# To Switch or Not To Switch: Understanding Social Influence in Online Choices

## Haiyi Zhu[1] and Bernardo A. Huberman[2]

## Abstract

The authors designed and ran an experiment to measure social influence in online recommender systems, specifically, how often people's choices are changed by others' recommendations when facing different levels of confirmation and conformity pressures. In this experiment, participants were first asked to provide their preference from pairs of items. They were then asked to make second choices about the same pairs with knowledge of other people's preferences. The results show that other people's opinions significantly sway people's own choices. The influence is stronger when people are required to make their second decision sometime later (22.4%) rather than immediately (14.1%). Moreover, people seem to be most likely to reverse their choices when facing a moderate, as opposed to large, number of opposing opinions. Finally, the time people spend making the first decision significantly predicts whether they will reverse their decisions later on, whereas demographics such as age and gender do not. These results have implications for consumer behavior research as well as online marketing strategies.

## Keywords

social influence, social media and choices, conformity theory

## Introduction

Picture yourself shopping online. You already have an idea about what product you are looking for. After navigating through the website, you find that particular item as well as several similar items and other people's opinions and preferences about them provided by the recommendation system. Will other people's preferences reverse your own?

[1]Carnegie Mellon University, Pittsburgh, PA, USA
[2]HP Laboratories, Palo Alto, CA, USA

**Corresponding Author:**
Haiyi Zhu, Carnegie Mellon University, Pittsburgh, PA, USA.
Email: haiyiz@cs.cmu.edu

Notice that in this scenario, there are two contradictory psychological processes at play. On one hand, when learning of other people's opinions, people tend to select those aspects that confirm their own existing ones. Prior literature suggests that once one has taken a position on an issue, one's primary purpose becomes defending or justifying that position (e.g., Nickerson, 1998). From this point of view, if the recommendations of others contradict our own personal opinions, we tend to not take this information into account and stick to our own choices. But research on social influence and conformity theory (Cialdini & Goldstein, 2004) suggests that even when not directly, personally, or publicly chosen as the target of others' disapproval, individuals may choose to conform to others and reverse their own opinions in order to restore their sense of belonging and self-esteem.

To investigate whether online recommendations can sway people's own opinions, we designed an online experiment to test how often people's choices are reversed by others' preferences when facing different levels of confirmation and conformity pressures. We used Rankr (Luon, Aperjis, & Huberman, 2012) as the study platform, which provides a lightweight and efficient way to crowdsource the relative ranking of ideas, photos, or priorities through a series of pairwise comparisons. In our experiment, participants were first asked to provide their preferences between pairs of photos. Then, they were asked to make a second choice about the same pairs with the knowledge of others' preferences. To measure the pressure to confirm people's own opinions, we manipulated the time between the participants' two decisions about the same pair of photos. To determine the effects of social pressure, we manipulated the number of opposing opinions that the participants saw when making the second decision. Finally, we tested whether other factors (i.e., age, gender, and decision time) affect the tendency to revert.

Our results show that other people's opinions significantly sway choices. The influence is stronger when people are required to make their second decision later (22.4%) rather than immediately (14.1%) after their first decision. Furthermore, people are most likely to reverse their choices when facing a moderate number of opposing opinions. Last but not least, the time people spend making the first decision significantly predicts whether they will reverse their decisions later on, whereas demographics such as age and gender do not.

The main contribution of this article is that we designed and ran an experiment to understand the mechanisms of social influence in online recommender systems. Specifically, we measured the effect of others' preferences on people's own choices under different conditions. The results have implications for consumer behavior research and online marketing strategies.

## Related Work

*Confirming existing opinions.* Confirmation of existing opinions is a long-recognized phenomenon (Nickerson, 1998). As Francis Bacon (1939) stated several centuries ago,

> The human understanding when it has once adopted an opinion (either as being received opinion or as being agreeable to itself) draws all things else to support and agree with it.

> Although there be a greater number and weight of instances to be found on the other side, yet these it either neglects and despises, or else by some distinction sets aside and rejects. (p. 36)

This phenomenon (often referred to as confirmation bias) can be explained by Festinger's (1957) dissonance theory: As soon as individuals adopt a position, they favor consistent over inconsistent information to avoid dissonance.

A great deal of empirical studies supports this idea (see Nickerson, 1998, for a review). Many of these studies use a task invented by Wason (1960), in which people are asked to find the rule that was used to generate specified triplets of numbers. The experimenter presents a triplet, and the participant hypothesizes the rule that produced it. The participants then test the hypothesis, by suggesting additional triplets and being told whether it is consistent with the rule to be discovered. Results show that people typically test hypothesized rules by producing only triplets that are consistent with the hypotheses, indicating hypothesis-determined information seeking and interpretation. Confirmation of existing opinions also contributes to the phenomenon of belief persistence. Ross, Lepper, and Hubbard (1975) showed that once a belief or opinion has been formed, it can be very resistant to change, even after learning that the data on which the beliefs or opinions were originally based were fictitious.

