Data from *Success and Failure in Cultural Markets*[*]

Prepared by Matthew J. Salganik

July 30, 2008

Abstract

The write-up documents the data files that are need to reproduce, and hopefully expand upon, the analysis conducted in the dissertation *Success and Failure in Cultural Markets*, written by Matthew J. Salganik, supervised by Duncan J. Watts, and conducted at the Department of Sociology at Columbia University between 2004 and 2007. The project was motivated by a puzzling aspect of contemporary cultural markets: successful cultural products, such as hit songs, bestselling books, and blockbuster movies, are orders of magnitude more successful than average; yet which particular songs, books, and movies will become the next "big thing" appears impossible to predict. The dissertation proposed that both of these features, which appear to be contradictory at the collective level, can arise from the process of social influence at the individual level. To explore this possibility empirically we constructed a website where participants could listen to, rate, and download new music, and more importantly, where we could control the information that these participants had about the behavior of others. Using a "multiple-worlds" experimental design we found support for our ideas in a series of four experiments involving a total of 27,267 participants. Included in this data release are 167 files containing the experimental results and the mp3 files from the 48 songs.

Contents

[*]We would like to thank Peter Hausel for developing the MusicLab Web site; Jason Booher-Jennings for design work; Alison Binkowski and Robert Stewart for computer support; Bolt Media for their assistance in recruiting participants; and Peter Dodds and the rest of the Collective Dynamics Group for many helpful discussions. This work was supported in part by an NSF Graduate Research Fellowship, the Institute for Social and Economic Research and Policy at Columbia University and grants to Duncan Watts from the NSF (SES-0094162 and SES-0339023), the James S. McDonnell Foundation, and Legg Mason Funds. Please direct any questions to Matt Salganik.

	Participants	
	`www.bolt.com`	Small-world experiment
Weaker influence	Experiment 1 ($n = 7,149$)	
Stronger influence	Experiment 2 ($n = 7,192$)	Experiment 3 ($n = 2,930$)
Deception		Experiment 4 ($n = 9,996$)

Table 1: The four experimental studies that were performed. Experiments 1 and 2 can be compared to understand the effect of increasing the strength of the social influence. Experiments 2 and 3 can be compared to see if the aggregate-level features of the market are robust to the population of participants. Finally, experiments 3 and 4 can be compared to explore the extent to which beliefs about the success of the songs can become a self-fulfilling prophecy.

1 Background

Before describing the data files we will briefly review the experimental set-up. Readers already familiar with the set-up can skip directly to section 2. A more complete description of the experimental set-up, the theoretical and substantive motivations behind the research, and the results are available elsewhere:

- Salganik, Matthew J. 2007. *Success and Failure in Cultural Markets*. Doctoral dissertation. Department of Sociology. Columbia University.
- Salganik, Matthew J., Peter S. Dodds, and Duncan J. Watts. 2006. "Experimental study of inequality and unpredictability in an artifical cultural market." *Science*, 311:854-856.
- Salganik, Matthew J. and Duncan J. Watts. 2008. "Leading the herd astray: An experimental study of self-fullling prophecies in an artificial cultural market." *Social Psychology Quarterly*, in press.
- Salganik, Matthew J. and Duncan J. Watts. 2009. "The puzzling nature of success in cultural markets." in *The Oxford Handbook of Analytic Sociology* edited by Peter Bearman and Peter Hedstrom. in press.
- Salganik, Matthew J. and Duncan J. Watts. "An experimental approach to the study of collective behavior in cultural markets." Under review.

The four experiments used the design presented in figure 1 and involved a total of 27,267 participants (table 1).[1] In real-time, participants arriving at the experiment were randomly assigned to one of two experimental conditions—*independent* and *social influence*—which differed only by the availability of information on the past behavior of others. Furthermore, participants in the social influence condition were randomly assigned to one of a number of independently evolving "worlds." Participants in the independent condition chose which songs to listen to based solely on the names of the bands and their songs, while participants in the social influence condition could also see how many times each song was downloaded by previous participants in their world. Thus, these social influence worlds may be thought of as multiple, parallel "histories" or "realizations."

Participants were unaware of this experimental design. Upon entering our website (`http://musiclab.columbia.edu`) participants were presented with a welcome screen telling them that they were about to participate in a study on musical tastes and that in exchange for participating they would be offered the chance to download free songs by up-and-coming artists. Participants next gave their informed consent, filled out a brief survey, and were shown a page of instructions. Finally participants were presented with a menu of 48 songs. These songs were randomly sampled from the music website `www.purevolume.com` and screened to insure that they would be unknown to the participants; the final list of songs along with more detailed information about the sampling and screening is presented in section 4. The use of 48 songs—the

[1] All experimental protocols were approved by the Columbia University Institutional Review Board. Experiments 1, 2, and 3 operated under protocol IRB-AAAA5286; Experiment 4 operated under protocol IRB-AAAB1483.

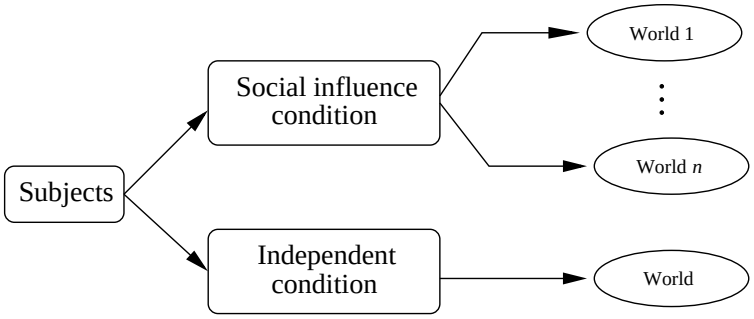


Figure 1: Schematic of the experimental design. This design has two main features. First, it allows researchers to isolate the difference in aggregate outcomes when social influence is present or absent. Second, the design allows researchers to observe multiple realizations of the same process, and thus understand the role of chance in collective outcomes.

maximum number that could fit onto a computer screen under the design used in experiment 1—was an attempt to mimic the choice overload that exists in real cultural markets.

When presented with this song menu, participants in the influence condition were shown the song download counts in their world, while participants in the independent condition were not presented with any download count information. In experiment 1 the songs were presented in a 3×16 grid, unsorted by popularity. In experiments 2, 3, and 4, the songs were presented in a single column; in the influence worlds these songs were sorted by popularity and in the independent world they were randomly ordered for each participant. If a participant clicked on a specific song, she was taken to a new screen where the song automatically began playing. All songs were played using Macromedia Flash Player, streamed in the mp3 format, and encoded at 96kbps. While the participant listened to a song they were asked to rate it on a scale from 1 star ("I hate it") to 5 stars ("I love it"). After rating the song, participants were offered a chance to download the song and were then returned to the song menu where they were able to choose again. Once participants had listened to as many songs as they wished, they could click "log off" and were taken to a screen thanking them for participating and providing them links to the webpages of all 48 bands. Participants who returned to the site while the experiment they participated in was still underway were automatically returned to their world and taken to the appropriate song menu without the need to re-register. Participants who returned to the site after their experiment was complete were prevented from participating in new experiments. Because all participants provided informed consent, they were all aware that they were in a research study, but they were never told about the experimental design or that there were multiple realizations running at the same time.

Screenshots from all steps of the experiment and presented in figures 2 to 11. The consent forms, screen text, and survey questions are presented in the dissertation itself.

Figure 2: Splash screen from the website.

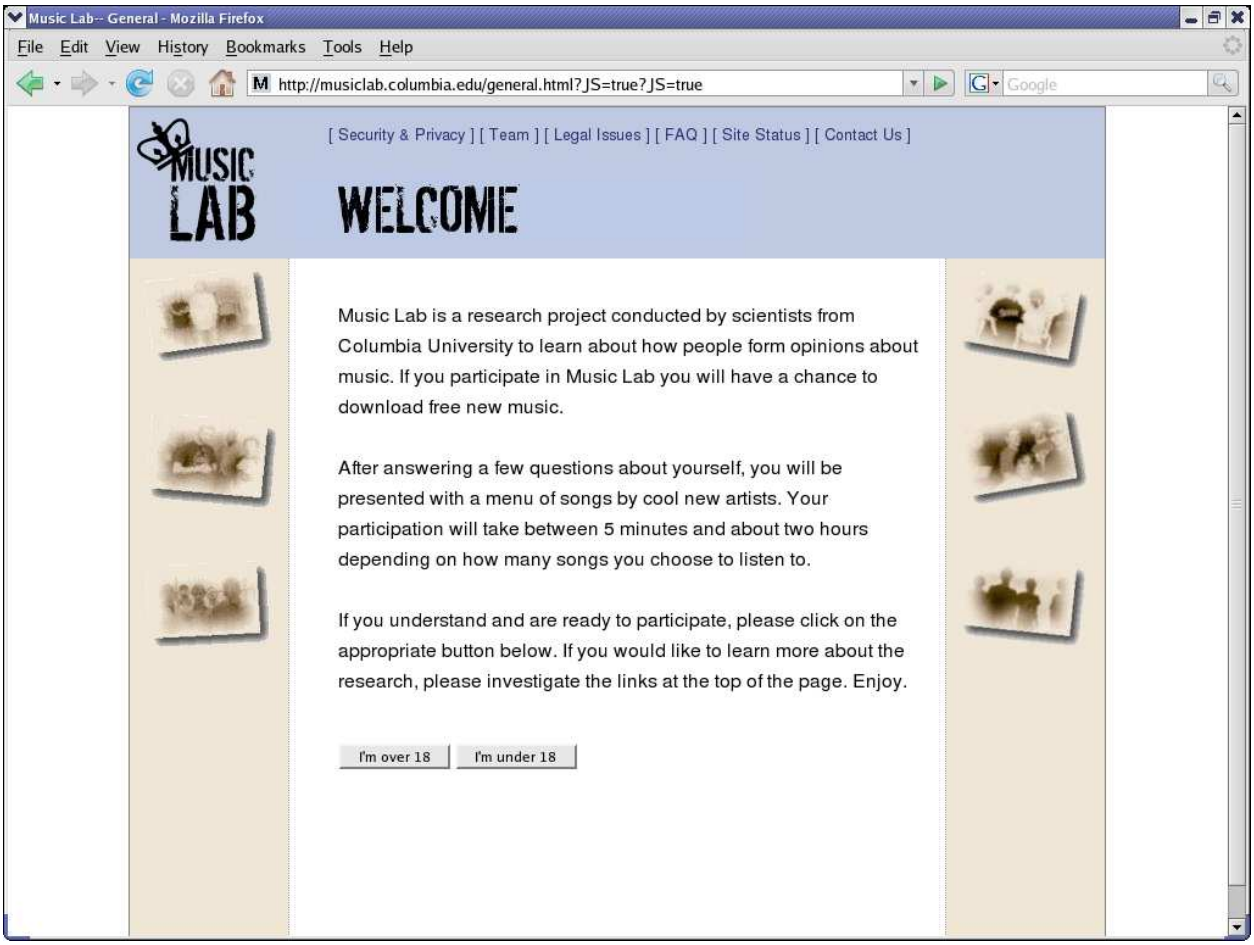


Figure 3: Welcome screen from the website.

