Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018)

Songs of a Future Past — An Experimental Study of Online Persuaders

Marzia Antenore
Communication and Social Research Department
Sapienza University of Rome
Via Salaria 113, 00198 Rome, Italy
marzia.antenore@uniroma1.it

Alessandro Panconesi, Erisa Terolli
Computer Science Department
Sapienza University of Rome
Via Salaria 113, 00198 Rome, Italy
{ale,terolli}@di.uniroma1.it

Abstract

In this paper, we present the results of an extensive experimental study on users decisions inside an online setting. In the experiment, participants purchase songs using real money while having enough time to explore them at leisure before buying. In such a set up, surpisingly, common social influence signals such as star ratings, download counts and recommendations had no influence. However, as soon as the exploration was made slightly more cumbersome market inequality appeared. This is an indication that it is decision-making shortcuts, rather then social influence, to trigger distorting market effects.

Introduction

In order to help users navigate the bewildering cornucopia of online choices, websites nowadays make use of sophisticated user interfaces. These platforms deploy a vast array of ergonomic digital cues, from simple ones such as star ratings, download counts, and number of likes/dislikes, to the tailor-made suggestions given by complex recommender systems. The conventional wisdom is that these online signals must surely be very effective, and a host of academic studies, several of which are reviewed in this article, support this belief. In this paper we discuss a lab experiment whose outcome is somewhat unexpected and counterintuitive when seen against this backdrop. In the experiment people were asked to explore a small collection of high-quality songs and download the two they liked the most. A rather unique feature of the experiment was that the songs came from a well-known online store and had a real cost of 0.99 euros each. Each participant was given a budget of 1.98 euros so that up to two songs could be bought. Users were put under the influence of two classes of signals. The first class consisted of common social influence cues such as star ratings, download counts and recommendations. The second was related to placement on the screen and consisted of two different types of cues. The first such cue was simply positioning on the screen, while the second was a "small amount" of information overload. This effect was induced by presenting the songs in a scrollable list, as opposed to a more ergonomic layout where all the songs were presented on the screen. The reason to commit to a small number of songs in the experiment was to avoid as much as possible interferences stemming from information overload (except when this feature was purposely introduced) and other effects. The outcome of this simple experiment was somewhat surprising. First, social influence signals were analysed along two different dimensions. On the one hand, somewhat unexpectedly, they had essentially zero effect as far as influencing user choices. On the other, while the success of a song remained the same whether or not social influence cues were used, user engagement, measured as the number of listens, increased. In other words, when participants were in the "virtual presence of others", they listened to the songs much more than in the other conditions. Second, position on the screen had a strong influence even if participants spent a lot of time exploring the songs. Likewise, even the modicum amount of information overload induced by a scrollable list had a significant impact on user choices.

In the remainder of this introduction we describe more precisely our experiment, motivations and findings. We begin by discussing the digital cues whose power to influence we have investigated.

Bandwagons, digital shop assistants and hidden persuaders. Online users are quite familiar with a vast array of icons representing the degree of success of an item. Star ratings, download counts, number of likes/dislikes, and several others, are commonly understood as indicators of popularity and have become widespread if not ubiquitous. Cues of this type try to take advantage of the so-called "bandwagon effect" which, as suggestively put by Go et al. (Go, Jung, and Wu 2014), operates implicitly under the motto

> *"if others think that something is good, then I should, too"*.

Seen in this light, the bandwagon effect can be considered as an instance of conformity, the well-known "tendency to align your attitudes, beliefs, and behaviors with those around you" (Crutchfield 1955). Conformity is a powerful form of social pressure whose systematic study goes back to (at least) the classic work in psychology of Solomon Asch (Asch 1951). Indeed, bandwagon cues in the online domain have been studied in different situations under a variety of conditions and are typically, if not invariably, reported to produce significant effects (Salganik, Dodds, and Watts 2006; Salganik and Watts 2008; Muchnik, Aral, and Taylor 2013; Van de Rijt et al. 2014; Chen 2008; Zhu and Huberman 2014; Abeliuk et al. 2017). In our study we picked star ratings and download counts as the champions of conformity-type cues.

In recent years recommender systems (henceforth RS's) have emerged as a very useful and effective addition to the landscape of "digital nudges". Before the advent of the internet, a regular customer of, say, a music shop, in order to overcome the problem of cognitive overloading due to the large number of products available, could seek the advice of a knowledgeable shop assistant with whom s/he had developed a relationship of trust. Based on the knowledge of the customer's taste and that of the music world, the shop assistant could offer insightful suggestions, providing welcome advice in a friendly environment. Recommender systems are, in a sense, the algorithmic analog of the shop assistants of yesteryear. On the basis of the past online behavior of the current customer and of the entire collective behavior of online visitors, they help navigate the huge catalogue of online choices by providing tailor-made suggestions in a purely algorithmic fashion. Thus, a visitor to the YouTube home site will be presented with a list of videos that, hopefully, will match his/her interests, while a person looking for a book on Amazon or a movie on Netflix will likewise see a list of other interesting items to buy or rent, presumably tailored on his/her needs and taste.

RS's can be seen as a further refinement of approaches based on the bandwagon effect in that they provide "bespoke nudges". Their suggestions are based not only on aggregate information mined from the online choices of the entire population of web users but also, crucially, on user profiling at the individual level. In particular, collaborative filtering makes suggestions based on the choice similarities between users. Their intended mission can thus be captured by the "improved" maxim

> *"if others like me think that something is good, then I should, too"*.