*Social conformity.* In contrast to confirmation theories, social influence experiments have shown that often people change their own opinion to match others' responses. The most famous experiment examining this is Asch's (1956) line-judgment conformity experiments. In the series of studies, participants were asked to choose which of a set of three disparate lines matched a standard, either alone or after 1 to 16 confederates had first given a unanimous incorrect answer. Meta-analysis showed that, on average, 25% of the participants conformed to the incorrect consensus (Bond & Smith, 1996). Moreover, the conformity rate increased with the number of unanimous majority. More recent, Cosley, Lam, Albert, Konstan, and Riedl (2003) conducted a field experiment on a movie rating site. They found that by showing manipulated predictions, users tended to rate movies toward the shown prediction. Researchers have also found that social conformity leads to multiple macro-level phenomena, such as group consensus (Asch, 1956), inequality and unpredictability in markets (Salganik, Dodds, & Watts, 2006), unpredicted diffusion of soft technologies (Bendor, Huberman, & Wu, 2009), and undermined group wisdom (Lorenz, Rauhut, Schweitzer, & Helbing, 2011).

Latané (1981) proposed a theory to quantitatively predict how the effect of social influence will increase as a function of the size of the influencing social source. The theory states that the relationship between the effect of the social influence (I) and the size of the influencing social source (N) follows a negative accelerating power function, $I = sN^t, 0 < t < 1$ (Latané, 1981). The theory has been empirically supported by a meta-analysis of conformity experiments using Asch's line-judgment task (Bond & Smith, 1996).

There are informational and normative motivations underlying social conformity, the former based on the desire to form an accurate interpretation of reality and behave

correctly, and the latter based on the goal of obtaining social approval from others (Cialdini & Goldstein, 2004). However, the two are interrelated and often difficult to disentangle theoretically as well as empirically. In addition, both goals act in service of a third underlying motive to maintain one's positive self-concept (Cialdini & Goldstein, 2004).

Both self-confirmation and social conformity are extensive and strong and they appear under many guises in life, both online and in physical interactions. In what follows, we consider both processes to understand the users' reactions to online recommender systems.

*Online recommender systems.*  Compared to traditional sources of recommendations—peers such as friends and coworkers, experts such as movie critics, and industrial media such as Consumer Reports—online recommender systems combined personalized recommendations sensitive to people's interests and independently reporting other peoples' opinions and reviews. One popular example of a successful online recommender system is the Amazon product recommender system (Linden, Smith, & York, 2003).

*Understanding users in online recommender systems.* In computer science and the human–computer interaction (HCI) community, for a long time, most research in recommender systems has focused on creating accurate and effective algorithms (e.g., Breese, Heckerman, & Kadie, 1998). Recently, researchers have realized that recommendations generated by standard accuracy metrics, although generally useful, are not always the most useful to users (McNee, Riedl, & Konstan, 2006). Researchers started building new user-centric evaluation metrics (Pu, Chen, & Hu, 2011; Xiao & Benbasat, 2007). There are few empirical studies investigating the basic psychological processes underlying the interaction of users with recommendations, and none of them addresses both self-confirmation and social conformity. As mentioned above, Cosley and his colleagues (2003) studied conformity in movie rating sites and showed that people's ratings are significantly influenced by other users' ratings. But they did not consider the effects of self-confirmation or the effects of different levels of social conformity pressures. Schwind, Buder, and Hesse (2011) studied how to overcome users' confirmation bias by providing preference-inconsistent recommendations. However, they represented recommendations as search results rather than recommendations from humans and thus did not investigate the effects of social conformity. Furthermore, their task was more related to logical inference rather than purchase decision making.

In the area of marketing and customer research, studies about the influence of recommendations are typically subsumed under personal influence and word-of-mouth research (Senecal & Nantel, 2004). Past research has shown that word-of-mouth plays an important role in consumer buying decisions, and use of the Internet brings new threats and opportunities for marketing (Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004; Senecal & Nantel, 2004; Stauss, 1997). There were several studies specifically investigating social conformity in product evaluations (Burnkrant & Cousineau, 1975; Cohen & Golden, 1972; Pincus & Waters, 1977). Although they found substantial

effects of others' evaluations on people's own judgments, the effects were not always significantly stronger when the social conformity pressures are stronger.[1] In Burnkrant and Cousineau's (1975) and Cohen and Golden's (1972) experiments, participants were exposed to evaluations of coffee with high uniformity or low uniformity. Both results showed that participants did not exhibit significantly increased adherence to others' evaluation in the high uniformity condition (although in Burnkrant and Cousineau's experiments, the participants recognized that the difference between high and low uniformity was significant). On the other hand, in Pincus and Waters's (1977) experiments (college students rated the quality of one paper plate while exposed to simulated quality evaluations of other raters), it was found that conformity effects are stronger when the evaluations are more uniform.