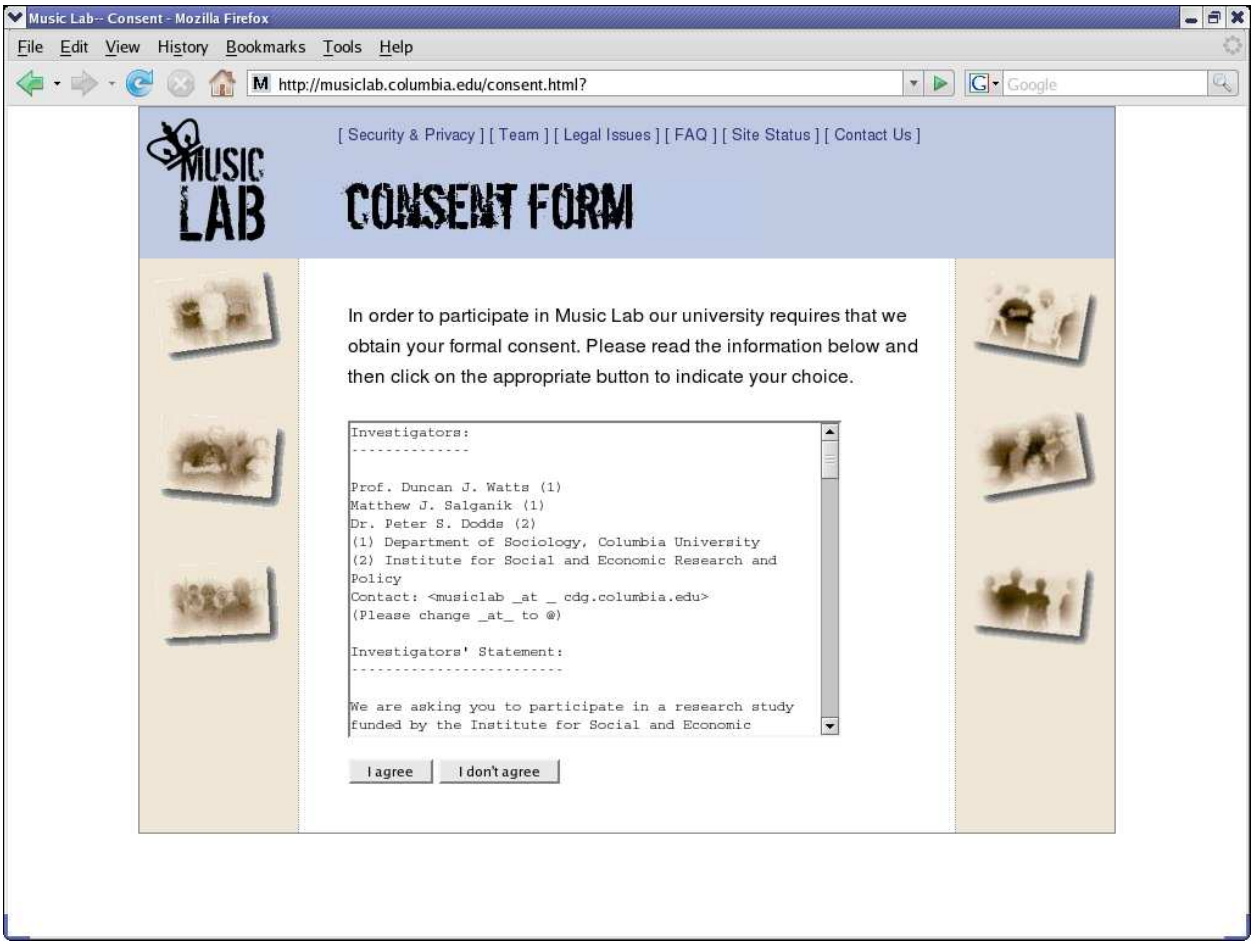


Figure 4: Consent screen from the website.

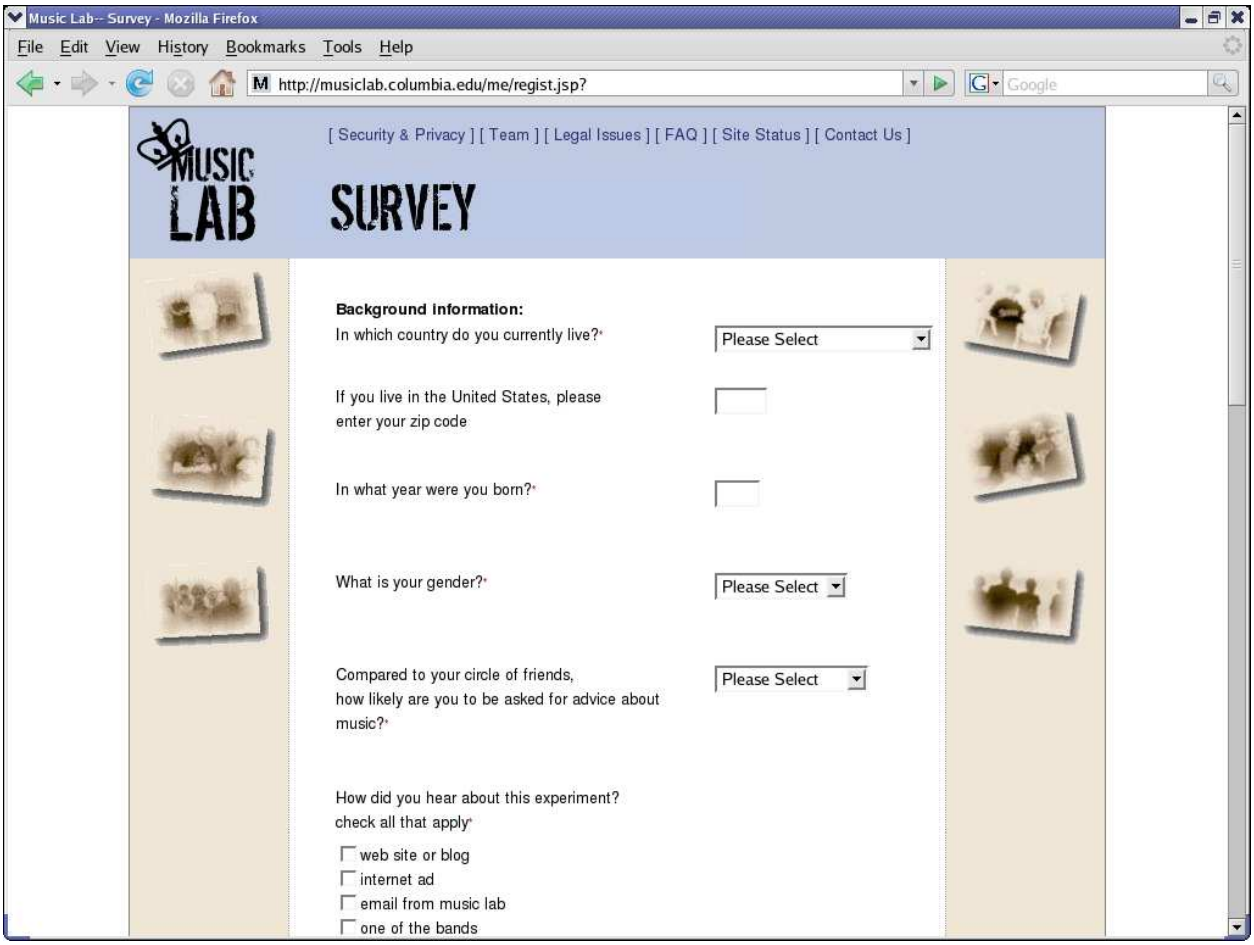


Figure 5: Survey screen from the website.