While the workings of RS's at the psychological level too can be seen as an example of conformity, they can also be understood in terms of another powerful psychological demand highlighted by classic studies in psychology – cognitive consistency. Starting from the seminal work of Leon Festinger the need to avoid so-called "cognitive dissonance" has been recognized as an important factor shaping people's choices (Festinger and Carlsmith 1959). In the case at hand, RS's can be seen not only to provide a push toward conformity but also, when successfully executed, as a skillful way to reconstruct a representation of a person's interests and tastes, so that the choice to be made – e.g the book to buy or the movie to download – becomes constrained to be consistent with one own's worldview. In our study we have implemented and deployed a simple collaborative filtering based recommendation, which is a widespread and very effective type of recommendation system (Adomavicius and Tuzhilin 2005; Resnick and Varian 1997). The intended mission of collaborative filtering is precisely that of providing

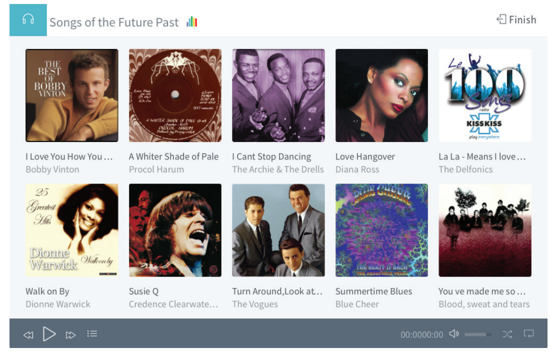


Figure 1: Control group layout.

suggestions based on the choices of "people like us" and for this reason we have chosen it as the champion for cognitive consistency (Linden, Smith, and York 2003).

The types of cues we have discussed so far fall under the umbrella of social influence. In our study we have compared them against another type of signals commonly deployed online– screen placement– and did this in two different ways. In all conditions but one, the ten songs were arranged in a flat layout assigning to each one of them one out of ten slots (See Figure 1). In this way we could investigate whether different slots put songs at an advantage or a disadvantage. In a different condition, the songs were organized in a scrollable list, in order to make their exploration slightly more cumbersome. The reason to do this was to be able to trigger decision-making shortcuts, which we discuss next.

Cognitive shortcuts. The idea that people resort to mental shortcuts, or heuristics, when confronted with complex decision-making tasks has wide currency in psychology and behavioural economics (Tversky and Kahneman 1974; Lewis 2012; Gigerenzer and Gaissmaier 2011; Hastie and Dawes 2010). When confronted with a bewildering array of similarly looking choices, rather than engaging in a process of systematic evaluation that would be very costly from a cognitive point of view, people often resort to mental shortcuts (Sundar, Knobloch-Westerwick, and Hastall 2007; Chaiken 1980). This poses a problem if one wants to ascertain the effects of social influence online. If our experimental platform created a situation of information overload, digital cues could trigger decision-making heuristics and it would become difficult to disentangle this effect from social influence. Therefore, in order to isolate its effects we have opted for a small number of songs. In this way, participants in the experiment had the possibility to systematically explore the choices before them and come to a conscious decision, avoiding decision-making shortcuts (Russo, Schoemaker, and Russo 1989). In such a scenario, will the deep-rooted psychological demands of conformity and cognitive consistency exert their power? And, likewise, will the subtle effects of placement continue to offer a discrete but powerful advantage?

Our results. We can now finally describe with some precision our findings. We have conducted a simple experiment under laboratory conditions in which participants could connect to a website to listen to a small number of songs (ten), rate them, and download the two they liked the most. Each download had a real cost of 0.99 euros and participants were given a budget to download two songs. As discussed, the small number of songs made it possible their systematic, conscious exploration so as to avoid the triggering of decision-making shortcuts. The role of money is to further facilitate such a conscious exploration– since downloading has a value people will choose carefully. Such a systematic exploration did take place. Participants spent an average of 16.68 minutes listening to the songs.

Participants were assigned to different scenarios, in each of which they were subject to a specific online cue or, in the case of the control group, none at all. As discussed, there were two classes of cues. Social influence, in particular conformity and cognitive consistency, was championed by star ratings, download counts and recommendations. While placement was championed by the flat layout and a scrollable list. The outcome of the experiment was the following:

- Social influence showed to have essentially zero effect in altering the market share of the songs.
- However, it increased user participation, measured as the number of listens.
- Placement had a noticeable effect on user choices.

The first outcome is somewhat surprising and stands in some contrast with the reported outcome of many studies. There is no direct contradiction since the conditions under which bandwagon cues have been studied are different, but it does indicate that their power to influence may be not as strong as commonly believed. The second effect too is unexpected. A possible interpretation of the results is that, in absence of choice overload, while it fails to influence market shares, social influence still increases market volume (in our case the number of listenings). The third results are not particularly surprising per se, but provide a useful benchmark against which social influence can be evaluated.

Literature Review

The study of hidden persuasion goes back at least to the classical and compelling work of Solomon Asch on conformity (Asch 1951) (see, for instance, (Aronson 1988)). With the advent of the Internet, the attention has shifted quite naturally to the online world. In what follows we will focus on what seems to be more directly relevant to our present work. An influential study in this context is MusicLab (Salganik, Dodds, and Watts 2006). It simulates an online cultural market where participants could access a web site via the internet (so that this was *not* a controlled lab experiment) containing a collection of 48 songs by new, emerging groups. Songs could be listened to, and any number of them rated and downloaded for free. So, this experiment lacks the fundamental ingredient of a market– money— that is one of the main ingredients of our experiment. Each participant was randomly assigned to one of two experimental conditions: the independent condition, or control group, and the social influence condition, in which they were given feedback in the form of download counts for each song. Seemingly in contrast with our findings, the social-influence scenario was recorded to present large market inequality and unpredictability. We will try to give a comprehensive explanation of this apparent discrepancy in Section "Interpretations".

