behavior, we first performed regression analysis on the raw data per room and overall. Then in order to control for outlier effects, for each room we divided the users into deciles based on the number of votes casted. Afterwards, we analyzed the number of users by interface in each decile.

Table 2 shows the descriptives for each interface. Note that these descriptives fail to highlight the fact that a lot of the observations (N) are in fact multiple observations from the same user, who has been tracked and participated in multiple sessions (the system tracks individual users based on the id provided by the device without revealing user identity).

The cluster regression in Table 3 corrects this issue. A cluster-robust regression shows that users with the thumb interface cast significantly more votes than users with both other interfaces pooled together ($F(2, 281) = 7.902, p < 0.01$). We also ran a regression accounting for the effects of each interface individually, and as shown in Table 3, we could not find a substantial difference between any of the three conditions. However, the users in the thumb interface condition showed significantly more voting interaction than the users in the plus interface condition ($t = 2.92, p < 0.01$). We ran the same analysis using the total number of interactions, including vote switch, did not significantly change the results.

Do content rating icons influence the number of posts? The number of posts per user are shown in Table 4 and does not appear to be influenced by the interface.

Table 4: Posts per user grouped by week and per interface.

	Week 1	Week 2	Week 3	Week 4	Week 5
Thumbs	1.09	2.39	0.76	0.42	0.24
Arrows	1.51	2.20	1.91	1.27	0.57
Plus	1.98	2.02	1.47	0.65	0.38

Do content rating icons influence the direction of the vote (positive or negative)? Overall, according to the descriptive data represented in Table 5, in each of the conditions the votes are generally more positive than negative, with a ratio around 70 to 30 in favor of the positive votes. Statistical analyses failed to show a substantial difference between any of the three conditions hence the different interfaces do not appear to influence the positive/negative ratio.

Table 5: Positive vote ratios per week and per interface.

	Week 1	Week 2	Week 3	Week 4	Week 5
Thumbs	74%	60%	66%	70%	71%
Arrows	73%	72%	79%	89%	66%
Plus	66%	64%	72%	83%	68%

Survey analysis

In order to understand what users meant when they up or down voted a message in SpeakUp with the rating icon they were assigned, we have performed a voluntary survey to ask users about their attitudes towards voting. In general, the users appreciated being able to vote as over half of the survey respondents strongly agreed to the following three statements: I like being able to vote in SpeakUp, I like being able to vote positively in SpeakUp, and I like being able to vote negatively in SpeakUp.

Furthermore, we inquired about the meaning of positive and negative votes. According the Reddiquette, Reddit's etiquette: *If you think something*

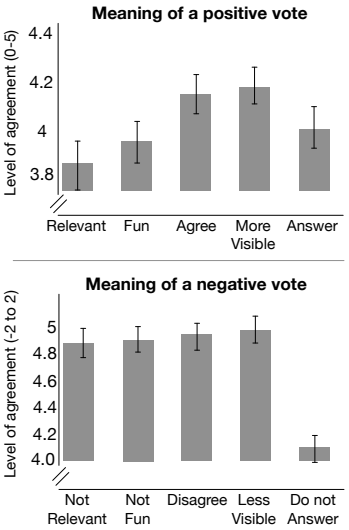

Figure 3: Mean and standard error of the level of agreement with different meanings of a positive or a negative vote on a message. The score ranges from 0 (strongly disagree) to 5 (strongly agree)

contributes to conversation, upvote it. If you think it does not contribute to the subreddit it is posted in or is off-topic in a particular community, downvote it. However, the user community on Reddit has a diversity of views concerning the issue, as is demonstrated in a popular thread [22] debating the issue (the thread was upvoted 1222 times and 777 comments were posted). This thread points to several potential reasons such as to promote visibility of a post, to convey agreement and disagreement, to label posts as relevant or trolls, or to indicate that the post is funny. Thus, in our survey we asked the users about the meaning they associated with a positive and a negative vote. On the 5-point Likert scale we asked them if they agreed with the claims that they rated a message positively because it was (1) relevant to the class, (2) because it was fun, (3) because they wanted to show that they agreed with it, (4) because they wanted to make the message more visible, (5) because they wanted to get an answer from the lecturer. We also asked the opposite questions for negative votes.

We obtained 108 results (age 18-22, 18 females). Figure 3 shows the mean of the level of agreement for each of the meanings (0 = strongly disagree, 5 = strongly agree) with standard error bars for both positive and negative votes for all 108 users. Using an ordinary least squares regression, we found a significantly different approval rating for both visibility and agreement interpretations, compared to relevance interpretation of the positive vote $(F(1, 108) = 8.86, p < .01)$. For the negative vote interpretation, all meanings fared equally well except the meaning that stated that people would vote negatively in order to indicate that they would not want the teacher to answer the question. This interpretation scored much lower than all other interpretations $(F(1, 108) = 52.43, p < 0.001)$. In the context of SpeakUp, this finding indicates that negative votes might not be a good filter for teachers when they select messages to answer. The analysis also failed to show any substantial difference between interfaces in terms of the meaning of positive and negative votes. The familiarity principle of attraction [18] could provide a possible explanation, where the familiarity of the thumbs icon leads to more interaction. However, this principle would also predict that users would find familiar interfaces more attractive, which was not supported by our data.

Conclusion

This paper investigated the difference in engagement caused by the icon (thumb, arrow, plus one) of social buttons. It presented a controlled experiment in the context of the regular usage of a social classroom interaction app called SpeakUp, where 173 users could post anonymous messages during five lessons lasting 90 minutes each. The users were randomly assigned one of the three vote icon representations and we analysed how these conditions affected their voting behaviour. Furthermore, we conducted surveys to understand possible reasons for behaviour differences.