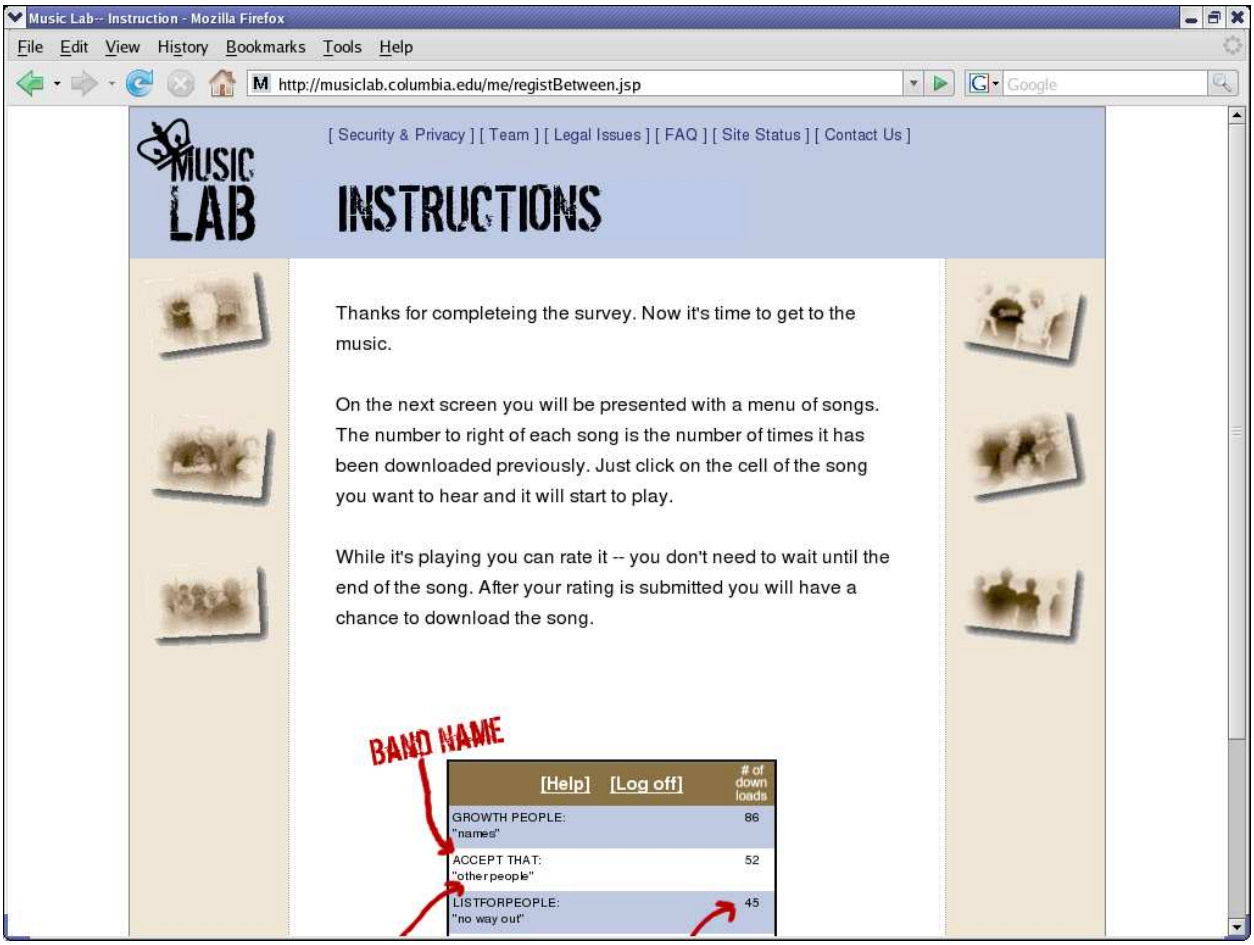


Figure 6: Instructions screen from the website.

Figure 7: Screenshot of the song menu from a social influence world in experiment 1. The song menu in the independent condition (not shown) was identical except that the download counts to the right of each song were not present. In both conditions songs were presented to each participant in a random order.

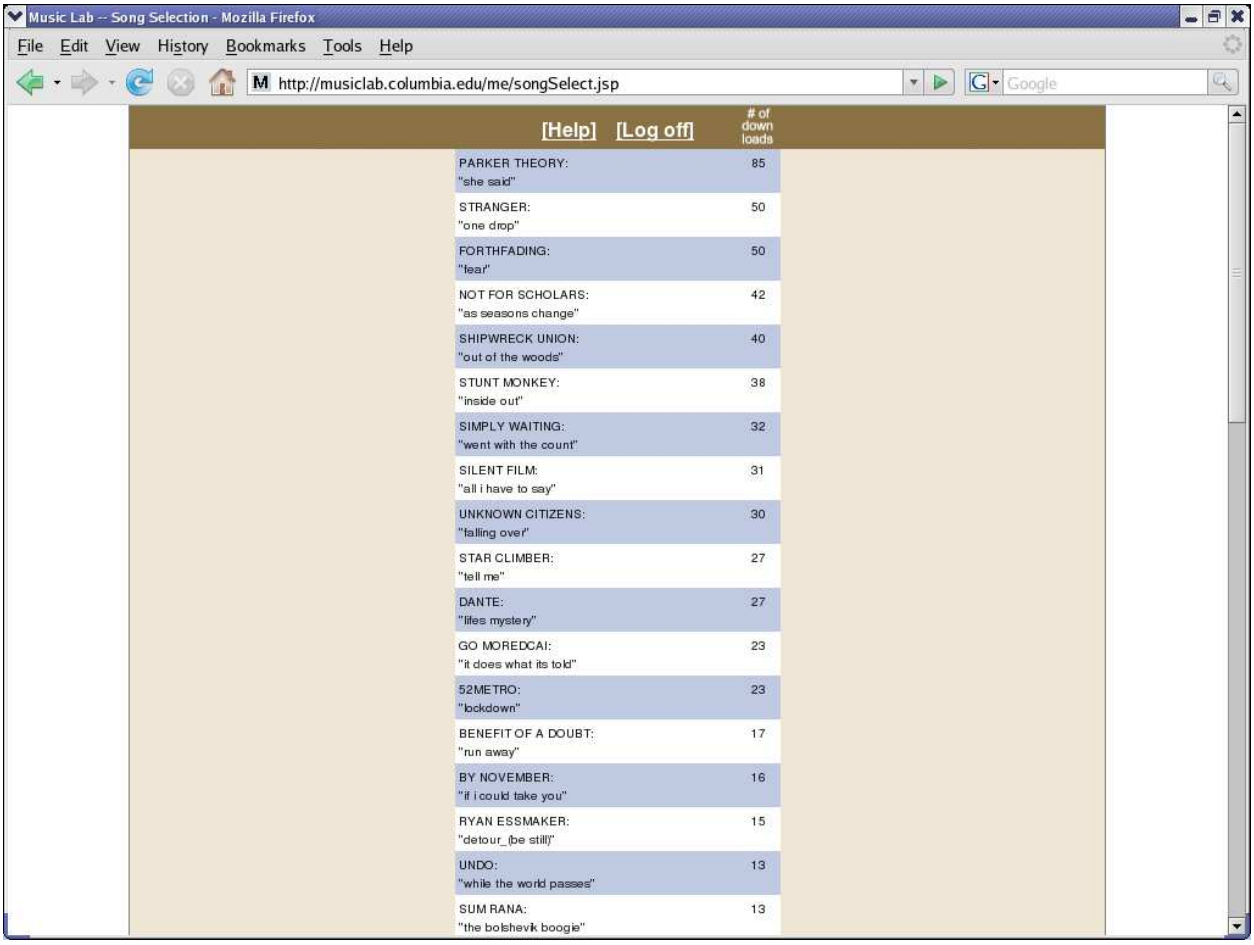


Figure 8: Screenshot of the song menu in a social influence world in experiments 2, 3, and 4. The song menu in the independent condition (not shown) was identical except that the download counts to the right of each song were not present. In the social influence worlds the songs were sorted by popularity and in the independent condition they were ordered randomly.

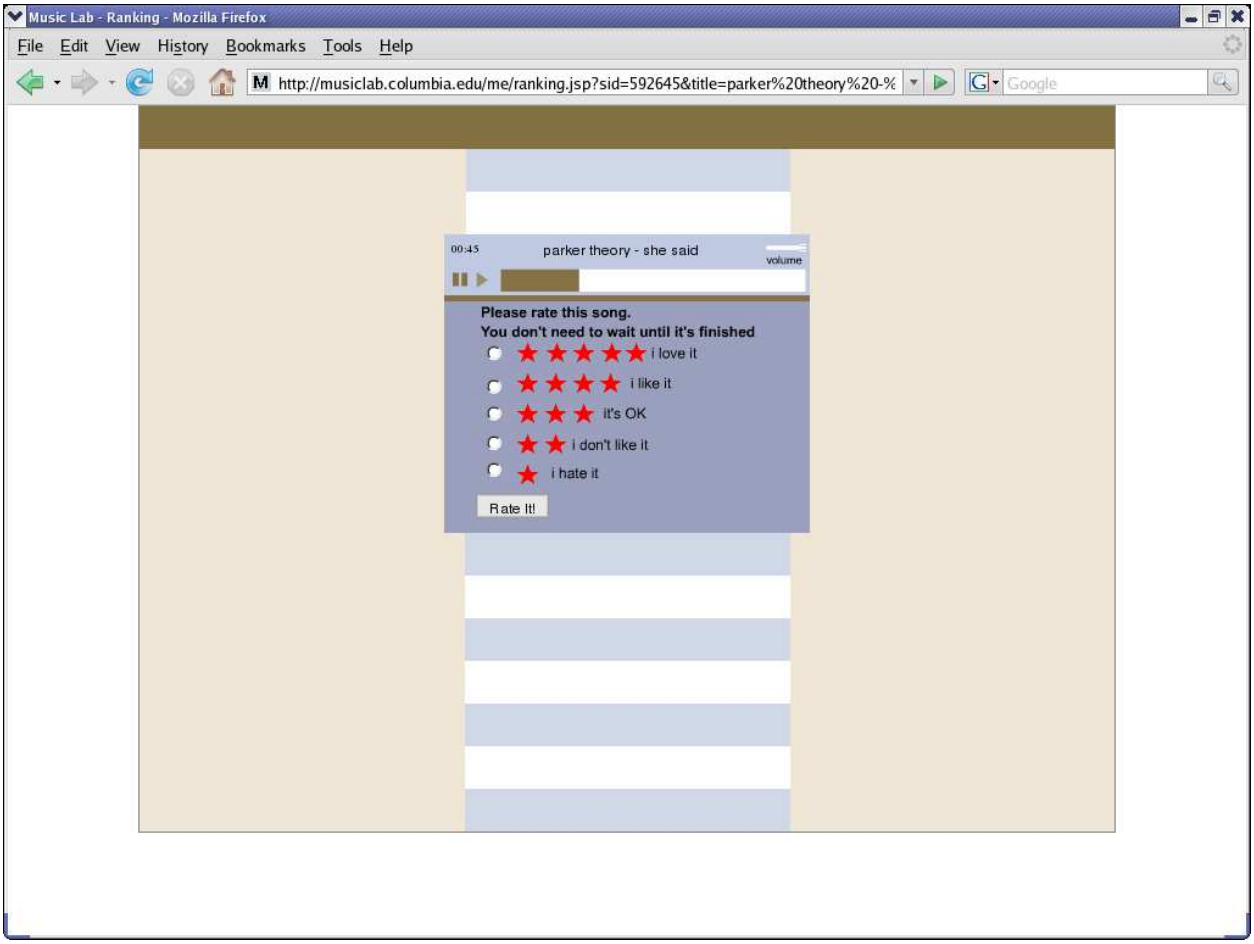


Figure 9: Screenshot of the listening screen. While a song was playing, subjects were required to rate it on a scale of 1 to 5 stars. This rating could be submitted before the song was finished playing.