Other web-based experiments in the same vein are (Salganik and Watts 2008), which analysed the impact of social influence on music, (Muchnik, Aral, and Taylor 2013) exploring news aggregation, (Van de Rijt et al. 2014) for human rewarding systems, (Chen 2008) for online book purchasing , (Zhu and Huberman 2014) for images, (Hou 2017) for food consumption and (Abeliuk et al. 2017) for scientific articles. All these studies show that social influence tend to create unequal and unpredictable markets. However, they share some common features that stand in opposition with the environment created by our setup, as discussed in Section "Interpretations".

Other papers limit themselves to study intent as opposed to actual behaviour. For instance, in some studies users are presented with the description of two products (books) and asked which one *they would buy* if given the opportunity (Chen 2008). Many studies confirm what is intuitively very clear. Namely that stated intention is not always (if ever) a precise indicator of actual behaviour (LaPiere 1934; Kutner, Wilkins, and Yarrow 1952; Sheeran 2002; Linn 1965; Ajzen, Brown, and Carvajal 2004; Senecal and Nantel 2004; Chen 2008). In our experiment we measure real behaviour.

Songs of a Future Past

Songs of the Future Past is a web platform simulating an online cultural market. The experiment is *not* a web based experiment and took place in a computer lab under the supervision of the experimenters. Each session involved many participants at the same time, each one having access to a separate PC and headphones. The main goal of the experiment was to see if social influence signals would alter people's choices. To measure this properly we made sure that the exprimental set up was devoid of choice overload effects, which would make it difficult to disentangle the effects of social influence from those of decision-making mental shortcuts. The effects of social influence was compared against that of placement (screen positioning) and that of a "modicum quantity" of choice overload (scrollable list).

Participants

The experiment took place from February 16, 2016 to December 02, 2016 and involved 1107 participants[1] (see Table 1) from 28 different nationalities. Except for a very small group of high school students, participants were university students most of which were recruited from various Italian universities in Perugia, Salerno and Rome, and the remaining ones from Albanian univeristies. As a consequence, participants were mostly young ($M = 22.04$, $SD = 3.74$). Detailed demographic information is summarised in Table 2.

[1]A total of 36 participants did not conclude the experiment and left after giving their informed consent for participation.

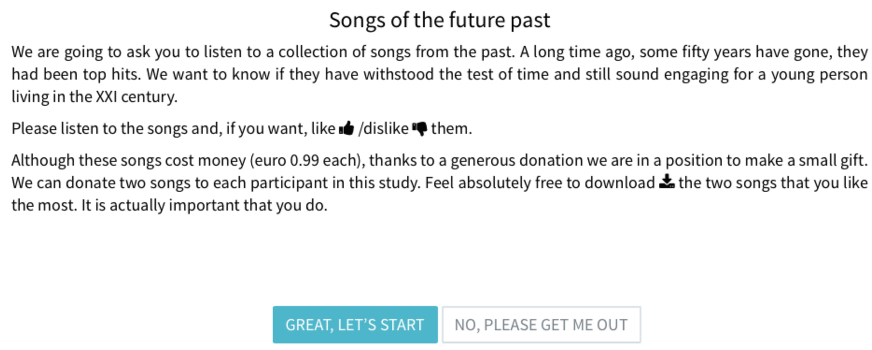


Figure 2: Cover story describing the experiment.

	Participants	Females	Males	Other
Control Group	290	151	137	2
Stars	259	108	150	1
Downloads	223	65	156	2
Linear Layout	108	52	50	6
Collaborative Filtering	227	115	109	3

Table 1: Participants statistics for all scenarios.

Category	Number of participants	%
Gender		
Female	487	44%
Male	597	54%
Other	23	2%
Nationality		
Italians	885	80%
Albanians	166	15%
Other	56	5%
Age		
< 18	12	1%
18 to 24	929	84%
> 24	166	15%

Table 2: Demographics. Participants were all university students, with the exception of 15 high school students.

Songs

The songs were selected from the USA hit parade list of 1968 and are listed in Table 3. These songs were chosen intentionally to be of high quality content but likely to be unfamiliar to young people. As discussed, the reason we committed to a small number of songs (ten) in the experiment was to avoid as much as possible the interferences attributable to choice overload. This phenomenon occurs as a result of too many choices presented to the consumers (Scheibehenne, Greifeneder, and Todd 2010). Several studies have associated choice overload with decrease in the motivation to decide or avoiding making a decision (Iyengar and Lepper 2000) and consumers' dissatisfaction (Schwartz 2004; Iyengar, Wells, and Schwartz 2006).

Band	Song
Diana Ross	*Love hangover*
Blue Cheer	*Summertime blues*
Blood, sweat and tears	*You have made me so very happy*
Procol Harum	*A whiter shade of pale*
The Delfonics	*La la - means I love you*
The Vogues	*Turn around, look at me*
The Archie and The Drells	*I can't stop dancing*
Dionne Warwick	*Walk on by*
Credence Clearwater Revival	*Susie Q*
Bobby Vinton	*I Love you how you love me*

Table 3: List of songs used in the experiment.