In summary, although previous research showed that others' opinions can influence people's own decisions, none of that research addresses both the self-confirmation and social conformity mechanisms that underlie choice among several recommendations. In addition, regarding the effects of increasing social conformity pressures, experiments using Asch's line-judgment tasks supported that people are more likely to be influenced when facing stronger social pressures, whereas the findings of studies using product evaluation tasks were mixed.

Our experiments address how often people reverse their own opinions when confronted with other people's preferences, especially when facing different levels of confirmation and conformity pressures. The hypothesis is that people are more likely to reverse their opinions when the reversion causes less self-inconsistency (the confirmation pressure is weaker) or the opposing social opinions are stronger (the conformity pressure is stronger).

## Method

### Experimental Design

We conducted a series of online experiments. All participants were asked to go to the website of Rankr (Luon et al., 2012) to make a series of pairwise comparisons with or without knowing other people's preferences (Figure 1). The pictures were collected from Google Images. We wanted to determine whether people reverse their choices by seeing others' preferences.

*Basic idea of the experiment.* Participants were asked to provide their preferences between the same pair of items twice. The first time the participant encountered the pair, he or she made a choice without the knowledge of others' preferences. The second time the participant encountered the same pair, he or she made a choice with the knowledge of others' preferences. Social influence was measured according to whether or not people switched their choices between the first time and second time they encountered the same pair.

To manipulate the pressure to confirm people's own opinions, we changed the time between the two decisions. In the short interval condition, people first compared two

**Figure 1.** Example of pairwise comparisons in Rankr.

pictures on their own and they were then immediately asked to make another choice with available information about others' preferences. When their memories were fresh, reversion led to strong inconsistency and dissonance with other people's choices (strong confirmation pressure). However, in the long interval condition, participants compared pairs of items in the beginning of the test followed by several distractor pairs. This was followed again by the same pairs that the participant had previously compared, but this time with augmented information about others' preferences. In this case, participants' memories of their previous choices decay, so the pressure to confirm their own opinions is less explicit.

To manipulate the social pressure, we changed the number of opposing opinions that the participants saw when making the second decision. We selected four levels: the opposing opinions were 2 times, 5 times, 10 times, or 20 times as many as the number of people who supported their opinions.

In the following section, the details of experimental conditions are discussed.

*Conditions.* The experimental design was 2 (baby pictures and loveseat pictures) × 3 (short interval, long interval, and control) × 4 (ratio of opposing opinions to supporting opinions: 2:1, 5:1, 10:1, and 20:1). Participants were recruited from Amazon's Mechanical Turk (mTurk) and were randomly assigned into one of six conditions (baby-short, baby-long, baby-control, loveseat-short, loveseat-long, and loveseat-control) and made four choices with different levels of conformity pressure.

In the baby condition, people were asked to compare 23 or 24 pairs of baby pictures by answering the question, "Which baby looks cuter on a baby product label?" Note that the Caucasian baby pictures in Figure 1 are examples. We displayed baby pictures from different races in the experiment. In the loveseat condition, the question was, "Your close friend wants your opinion on a loveseat for their living room. Which one do you suggest?" People needed to make 23 or 24 choices.

In the short interval condition, people first compared two pictures on their own and they were then immediately asked to make another choice with available information
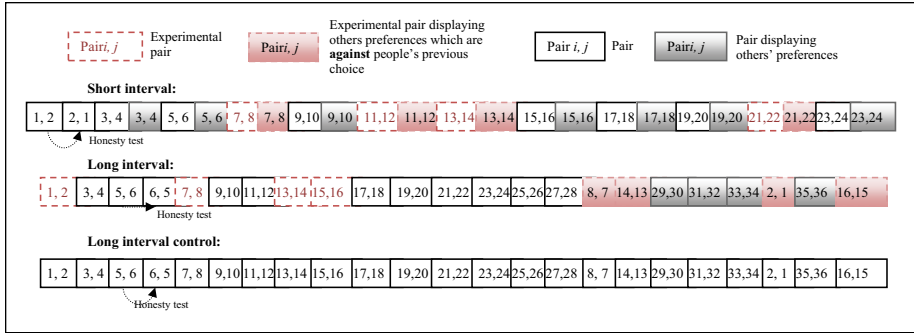
Figure 2 legend boxes:

Pair *i, j* — Experimental pair

Pair *i, j* — Experimental pair displaying others preferences which are **against** people's previous choice