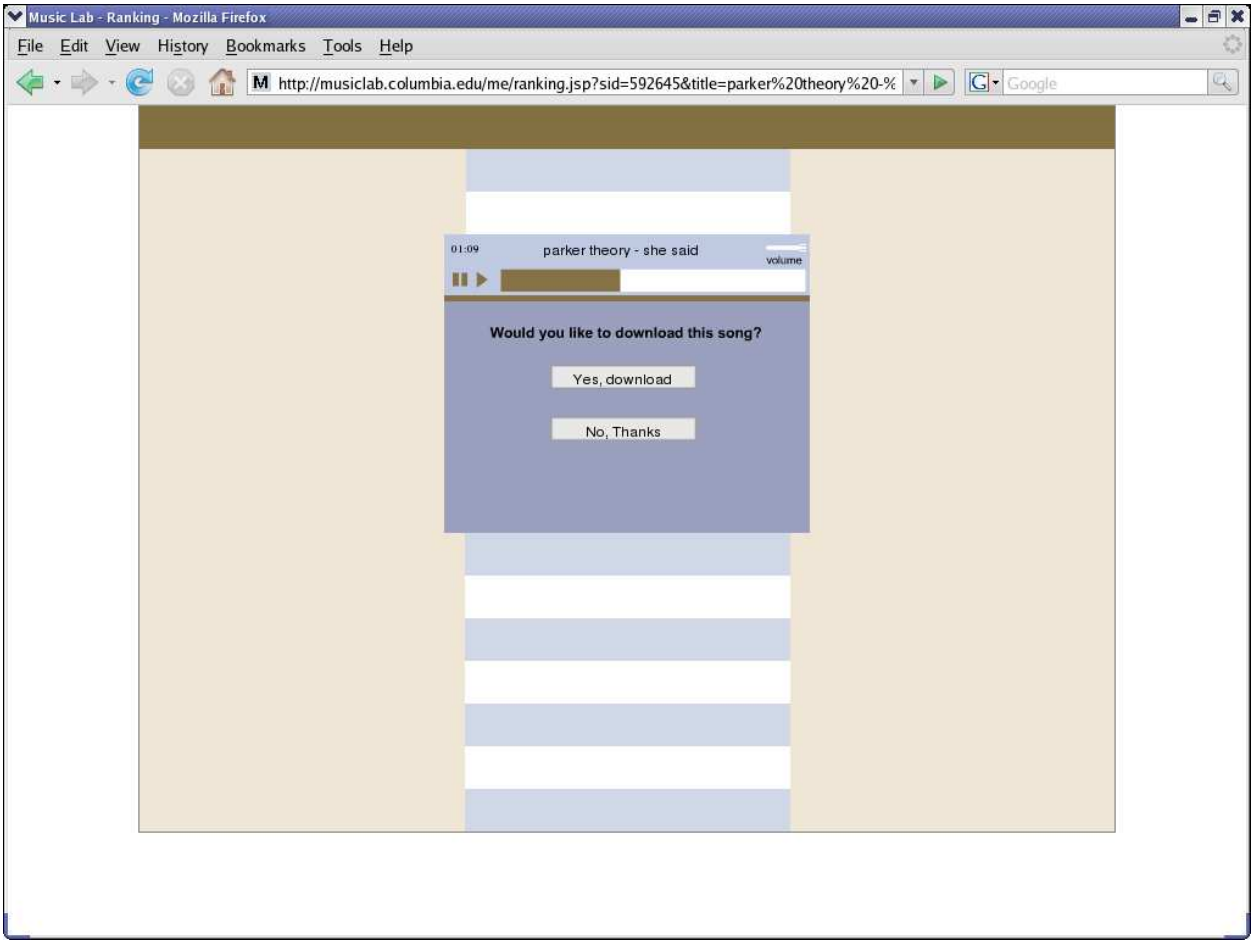


Figure 10: Screenshot of the download decision screen. After rating the song, subjects had to decide to download the song or not.

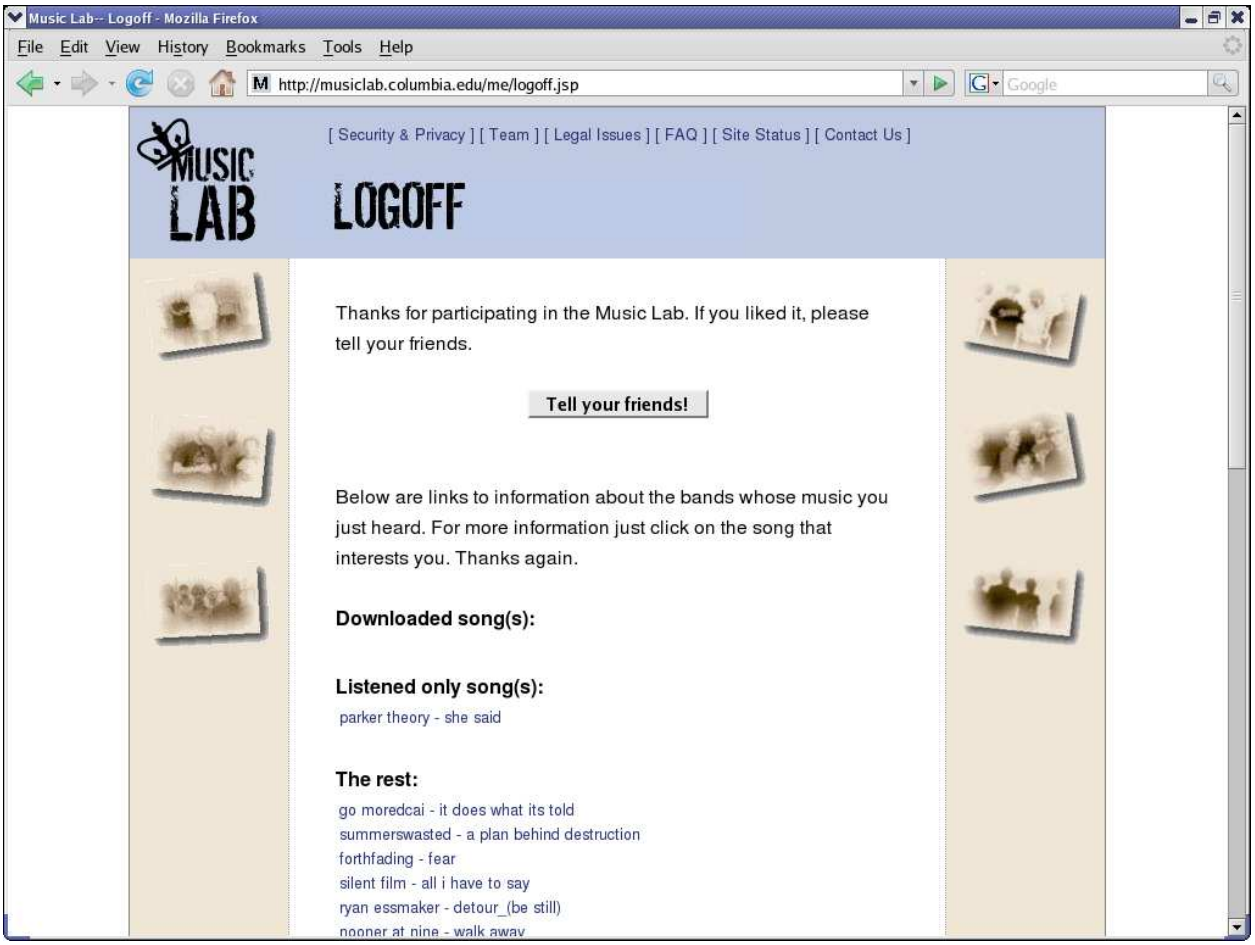


Figure 11: Logoff screen from the website.

2 Data files

There are a variety of data files which store the information from the experiments in different ways. Some files contain information from an entire experiment while some files just contain information from a specific world within an experiment. There were 9 worlds in experiment 1, 9 worlds in experiment 2, 3 worlds in experiment 3, and 4 worlds in experiment 4. Descriptive statistics from these experiments are presented in tables 2 to 5.

- `listens_v*.txt` [4 files (1 from each experiment)]: These files contains the total listen counts for each song in each world in a particular experiment. The first column is the song id and the next columns are the listen counts in the different worlds. The final column is the listen count in the independent world. The songs are sorted in lexicographic order by song id. For example, consider the first line of `listens_v1.txt`:
  100102, 61, 48, 60, 124, 64, 74, 92, 84, 164
  This says that song id 100102 (which table 7 tells us is "This Upcoming Winter" by *Ember Sky*) was listened to 61 times in world 1, 48 times in world 2, etc. It was listened to 164 in the independent world.