Procedure

Participants were welcomed and told what they were asked to do. Namely, connect to a web site containing ten songs from the 60's. These songs had been a hit in the past and the aim of the experiment was described as that of finding out if their sound was still captivating for the "young generations" (See Figure 2). The actual goal of the experiment was kept hidden and described in detail after the conclusion of the experiment in a debriefing session. The experiment was done with the informed consent of the participants and in full compliance with the relevant rules and regulations on psychological research that apply in our country.

Upon entering the web site,

1. Participants were shown a message asking for informed consent;
2. Once they agreed to participate they were asked for some demographic data such as gender, age and nationality.
3. Finally, they were instructed on what to do, namely listen to a collections of ten songs as much as they wanted, like/dislike any number of them, and download up to two songs. These instructions had already been given them orally before the beginning of the experiment, but subjects were asked to read them carefully before starting (Figure 2). Every song cost euro 0.99 and they were given a budget of 1.98 euros. Participants were informed that *downloaded songs would actually be bought* by us from

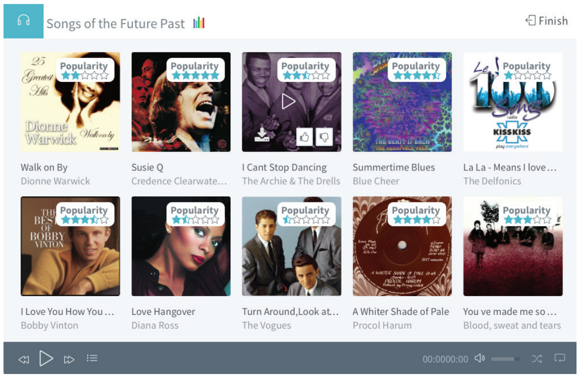


Figure 3: Popularity with stars layout.

a well-known online store and sent to them via email as a gift.

Participants could end the experiment at any time by clicking the "Finish" button found on the top right corner of the web site. As a final step, they were asked whether they already knew the songs. The last questionnaire about the familiarity with the content proved that the selected songs were unknown to almost all participants.

Experimental Conditions

Subjects entering the experiment were assigned to one of five different experimental conditions, or scenarios.

- SCENARIO 1: THE CONTROL GROUP. Subjects were shown a layout consisting of a $5x2$ array of slots. Each song was assigned to a slot uniformly at random as shown in Figure 1. *This scenario was the first to be concluded* and provided a benchmark for the other scenarios. We will refer to the number of downloads obtained at the end of this scenario as the *download count*.
- SCENARIO 2: STARS. The layout of this scenario is the same as in the CONTROL GROUP, with songs assigned randomly to slots. The only difference consisted in a clearly visible "star rating" assigned as follows. Songs were ranked according to the download count, with the number of stars reflecting their ranking (the most downloaded song would get 5 stars, the second $4^1/_2$ and so on, until the last song receiving $^1/_2$ stars. See Figure 3). The stars assignment to songs was the same for all participants in this experimental condition.
- SCENARIO 3: DOWNLOADS. Same as the previous scenario, but with the actual number of downloads replacing the stars. The number of downloads used was the one given by the download count and so it was the same for all participants in this experimental condition.
- SCENARIO 4: LINEAR LAYOUT. In this scenario songs were arranged in a vertical, scrollable list, ranked from top to bottom according to the download count. Only the first three songs were displayed on the screen, with the remaining ones accessible via scrolling. The number of downloads used was the one given by the download count and so it was the same for all participants in this experimental condition.
- SCENARIO 5: COLLABORATIVE FILTERING. Same layout as in the CONTROL GROUP. After each participant's first download, a collaborative filtering based recommendation box would pop up and suggest a second song to download as shown in Figure 4. The filter operates as follows. Suppose that in scenario 5 a user downloads song a. We look at all the people in the CONTROL GROUP who also downloaded a, and form the set of songs that they downloaded. Call this set S, and let b be the most frequent song in it. Then, b is the song recommended to the user.

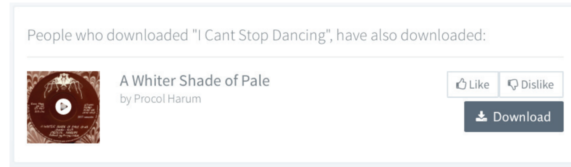


Figure 4: Collaborative Filtering based recommendation box.

The Experimental Outcome

User Engagement

In our experiments, user engagement is measured according to the following metrics:

- Exploration Time: Total amount of time a participant spends on the experiment, starting from the moment s/he gives his/her consent to participate until s/he exits the platform, i.e. terminates the experiment.
- Listens: Number of songs a participant has listened to.
- Downloads: Number of songs a participant has downloaded (bought).
- Votes: Number of likes/dislikes a participant has given.

Table 4 illustrates an overview about participant engagement in each scenario. The main take away is that participants willingly spent a considerable amount of time on the experiment (an average of 16.68 minutes in total). This is important, because it shows that people took their time before making a choice, which is a strong indication that they made a conscious decision avoiding decision-making mental shortcuts.

The analysis of variance shows that there is a significant difference in the mean length of time spent by participants among the experimental conditions ($F(4, 1066) = 11.73, p < 0.001$). A Tukey post-hoc test reveals that participants in the LINEAR LAYOUT scenario took statistically significantly more time ($M = 20.683$, SD = 8.1 mins) to complete the experiment compared to the other scenarios. This can be simply explained by the presence of the scrollable list, which makes the interaction with the website more cumbersome and time consuming.