Pair *i, j* — Pair

Pair *i, j* — Pair displaying others' preferences

**Short interval:**

| 1, 2 | 2, 1 | 3, 4 | 3, 4 | 5, 6 | 5, 6 | 7, 8 | 7, 8 | 9,10 | 9,10 | 11,12 | 11,12 | 13,14 | 13,14 | 15,16 | 15,16 | 17,18 | 17,18 | 19,20 | 19,20 | 21,22 | 21,22 | 23,24 | 23,24 |

Honesty test

**Long interval:**

| 1, 2 | 3, 4 | 5, 6 | 6, 5 | 7, 8 | 9,10 | 11,12 | 13,14 | 15,16 | 17,18 | 19,20 | 21,22 | 23,24 | 25,26 | 27,28 | 8, 7 | 14,13 | 29,30 | 31,32 | 33,34 | 2, 1 | 35,36 | 16,15 |

Honesty test

**Long interval control:**

| 1, 2 | 3, 4 | 5, 6 | 6, 5 | 7, 8 | 9,10 | 11,12 | 13,14 | 15,16 | 17,18 | 19,20 | 21,22 | 23,24 | 25,26 | 27,28 | 8, 7 | 14,13 | 29,30 | 31,32 | 33,34 | 2, 1 | 35,36 | 16,15 |

Honesty test

**Figure 2.** Example displaying orders in each condition.

about others' preferences. Furthermore, we tested whether people would reverse their first choice under four levels of social pressure: when the number of opposing opinions was 2 times, 5 times, 10 times, and 20 times as many as the number of people who supported their opinions. The numbers were randomly generated.[2] In addition to these 8 experimental pairs, we added 14 noise pairs and an honesty test composed of 2 pairs (24 pairs in total; see Figure 2 for an example). In this condition, noise pairs also consisted of consecutive pairs (a pair with social information immediately after the pair without social information). However, others' opinions were either indifferent or in favor of the participants' choices. We created an honesty test to identify participants who cheated the system and quickly clicked on the same answers. The test consisted of 2 consecutive pairs with the same items but with the positions of the items exchanged. Participants needed to make the same choices among these consecutive 2 pairs in order to pass the honesty test. The relative orders of experimental pairs, noise pairs, and honesty test in the sequence and the items in each pair were randomly assigned to each participant.

In contrast with the short interval condition, where people were aware that they reversed their choices, in the long interval condition, we manipulated the order of display and the item positions so that the reversion was less explicit. People first compared pairs of the items without knowing others' preferences, and then after 11.5 pairs later, on average, we showed the participants the same pair (with the positions of items in the pair exchanged) and others' opinions. Similarly, with the short interval condition, we showed 8 experimental pairs to determine whether people reversed their previous choices with increasing pressures of social influence. In addition, we showed 13 noise pairs (9 without others' preferences and 4 with others' preferences) and performed an honesty test (see Figure 2 for an example).

By increasing the time between two choices, we blurred the people's memories of their choices so as to exert a subtle confirmation pressure. However, as people proceeded with the experiment, they were presented with new information to process. This new information might lead them to think in a different direction and change their own opinions regardless of social influence. To control for this confounding factor, we

added a long interval control condition, where the order of the pairs was the same as with the long interval condition but without showing the influence of others.

*Procedures.* We conducted our experiment on Amazon's mTurk (Kittur, Chi, & Suh, 2008). The recruiting messages stated that the objective of the study was to do a survey to collect people's opinions. Once mTurk users accepted the task, they were asked to click the link to Rankr, which randomly directed them to one of the six conditions. This process was invisible to them.

First, the participants were asked to provide their preferences about 23 or 24 pairs of babies or loveseats. They were then directed to a simple survey. They were asked to report their age and gender and answer two 5-Likert scale questions. The questions were as follows: "Is showing others' preferences useful to you?" and "How much does showing others' preferences influence your response?" After filling out the survey, a unique confirmation code was generated and displayed on the webpage. Participants needed to paste the code back to the mTurk task. We matched mTurk users with the participants of our experiments using the confirmation code, allowing us to pay mTurk users according to their behaviors. We paid $0.35 for each valid response.

*Participants.* We collected 600 responses. Of this number, we omitted 37 responses from 12 people who completed the experiment multiple times; 22 incomplete responses; 1 response from someone who did not conform to the participation requirements (i.e., being at least 18 years old); and 107 responses from those who did not pass the honesty test. These procedures left 433 valid participants in the sample—about 72% of the original number. According to participant self-reporting, 40% were women; ages ranged from 18 to 82 years with a median age of 27 years. Geocoding[3] the IP addresses of the participants revealed that 57% were from India, 25% were from the United States, and the remaining 18% of participants came from more than 34 different countries.