- `downloads_v*_lexorder.txt` [4 files (1 from each experiment)]: These files contain the total download counts for each song in each world in a particular experiment. The first column is the song id and the next columns are the download counts in the different worlds. The final column is the download count in the independent world. The songs are sorted in lexicographic order by song id. For example, consider the first line of `downloads_v1_lexorder.txt`:
  100102, 15, 10, 18, 40, 14, 25, 37, 28, 44
  This says that song id 100102 (which table 7 tells us is "This Upcoming Winter" by *Ember Sky*) was downloaded 15 times in world 1, 10 times in world 2, etc. It was downloaded 44 times in the independent world.

- `dynamics_listens_w*_v*.txt` [25 files (9 from experiment 1, 9 from experiment 2, 3 from experiment 3, 4 from experiment 4)]: These files record whether each participant listened to each song in each world in each experiment. The first column is the user id, the second column is the world (this should be the same for everyone in each file), and the next 48 columns record whether the participant listened to the song. Songs are sorted in lexicographic order by song id. For example, consider the first line of `dynamics_listens_w1_v1.txt`:
  4, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0
  This says that the participant with user id 4 was in world 1 and listened to 7 songs: the 10th, 12th, 26th, 29th, 30th, 42nd, and 45th (when the songs are sorted in lexicographic order by song id). By referring to table 7 we can see that this is "Father to Son" by *Beerbong*, "Til Death do us Part (I Don't)" by *The Fastlane*, etc. Note that this file does not have information about the order in which a participant listened to the songs; for that information a research needs to look at `autoplay_w*_v*.txt`.

- `dynamics_ratings_w*_v*.txt` [25 files (9 from experiment 1, 9 from experiment 2, 3 from experiment 3, 4 from experiment 4)]: These files record each participant's ratings in each world in each experiment. The first column is the user id, the second column is the world (this should be the same for everyone in each file), and the next 48 columns record the participant's rating: 0 means no rating is provided, 1 is 1 star ("i hate it"), 2 is 2 stars ("i don't like it"), 3 is 3 stars ("it's OK"), 4 is 4 stars ("i like it"), and 5 is 5 stars ("i love it"). Songs are sorted in lexicographic order by song id. For example, consider the first line of `dynamics_ratings_w1_v1.txt`:
  4, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 4, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 2, 0, 0, 0
  This says that the participant with user id 4 was in world 1 and listened to 7 songs: the 10th, 12th, 26th, 29th, 30th, 42nd, and 45th (when the songs are sorted in lexicographic order by song id). By referring to table 7 we can see that this participant rated "Father to Son" by *Beerbong* 3 stars. This participant rated "Til Death do us Part (I Don't)" by *The Fastlane* 2 stars, etc.

- `dynamics_downloads_w*_v*.txt` [25 files (9 from experiment 1, 9 from experiment 2, 3 from experiment 3, 4 from experiment 4)]: These files record whether each participant downloaded each song in each world in each experiment. The first column is the user id, the second column is the world (this should be the same for everyone in each file), and the next 48 columns record whether the participant downloaded the song. Songs are sorted in lexicographic order by song id. For example, consider the first line of `dynamics_download_w1_v1.txt`:
  4, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
  This says that the participant with user id 4 was in world 1 and download one song: the 29th (when the songs are sorted in lexicographic order by song id). By referring to table 7 we can see that this is "Out of the Woods" by *Shipwreck Union*.

- `screenorder_w*_v*.txt` [25 files (9 from experiment 1, 9 from experiment 2, 3 from experiment 3, 4 from experiment 4)]: These files record the order that songs appeared on the screen for each participant. The first column records the user id and the next 48 columns record the order that the songs were presented. In experiment 2, 3, and 4, the first song was presented in the top spot on the list, the second song in the second spot, etc.[2] In experiment 1, the 1st song was presented in the upper-left corner, the second song was presented in the 1st row, 2nd column, the 3rd song was presented in the 1st row, 3rd column, the 4th song was presented in the 2nd row, 1st column, etc. For example, consider the first line of `screenorder_w1_v1.txt`:
  4, 995601, 2814479, 765447, 4123311, 994601, 311089, 2814575, 3123313, 806122, 225647, 946646, 8814579, 100102, 911789, 744601, 501405, 123111, 592645, 792647, 3124513, 641126, 331122, 395650, 306121, 781123, 411241, 595655, 777561, 7014523, 165411, 144102, 165444, 911249, 440341, 523645, 7814573, 846626, 995651, 744101, 806126, 241124, 881121, 865431, 646341, 5014503, 131405, 326122, 4004513
  This says that the participant with user id 4 saw song 995601 ("Trapped in an Orange Peel" by *The Calefaction*) in the upper left corner of the grid, song 2814479 ("Went with the Count" by *Simply Waiting*) in the 1st row/2nd column, etc.

- `dynamics_downloads_download_w*_v*.txt` [25 files (9 from experiment 1, 9 from experiment 2, 3 from experiment 3, 4 from experiment 4)]: These files record the time at which all downloads happened in each world in each experiment. The first column records the user id, the second column records the world (this should be the same for everyone in each file), the next 48 columns record whether the song was downloaded or not and the final column is a timestamp. For example, consider the first 6 records in `dynamics_downloads_download_w1_v1.txt`:
  4, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2004-10-07 13:37:24
  9, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2004-10-07 13:59:00
  62, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2004-10-07 18:51:06
  62, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2004-10-07 18:53:30
  62, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2004-10-07 18:56:27
  62, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2004-10-07 19:21:48
  This shows that the participant with user id 4 was in world 1 and downloaded the 29th song (when the songs are sorted in lexicographic order by song id). By referring to table 7 we can see that this is "Out of the Woods" by *Shipwreck Union*. The download was initiated on October 7, 2004 at 1:37pm and 24 seconds (all times are New York City times). The next line records the download of the participant with user id 9. The next 4 lines record the 4 songs that were downloaded by the participant with user

[2]Note that for an extremely small number of participants in experiments 2, 3, and 4 these screen locations may have actually changed slightly during the course of the visit if the song changed popularity while a user was participating. We do not think this is a cause for concern, but one could calculate this for sure using the files `dynamics_downloads_download_w*_v*.txt`.

id 62 and the times that these downloads occurred. Note that this file structure is somewhat inefficient because it stores lots of unnecessary 0's, but it was an easy data structure to work with for analysis.

- `autoplay_w*_v*.txt` [25 files (9 from experiment 1, 9 from experiment 2, 3 from experiment 3, 4 from experiment 4)]: These files record the time at which all listens happened in each world and experiment. The first column is the user id, the second column is the world (should be the same for everyone in each file), the third column is the song id, the fourth column is the band number when the songs are sorted in lexicographic order (ie the first column in table 7), and the fifth column is the timestamp. For example, consider the first record in `autoplay_w1_v1.txt`:
  4, 1, 2814575, 10, 2004-10-07 13:29:32
  This shows that a participant with user id 4 was in world 1 and listened to song 2814575 ("Father to Son" by *Beerbong*) at 1:29pm and 32 seconds (all times are New York City times).

- `subject_recruitment_v*.txt` [4 files (1 from each experiment)]: These files record the number of participants that were recruited on each day. The first column records the date, the second column records the day of the week (1 = Monday, 2 = Tuesday, etc.), the third column is the number of new participants. For example, consider the first record in `subject_recruitment_v1.txt`:
  2004-10-07, 4, 34
  This shows that the experiment began on Thursday, October 10, 2004 and 34 people participated. Note that because of a database error during experiment 1, we don't know the date of registration for participants who registered after November 10, 2004. This error was corrected in subsequent experiments and did not affect any other data.[3]

- `user_demographics_v*.txt` [4 files (1 from each experiment)]: These files record the information that each participant reported during the registration survey. The first column records the user id, the second column is the world to which they were assigned, the third column is a timestamp. For the meaning of the rest of the columns, see appendix A.
  1,2,2004-10-07 12:50:40,2,United Kingdom,00000,59,1,3,0,0,1,0,0,0,0,0,0,1,2,3,2,1,0,1,1,1,3,1,1
  Note that the registration dates for all users in experiment 1 after November 10, 2004 are incorrect (see note about `subject_recruitment_v*_w*.txt`).

- `downloads_ic4_lexorder.txt` [1 file]: This file records the initial conditions for the four worlds in experiment 4. The first column is the song id, the second column is the initial download count in world 1 (the unchanged world), the third column is the initial download count in world 2 (an inverted world), the fourth column is the initial download count in world 3 (another inverted world), and the fifth column is the initial download count in the independent world. For example, consider the first two records:
  100102, 13, 13, 13, 0
  123111, 10, 31, 31, 0
  This shows that song 100102 ("This Upcoming Winter" by *Ember Sky*) initially had 13 downloads in the unchanged world and 13 in the inverted worlds. Song 123111 ("It does what its Told" by *Go Mordecai*) had 10 downloads in the unchanged world and 31 in the inverted worlds.

3 Recruiting the participants

The majority of our 27,267 participants who came from two sources: `www.bolt.com` (for experiments 1 and 2) and emails to participants of the electronic small-world experiment (experiments 3 and 4). In addition to these two sources, web-posting generated additional traffic. Demographics about these subjects are presented in table 6.

Experiment 1 took place from October 7, 2004 to December 15, 2004 (69 days) and involved 7,149 subjects. Recruitment dynamics from this experiment are presented in figure 13. The largest spike in traffic during

[3]The source of this error was confusion about the `timestamp` field in mySQL. We did not know that this field automatically updated every time a record was changed. After we discovered this problem we created two fields `timestamp` and `creation_timestamp`.

	Experiment 1		
	Influence 8 worlds ($n = 5,708$)	Independent 1 world ($n = 1,441$)	Total 9 worlds ($n = 7,149$)
Number of listens	21,971	5,394	27,365
Mean per subject	3.8	3.7	3.8
Number of downloads	6,625	1,578	8,203
Mean per subject	1.2	1.1	1.1
Pr[download | listen]	0.302	0.293	0.300
Average rating (# of stars)	3.0	2.9	3.0

Table 2: Descriptive statistics of subject behavior in experiment 1.