Scenario	Exploration Time (minutes)	Listening	Downloads	Votes
Control Group	15.21	13.00	1.33	8.54
Stars	17.59	13.12	1.09	8.59
Downloads	16.91	14.57	1.43	8.25
Linear Layout	20.68	16.82	1.23	8.70
Collaborative Filtering	15.55	14.01	1.42	8.12

Table 4: Mean values for user engagement metrics

The listen data show an interesting outcome. According to a post-hoc analysis using Tukey test, the number of listens under the conditions DOWNLOADS and LINEAR LAYOUT (whose mean and standard deviations are, respectively $\{M = 14.57, SD = 6.0\}$ and $\{M = 16.82, SD = 7.1\}$) is more than that of CONTROL GROUP ($M = 13.00, SD = 4.0$). This difference was found to be statistically significant with a one way Anova test ($F(4, 1066) = 212.53, p < 0.001$). Both conditions use the number of downloads as a cue, which is a social influence signal. Recommendations too increased the number of listens, but this effect is only marginally significant in statistical terms ($p = .06$).

Market Distortion

There were three possible actions that a participant in the experiment could take: listening to a song, downloading, and casting a vote (like/dislike). This gives rise to three possible "markets", denoted as L, D, and V respectively.

In order to define market distortion we define first the following quantities. Fix a scenario $\sigma \in \{1, \dots, 5\}$ (e.g scenario 1: the control group, scenario 2: stars, etc.) and a specific market $M \in \{L, D, V\}$, and let S denote the set of songs and s a generic song. Let us denote by $n_s^{\sigma,M}$ the number of times that song s was selected in market M and scenario σ (to "select" a song in market L means to listen to it, while in market D it means to download it, and in market V it means to vote for it). The *market share* of song s in market M and scenario σ is defined as

$$m_s^{\sigma,M} := \frac{n_s^{\sigma,M}}{\sum_{x \in S} n_x^{\sigma,M}}.$$

We say that an online signal distorts any specific market if it can significantly change the portion of listens/downloads each of the songs has accumulated, i.e. alter market shares. The "natural state" of each of the markets is given by the CONTROL GROUP, where there are no signals based on social influence and slot position is made ineffective by randomisation. Therefore the CONTROL GROUP will serve as a benchmark for computing the distortion introduced in the market.

For every song, we compared the outcome (say, the number of downloads) in the control group versus that under a different experimental condition (say, star ratings). Since there are ten songs, we used a Bonferroni test with adjusted alpha levels of .005 per test (.05/10). The results indicate that the differences due to social influence cues (star ratings, download counts and recommendations) are not statistically significant and therefore the hypothesis that they induce no change cannot be rejected (see Figure 5).

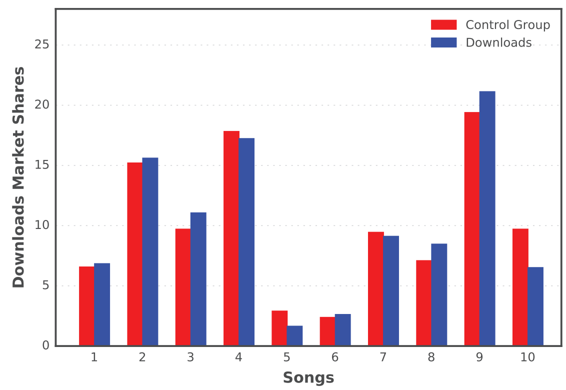


Figure 5: Market shares for CONTROL GROUP and DOWNLOADS scenarios: the natural state of the market is not altered by number of downloads.

In contrast, the data reveal that even a "small dose" of information overload induced by a scrollable list (LINEAR LAYOUT scenario) causes a significant distortion in the listen market. Each song in the LINEAR LAYOUT scenario has a statistically significant difference in the portion of listenings with its correspondent in the control group ($p < .005$). Figure 6 compares the distribution of the songs (market shares) under the four experimental conditions against the control group. The distance between the distribution of the LINEAR LAYOUT and the control group (as measured by the ℓ_1-norm) is three times the distance between the control group and the other distributions.

Market Inequality

Market inequality is another interesting quantity worth analysing. A well-known measure for it is the GINI COEFFICIENT (Wikipedia 2015). Given a specific scenario $\sigma \in \{1, \dots, 5\}$, and a market $M \in \{L, D, V\}$, the corresponding GINI COEFFICIENT is given by,

$$\text{GINI COEFFICIENT}(\sigma, M) := \frac{\sum_{i,j \in S} |m_i^{\sigma,M} - m_j^{\sigma,M}|}{2|S| \sum_{k \in S} m_k^{\sigma,M}}. \tag{1}$$

To exemplify, GINI$(1, L)$ is the GINI COEFFICIENT of the listening market in scenario 1 (the control group), GINI$(5, D)$ is the GINI COEFFICIENT of the market of downloads under the collaborative filtering regime, etc.

A GINI COEFFICIENT of 0 denotes perfect equality, when all songs have the same market share. At the other end of

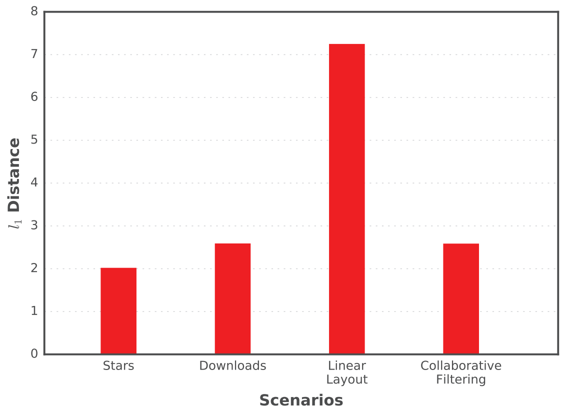


Figure 6: Market shares distance with respect to the control group.