The numbers of participants in each condition were as follows: baby-short, 72; baby-long, 91; baby-control, 49; loveseat-short, 75; loveseat-long, 99; and loveseat-control, 47.[4]

People spent a reasonable amount of time on each decision (average = 6.6 seconds; median = 4.25 seconds).

Among the 433 responses, 243 left comments in the open-ended comments section at the end of the experiments. Most of them said that they had a good experience when participating in the survey. (They were typically not aware that they were in an experiment.)

*Measures.* The measures are as follows:

1) Reversion: whether people reverse their preferences after knowing others' opinions
2) Social conformity pressures: the ratio of opposing opinions to supporting opinions
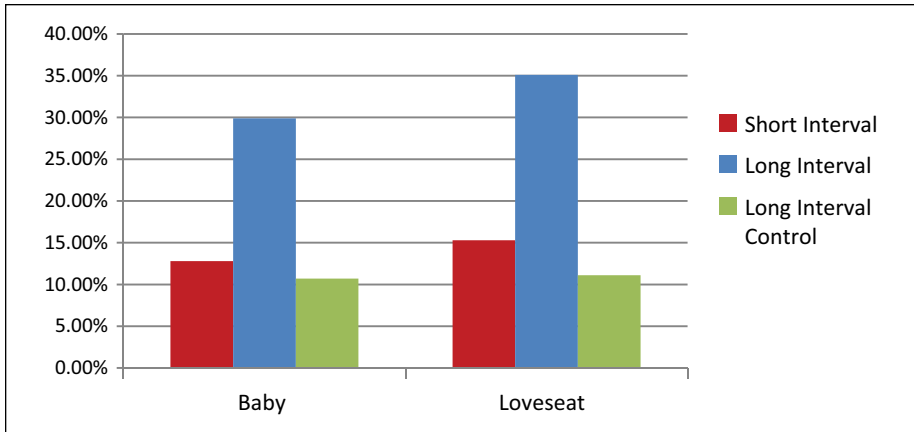3) Decision time: the time (in seconds) people spent in making each decision

**Figure 3.** Reversion rate by conditions.

4)  Demographic information: age and gender
5)  Self-reported usefulness of others' opinions
6)  Self-reported level of being influenced

## Results

### 1. Did people reverse their opinions by others' preferences when facing different confirmation pressures?

Figure 3 shows the reversion rate as a function of the conditions that we manipulated in our experiment. First, we found out that content does not matter, that is, although baby pictures are more emotionally engaging than loveseat pictures, the patterns are the same. The statistics test also shows that there is no significant difference between the baby and the loveseat results, $t(431) = 1.35$, $p = .18$.

Second, in the short interval condition, the reversion rate was 14.1%, which is higher than zero [the results of the $t$ test are $t(146) = 6.7$, $p < .001$].

Third, the percentage of people who reversed their opinions was as high as 32.5% in the long interval condition, significantly higher than the long interval control condition (10.1%), which measures the effects of other factors leading to reversion regardless of the social influence during the long interval. $t$ test shows that this difference is significant, $t(284) = 6.5$, $p < .001$. We can therefore conclude that social influence contributes to approximately 22.4% of the reversion of opinions observed.

To summarize the results, in both the long and the short interval conditions, others' opinions significantly swayed people's own choices (22.4% and 14.1%[5]). The effect size of social influence was larger when the self-confirmation pressure was weaker (i.e., the time between the two choices is larger).

**Table 1.** Linear Regression Predicting the Reversion Percentage.

| Predictor | Coefficient | Standard Error | P Value |
|---|---|---|---|
| Condition[a] | .839 | .059 | < .001 |
| Ratio of opposing opinions | .563 | .184 | .038 |
| Square ratio of opposing opinions | −.142 | .049 | .045 |
| Intercept | −2.42 | .151 | < .001 |
| Adjusted R2 | | .96 | |

*Note.* The squared ratio of opposing opinions has a significant negative value (–0.142, p = .045).
[a]1 = long interval; 0 = short interval.

## 2. Were people more likely to reverse their own preferences when more people were against them?

It is interesting that we saw an increasing and then decreasing trend when the opposing opinions became exponentially stronger (from 2×, 5×, 10×, to 20×). The condition with the most uniform opposing opinions (20×) was not more effective in reversing people's own opinions than the moderate opposing opinions (5× and 10×). The regression results are shown in Table 1. Note that the squared ratio of opposing opinions has a significant negative value (Coef. = −0.142, $p < .05$), suggesting that the returning effect is statistically significant.