	Experiment 2		
	Influence 8 worlds ($n = 5,746$)	Independent 1 world ($n = 1,446$)	Total 9 worlds ($n = 7,192$)
Number of listens	20,217	5,643	25,860
Mean per subject	3.5	3.9	3.6
Number of downloads	8,106	2,192	10,298
Mean per subject	1.4	1.5	1.4
Pr[download | listen]	0.401	0.388	0.398
Average rating (# of stars)	3.2	3.2	3.2

Table 3: Descriptive statistics of subject behavior in experiment 2.

	Experiment 3		
	Social influence 2 worlds ($n = 1,471$)	Independent 1 world ($n = 1,459$)	Total 3 worlds ($n = 2,930$)
Number of listens	10,591	11,844	22,435
Mean per participant	7.2	8.1	7.7
Number of downloads	2,040	1,691	3,731
Mean per participant	1.4	1.2	1.3
Pr[download | listen]	0.193	0.143	0.166
Mean rating (# of stars)	2.70	2.55	2.62

Table 4: Descriptive statistics of subject behavior in experiment 3.

	Experiment 4				
	Unchanged ($n = 2,015$)	Inverted, # 1 ($n = 2,014$)	Inverted, # 2 ($n = 1,970$)	Independent ($n = 3,997$)	Total ($n = 9,996$)
Number of listens	14,430	12,498	12,633	30,142	69,703
Mean per subject	7.2	6.2	6.4	7.5	7.0
Number of downloads	2,898	2,197	2,160	5,089	12,344
Mean per subject	1.4	1.1	1.1	1.3	1.2
Pr[download | listen]	0.201	0.176	0.171	0.169	0.177
Average rating (# of stars)	2.71	2.64	2.60	2.63	2.64

Table 5: Descriptive statistics of subject behavior in experiment 4. These download counts do not include the initial conditions.

Figure 12: Banner advertisement used to recruit subjects from `http://www.bolt.com` for experiment 2.

experiment 1 occurred after the experiment was mentioned on the popular blog `www.kottke.org` (October 19, 2004). Other spikes in traffic were largely driven by the prominence that we were given on `www.bolt.com`.

Immediately after completing experiment 1, we began experiment 2 which ran from December 15, 2004 to March 8, 2005 (83 days) and involved 7,192 subjects. Figure 14 shows the recruitment dynamics. As far as we know, spikes in traffic during experiment 2 were largely driven by the prominence that we were given on `www.bolt.com`.

In table 6 we note that there was a change in percentage of females from experiment 1 to experiment 2. Subjects in both experiments were drawn from `www.bolt.com`, but they were drawn from different parts of the website. A majority of the subjects in experiment 1 were likely drawn from the "music" and "free-stuff" sections while a majority of the subjects in experiment 2 were likely drawn from a special email sent to a set of Bolt users and from banner ads in all sections of the site (for example, figure 12). Another potential reason for the difference is that while experiment 1 was underway, the project was mentioned on the popular blog `www.kottke.org` which probably has an older, more male readership. Ideally these differences in recruitment between experiments would not have occurred, but we do not believe that they had a substantial effect on our findings.

Experiment 3 took place from March 14, 2005 to April 7, 2005 (24 days) during which time we sent 13,546 emails to participants in the electronic small-world experiment; see Dodds, Muhamad, and Watts (2003). Recruitment dynamics are presented in figure 15, and the large spike in traffic was caused by a mention of the experiment on the popular website `www.boingboing.net` (April 5, 2005).

Immediately after completing experiment 3, we began experiment 4 which ran from April 7, 2005 to August 10, 2005 (126 days) during which time we sent emails to all remaining participants of the electronic small-world experiment who had not been contacted during experiment 3 ($n = 50,800$). The large spike in traffic at the beginning of the experiment was because we sent out a very large number of emails very quickly.[4] The source of the spike around day 60 is unknown.[5] Experiment 4 ended on August 10, 2005 so that the results could be presented at the American Sociological Association Annual Meeting. Once the results were presented, the manipulation was no longer secret and we could not be confident that future data would continue to be clean. This may seem to be a somewhat artificial endpoint, but by that time recruitment had slowed to a trickle with only about 10 new participants per day.

As with experiments 1 and 2, there were some differences in the demographics between experiments 3 and 4 (table 6). For example, there was a large increase in the number of Brazilians which was caused by a mention of the experiment on the popular Brazilian website `www.estadio.com.br`. However, other than this difference, the demographics across the experiments were similar, and we don't think this large increase in Brazilians affected our results.

4 The songs

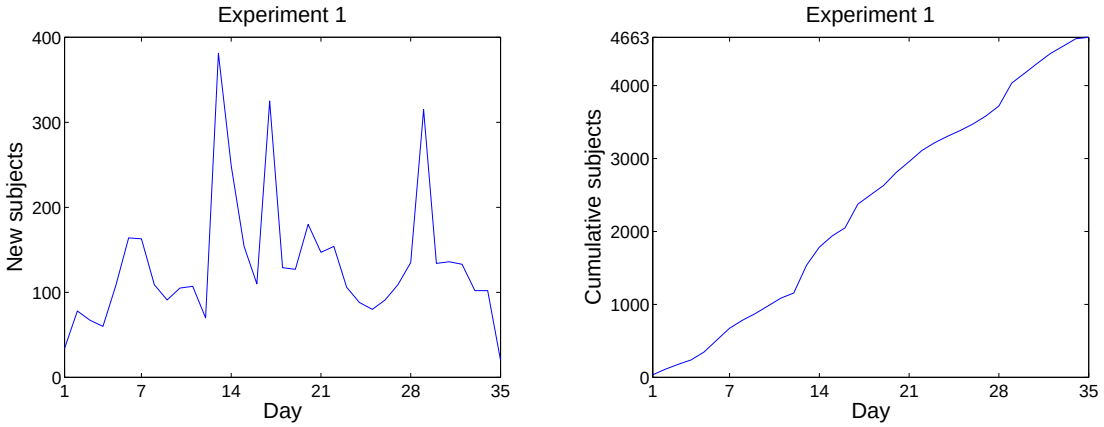
The music for the experiment (see table 7) comes from `www.purevolume.com`, a website where bands can create homepages and post their music for download. In July 2003 there were approximately 42,000 bands with homepages. Preliminary research revealed that many of the song recordings had extremely poor audio quality. Therefore, we restricted our sample to the approximately 1,000 premium member bands—those

[4]We sent these emails so quickly because at that time Peter Hausel, the programmer of the site, told us that he was moving to a new job soon. Therefore, we wanted to finish the experiment as quickly as possible. The heaviest emailing was from April 12th to the 20th.

[5]Information to help us locate the source of this spike might have been available in our web-server logs, but these have been lost. In the future, these server-logs should be treated as data and therefore achieved.

Category	www.bolt.com		Small-world experiment	
	Experiment 1 ($n = 7,149$) (% of participants)	Experiment 2 ($n = 7,192$) (% of participants)	Experiment 3 ($n = 2,930$) (% of participants)	Experiment 4 ($n = 9,996$) (% of participants)
Female	36.4	73.9	38.0	43.9
Broadband connection	74.1	69.0	90.6	89.4
Has downloaded music from other sites	60.4	62.4	69.3	65.3
Country of Residence				
United States	79.8	81.8	68.4	54.7
Brazil	0.3	0.0	1.2	12.5
Canada	4.5	4.4	6.3	4.9
United Kingdom	4.4	4.7	6.6	6.9
Other	11.0	9.1	17.5	21.0
Age				
14 and younger	11.5	16.0	1.5	2.3
15 to 17	27.8	34.9	5.7	5.6
18 to 24	38.5	39.2	29.8	26.6
25 and older	22.3	9.9	63.1	65.6

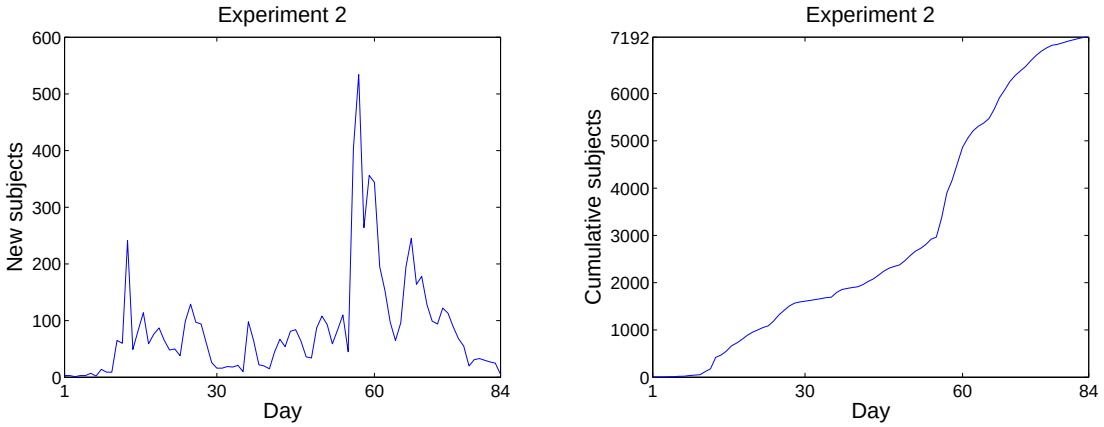
Table 6: Descriptive statistics about the participants in the four experiments. Most participants from experiments 1 and 2 were recruited from www.bolt.com. Most participants from experiments 3 and 4 were recruited by emails to participants in the electronic small-world experiment and the subsequent web postings these emails generated. Participants in experiments 3 and 4 were older and more international.