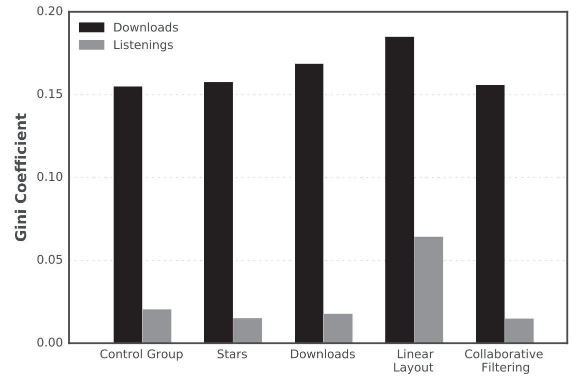


Figure 7: Market inequality for listening and downloads for each scenario.

the spectrum, a GINI COEFFICIENT of 1 expresses perfect inequality, a situation where only one of the songs has 100% of the market.

To compute the GINI COEFFICIENT in a statistical meaningful way, we randomly split the participants to each of the experimental conditions into four groups of the same size, and calculated the GINI COEFFICIENT for each of the groups. The resulting GINI COEFFICIENTS are the mean values for each condition.

Figures 7 tells us what happens in the "attention market" L, i.e. the number of times a song is listened to. We can see that in all scenarios except for the LINEAR LAYOUT, the GINI COEFFICIENT does not change. This means that no signal (except for the LINEAR LAYOUT,) introduces market inequality. With a value around 0.04, which is quite more close to zero, the GINI COEFFICIENT also tells us that this particular market is very balanced– no song was listened to much more often than the others. Our raw data further confirm this– out of the collection of ten songs, every one got nearly 10% of the listening market share.

The presence of the outlier can be easily explained ($F(4, 15) = 33.86, p < 0.001$). Recall that in the LINEAR LAYOUT, the most downloaded songs in the CONTROL GROUP got the top positions on the screen. Since scrolling is a hassle, the songs at the top are listened to more often. In our opinion, this effect also explains to a large extent the findings of (Salganik, Dodds, and Watts 2006), where the list of songs was even longer than ours (48 songs instead of 10). We will analyse this a bit more closely in the "Interpretations" section.

So, except for the linear layout, each song got the same level of attention. What about the number of downloads? Again, the GINI COEFFICIENT gives us the information we need. At a value of nearly 0.3 the GINI COEFFICIENT tells us that the market is somewhat unequal and hence some songs got a larger share of the market. Notice however, that the degree of market inequality stays the same in all scenarios (Figure 7). All scenarios have the same degree of inequality of the control group, and so there is no amplification effect of the sort reported in (Salganik, Dodds, and Watts 2006) (the p- value is greater than 0.05 so the differences are not statistically significant).

This discrepancy calls for an explanation, which we attempt in the "Interpretations" section.

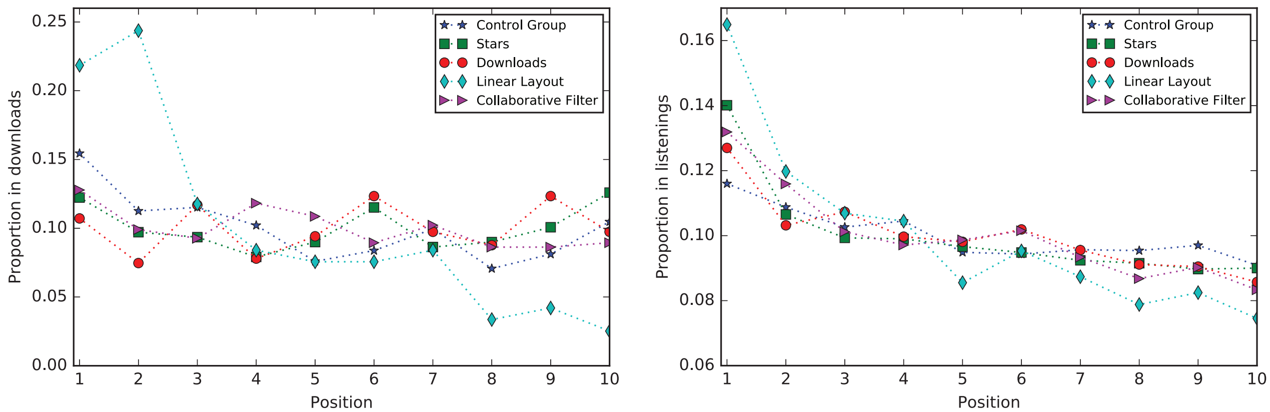
Hidden persuaders

What we have seen so far is that widely used social influence cues do not induce market distortion in our set up. In contrast, slot position turned out to have measurable distorting effects. Let us refer to the 5×2 layout as the *flat* layout, and let us use the following convention. We will refer to the slots of the top row as slots 1 to 5, going from left to right, and likewise as slots 6 to 10 in the bottom row. The following turned out to be statistically significant.

- In all scenarios that use the flat layout, slot 1 commands between 1.5% and 4% more listens compared to the other slots.
- In all scenarios using the flat layout, the first slot accumulated up to 5% more downloads than other slots
- In the linear layout, the two topmost slots (the most visible ones) have a 10% increase in the listening market share, and 7% more listens when compared to other scenarios.
- In the linear layout, the combined download market share of the two slots is 45% and they obtain up to 16% more downloads compared to other scenarios.