Our results show that the thumb interface demonstrates significantly higher level of interaction (number of votes) compared to the plus interface. This finding indicates that designers should pay special attention to social icons in a context where every like is counted and monetized. At the same time it did not influence significantly the outcome of the votes (i.e., there were not more positive or negative votes with any interface). This finding conveys the fact that the change

in icons in this research appears to increase interaction but not bias the outcome of the votes. In order to explain why thumbs drew more votes that the other designs, we conducted a survey to assess the meaning users attributed to their votes. However, our results fail to highlight a specific difference in the meaning users conferred to their votes.

Our research has several limitations inherent to any field study that limit the generalizability of our findings. First, as noted above, even though we conducted qualitative surveys to understand the observed difference in behaviour, at this stage of the research, we failed to provide an adequate explanation for it. Second, our research has been conducted in the classroom setting for a particular interaction application. The study was done within a single instance of a single course (although across 5 weeks). Also, as in any social experiment outside of a laboratory, there are unobserved factors that can also contribute to explain the variance of the dependent variable, hence the relatively small effect size. Third, some factors affecting the study could be related to the cultural or educational context, particular class size and dynamics, or self report bias. Nevertheless, by design, because of the random assignment of experimental conditions there is no endogeneity issues regarding the effect of interface on number of interactions. Other factors might have stronger influence on the dependent variable, but the thumb interface still significantly increases the number of interaction by 20%.

Future research should aim at replicating our results and should investigate the reasons behind the observed behaviour in order to inform designers. Furthermore, future research could investigate if these findings also apply to other voting context than class interactions, such as computer mediated elections, where the number of votes could be potentially increased. Moreover, future research could be extended to include other shapes of icons (e.g., stars, hearts) or colors, sizes, placements.

As with all persuasive technology, acting as choice architect and nudging behavior is a delicate endeavour where there is a fine line between improving interaction and manipulating users. Designers should thus assess and potentially mitigate the issue if deemed necessary (e.g., disclosing intent, requesting consent).

References

[1] Alonso, O., and Kandylas, V. A study on placement of social buttons in web pages. *arXiv preprint arXiv:1410.2828* (2014).

[2] Bangor, A., Kortum, P. T., and Miller, J. T. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction 24*, 6 (2008), 574–594.

[3] BBC. 'dislike' button coming to facebook, 2015.

[4] Bobadilla, J., Ortega, F., Hernando, A., and Gutiérrez, A. Recommender systems survey. *Knowledge-Based Systems 46* (2013), 109–132.

[5] Business Insider. What is a facebook 'like' actually worth in dollars?, 2013.

[6] Coetzee, D., Fox, A., Hearst, M. A., and Hartmann, B. Should your mooc forum use a reputation system? In *CSCW'14*, ACM (2014), 1176–1187.

[7] Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., and Riedl, J. Is seeing believing?: how recommender system interfaces affect users' opinions. In *CHI'03*, ACM (2003), 585–592.

[8] D'alessio, D. Use of the world wide web in the 1996 us election. *Electoral Studies 16*, 4 (1997), 489–500.

[9] Dearman, D., and Truong, K. N. Why users of yahoo!: answers do not answer questions. In *CHI'10*, ACM (2010), 329–332.

[10] Dreamgrow. Top 10 social networking sites by market share of visits [august 2015], 2015.

[11] Fogg, B. J. Persuasive technology: Using computers to change what we think and do. *Ubiquity 2002* (Dec. 2002).

[12] Gerlitz, C., and Helmond, A. The like economy: Social buttons and the data-intensive web. *New Media & Society 15*, 8 (2013), 1348–1365.

[13] Gerlitz, C., Helmond, A., et al. Hit, link, like and share. organising the social and the fabric of the web. In *DMI'11* (2011), 1–29.

[14] Holzer, A., Govaerts, S., Ondrus, J., Vozniuk, A., Rigaud, D., Garbinato, B., and Gillet, D. Speakup–a mobile app facilitating audience interaction. In *ICWL'13*. 2013, 11–20.

[15] Holzer, A., Govaerts, S., Vozniuk, A., Kocher, B., and Gillet, D. Speakup in the classroom: anonymous temporary social media for better interactions. In *CHI'14 EA*, ACM (2014), 1171–1176.

[16] Jin, X., Wang, C., Luo, J., Yu, X., and Han, J. Likeminer: a system for mining the power of'like'in social media networks. In *KDD'11*, ACM (2011), 753–756.

[17] Lampe, C. The role of reputation systems in managing online communities. *H. Masum, M. Tovey, eds* (2012), 77–88.

[18] Le, B. *Familiarity Principle of Attraction*, 0 ed. SAGE Publications, Inc., 2009, 597–598.

[19] Lehmann, J., Lalmas, M., Yom-Tov, E., and Dupret, G. Models of user engagement. In *User Modeling, Adaptation, and Personalization*. Springer, 2012, 164–175.

[20] New York Times. Putting a bolder face on google, 2009.

[21] PCWorld. Google wants you to 'like' +1: But will you?, 2011.

[22] Reddit. What do you think the up/down arrows mean?

[23] Techcrunch. Facebook activates "like" button; friendfeed tires of sincere flattery, 2009.

[24] Thaler, R. H., and Sunstein, C. R. *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press, 2008.

[25] The Guardian. What is the 1% rule?, 2006.

[26] Tsikerdekis, M. Dynamic voting interface in social media: Does it affect individual votes? In *Springer SOFSEM'13*. 2013, 552–563.

[27] Turnbull, D. Rating, voting & ranking: designing for collaboration & consensus. In *CHI'07 EA*, ACM (2007), 2705–2710.