(a) Subjects per day

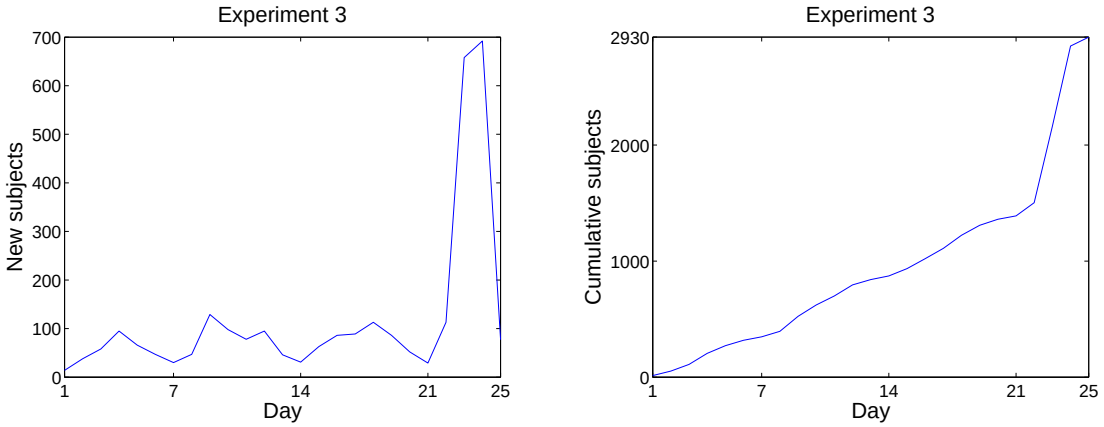
(b) Cumulative subjects

Figure 13: Recruitment dynamics for experiment 1 (October 7, 2004 to December 15, 2004). The largest spike in traffic during version 1 occurred after the experiment was mentioned on the popular blog www.kottke.org (October 19, 2004). Other spikes in traffic were largely driven by the prominence that we were given on www.bolt.com. We do not have data on when subjects registered after November 10, 2004 because of a database error; hence, the cumulative total in figure (b) is less than the total number of subjects in experiment 1.



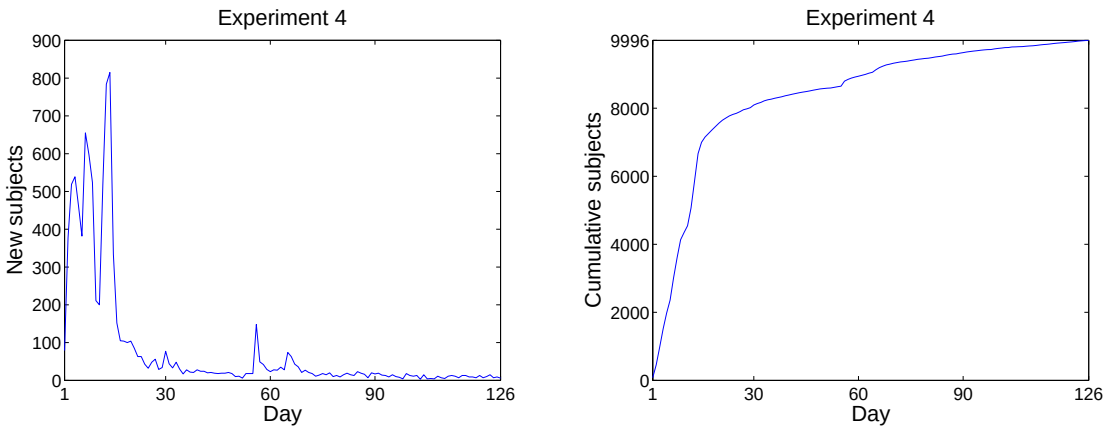
(a) Subjects per day (b) Cumulative subjects

Figure 14: Recruitment dynamics for experiment 2 (December 15, 2004 to March 8, 2005). Spikes in traffic were largely driven by the prominence that we were given on `www.bolt.com`.



(a) Subjects per day (b) Cumulative subjects

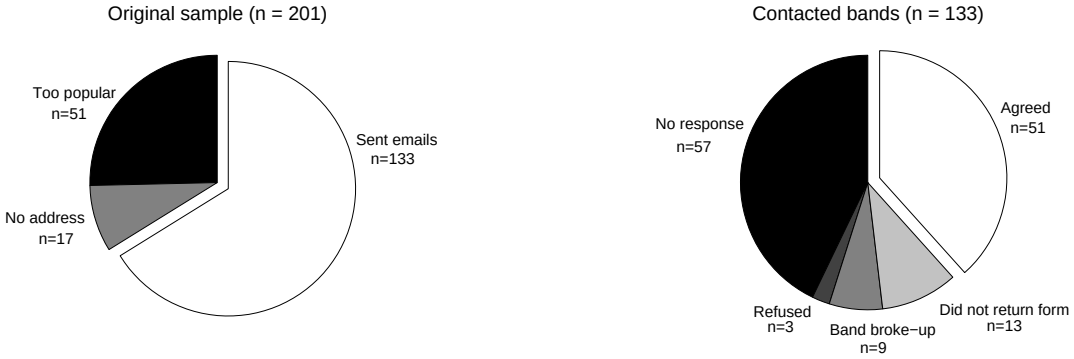
Figure 15: Recruitment dynamics for experiment 3 (March 14, 2005 to April 7, 2005). The periodicity in this graph is because we did not send recruitment emails on the weekend and these recruitment emails where the main source of traffic. The large spike in traffic was probably caused by a mention on the popular website `www.boingboing.net` (April 5, 2005).



(a) Subjects per day

(b) Cumulative subjects

Figure 16: Recruitment dynamics for experiment 4 (April 7, 2005 to August 10, 2005). The large spike in traffic at the beginning of the experiment was because we sent out a very large number of emails very quickly.



(a) Original sample (n = 201)

(b) Contacted bands (n = 133)

Figure 17: Pie charts showing various aspects of attrition for the sample of bands selected from the music website `www.purevolume.com`. Approximately, 40% of the contacted bands agreed to be in the study.

who paid approximately $10 per month for additional features on their homepages—whose audio quality was generally better.

Initially, about 200 bands were selected for consideration. Because the experiments required bands that were unknown to the participants, we screened out any band that had played in more than 10 states, or had played more than 15 concerts in the past 30 days, or had appeared on the Warped Tour, or had 30,000 or more hits on their purevolume page. These screening criteria are ultimately arbitrary, but they are reasonable. We have no reason to believe that the results would be any different if other reasonable criteria were used. In all, these criteria removed 51 bands. In addition, 17 bands could not be contacted because they did not have a publicly available email address. The remaining 133 bands were contacted via email (results summarized in figure 17(a)). In order to minimize non-response bias, all non-responding bands received two follow-up emails spaced at one week intervals. In the end, 51 of these bands agreed to be in the study and provided us with a song of their choice, the other bands becoming ineligible for a variety of reasons (results summarized in figure 17(b)). The email to the bands and band consent form are available in the dissertation.

Preliminary pilot testing revealed that, for the song menu used in experiment 1 (figure 7), the maximum number of songs that could be legibly presented on a typical computer screen was 48. Thus, we took a sample of 48 of the 51 bands to be in the experiments. In order to check that our initial screening criteria filtered out music that might be known to the participants, we presented the list of bands and songs to

Order	Song ID	Band name	Song name
1	100102	ember sky	this upcoming winter
2	123111	go mordecai	it does what its told
3	131405	post break tragedy	florence
4	144102	deep enough to die	for the sky
5	165411	hall of fame	best mistakes
6	165444	up for nothing	in sight of
7	225647	stunt monkey	inside out
8	241124	miss october	pink aggression
9	2814479	simply waiting	went with the count
10	2814575	beerbong	father to son
11	306121	silverfox	gnaw
12	311089	the fastlane	til death do us part (i don't)
13	3123313	forthfading	fear
14	3124513	silent film	all i have to say
15	326122	this new dawn	the belief above the answer
16	331122	by november	if i could take you
17	395650	hartsfield	enough is enough
18	4004513	ryan essmaker	detour_(be still)
19	411241	undo	while the world passes
20	4123311	summerswasted	a plan behind destruction
21	440341	cape renewal	baseball warlock v1
22	501405	up falls down	a brighter burning star
23	5014503	nooner at nine	walk away
24	523645	dante	life's mystery
25	592645	parker theory	she said
26	595655	the thrift syndicate	2003 a tragedy
27	641126	hydraulic sandwich	separation anxiety
28	646341	salute the dawn	i am error
29	7014523	shipwreck union	out of the woods
30	744101	benefit of a doubt	run away
31	744601	52metro	lockdown
32	765447	fading through	wish me luck
33	777561	moral hazard	waste of my life
34	781123	unknown citizens	falling over
35	7814573	sibrian	eye patch
36	792647	art of kanly	seductive intro, melodic breakdown
37	806122	far from known	route 9
38	806126	drawn in the sky	tap the ride
39	846626	star climber	tell me
40	865431	selsius	stars of the city
41	881121	a blinding silence	miseries and miracles
42	8814579	the broken promise	the end in friend
43	911249	not for scholars	as seasons change
44	911789	stranger	one drop
45	946646	sum rana	the bolshevik boogie
46	994601	evan gold	robert downey jr.
47	995601	the calefaction	trapped in an orange peel
48	995651	secretary	keep your eyes on the ballistics

Table 7: The 48 bands and songs used in the experiments sorted in lexicographical order by song id.

		How familiar are you with the following bands?		
Band type	Name	Don't know it at all (% of participants)	Heard of it (% of participants)	Know it pretty well (% of participants)
Real	Guys on Couch	91.0	8.1	1.0
Real	Grover Dill	91.2	7.8	0.9
Fake	Peter on Fire	88.1	10.5	1.4
Real	U2	4.6	24.6	70.9
Real	Remnant Soldier	83.2	14.7	2.1

Table 8: Comparing the popularity of the potential bands from our sample to a fake band. Participants reported being about as familiar with an fake band (Peter on Fire) as three potential bands from our sample. The higher recognition rate for Remnant Soldier is likely a question ordering effect—it was asked immediately after the well known band U2. Totals may not sum to 100 because of rounding.