These results might be explained by Brehm's (1966) finding of psychological reactance. According to Brehm, if an individual perceives one's freedom as being reduced or threatened with reduction, one will become aroused to maintain or enhance one's freedom. The motivational state of arousal to reestablish or enhance one's freedom is called psychological reactance. Therefore, if the participants perceived the uniform opposing opinions as a threat to their freedom to express their own opinions, their psychological reactance might be aroused to defend and confirm their own opinions.

These results can also be explained in terms of Wu and Huberman's (2010) findings about online opinion formation. In their work, they used the idea of maximizing the effect that individuals have on the average rating of items to explain the phenomenon that later reviews tend to show a big difference from earlier reviews on Amazon.com and IMDB.com.

We can use the same idea to explain our results. Social influence in product recommendations is not just a one-way process. People are not just passively influenced by others' opinions but also want to maximize their effect on other people's future decision making (e.g., in our experiments, according to our recruiting messages, participants would assume that their choices would be recorded in the database and shown to others; in real life, people like to influence their friends and family). We assume that the influence of an individual on others can be measured by how much his or her expression will change the average opinion. Suppose there are $X_1$ supporting opinions and $X_2$ opposing opinions and that $X_2 > X_1$. A person's choice $c$ (0 indicates confirming his or her

**Table 2.** Logistic Regression Predicting the Reversion.

| Predictor | Coefficient | Standard Error | P Value |
| --- | --- | --- | --- |
| Condition[b] | 1.26 | .152 | < .001 |
| Age | −.006 | 6.89e-3 | .407 |
| Gender | .067 | .143 | .642 |
| Self-reported usefulness | .164 | .070 | .020 |
| Self-reported influence level | .334 | .072 | < .001 |
| Standardized first decision time | .323 | .065 | < .001 |
| Log likelihood | | −657.83 | |

[b]1 = long interval; 0 = short interval.

own choice; 1 indicates conforming to others) can move the average percentage of opposing opinions from $X_2 / (X_1 + X_2)$ to $(X_2 + c)/(X_1 + X_2 + 1)$. So, the influence on the average opinion is $\left| \dfrac{X_2}{X_1 + X_2} - \dfrac{X_2 + c}{X_1 + X_2 + 1} \right|$. A simple derivation shows that to maximize the influence on average opinion, people need to stick to their own choices and vote for the minority. Then, their influence gain will be stronger when the difference between existing majority opinions and minority ones is larger. Therefore, the motivation to exert influence on other people can play a role in resisting the social conformity pressure and lead people to confirm their own decisions, especially when facing uniform opposing opinions.

### 3. What else predicts the reversion?

We used a logistic regression model to predict the decision-level reversion with the participants' age, gender, self-reported usefulness of the recommendation system, self-reported level of being influenced by the recommendation system, and standardized first decision time (as shown in Table 2). Note that standardized first decision time = (time in this decision – this person's average decision time) / this person's standard deviation. So, "first decision time" is an intrapersonal variable.

The results showed that age and gender do not significantly predict reversion ($p$ = .407, $p$ = .642). Self-reported influence level has a strong prediction power (Coef. = 0.334, $p$ < .001), which is reasonable. The interesting fact is that decision time, a simple behavioral measure, also predicts reversion very well (Coef. = 0.323, $p$ < .001). The longer people spent on the decisions, the more equivalent the two choices were for them. According to Festinger's (1954) theory, the more equivocal the evidence, the more people rely on social cues. Therefore, the more time people spend on a choice, the more likely they are to reverse this choice and conform to others later on.

## Discussion

On one hand, the phenomenon we found (i.e., the returning effects of strong influence pressure) is quite different from the classic line-judgment conformity studies (Asch,

1956; Latané, 1981). Notice that these experiments used questions with only one single correct answer (Asch, 1956). In contrast, in our experiment we examined the social influence on people's subjective preferences among similar items, which might result in such a different phenomenon. On the other hand, our findings reconcile the mixed results of studies investigating social conformity in product evaluation tasks (which are often subjective tasks) (Burnkrant & Cousineau, 1975; Cohen & Golden, 1972; Pincus & Waters, 1977). Due to the returning effects, strong conformity pressure is not always more effective in influencing people's evaluations compared to weak conformity pressure.

In addition, in our experiment we did not look at social influence in scenarios where the uncertainty level is quite high. For example, in settings such as searching recommendations for restaurants or hotels where people have never been, the level of uncertainty is high and people need to rely on other cues. However, in our experimental setting, people can confidently make their choice by comparing the pictures of the settings. Therefore, some caution is needed when trying to generalize our results to other settings with high uncertainty.

Regarding the tasks used in the experiment, the choice of "which baby looks cuter on a product label" involves emotions and subjective feelings. Alternatively, similar choices include the preferences of iTunes music or YouTube's commercial videos. In contrast, the other type of question, "which loveseat is better," is less emotionally engaging. In this case, people are more likely to consider usability factors, such as color and perceived comfort, when making these types of choices. Results showed that the effect of social influence on these two different types of choices is similar and consistent, which suggests the general applicability of the results.