More nuanced information can be derived from Figure 8.

While the effect of the linear layout can be easily explained in terms of navigability (it makes it cumbersome and time consuming) the other effect, although well-known, remains quite startling. Informal discussions with participants during the debriefing of the experiment indicate that slot position seem to act at a subliminal, unconscious level. In contrast, participants seemed to have been well aware of the presence of other signals, including the fact that in a scrollable list the top positions stay at the top. This cannot be taken as solid evidence, but it is a compelling working hypothesis for future research.



(a) Proportion of downloads (b) Proportion of listens

Figure 8: The effect of slot position on market shares.

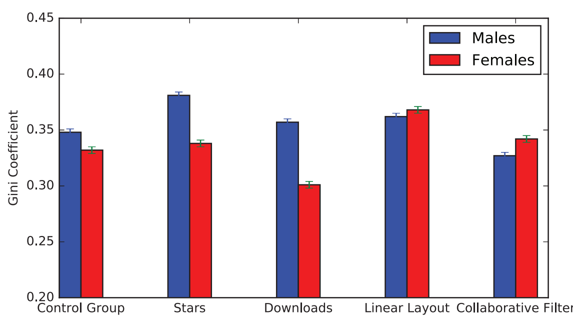


Figure 9: Market inequality of downloads.

Gender Inequalities

Research studies have shown that there exist gender differences in customer behaviour while shopping online (Rodgers and Harris 2003).

In our experiment, on average, males and females spent, respectively, 16.8 mins and 16.6 mins on the website but this difference is not statistically significant . Females ($M = 9.65, SD = 5.67$) however voted more songs compared to males ($M = 8.44, SD = 5.13$), giving 1.21 more votes on average ($F(1, 1075) = 13.49, p < 0.001$).

Let us now see whether gender has a tendency to create balanced or unequal markets (please refer to Figure 9). In the "attention market" (# listens) the GINI COEFFICIENT of the two markets (male vs. female) is essentially the same, with no statistically significant difference. In contrast, the purchasing behaviour based on gender shows interesting differences.

Figure 9 shows the average GINI COEFFICIENT (together with the standard deviation) in the market of downloads created by males and females for all five experimental conditions.

Note that in the control group, in the presence of stars and of download counts, males created more concentrated markets compared to females. The differences in Figure 9 are all statistically significant ($p < .001$) except that in the LINEAR LAYOUT condition where $p = .282$.

Interpretations

Our results concerning social influence feedback given by indicators of popularity such as download counts, star ratings and recommendations, seem to be at odds with previous studies where they were reported to introduce significant distorting effects (Salganik, Dodds, and Watts 2006; Salganik and Watts 2008; Muchnik, Aral, and Taylor 2013; Van de Rijt et al. 2014). In this section we will try to reconcile our findings with what was previously reported.

A possible explanation is that in situations of choice overload, people resort to decision-making shortcuts, or heuristics, in order to avoid the high cognitive cost of systematic exploration. In our setup, the collection of songs was limited to ten and the conditions of the experiment made it possible such a systematic analysis of the entire collection. Indeed, participants found the experiment enjoyable and took their time. The fact that it was taking place in a lab under the coordination of "authoritative" figures such as university professors, and the fact that they were given a budget of real money, helped them focus on what they were doing, making them more impermeable to the lure of decision-making shortcuts. In other words, people were focused, engaged and thus could "make up their own mind".

At the other end of the spectrum, consider a collection of songs that, because of its sheer size, cannot be explored systematically. Or, consider a collection of songs that, although not very large, is made difficult and quite cumbersome to explore. In cases such as these, it is reasonable to assume that a person would rely on cues such as the opinion and the experience of others (embodied by proxies like download counts and recommendations) in order to make a decision on what to buy or listen to.

This interpretation seems to explain quite well the apparent discrepancy between our study and previous findings. Our setup fits the description of a scenario where people can “make up their own mind”, while those of (Salganik, Dodds, and Watts 2006; Salganik and Watts 2008; Chen 2008; Zhu and Huberman 2014; Abeliuk et al. 2017) fall in the second category. For instance in (Salganik, Dodds, and Watts 2006), a collection of 48 songs was used. This is quite cumbersome to navigate, both with a flat and a linear layout. Not surprisingly, people relied on social influence cues to decide what to download. This conclusion is also consistent with what we observed. Even with a small collection of ten songs, a scrollable list constitutes already a hassle and the distorting effects on the market became immediately evident.

Slot position provided a very useful benchmark. While the absence of choice overload made social influence ineffective, this cue retained its power.

Finally, we observed that while social influence cues did not modify market shares, they modified market volume in the sense that the number of listens increased. This is quite surprising and unexpected and warrants further investigations.

Concluding remarks

There are several possibilities to improve and carry forward the research presented in this paper. One issue that could possibly be addressed more satisfactorily is the role played by monetary costs and its interplay with social influence. For instance, it could be worthwhile to repeat the experiment by making the songs freely downloadable to the participants, and see whether their behaviour changes. Another interesting issue is how to make a study like this more similar to a real-life situation. As discussed this presents an interesting conundrum because if we make the set of songs larger other effects such as information overload might creep in. To determine whether this actually happens appears to be an interesting problem in itself. Another problem that we do not know how to solve satisfactorily at the moment is how to use high quality songs (which unfortunately essentially means commercial songs) in a large scale web replica.

Although our group of participants was rather large for a laboratory experiment the population was skewed, consisting solely of university students. There was also an age bias— participants were mostly young– but this is mitigated by the fact that this age group is of particular interest by itself for the object of study.