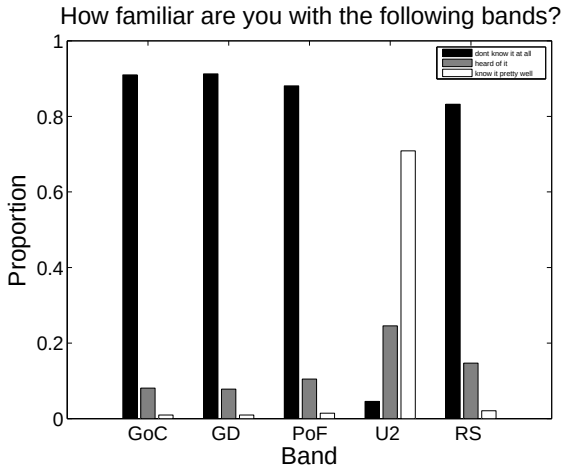


Figure 18: Comparing the popularity of the potential bands from our sample to a fake band. Participants reported being about as familiar with an fake band (Peter on Fire [PoF]) as three potential bands from our sample: Guys on Couch [GoC], Grover Dill [GD], and Remnant Soldier [RS]. The higher recognition rate for Remnant Soldier is likely a question ordering effect—it was asked immediately after the well known band U2.

two different experts in popular music: a DJ at the Barnard College student radio station and the music editor for `www.bolt.com`. Neither expert recognized any of the bands or songs. As an additional test, on our registration survey we asked subjects about their familiarity with five bands: the three potential bands who agreed to participate, but were ultimately not included (Guys on Couch, Grover Dill, and Remnant Solder), an imaginary band (Peter on Fire), and an extremely well known band (U2).[6] Table 8 and figure 18 show that some subjects reported being familiar with the three potential bands, but these recognition rates were no higher than for the imaginary band.[7] Further, the extremely different results observed for the band U2 suggest that respondents were actually reading the question and not simply reporting "don't know it at all" for all bands. These survey results, together with our screening and queries to two experts, lead us to believe that the music used in the experiment was essentially unknown. Also, while the experiment was in progress, we monitored the success of the bands and found nothing indicating any significant changes.

[6]We chose to ask only about bands that were ultimately not included because having the same bands in the survey and experiment might have biased subjects' music preferences, as is suggested by work on the recognition heuristic.

[7]The slightly higher recognition rate for the band Remnant Soldier is probably a question ordering effect; this question was asked immediately after a question about familiarity with the very popular band U2. In future studies we recommend randomization of question ordering to avoid this problem.

5 Data quality protocols

The data presented in these files are a subset of all of the data that we collected; all data that have been potentially contaminated have been excluded. More specifically, in all experiments researchers must take steps to ensure that data are generated by the appropriate set of participants in situations that match the experimental design, and that the participants have no malicious intent. These problems can be more difficult to deal with in web-based experiments where researchers have less control over participant recruitment and behavior than they would have in a standard laboratory-based experiment. Because of this limited control, some of the data from our experiments are possibly unsound. Instead of preventing this unsound data generation, and hence giving participants incentive to provide us with false information, we allowed all participants to participate in all situations, but flagged data that could have been unsound and excluded them from our analysis and these files.

For example, our experimental design required that a participant's information about the behavior of others be limited to what we provided them (or did not provide them). Information contamination leading to unsound data could have occurred a number of ways: 1) between two participants from two different social influence worlds 2) between two participants from the independent condition and 3) between a participant in the independent condition and a participant in a social influence world. Unlike in a laboratory-based experiment, we were not able to physically isolate the participants to prevent this information contamination. As such, we flagged for exclusion data generated in several cases where the participant behavior could have possibly been influenced by information that was outside of the experimental design.

The first step in this data-flagging process was based on a survey that all participants completed. On this survey participants were asked to select, from a list of choices, all of the ways that they heard about the experiment. If a participant reported "friend told me about a specific song" or "friend told me about a specific band" all data generated by that participant were flagged. However, data generated by participants who reported "friend told me about the experiment in general" were not flagged. We also flagged all data generated after either the participant clicked "log-off" or 2 hours had passed since the participant registered. These data were flagged in order to exclude data where the participant could have participated, discussed the music with friends, and then returned with outside information. Our flagging criteria were quite strict and so we probably flagged data which was not contaminated. However, we cannot rule out the possibility that some contaminated data was not flagged.

In addition, to prevent information contamination within and between experiments, we placed several cookies—small pieces of information—into the participant's web browser. These cookies ensured that if a participant returned to the experiment, the participant would be placed in the same condition and same world without having to re-complete the registration process. The cookies also prevented participants who returned to the site after their experiment was completed from participating in future experiments.

When doing a web-based experiment, or any other experiment, one has to take a number of steps to guard against the possibility of malicious participants who intend to disrupt the experiment. This problem, while not limited to web-based experiments, is perhaps a larger issue in this set of experiments than in most. For example, members of one of the bands might have tried to artificially inflate the download count of their song. To prevent this possibility, each participant was allowed to download a specific song as many times as they liked, but could only add one to the displayed download count for that song. Members of the bands might have also tried to manipulate the results by sending their fans to the experiment. As such, we flagged all data generated by people who reported on our survey that they heard about the experiment from "one of the bands." We also checked our web-server log to ensure that we were not receiving participants from the websites of any of the bands. In two cases, links to the experiment was posted on bands' websites, but these links were detected quickly and both bands complied with our email request to remove the link.

An additional class of malicious participants could have simply wished to disrupt the experiment for no specific reason. To prevent against these participants, the experiment was run appropriate security precautions using the latest software at the time (Apache 2.0, MySQL 4.0, and Tomcat 5.0) with strict firewall settings.

Despite all of our security precautions, it was still possible for a participant to manipulate our results. For example, there is no way that we could prevent the same person from registering from several different computers and providing us with false information each time. However, given that participants have little incentive to undertake this behavior, we think that this probably did not occur. Taken together, our data-

quality measures give us confidence that our data are reasonably clean. Of course we cannot rule out all possible problems, but we have not seen any patterns in the data that indicate data contamination or malicious manipulation occurred.

A Survey

This is documentation for the demographics file.
Field 1: User ID
Field 2: World ID

- Range 1-9 for experiments 1 and 2, 1-3 for experiment 3, and 1-4 for experiment 4
- Highest group ID in given experiment is the independent condition (ie 9 in experiments 1 and 2, 3 in experiment 3, and 4 in experiment 4.

Field 3: Creation timestamp
Field 4: In which country do you currently live? (numerical code)

- 1 = United States
- 2 = United Kingdom
- 3 = Canada
- 4 = Brazil
- 5 = Other

Field 5: In which country do you currently live? (string)
Field 6: If you live in the United States, please enter your zip code.

- 00000 = missing value (non-US residents)

Field 7: In what year were you born? (converted to age)
Field 8: What is your gender? (numeric code)

- 0 = female
- 1 = male

Field 9: Compared to your circle of friends, how likely are you to be asked for advice about music? (Opinion leader score)

- 1 = much less likely
- 2 = less likely
- 3 = more likely
- 4 = much more likely

How did you hear about the experiment?
Field 10: web site or blog (1 = yes, 0 = no)
Field 11: internet ad (1 = yes, 0 = no)
Field 12: email from musiclab (1 = yes, 0 = no)
Field 13: one of the bands (1 = yes, 0 = no)
Field 14: friend told me about a specific song (1 = yes, 0 = no)
Field 15: friend told me about a specific band (1 = yes, 0 = no)
Field 16: friend told me about the site in general (1 = yes, 0 = no)
Field 17: search engine (1 = yes, 0 = no)
Field 18: other (1 = yes, 0 = no)
Field 19: What type of Internet connection are you currently using? (numerical code)

- 1 = broadband
- 2 = dial-up

Field 20: Where are you while you are participating in this experiment? (numerical code)

- 1 = home
- 2 = office
- 3 = public computer at school
- 4 = other

Field 21: How would you rate your ability to use the World Wide Web? (numerical code)

- 1 = excellent
- 2 = good
- 3 = fair
- 4 = poor
- 5 = very poor

Field 22: In the past 30 days, how often have you visited a site for information about music or concerts? (numerical code)

- 0 = never
- 1 = 1 - 2 times
- 2 = 3 - 5 times
- 3 = more than 5 times

Field 23: Have you ever purchased a record as a result of hearing it on the web? (numerical code)

- 0 = no
- 1 = yes

Field 24: Approximately, how many songs have you downloaded in the past 30 days?

- self-reported value

How familiar are you with the following bands?
Field 25: Guys on Couch (real) (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
Field 26: Grover Dill (real) (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
Field 27: Peter on Fire (fake) (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
Field 28: U2 (real) (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
Field 29: Remnant Soldier (real) (1 = don't know it at all, 2 = heard of it, 3 = know it pretty well)
Field 30: Please provide your email address so that we can tell you about the results of Music Lab (optional)

- 0 = no email address given
- 1 = email address given

B Miscellaneous notes

As with all data collected in the real-world, there are some "quirky" features of these data. Here are some things that you might find and possible explanations.

- There are a very small number of users who appear to have done things that are not possible. For example, user 608 rated song 131405 without appearing to listen to it. This rating without listening happens 36 times in experiment 1 which is about 0.01% of the overall number of listens. There are three likely sources for these unusual sequences. First, the back-button is treated differently by different browsers and potentially could cause unusual patterns. Second, extremely outdated versions of Flash running on outdated operating systems were found to occasionally cause problems, but these situations were so rare it was hard to precisely understand the nature of the problems. Third, a user who directly manipulates the url attempting to understand our site had the potential to create strange paths through the site. Again, these strange paths are very rare, but if you find them, this might be why.
- The song ids are sorted in lexicographic order, because we initially treated the song ids as strings. If you read the song ids as numbers and sort them they will be in the wrong order.
- Not all user ids are used. For example, in experiment 1, the released files do not include user 2, 7, 10 etc. This is because whenever we the experimenters registered at the site for testing or demonstration purposes, we were assigned a user id. All the actions by the experimenters were not included in the analysis and were removed from these files. Also, the user ids in experiment 2 start at 51. This is because we forgot to reset the database counter after a short pilot test. The first 50 records are not missing.