There might be concern about cultural differences between the global nature of the participants and the Western-style nature of the tasks. We point out that this intercultural preference difference is inherent in the participant. Therefore, it is assumed to be consistent throughout the test and should not affect a particular person's preference changes and will not confound the results.

During our research, we invented an experimental paradigm to easily measure the effects of social influence on people's decision making (i.e., the reversion) and manipulate conditions under which people make choices. This paradigm can be extended to scenarios beyond those of binary choices, to the effect of recommendations from friends as opposed to strangers, and to social influence, which could be varied with different visualizations for the recommendations (e.g., displaying the average ratings from all the users, or showing the ratio of opposed opinions, or showing only the number of positive reviews such as "likes").

## Limitations & Future Work

In our experiments, we examined whether people reverse their choices when facing different ratios of opposing opinions versus supporting opinions (2×, 5×, 10×, and 20×). To further investigate the relationship between the ratio of the opposing opinions and the tendency to revert, it would be better to include more fine-grained conditions in

the ratio of opposing opinions. The ideal situation would be a graph with the continuous opposing versus supporting ratio as the x-axis and the reversion rate as the y-axis.

Also, additional manipulation checks or modification of the design of the experiment would be needed to establish whether processes such as psychological reactance or the intent to influence others have been operating. For example, we could measure the degree of perceived freedom in a given task. In addition, it would be revealing to manipulate the visibility of people's choices to other participants to see if the intent of influencing others exists.

Regarding the usage of mTurk as a new source of experimental data, we agree with Paolacci, Chandler, and Ipeirotis (2010) that "workers in Mechanical Turk exhibit the classic heuristics and biases and pay attention to directions at least as much as subjects from traditional sources" (p. 417). In particular, compared to recruiting people from a college campus, we believe that the use of mTurk has a lower risk of introducing the Hawthorne Effect (i.e., people alter their behaviors due to the awareness that they are being observed in experiments), which is itself a form of social influence and might contaminate our results.

In our experiment, we used several methods, such as an honesty test and IP address check, to further ensure that we collected earnest responses from mTurk workers. The average time they spent on the task, the statistically significant results of the experiment, and the comments that participants left all indicate that our results are believable. However, there is still a limitation to our honesty test. On one hand, the presentation of two consecutive pairs with the positions of items switched was unable to identify all the users who tried to cheat the system by randomly clicking on the results, which added noise to our data. On the other hand, the honesty test might also exclude some earnest responses. Thus, it is possible that immediately after people made a choice, they regretted it.

## Conclusion & Implications

In this article, we presented results of a series of online experiments designed to investigate whether online recommendations can sway people's own opinions. These experiments exposed participants making choices to different levels of confirmation and conformity pressures.

Our results show that people's own choices are significantly swayed by the perceived opinions of others. The influence is weaker when people have just made their own choices. In addition, we showed that people are most likely to reverse their choices when facing a moderate, as opposed to large, number of opposing opinions. And last but not least, the time people spend making the first decision significantly predicts whether they will reverse their own later on.

Our results have three implications for consumer behavior research as well as online marketing strategies. (a) The temporal presentation of the recommendation is important; it will be more effective if the recommendation is not provided immediately after the consumer has made a similar decision. (b) The fact that people can reverse their choices when presented with a moderate countervailing opinion suggests that rather

than overwhelming consumers with strident messages about an alternate product or service, a more gentle reporting of a few people having chosen that product or service can be more persuasive than stating that thousands have chosen it. (c) Equally important is the fact that a simple monitoring of the time spent on a choice is a good indicator of whether or not that choice can be reversed through social influence. There is enough information in most websites to capture these decision times and act accordingly.