Acknowledgement

This work was supported by a Google Focused Research Award, by the Sapienza Grant C26M15ALKP, by the SIR Grant RBSI14Q743, and by the ERC Starting Grant DMAP 680153.

Many people helped us recruiting students to participate in the research. We are very grateful to them: Alberto Mattiacci, Andrea Ranzato, Angela Scaringella, Carlo Alberto Pratesi, Carlo D’Ippoliti, Cristina Sofia, Elena Valentini, Erind Bedalli, Fabrizio Martire, Franca Faccioli, Francesca Comunello, Gaia Peruzzi, Giovanni Ciofalo, Igli Hakrama, Igor Melatti, Julinda Stefa, Marcella Corsi, Marco Bruno, Mario Morcellini, Marta Tedesco, Mihaela Gavrila, Mirzia Bianca, Moira Colantoni, Nerio Naldi, Paolo Montesperelli, Romana Andò, Sergio de Agostino, Silvia Leonzi, Simona Romani, Valentina Martino, Vincenzo Auletta and Walter Didimo.

References

Abeliuk, A.; Berbeglia, G.; Van Hentenryck, P.; Hogg, T.; and Lerman, K. 2017. Taming the unpredictability of cultural markets with social influence. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee.

Adomavicius, G., and Tuzhilin, A. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering* 17(6):734–749.

Ajzen, I.; Brown, T. C.; and Carvajal, F. 2004. Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation. *Personality and social psychology bulletin* 30(9):1108–1121.

Aronson, E. 1988. *The social animal*.

Asch, S. E. 1951. Effects of group pressure upon the modification and distortion of judgments. *Groups, leadership, and men. S* 222–236.

Chaiken, S. 1980. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology* 39(5):752.

Chen, Y.-F. 2008. Herd behavior in purchasing books online. *Computers in Human Behavior* 24(5):1977–1992.

Crutchfield, R. S. 1955. Conformity and character. *American Psychologist* 10(5):191.

Festinger, L., and Carlsmith, J. M. 1959. Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology* 58(2):203.

Gigerenzer, G., and Gaissmaier, W. 2011. Heuristic decision making. *Annual review of psychology* 62:451–482.

Go, E.; Jung, E. H.; and Wu, M. 2014. The effects of source cues on online news perception. *Computers in Human Behavior* 38:358–367.

Hastie, R., and Dawes, R. M. 2010. *Rational choice in an uncertain world: The psychology of judgment and decision making*. Sage.

Hou, J. 2017. Can interface cues nudge modeling of food consumption? experiments on a food-ordering website. *Journal of Computer-Mediated Communication* 22(4):196–214.

Iyengar, S. S., and Lepper, M. R. 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology* 79(6):995.

Iyengar, S. S.; Wells, R. E.; and Schwartz, B. 2006. Doing better but feeling worse: Looking for the ?best? job undermines satisfaction. *Psychological Science* 17(2):143–150.

Kutner, B.; Wilkins, C.; and Yarrow, P. R. 1952. Verbal attitudes and overt behavior involving racial prejudice. *The Journal of Abnormal and Social Psychology* 47(3):649.

LaPiere, R. T. 1934. Attributes versus actions. *Social Forces* 230–237.

Lewis, A. 2012. *The Cambridge handbook of psychology and economic behaviour*. Cambridge University Press.

Linden, G.; Smith, B.; and York, J. 2003. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing* 7(1):76–80.

Linn, L. 1965. Verbal attitudes and overt behavior involving racial prejudice. *Social Forces* 353–364.

Muchnik, L.; Aral, S.; and Taylor, S. J. 2013. Social influence bias: A randomized experiment. *Science* 341(6146):647–651.

Resnick, P., and Varian, H. R. 1997. Recommender systems. *Communications of the ACM* 40(3):56–58.

Rodgers, S., and Harris, M. A. 2003. Gender and e-commerce: an exploratory study. *Journal of advertising research* 43(3):322–329.

Russo, J. E.; Schoemaker, P. J.; and Russo, E. J. 1989. *Decision traps: Ten barriers to brilliant decision-making and how to overcome them*. Doubleday/Currency New York, NY.

Salganik, M. J., and Watts, D. J. 2008. Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social psychology quarterly* 71(4):338–355.

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

Scheibehenne, B.; Greifeneder, R.; and Todd, P. M. 2010. Can there ever be too many options? a meta-analytic review of choice overload. *Journal of Consumer Research* 37(3):409–425.

Schwartz, B. 2004. The paradox of choice.

Senecal, S., and Nantel, J. 2004. The influence of online product recommendations on consumers? online choices. *Journal of retailing* 80(2):159–169.

Sheeran, P. 2002. Intention?behavior relations: A conceptual and empirical review. *European review of social psychology* 12(1):1–36.

Sundar, S. S.; Knobloch-Westerwick, S.; and Hastall, M. R. 2007. News cues: Information scent and cognitive heuristics. *Journal of the Association for Information Science and Technology* 58(3):366–378.

Tversky, A., and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *science* 185(4157):1124–1131.

Van de Rijt, A.; Kang, S. M.; Restivo, M.; and Patil, A. 2014. Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences* 111(19):6934–6939.

Wikipedia. 2015. Gini coeffient. [Online; accessed 15-December-2015].

Zhu, H., and Huberman, B. A. 2014. To switch or not to switch: understanding social influence in online choices. *American Behavioral Scientist* 58(10):1329–1344.