## Declaration of Conflicting Interests

## Funding

## Notes

1. Unlike other conformity experiments such as line-judgment where the pressure of social conformity is manipulated by increasing the number of "unanimous" majority, experiments about social influence in product evaluation (Burnkrant & Cousineau, 1975; Cohen & Golden, 1972; Pincus & Waters, 1977) usually manipulate the pressure of social influence by changing the degree of uniformity of opinions. As discussed in Cohen and Golden (1972), since it is seldom that no variation exists in the advice or opinions in reality, the latter method is more likely to stimulate participants' real reactions. We also use the latter method in our experiment by manipulating the ratio of opposing opinions versus supporting opinions.
2. We first generated a random integer from 150 to 200 as the total participants. Then, we generated the number of people holding different opinions according to the ratio. Here are a few examples: 51 versus 103 (2×), 31 versus 156 (5×), 16 versus 161 (10×), and 9 versus 181 (20×).
3. MaxMind GeoLite (http://dev.maxmind.com/geoip/legacy/geolite/) was used to geocode the IP addresses and self-reports a 99.5% accuracy rate.
4. Among the 600 responses, originally 20% were assigned for baby-strong; 20% for baby-weak; 10% for baby-control; 20% for loveseat-strong; 20% for loveseat-weak; and 10% for loveseat-control. The valid responses in short interval conditions were fewer than the ones in long interval conditions because the short interval condition had a higher failure rate in the honesty test. The reason might be that the short interval condition had more repetitive pairs, fewer new items, and more straightforward patterns, leading to boredom and casual decisions, which in turn caused failure in the honesty tests.
5. To calibrate the magnitude of our results, we point out that our results are of the same magnitude as the classic line-judgment experiments. According to a 1996 meta-analysis of line-judgment experiment consisting of 133 separate experiments and 4,627 participants, the average conformity rate is 25% (Bond & Smith, 1996).

## References

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs*, *70*(9), 1-70.

Bacon, F. (1939). Novum organum. In E. A. Burtt (Ed.), *The English philosophers from Bacon to Mill* (pp. 24-123). New York: Random House. (Original work published 1620)

Bendor, J., Huberman, B. A., & Wu, F. (2009). Management fads, pedagogies, and other soft technologies. *Journal of Economic Behavior & Organization*, *72*(1), 290-304.

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, *119*(1), 111-137.

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th conference on uncertainty in artificial intelligence*. San Francisco, CA: Morgan Kaufmann.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York, NY: Academic Press.

Burnkrant, R. E., & Cousineau, A. (1975). Informational and normative social influence in buyer behavior. *Journal of Consumer Research*, *2*(3), 206-215.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591-621.

Cohen, J. B., & Golden, E. (1972). Informational social influence and product evaluation. *Journal of Applied Psychology*, *56*(1), 54-59.

Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., & Riedl, J. (2003). Is seeing believing? How recommender system interfaces affect users' opinions. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '03)* (pp. 585-592). New York, NY: ACM.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117-140.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38-52.

Kittur, A., Chi, E. H., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the twenty-sixth annual SIGCHI conference on human factors in computing systems (CHI '08)* (pp. 453-456). New York, NY: ACM.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343-356.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *Internet Computing*, *7*(1), 76-80.

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *PNAS*, *108*(22), 9020-9025.

Luon, Y., Aperjis, C., & Huberman, B. A. (2012). Rankr: A mobile system for crowdsourcing opinions. In *Mobile computing, applications, and services* (pp. 20-31). Berlin: Springer.

McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Being accurate is not enough: How accuracy metrics have hurt recommender systems. *CHI Extended Abstracts*, pp. 1097-1101.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175-220.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411-419.

Pincus, S., & Waters, L. K. (1977). Informational social influence and product quality judgments. *Journal of Applied Psychology*, *62*(5), 615-619.

Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on recommender systems (RecSys '11)* (pp. 157-164). New York, NY: ACM.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, *32*, 880-892.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854-856.

Schwind, C., Buder, J., & Hesse, F. W. (2011). I will do it, but I don't like it: User reactions to preference-inconsistent recommendations. In *Proceedings of the 2011 annual conference on human factors in computing systems (CHI '11)* (pp. 349-352). New York, NY: ACM.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, *80*(2), 159-169.

Stauss, B. (1997). Global word of mouth. Service bashing on the Internet is a thorny issue. *Marketing Management*, *6*(3), 28-30.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology*, *12*, 129-140.

Wu, F., & Huberman, B. A. (2010). Opinion formation under costly expression. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *1*(1), 5.

Xiao, B., & Benbasat, I. (2007). Ecommerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly*, *31*(1), 137-209.

## Author Biographies

**Haiyi Zhu** is a PhD candidate in the Human-Computer Interaction Institute at Carnegie Mellon University. Her research interests include human-computer interaction, social computing, and online communities. She is specifically interested in conducting large-scale data analysis and field experiments to understand the underlying principles in online social computing systems. She received a BA in computer science at Tsinghua University in 2009 and a master's in human computer interaction from Carnegie Mellon University in 2012.

**Bernardo A. Huberman** is a Senior HP Fellow and director of the Social Computing Research Group at HP Labs, which focuses on methods for harvesting the collective intelligence of groups of people in order to realize greater value from the interaction between users and information. Huberman received his PhD in physics from the University of Pennsylvania and is currently a consulting professor in the Department of Applied Physics at Stanford University. He has been a visiting professor at the Niels Bohr Institute in Denmark, the University of Paris, and Insead, the European School of Business in Fontainebleau, France.